Alexander Soifer

The New Mathematical Coloring Book

Mathematics of Coloring and the Colorful Life of Its Creators

Second Edition

Springer

The New Mathematical Coloring Book


$W(k) < 2^{2^k}$?
$S(n) = n!e$
$m_0 = R(r,n,k)$
$W = W(k,l)$
$S_n' = \lambda S_n + a$

Alexander Soifer

The New Mathematical Coloring Book

Mathematics of Coloring and the Colorful Life of Its Creators

Second Edition

Forewords by Peter D. Johnson Jr., Geoffrey Exoo, Branko Grünbaum, and Cecil Rousseau

Springer

Alexander Soifer
College of Letters, Arts, and Sciences
University of Colorado at Colorado Springs
Colorado Springs, CO, USA

Foreword by
Peter D. Johnson
Department Mathematics and Statistics
Auburn University College of Sciences and Mathematics
Auburn, AL, USA

Foreword by
Geoffrey Exoo
Department Computer Science
Indiana State University
Terre Haute, IN, USA

Foreword by
Branko Grünbaum
University of Washington Department Mathematics
Seattle, WA, USA

Foreword by
Cecil Rousseau
University of Memphis Department Mathematical Sciences
Memphis, TN, USA

ISBN 978-1-0716-3596-4 ISBN 978-1-0716-3597-1 (eBook)
https://doi.org/10.1007/978-1-0716-3597-1

Mathematics Subject Classification: 01-Axx; 03-XX; 05-XX

© Alexander Soifer 2009, 2024
Originally published with the title: The Mathematical Coloring Book
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

The Frontis: An Artistic Summary of *The Mathematical Coloring* Book by Leto Quarles, ink on paper, 2009. Depicted from the left are Bartel Leendert Van der Waerden, Frank Plumpton Ramsey, Tibor Gallai, Saharon Shelah, Issai Schur, Paul Erdos, with Alexander Soifer depicted above them.

This Springer imprint is published by the registered company Springer Science+Business Media, LLC, part of Springer Nature.
The registered company address is: 1 New York Plaza, New York, NY 10004, U.S.A.

Paper in this product is recyclable.

This coloring book is for my late father Yuri Soifer,
a great painter, who introduced colors into my life.

To Paint the Portrait of a Bird

by Jacques Prévert[1]

First paint a cage
With wide open door,
Then paint something
Beautiful and simple,
Something very pleasant
And much needed
For the bird;
Then lean the canvas on a tree
In a garden or an orchard or a forest –
And hide behind the tree,
Do not talk
Do not move. . .
Sometimes the bird comes quickly
But sometimes she needs years to decide
Do not give up,
Wait,
Wait, if need be, for years,
The length of waiting –
Be it short or long –
Does not carry any significance
For the success of your painting
When the bird comes –
If only she ever comes –
Keep deep silence,
Wait,
So that the bird flies in the cage,

---

[1] [Pre]. Prévert dedicated this poem to Elsa Henriquez. Translation by Alexander Soifer and Maurice Stark.

And when she is in the cage,
Quietly lock the door with the brush,
And without touching a single feather
Carefully wipe out the cage.
Then paint a tree,
And choose the best branch for the bird
Paint green leaves
Freshness of the wind and dust of the sun,
Paint the sounds of animals of the grass
In the heat of summer
And wait for the bird to sing
If the bird does not sing –
This is a bad omen
It means that your picture is of no use,
But if she sings –
This is a good sign,
Sign that you can be
Proud and sign,
So you very gently
Pull out one of the feathers of the bird
And you write your name
In a corner of the picture.

Foreword to the Second Edition

On reading my preface to the first *Mathematical Coloring Book*, I was pleasantly surprised by what my younger self had written, and very much in agreement with it. A new thing in the world at the time, *TMCB I* is now joined by a colossal sibling containing more than twice as much of what only Alexander Soifer can deliver: an interweaving of mathematics with history and biography, well-seasoned with controversy and opinion.

Let me expand somewhat on a feature of A. Soifer's methods of investigation that I see mentioned briefly in my earlier preface, and in the preface of Branko Grünbaum, but which evidently has played a much larger role in the construction of *TMCB II* than in the earlier work. To a great extent, Soifer bases his accounts of mathematical and human events post-1920 on conversations and correspondence with the mathematicians involved and, quite often, with their family, friends, and other contemporaries. He is ferociously meticulous, and entertainingly concerned with whom should get credit for what. This feature of his method, direct communication with those who were there, or who know somebody that was there, is made explicit in the sections of *TMCB II* on B. L. van der Waerden and various other mathematicians active in the 30-year interval around World War II, for he reports conversations and exhibits copies of actual letters. He also traces his path to the discovery of the true origin of the chromatic-number-of-the-plane problem, which surely would never have been found had he not undertaken the search. I mention this trait of journalistic investigation in Soifer's character in hopes that future writers of mathematics-plus-history might be encouraged to emulate him.

But this brings me to a slight disagreement with what I wrote in my preface to *TMCB I*. There it is intimated that the singularity of the book might be entirely attributed to the special mix of talents and purposes of its author, and that once the singularity had exploded, the writing of mathematical history would be greatly affected. The insinuation was that Alexander Soifer might be to the writing of mathematical history as Julius Erving (Dr. J!) or Pete Maravich were to professional basketball; those two changed how the game was played by introducing tricks and moves never before seen, that then were ever after added to the repertoires of promising young players.

I still believe that *TMCB I* would not and could not have been written by anyone on earth but A. Soifer, but now I see something else to which the book owes its character: the peculiarity of its subject.

My view is that progress in mathematics is an evolutionary process that sorts itself into "strands". There are moments of creation with no antecedents, and from these little "big bangs" come the strands. This process is not very evident to us moderns because the moments of creation from which descend the strands comprising what we regard as the main "branches" of mathematics occurred either in prehistory (counting) or in ancient historical times—for instance, the invention of arithmetic, perhaps a byproduct of the invention of money.

Of course, creation does occur within existing strands (think of topology, or modern logic), but in these instances we can see antecedents; the new developments do not come out of nowhere.

The subject of *TMCB I* and *II* is one of two strands descending from a small big bang, a moment of creation, that did come out of nowhere: The Four-Colour Conjecture, due to Francis Guthrie, in London, in 1852, formulated in conversation with his brother Frederick. (Apparently Francis did not express the conjecture as a conjecture—not at first, anyway.) Two major mathematical strands descended from this eruption: (1) graph theory, and (2) the idea of *coloring* objects in a mathematical structure so that certain requirements are satisfied. Of course, there are strong bonds between the two strands. The objects being colored are often the vertices, or the edges, or both, of a graph or hypergraph. But this is not always the case, so the chromatic strand is not contained in graph theory.

I conjecture that it is the novelty of the chromatic strand and our proximity in time to that moment of creation, in 1852, that made *TMCB I* and *II* possible (although not inevitable; thanks again, Professor Soifer!), and that we are not likely to soon see another such event.

Still, there are plenty of mathematical strands coming from little big bangs that do have recognizable antecedents; perhaps *TMCB* will inspire someone in the future to attempt to give one of these strands the Soifer treatment. For example, let's take *representation of functions by infinite series.* Here's my outline: I don't think such representations existed before the 16th century—I could be wrong, but they certainly didn't exist in the 13th century. Then, suddenly, we had power series—probably familiarity with polynomials gave somebody the idea. Then, around 1816, J.-B. Joseph Fourier presented his idea for trigonometric series to the French Academy, leading to an explosion of activity across the 19th century and into the 20th. After Schauder bases in Banach spaces and series of wavelets in the mid-to-late 20th century, interest in research on series representations seems to have died out—the great developments of the 19th and 20th centuries have been carved in stone and consigned to the repository of the Known, like calculus, or trigonometry.

But let's have some details! When and where did power series first appear? Who had the idea? What was their story? Tell us about Fourier! After that, who did what? Won't somebody please give the subject the Soifer treatment?

Because we have made no progress in communicating directly with the dead, giving this particular subject a full Soifer treatment is impossible. But because *The Mathematical Coloring Books* exist, we have a model for trying. How would the Soifer treatment differ from the usual work of historians of mathematics? It's a matter of sticking to a strand of mathematics, reporting the disputes and scuffles of the mathematicians involved, and venturing judgments on the relative importance of various developments within the strand. This will involve modest changes in the kind of knowledge the historians aim to collect, and in the narrative structure they will aim to impose on what they find. It may take a while.

Auburn University, Auburn, AL, USA Peter D. Johnson Jr.
June 26, 2023

Foreword to the Second Edition

I first became aware of the first *Mathematical Coloring Book* (TMCB I) in March 2009 during the 40th Southeastern International Conference on Combinatorics, Graph Theory and Computing in Boca Raton. The publisher of the book Springer brought a couple of boxes of its new book to the conference and set up a display on a table in one of the meeting areas. Attendees were crowded around the table, clamoring for copies. I'd never seen a group of Mathematicians quite this excited about a book. I really wasn't sure what to make of it, and the copies were gone by the time I got to front of the melee. Six years passed before I had a chance to learn what all the excitement was about. I am privileged to be one of the first to read *The Second Mathematical Coloring Book* (TMCB II).

Like TMCB I, TMCB II is a unique combination of Mathematics, History, and Biography written by a skilled journalist who has been intimately involved with the story for the last half-century. As Mathematics, TMCB II deals with Ramsey Theory and Combinatorial Geometry, with a few other topics sprinkled about. The author does not cut corners when presenting the Mathematics. The nature of the subject makes much of the material accessible to students, but also of interest to working Mathematicians. Many old and new proofs of fundamental theorems have been rewritten and clarified by the author. In addition to learning some wonderful Mathematics, students will learn to appreciate the influences of Paul Erdős, Ron Graham, and others.

As History, TMCB II traces the development of the subject from the earliest results of Hilbert, Van der Waerden, Schur and Ramsey, through the tragedies of the Second World War, and thence through the career of Paul Erdős. It is a meticulously researched history which reads like a good novel. The adventure of the chase, both for Mathematical results and historical accuracy, has been captured. Professor Soifer was personally involved in much of the development of the subject over the past half century. During this time, he was in communication, one way or another, with most of the main characters in the story.

The TMCB II is twice as long as TMCB I, reflecting the mathematical progress of the past 20 years and the extensive scholarship on the early history of Ramsey Theory undertaken by Professor Soifer, particularly in relation to the life of B. L. van der Waerden. Professor Soifer has included new research on the origin of Van der Waerden's Theorem and on the Van der Waerden controversy. As always, Professor Soifer is unrelenting in his efforts to give Mathematical credit where it is due.

There has been considerable progress on the many of the important unsolved problems discussed in TMCB I. The central problem in both books is that of determining the chromatic number of the Euclidean plane, which involves finding the minimum numbers of colors needed so that each point in the plane can be assigned a color with no two points at distance one getting the same color. It had been known that this minimum number of colors is at least 4 and at most 7, bounds that stood for more than 60 years. Then in 2018 Aubrey de Grey showed that at least five colors are required. De Grey's result stimulated a flurry of activity on the problem, including the Polymath 16 project, which resulted in the construction of smaller examples along with a number of other results.

Professor Soifer presents de Grey's construction, a graph of order 1585, in detail and with wonderful clarity. An important and related problem, determining the chromatic number of the odd distance graph has been settled by James Davies within the past year. In addition to new results on old questions, some new topics have also been added. There is now a section on Schur numbers, including details on Marijn Heule's computation of the fifth Schur number. New historical research on the origin of Van der Waerden's Conjecture and Theorem are especially fascinating.

Finally, I would like to add a warning for the reader. Using the book to look up simple facts or references can be dangerous. I have often picked up TMCB I to check a reference or some simple fact, and found myself, hours later, lost in some aspect of the story. And to those who have never read either book and will be reading it for the first time, I envy you.

Indiana State University, Terre Haute, IN, USA Geoffrey Exoo
Monday, July 10, 2023

Foreword to the First Edition

This is a unique type of book; at least, I have never encountered a book of this kind. The best description of it I can give is that it is a mystery novel, developing on three levels, and imbued with both educational and philosophical/moral issues. If this summary description does not help understanding the particular character and allure of the book, possibly a more detailed explanation will be found useful.

One of the primary goals of the author is to interest readers – in particular, young mathematicians or possibly pre-mathematicians – in the fascinating world of elegant and easily understandable problems, for which no particular mathematical knowledge is necessary, but which are very far from being easily solved. In fact, the prototype of such problems is the following: If each point of the plane is to be given a color, how many colors do we need if every two points at unit distance are to receive distinct colors? More than half a century ago it was established that the least number of colors needed for such a coloring is either 4, or 5, or 6 or 7. Well, which is it? Despite efforts by a legion of very bright people – many of whom developed whole branches of mathematics and solved problems that seemed much harder – not a single advance towards the answer has been made. This mystery, and scores of other similarly simple questions, form one level of mysteries explored. In doing this, the author presents a whole lot of attractive results in an engaging way, and with increasing level of depth.

The quest for precision in the statement of the problems and the results and their proofs leads the author to challenge much of the prevailing historical "knowledge." Going to the original publications and drawing in many cases on witnesses and on archival and otherwise unpublished sources, Soifer uncovers many mysteries. In most cases, dogged perseverance enables him to discover the truth. All this is presented as following in a natural development from the mathematics to the history of the problem or result, and from there to the interest in the people who produced the mathematics. For many of the persons involved this results in information not available from any other source; in lots of the cases, the available publications present an inaccurate (or at least incomplete) data. The author is very careful in documenting his claims by specific references, by citing correspondence between the principals involved, and by accounts by witnesses.

One of these developments leads Soifer to examine in great detail the life and actions of one of the great mathematicians of the twentieth century, Bartel Leendert van der Waerden. Although Dutch, van der Waerden spent the years from 1931 to 1945 in the Nazi Germany.

This, and some of van der Waerden's activities during that time, became very controversial after World War II, and led Soifer to examine the moral and ethical questions relevant to the life of a scientist in a criminal dictatorship.

The diligence with which Soifer pursued his quests for information is way beyond exemplary. He reports exchanges with I am sure hundreds of people, via mail, phone, email, visits – all dated and documented. The educational aspects that begin with matters any middle-school student can understand, develop gradually into areas of most recent research, involving not only combinatorics but also algebra, topology, questions of foundations of mathematics, and more.

I found it hard to stop reading before I finished (in two days) the whole text. Soifer engages the reader's attention not only mathematically, but emotionally and esthetically. May you enjoy the book as much as I did!

University of Washington, Seattle, WA, USA Branko Grünbaum
February 28, 2008

Foreword to the First Edition

Alexander Soifer's latest book is a fully-fledged adult specimen of a new species, a work of literature in which fascinating elementary problems and developments concerning colorings in arithmetic or geometric settings are fluently presented and interwoven with a detailed and scholarly history of these problems and developments.

This history, mostly from the 20th century, is part memoir, for Professor Soifer was personally acquainted with some of the principals of the story (the great Paul Erdős, for instance), became acquainted with others over the 18-year interval during which the book was written (Dima Raiskii, for instance, whose story is particularly poignant), and created himself some of the mathematics of which he writes.

Anecdotes, personal communications, and biography make for a good read, and the readability in *The Mathematical Coloring Book* is not confined to the accounts of events that transpired during the author's lifetime. The most important and fascinating parts of the book, in my humble opinion, are Parts IV, VI, and VII, in which is illuminated the progress along the intellectual strand that originated with the Four-Color Conjecture and runs through Ramsey's Theorem via Schur, Baudet, and Van der Waerden right to the present day, via Erdős and numbers of others, including Soifer. Not only is this account fascinating, it is indispensable: it can be found nowhere else.

The reportage is skillful, and the scholarship is impressive – this is what Seymour Hersh might have written, had he been a very good mathematician curious to the point of obsession with the history of these coloring problems.

The unusual combination of abilities and interests of the author make the species of which this book is the sole member automatically endangered. But in the worlds of literature and mathematics and literature about mathematics, unicorns can have offspring, even if the offspring are not exactly unicorns. I think of earlier books of the same family as *The Mathematical Coloring Book* – G. H. Hardy's A *Mathematician's Apology*, James R. Newman's *The World of Mathematics*, Courant and Robbins' *What Is Mathematics?*, Paul Halmos' *I Want to Be a Mathematician: An Automathography*, or the books on Erdős that appeared soon after his death – all of them related at least distantly to *Mathematical Coloring Book* by virtue of the attempt to blend (whether successfully or not is open to debate) mathematics with history or personal memoir, and it seems to me that, whatever the merits of those works, they have all affected how mathematics is viewed and written about. And this will be a large part of the legacy of The Mathematical Coloring Book – besides providing inspiration and plenty of mathematics to work on to young mathematicians, and a priceless source to historians, and entertainment to those who are curious about the activities of mathematicians, *The Mathematical Coloring Book* will (we can hope) have a great and salutary influence on all writing on mathematics in the future.

Auburn University, Auburn, AL, USA
March 28, 2008

Peter D. Johnson Jr.

Foreword to the First Edition

What is the minimum number of colors required to color the points of the Euclidean plane in such a way that no two points that are one unit apart receive the same color? *Mathematical Coloring Book* describes the odyssey of Alexander Soifer and fellow mathematicians as they have attempted to answer this question and others involving the idea of partitioning (coloring) sets.

Among other things, the book provides an up-to-date summary of our knowledge of the most significant of these problems. But it does much more than that. It gives a compelling and often highly personal account of discoveries that have shaped that knowledge.

Soifer's writing brings the mathematical players into full view, and he paints their lives and achievements vividly and in detail, often against the backdrop of world events at the time. His treatment of the intellectual history of coloring problems is captivating.

Memphis State University, Memphis, TN, USA — Cecil Rousseau
March 2008

Acknowledgments—2023

I am grateful to the referees of this opus, Geoffrey Exoo and Peter D. Johnson, Jr., for their feedback and forewords. I thank the referees William Gasarch and Aubrey de Grey for reading the manuscript and providing valuable feedback.

A few years after the first edition was published, Branko Grünbaum sent me a list of corrections. I am so very thankful to him, for no one then, me included, thought that a new edition was possible.

I am grateful to my colleagues who contributed to this new book: Branko Grünbaum, Ronald L. Graham, Stanisław P. Radziszowski, Geoffrey Exoo, Jaan Parts, Oleg V. Borodin, Aubrey David Nicholas Jasper de Grey, Marienus Johannes Hendrikus Heule, Moritz Epple, Elena Nikolaevna Lambina, Andrei Lodkin, Tom C. Brown, and others.

I am in debt to my intellectual friends for their ideas and inspiration: Albert Camus, Dalai Lama, Jacques Prévert, Georges Brassens, Eli Wiesel, Edward R. Murrow, Eli Wiesel, Albert Einstein, Pablo Picasso, Michelangelo Antonioni, Muriel Rukeyser, Viktor Shklovsky, and others.

I thank my Springer Editor of many years, Elizabeth Loew, for offering me this rare opportunity of the second expanded edition. I thank Bob Ewell for converting some of my sketches into clean illustrations.

My deep apologies to those who helped, and the absent-minded professor—me—forgot to include them here.

Acknowledgments—2009

My first thank you goes to my late father Yuri Soifer, a great painter, who introduced colors into my life, and to whom this book is gratefully dedicated. As the son of a painter and an actress, I may have inherited the genes of understanding and appreciating the arts. Yet, it was my parents, Yuri Soifer and Frieda Hoffman, who inspired my development as a connoisseur and student of the arts. I have enjoyed mathematics only because it could be viewed as an art as well. I am deeply indebted to my kids Mark, Isabelle, and Leon, and to my cousin and fine composer Leonid Hoffman for the support their love has always provided. I thank my old friend Konstantin Kikoin and Leonid Hoffman for years of stimulating conversations on all themes of high culture. I thank Maya Soifer for restarting my creative engine when at times it worked on low rpm (even though near the end of my work on this book, she almost broke the engine by abandoning the car :). And I am grateful to Branko Grünbaum and Peter D. Johnson, Jr., who were first to read the entire manuscript, for their very kind words and valuable suggestions.

This is a singular book for me, a result of 18 years of mathematical and historical research and thinking over the moral and philosophical issues surrounding a mathematician in the society. The long years of writing have produced one immense benefit that quickly baked book would never fathom to possess. I have had the distinct pleasure to discuss mathematics and the history for this book with senior sages Paul Erdős, George Szekeres, Esther (Klein) Szekeres, Martha (Wachsberger) Svéd, Henry Baudet, Nicolaas G. de Bruijn, Bartel L. van der Waerden, Harold W. Kuhn, Dirk Struick, Hilde Brauer (Mrs. Alfred Brauer), Hilde Abelin–Schur, Walter Ledermann, Anne Davenport (Mrs. Harold Davenport), Victor Klee, Branko Grünbaum, and Harold W. Kuhn. Many of these great people are no longer with us; others are in their 80s. Their knowledge, their memories have provided blood to the body of my book. I am infinitely indebted to them all, as well as to the younger contributors Ronald L. Graham, Edward Nelson, John Isbell, Adriano Garsia, James W. Fernandez, and Renate Fernandez, who are merely in their 70s.

Harold W. Kuhn wrote a triptych on the economics of three beautiful minds, Frank P. Ramsey, John von Neumann, and John F. Nash, Jr. especially for this book – read it in Chapter 32. Steven Townsend wrote and illustrated a new version of his proof especially for this book (Chapter 25). Kenneth J. Falconer wrote a new clear exposition of his proof especially for this book (Chapter 9). Your contributions are so great, I thank you three so very much!

I thank several colleagues for their self-portraits written for this book: Nicolaas G. de Bruijn, Hillel Furstenberg, Vadim G. Vizing, Kenneth J. Falconer, Paul O'Donnell, Vitaly Bergelson, Alexander Leibman, and Michael Tarsi.

I am thrilled to be able to present results and writings of my colleagues I have known personally: Paul Erdős, George Szekeres, Esther Klein Szekeres, Paul O'Donnell, Rob Hochberg, Kiran Chilakamarri, Ronald L. Graham, Joel H. Spencer, Hillel Furstenberg, Paul Kainen, Peter D. Johnson, Jr., Nicholas Wormald; David Coulson, Paul Seymour, Neil Robertson, Robin J. Wilson, Stanisław P. Radziszowski, Ed Pegg, Jr.; and colleagues I got to know through correspondence: Vadim G. Vizing, Bruce L. Rothschild, Mehdi Behzad, Douglas Woodall, Bruce L. Rothschild, Jan Kyncl, and many others.

I am grateful to George Szekeres and MIT Press for their permission to reproduce here the most lyrical George's *Reminiscences* (part of Chapter 29). I thank Bartel Leendert van der Waerden and Academic Press, London, for their kind permission to reproduce here Van der Waerden's insightful *How the Proof of Baudet's Conjecture Was Found* (Chapter 33).

I thank all those who provided me with rare, early photographs of themselves and others: Paul Erdős (including a photograph with Leo Moser); George Szekeres and Esther Klein (including a rare photo of Paul Turán); Vadim G. Vizing; Edward Nelson; Paul O'Donnell. Hilde Abelin–Schur has kindly provided photographs of her great father Issai Schur. Dorith van der Waerden and Theo van der Waerden have generously provided rare photographs of their uncle Bartel L. van der Waerden and the rest of their distinguished family. Henry Baudet II generously shared with me photographs of his brilliant father Pierre Joseph Henry Baudet, and of his family with the legendary world chess champion Emanuel Lasker. Ronald L. Graham has kindly provided a photograph of him presenting Timothy Gowers with the check for $1000. Alice Bogdán has shared a photograph of her great brother-in-low Tibor Gallai.

I thank numerous archives for documents and photographs and permissions to use them. Dr. Mordecai Paldiel and Yad Vashem, The Holocaust Martyrs' and Heroes' Remembrance Authority, Jerusalem, provided documents related to the granting Senta Govers Baudet the title of a "Righteous Among the Nations." John Webb and Mathematics Department of the University of Cape Town supplied a photograph of Francis Guthrie. By kind permission of the Provost and Scholars of King's College, Cambridge, I am able to reproduce here two photographs of Frank Plumpton Ramsey, from the papers of John Maynard Keynes. The Board of Trinity College Dublin kindly allowed me to reproduce the correspondence between Augustus De Morgan and William R. Hamilton. Humboldt University of Berlin shared documents from the personnel file of Issai Schur.

The following archivists and archives provided invaluable documents and some photographs related to Bartel L. van der Waerden: John Wigmans, *Rijksarchief in Noord-Holland*; Dr. Peter J. Knegtmans, The University Historian, *Universiteit van Amsterdam*; Prof. Dr. Gerald Wiemers and Martina Geigenmüller, *Universitätsarchiv Leipzig*; Prof. Dr. Holger P. Petersson and his personal archive; Gertjan Dikken, *Het Parool*; Madelon de Keizer; Dr. Wolfram Neubauer, Angela Gastl and Corina Tresch De Luca, *ETH–Bibliothek, ETH*, Zürich; Dr. Heinzpeter Stucki, *Universitätsarchiv, Universität Zürich*; Drs. A. Marian Th. Schilder, *Universiteitsmuseum de Agnietenkapel*, Amsterdam; Maarten H. Thomp, *Centrale Archiefbewaarplaats, Universiteit Utrecht*; Nancy Cricco, University Archivist, and her graduate students assistants, New York University; James Stimpert, Archivist, Milton S. Eisenhower Library, Special Collections, The Johns Hopkins University; Dr. Helmut

Rechenberg, Director of the Werner Heisenberg Archiv, Munich; G. G. J. (Gijs) Boink, *Het Nationaal Archief*, Den Haag; Library of Congress, Manuscript Division, Washington D.C.; *Centrum voor Wiskunde en Informatica*, Amsterdam; Prof. Nicholas M. Katz, Chair, Department of Mathematics, Princeton University; and Mitchell C. Brown and Sarah M. LaFata, Fine Library, Princeton University.

My research on the life of Van der Waerden could not be based on the archival materials alone. I am most grateful to the individuals who provided much help in this undertaking: Dorith van der Waerden; Theo van der Waerden; Hans van der Waerden; Helga van der Waerden Habicht; Annemarie van der Waerden; Prof. Dr. Henry Baudet (1919–1998); Prof. Dr. Herman J. A. Duparc (1918–2002); Prof. Dr. Nicolaas G. de Bruijn; Prof. Dr. Benno Eckmann (1917–2008); Prof. Dr. Dirk van Dalen; Prof. Dirk J. Struik (1894–2000); Dr. Paul Erdős (1913–1996); Thomas Powers; Prof. Mark Walker, Prof. James and Dr. Renate Fernandez; Dr. Maya Soifer; and Princeton University Professors Harold W. Kuhn, Charles Gillispie, Simon Kochen, John H. Conway, Edward Nelson, Robert C. Gunning, Hale Trotter, Arthur S. Wightman, Val L. Fitch, Robert Fagles, and Steven Sperber.

I thank my wonderful translators from Dutch, Dr. Stefan van der Elst and Prof. Marijke Augusteijn; and from German: Prof. Simon A. Brendle; Prof. Robert von Dassanowsky; Prof. Robert Sackett; Prof. Dr. Heiko Harborth; and Prof. Dr. Hans–Dietrich Gronau.

I thank those who converted my pencil doodles into computer-aided illustrations: Steve Bamberger, Phil Emerick, and Col. Dr. Robert Ewell.

The research quarterly *Geombinatorics* provided a major forum for the essays related to the chromatic number of the plane. Consequently, it is cited numerous times in this book. I wish to thank the editors of *Geombinatorics* Paul Erdős, Branko Grünbaum, Ronald L. Graham, Heiko Harborth, Peter D. Johnson, Jr., Jaroslav Nešetřil, and János Pach.

My University of Colorado bosses provided support and opportunity to be away for long periods of time, on loan to Princeton and Rutgers Universities – I thank my Dean Tom Christensen and Chancellor Pam Shockley–Zalabak.

I am grateful to my Princeton–Math colleagues and friends for maintaining a unique creative atmosphere in the historic Fine Hall, and Fred Roberts for the tranquility of his DIMACS at Rutgers University. Library services at Princeton have provided an invaluable swift service: while working there for three years, I must have read thousands of papers and many books.

I thank my Springer Editor Mark Spencer, who initiated our contact, showed trust in me and this project, based merely on the table of contents and a single chapter, who in 2004 proposed to publish this book in Springer. At a critical time in my life, Springer's Executive Director Ann Kostant made me believe that what I am and what I do really matters – thank you from the bottom of my heart, Ann!

There is no better place to celebrate the completion of the book than the land of Pythagoras, Euclid, and Archimedes. I thank Panayiotis "Takis" Vlamos for inviting me to lecture on the Island of Corfu, Thessaloniki and Athens days after my completed manuscript was sent to Springer.

Greetings to the Reader 2023

The Universe is made of stories, not of atoms.

– Muriel Rukeyser[2]

In order to create new, one has to come with his human biography and with his new experience, while using collective experience and enriching it.

– Viktor Shklovsky[3]

To create today means to create dangerously. Every publication is a deliberate act, and that act makes us vulnerable to the passions of a century that forgives nothing.

– Albert Camus[4]

Does the evolution of ideas of mathematical kind makes for an exciting story? Of course, for it is full of heroism and cowardice, comedy and drama, success and failure, loyalty and betrayal, high moral grounds and service to evil. It is as vast as the universe, yet it could also be seen as a chamber affair. My task is to assemble the past, create a braid of mathematics and history, share aspirations of many and achievements of some, come to the frontier, and look into the future by presenting exciting open problems and conjectures. As a writer, I share the view expressed on December 10, 1957, at the Nobel Banquet by Albert Camus:

> The writer's role is not free from difficult duties. By definition he cannot put himself today in the service of those who make history; he is at the service of those who suffer it.

On the Election Day, November 4, 2008, I bid farewell to my *Mathematical Coloring Book*, as she reached her adulthood, 18 years from its conception. On that very day, we Americans elected Barack Obama, one of the brightest presidents of my Land, who returned integrity to politics, alas, not for long.

---

[2] Quoted from the book "A Speed of Darkness," 1968.

[3] "Repeating lessons of "Battleship Potemkin," in the book *Over 40 Years*, Iskusstvo Publisher, Moscow, 1965 (Russian).

[4] [Cam2], Lecture at Uppsala University, Sweden, December 14, 1957.

We, authors of scholarly monographs, ought to put everything we got in our books, for we seldom get a second chance. This book is an exception. Its central place in the field and brilliant recent breakthroughs inspired Springer to enable me to have this second shot. The first version was universally well received. In August 2009, Springer funded my travel to Portland for book signing of *The Mathematical Coloring Book* during MathFest. Our fantastic Ramsey Theory leader Ron Graham asked me to inscribe his copy. "To our captain, Ron Graham," I wrote on his copy. In fact, Ron told me that he bought three copies, and to my question "why" replied that he had three offices and wanted to have this book handy wherever he was. Our conversation was so special that I recorded the time: 4:29 p.m., August 6, 2009, and our dialog.

– Can you write another book like this? asked Ron.
– No, readily replied I, for I thought about this question myself for a while.
– Heart and soul, and blood, and sweat, and tears, and none left for the next one, remarked Ron.

Harold W. Kuhn, the Princeton double professor of mathematics and economics, wrote especially for the first *Coloring Book* an essay about the economics of John F. Nash, Jr., whom he nominated for the 1994 Nobel Prize. On January 9, 2009, at 8:40 a.m. MST, Harold sent me an email that made my eyes wet:

> Dear Sasha:
> Just now a postman came to the door with a copy of the masterpiece of the century. I thank you and the mathematics community should thank you for years to come. You have set a standard for writing about mathematics and mathematicians that will be hard to match.
> With very warm regards,
> Harold

The reception of the first *Coloring Book* was satisfying. It was called by many a standard text and quoted by nearly every publication in this field. The book inspired mathematicians and non-mathematicians to take on some of its problems, first of all the chromatic number of the plane, and this new wave of enthusiasm resulted in brilliant new results. There was, however, one person from mathematics who tried to silence me. I will address it on appropriate pages. Camus's warning in the epigraph about "the passions of a century that forgive nothing" proved prophetic. However, my detractors were naïve to imagine that someone who gave up everything earned in his Russian life to become a refugee in the Land of the Free could be silenced.

I view any written matter, be it history or mathematics, to be a genre of literary art. Archivally researching the past, finding eyewitnesses are essential components of my approach to writing. As Viktor Shklovsky wrote,[5]

> Art cannot reject the past, rethinking it and achieving deeper understanding, just as the language cannot reject its history.

[5] "Why cannot one convert a novel into a drama," in *Over 40 Years*, Iskusstvo, Moscow, 1965.

I am including in this book newly created history, my interactions with the authors of ideas, more rare photographs, and reminiscences of senior sages, biographies, and of course, more striking results. This new book is even more personable and more personal for me. I hope you will treat it softly on your hard drives, regularly dust hard copies on your shelves, and send me your ideas for the future editions and books. I really enjoy the opportunity to have guest writers in my books. Harold W. Kuhn, Harry Furstenberg, Kenneth Falconer, Steven Townsend, Marta Svéd, George Szekeres, Bartel L. van der Waerden, Hans van der Waerden, Dirk van Dalen, and others penned essays for this book or allowed me to reproduce their relevant writings. Paul Joseph Cohen, Robert Solovay, and Saharon Shelah responded to my questionnaire on the foundations of mathematics; their profound opinions are a treasure preserved between the covers of this book. And of course, I generously quote deep, often humorous, ideas of Einstein, Picasso, Camus, Wiesel, Murrow, Freud, Dreyer, Antonioni, Pasolini, Matisse, Kandinsky, Dalai Lama, Epple, Alonzo Church, Herbert Read, and other creators of high culture of our world. Following *Buch der Freunde* by Hugo von Hofmannsthal, I call all of them my friends, friends of intellectual and spiritual kind.

Nobel Laureate Joseph Brodsky's favorite American Poet Wystan Hugh Auden observes,

In our age, the mere making of a work of art is itself a political act.

Another Nobel Laureate, the great Albert Camus warns us in the epigraph, and I repeat:

To create today means to create dangerously. Every publication is a deliberate act, and that act makes us vulnerable to the passions of a century that forgive nothing.

Of course, I agree with them. However, deep in my heart I know that no one else can write *this* book, it is my duty, whatever the consequences. And I respond to the call of the Pulitzer Prize Laureate David Maraniss:

History writes people out of the story, and it's our job to write them back in.

As the Russian veteran of filmmaking and writer Viktor Shklovsky observes:[6]

Trees live longer than people.
Word lives longer than trees.

Recent years were not the easiest time in history. Epidemic of rare proportions, rise of populism with totalitarian flavor, reawakened hatred toward those who in any way differ from all-powerful majority, the unprovoked Russian war on Ukraine. Empires do not dissolve peacefully; they try to destroy what they cannot control. I am involved with the world and hope more of my colleagues will vacate Ivory Tower for the streets of our small, endangered planet. As Albert Camus put it in his December 14, 1957, lecture at Uppsala University [Cam2],

The age of irresponsible artists [and scholars] is over.

I am tracing in this book the evolution of a colorful 100-year-old field, thoughts of wisdom I acquired during my life (if any), and moral issues of the past and the present. I firmly believe that moral principles of a profession must lie in its foundation, not outside of it. Case in point: Grigory Perelman. This young genius mathematician achieved what no one else could: he

[6] Repeating lessons of "Battleship Potemkin," in *Over 40 Years*, Iskusstvo, Moscow, 1965.

settled the Poincaré and Geometrization Conjectures. This earned Grigory the Fields Medal and the Millennium Prize of $1,000,000. His refusal to accept these high awards (he did not wish to be "a poster boy" for the field where the majority tolerates immorality of its members) was "explained" as Perelman's insanity. IMHO, Perelman was the sane one, and the insane comprised the majority.

And so, *the most important unsolved problem in mathematics is establishing high moral principles of the profession and observing them*. I address this unsolved problem in this volume in three essays "Today I," "Today II," and "Today III."

A majority of publications in the field cite mathematics of the first edition [Soi44], with credit. Yet, they often lift my historical discoveries without credit. I comfort myself by recalling Picasso's statement, "*Plagiarism is the best compliment*."

The present book was originally (2018) envisioned as the expanded second edition of the original *Mathematical Coloring Book*. Now, looking at it, I believe that "second expanded edition" does not do justice to the magnitude of the new specimen, ca. twice as long, with lots of new material, and even Van der Waerden's chapters are all different from the first edition. So, I decided for this new book to use the new title:

The New
Mathematical Coloring Book.

I hope this book will inspire you, my contemporary, and those who will come in our stead, to think, to feel, and to create.

Colorado Springs, USA Alexander Soifer
July 4, 2023
The 247th Anniversary of American Independence

Greetings to the Reader 2009

I bring here all: what have I lived thru,
And that what keeps my soul alive,
My rectitude and aspirations,
And what have seen my own eyes.
– Boris Pasternak, *The Waves*, 1931[7]

When the form is realized, it is here to live its own life.
– Pablo Picasso

Pasternak's epigraph precisely describes my work on this book – I gave it all of myself, without reservation. August Renoir believed that just as many people read one book all their lives (the Bible, the Koran, etc.), so can the artist paint all his life one painting. Likewise, I could write one book all my life – in fact, I almost have, for I have been working on this book for 18 years.

It is unfair, however, to keep the book all to myself – many colleagues have been waiting for the birth of this book. In fact, it has been cited, and even reviewed many years ago. The first mention of it appears already in 1991 on page 336 of the book by Victor Klee and Stan Wagon [KW], where the authors recommend the book for "survey of later developments" of the chromatic number of the plane problem. On page 150 of their 1995 book [JT], Tommy R. Jensen and Bjarne Toft announced that "a comprehensive survey [of the chromatic number of the plane problem] . . . will be given by Soifer [to appear]." Once in the 1990s my son Mark S. Soifer told me that he saw my *Mathematical Coloring Book* available for $30 for a special

[7] [Pas], Russian History buffs would appreciate that in its first 1934 publication this 1931 poem was dedicated by Pasternak "To N.I. Bukharin," a prominent Soviet leader who fell victim to Stalin's terror of 1936–1938. Translated especially for this book by Ilya Hoffman. The original Russian text is:

Здесь будет все: пережитое,
И то, чем я еще живу,
Мои стремленья и устои,
И виденное наяву.

order at Borders Bookstore. I offered to buy a copy – to buy is easier and faster than to write a book!

I started writing this book when copies of my *How Does One Cut a Triangle?* [Soi1] arrived from the printer, in early 1990. I told my father Yuri Soifer then that this book will be dedicated to him, and so it is . . . This coloring book is for my late father, a fine artist and man. Yuri lived with his sketchpad and drawing utensils in his pocket, constantly and intensely looking at people and making sharp momentary sketches. He was a great painter and my lifelong example of searching for and discovering life around him, creating art that challenged "real" life herself. Yuri never taught me his trade, but during our numerous joint tours of art in museums and exhibitions, he pointed out beauties that only true artists could notice: a dream of harvest in Van Gogh's "Sower," Rodin's distortions in a search of greater expressiveness in "The Thinker." These timeless lessons allowed me to become a student of beauty, and discover subtleties in paintings, sculptures, and films throughout my life.

The book includes not just mathematics but the process of investigation, trains of mathematical thought, and psychology of mathematical invention. The book does not just include history and prehistory of Ramsey Theory and other related fields, but also conveys the process of historical investigation – the kitchen of historical research, if you will. It has captivated me, made me feel like a Sherlock Holmes – I hope you, my reader, my Doctor Watson, will enjoy the sense of suspense and the celebration of discovery as much as I have.

The epigraph for my book is an English translation of Jacques Prevert's genius and concise portrayal of a creative process – I know of no better. I translated it with the help of my friend Maurice Starck from *Nouvelle Calédonie*, the island in the Pacific Ocean to which no planes fly from America, but to paraphrase Rudyard Kipling, *I'd like to roll to Nouvelle Calédonie some day before I'm old!*

This book is dedicated to problems involving colored objects, and results about the existence of certain exciting and unexpected properties that occur regardless of how these objects (points in the plane, space, integers, real numbers, subsets, etc.) are colored. In mathematics, these results comprise *Ramsey Theory*, a flourishing area of mathematics, whose motto can be formulated as follows: *any coloring of a large enough system contains a monochromatic subsystem of given in advance structure*, or simply put, *absolute chaos is absolutely impossible*. Ramsey Theory thus includes parts of many fields of mathematics, such as combinatorics, geometry, number theory, topology, etc., and addresses new problems, often on the frontier of two or more traditional mathematical fields. The book also includes some problems that can be solved by inventing coloring, and results that prove the existence of certain colorings, most famous of the latter being, of course, The Four-Color Theorem.

Most books in the field present mathematics as a flower, dried out between the pages of an old dusty foliant, so dry that colors are faded and only the relentless theorem-proof-theorem-proof narrative survives. Along with my previous books, *The Mathematical Coloring Book* strives to become a live account of a live mathematics. I hope the book will present mathematics as a human endeavor: you should expect to find in it not only results, but also historical portraits of their creators; not only mathematical facts, but also open problems; not only new mathematical research, but also new historical investigations; not only mathematical aspirations, but also moral dilemmas of the times between and during the two tragic World Wars of the twentieth century. In my opinion, mathematics is done by human beings, and knowing their lives and cultures enriches our understanding of mathematics as a product of human activity, rather than an abstraction that exists outside and separately from us and comes to us exclusively as a catalog of theorems and formulas. Indeed, new facts and artifacts will be presented that are related to the history of the Chromatic Number of the Plane problem, the early history of Ramsey Theory, the lives of Issai Schur, Pierre Joseph Henry Baudet, Bartel Leendert van der Waerden, and other colorful personages.

I hope you will join me on a journey you would not forget, a journey full of passion, where mathematics and history are researched in the process of solving mysteries more exciting than fiction, precisely because those are mysteries of real affairs of human history. Can mathematics be received by all senses, like a vibrant flower, indeed, like life itself? One way to find out is to experience this book.

While much of the book is dedicated to results of Ramsey Theory, I did not wish to call my book "Introduction to Ramsey Theory," for such a title would immediately lose young talented readers. Somehow, the playfulness of *The Mathematical Coloring Book* appealed to me from the start, even though I was asked on occasion whether 5-year-olds would be able to enjoy my book and color in it between lines. To be a bit more serious, and on advice of Vickie Kern of the Princeton University Press, I created a subtitle *Mathematics of Coloring and the Colorful Life of Its Creators*. This is a faithful subtitle, for this book explores the birth of ideas and search for its creators. I discovered very quickly that in conveying "colorful lives of creators" I cannot always rely on encyclopedias and biographical articles but must conduct historical investigations on my own. It was a hard work to research some of the lives, especially that of B. L. van der Waerden, which alone took over 12 years of archival research and thinking over the assembled evidence. Fortunately, this produced a satisfying result: we have in this book some definitive biographies, of Bartel L. van der Waerden, Pierre Joseph Henry Baudet, Issai Schur, autobiographies of Hillel Furstenberg, Kenneth J. Falconer, and others.

I always aspire to understand who, when, and how made a discovery. Accordingly, this book tries to explore biographies of the discoverers and psychologies of their creative process. Every stone has been turned: numerous archives in Germany, the Netherlands, Switzerland, Ireland, England, United States helped with rarely if ever seen documents; invaluable and irreplaceable interviews were conducted with eyewitnesses; dialogues held with creators. I have read thousands of items in the process of writing this book. Cited bibliography alone includes over 800 titles. I was inspired by people I have known personally, such as Paul Erdős, James W. Fernandez, Harold W. Kuhn, and many others, and by people I have not personally met, such as Boris Pasternak, Pablo Picasso, Herbert Read—to name a few of the many influences. I agree with D. A. Smith, who in the discussion after Alfred Brauer's talk [Bra2, p. 36], writes:

> Mathematical history is a sadly neglected subject. Most of this history belongs to the twentieth century, and a good deal of it in the memories of mathematicians still living. The younger generation of mathematicians has been trained to consider the product, mathematics, as the most important thing, and to think of the people who produced it only as names attached to theorems. This frequently makes for a rather dry subject matter.

Milan Kundera, in his *The Curtain: An Essay in Seven Parts* [Kun], said about a novel what is true about mathematics as well:

> A novelist talking about the art of the novel is not a professor giving a discourse from his podium. Imagine him rather as a painter welcoming you into his studio, where you are surrounded by his canvases staring at you from where they lean against the walls. He will talk about himself, but even more about other people, about novels of theirs that he loves and that have a secret presence in his own work. According to his criteria of values, he will again trace out for you the whole past of the novel's history, and in so doing will give you some sense of his own poetics of the novel.

I was also inspired by the early readers of the book, and their feedback. Stanisław P. Radziszowski, after reviewing Chapter 27, e-mailed me on May 2, 2007:

> I am very anxious to read the whole book! You are doing great service to the community by taking care of the past, so the things are better understood in the future.

In his unpublished until recently letter, Ernest Hemingway in a sense defends my writing of this book for so very long:[8]

> When I make country, or a city, or a river in a novel it is slow work because you have to always *make* it, then it is alive. But nobody makes anything quickly nor easily if it is any good.

Branko Grünbaum, upon reading the entire manuscript, wrote in the February 28, 2008 e-mail:

> Somehow it seems that 18 years would be too short a time to dig up all this information!

This book will not treat you to completeness or most general results. Instead, it would give young active high school and college mathematicians an accessible introduction to the beautiful ideas of mathematics of coloring. Mathematics professionals, who may believe they know everything, would be pleasantly surprised by some unpublished or unnoticed mathematical gems. I hope young and not so young mathematicians alike would welcome an opportunity to try their hand – or mind? – on numerous open problems, all easily understood and not at all easily solved.

If the interest of my colleagues and friends at Princeton-Math is any indication, every intelligent reader would welcome an engagement in solving historical mysteries, especially those from the times of the Third Reich, World War II, and de-Nazification of Europe. Historians of mathematics would find much of new information and old errors corrected for the first time. And everyone will experience seeing, for the first time, faces one has not seen before in print, on rare photographs of the creators of mathematics presented herein, from Francis Guthrie to Issai Schur as a young man, from the teenager Edward Nelson to Paul O'Donnell, from Pierre Joseph Henry Baudet to Bartel Leendert van der Waerden and his family, and documents, such as the one where Adolph Hitler commits a "micromanagement" of firing a Jew, Issai Schur from his job of Professor at the University of Berlin.

This is a unique book, free from a strait jacket of a typical textbook, yet useable as a text for a host of various courses, two of which I have personally given to university seniors and graduate students at the University of Colorado: *What is Mathematics?*, and *Mathematical Coloring Course*, both presenting a "laboratory of a mathematician," a place where students

[8] From the unpublished 1937 letter. Quoted from *New York Times*, February 10, 2008, p. AR 8.

learn mathematics and its history by doing them, and realizing in the process what mathematics is and what mathematicians do.

Mathematics is an art. It is a poor man's art: Nothing is needed to conceive it, and only paper and pencil to convey.

This long work gave me so very much, in Aleksandr Pushkin's words,

The heavenly, and inspiration,
And life, and tears, and tender love.[9]

I have been raising this book for 18 years. Over the past few years, I felt as if the book herself had been dictating her content, as I obeyed her calling as a scribe. At 18, my book is now an adult, and deserves to separate from me to live her own life. As Picasso put it, "When the form is realized, it is here to live its own life." Farewell, my child, let the world love you, as I have and always will.

[9] In the original Russian, it sounds much better:

И божество, и вдохновенье,
И жизнь, и слезы, и любовь.

Contents

Part V Colored Graphs

Part VI The Ramsey Principles

Part VII Colored Integers: Ramsey Theory Before Ramsey and Its AfterMath

Part X Ask What Your Computer Can Do for You

Part XIV Farewell to the Reader

About the Author

Alexander Soifer is a Russian-born American mathematician and prolific author. His works include over 400 articles and 13 books. He was born in Moscow to the family Yuri Soifer, a fine artist, and Frieda Hoffman Soifer, an actress. At 6, he enrolled in music school in the class of piano and composition. At 14, Alexander chose to pursue a career in mathematics over music.

Soifer wrote his dissertation in 1971 and was granted a Ph.D. degree by the government in 1973. In 1978, he terminated his Soviet citizenship and left in search of freedom of speech and conscience.

He has been a professor of mathematics at the University of Colorado since 1979. He was a visiting fellow at Princeton University from 2002 to 2004, and again in 2006–2007. Soifer also teaches courses on art history and European and Japanese cinema.

In 1984 he founded and has been running the Colorado Mathematical Olympiad (CMO) at the University of Colorado at Colorado Springs. Soifer creates original problems for CMO, often inspired by contemporary and historical research publications. For the Olympiad's 30th anniversary, the university produced a film about it (see it on YouTube https://www.youtube.com/watch?v=EOTiA4YWU-k). In May 2018, in recognition of 35 years of leadership, the judges and winners decided to rename the Colorado Mathematical Olympiad as the Soifer Mathematical Olympiad.

In 1991, Soifer founded the research quarterly *Geombinatorics* and publishes it with the exceptional editorial board that used to include Paul Erdős, Ron Graham, and Branko Grünbaum. In the 33 years of *Geombinatorics*, 130 issues have been published to date. The journal is indexed in ZbMATH, Excellence in Research for Australia, and MathSciNet.

In 1992, Soifer was presented with the University of Colorado President's Service Award. In 1993–1995, Soifer served on the

Human Relations Commission of Colorado Springs and presided over its Civil Rights Ordinance Study Group.

In July 2006 at the University of Cambridge, Soifer was presented with the Paul Erdős Award by the World Federation of National Mathematics Competitions (WFNMC). He was elected and served as President of WFNMC from 2012 to 2018. Soifer's Erdős number is 1. His coauthors include Paul Erdős, Saharon Shelah, John H. Conway, Vladimir Boltyanski, Dmytro Karabash, Matthew Kahle, and Francisco Martinez-Figueroa.

The first four books that Soifer composed were published on his own without submitting them to any publisher. Springer published his last ten books, including *The Mathematical Coloring Book: Mathematics of Coloring and the Colorful Life of Its Creators* (2009) and *The Scholar and the State: In Search of Van der Waerden* (Springer 2015). Some of Soifer's books were translated into French and Russian.

Now that *The New Mathematical Coloring Book* has published, Soifer will work on two more books to publish by Springer: *Memory in Flashback: A Mathematician's Adventures on Both Sides of the Atlantic* and *Problems of pgom Erdős* (a joint book with Paul Erdős).

On a lighter side of life, Alexander enjoys Colorado's majestic mountains and badminton at the university gymnasium. His best friend, collie Bellissimo, takes him for a stroll around nearby Quail Lake no matter rain or snow.

Part I
Merry-Go-Round

Chapter 1
A Story of Colored Polygons and Arithmetic Progressions

"Have you guessed the riddle yet?" the Hatter said, turning to Alice again.
"No, I give it up," Alice replied. "What's the answer?"
"I haven't the slightest idea," said the Hatter.
"Nor I," said the March Hare.

– Lewis Carroll, *A Mad Tea-Party*
Alice's Adventures in Wonderland, 1866

1.1 The Story of Creation

I recall April of 1970. The 30 jury members of the Fourth Soviet Union National Mathematical Olympiad, of which I was one, stayed at a fabulous white castle with a white watch tower, halfway between the cities of Simferopol and Alushta, nestled in the sunny hills of Crimea, surrounded by the Black Sea. This castle should be familiar to movie buffs: in 1934, the Russian classic film *Vesyolye Rebyata* (*Jolly Fellows*) was filmed here by Sergei Eisenstein's long-term assistant, director Grigori Aleksandrov. The problems had been selected and sent to printers. The Olympiad was to take place a day later when something shocking occurred.

A mistake was found in the only solution the jurors had for the problem created by Nikolai (Kolya) B. Vasiliev, the vice-chair of the Soviet Union Mathematical Olympiad, a fine problem creator, and the head of the Problems Section of the journal *Kvant* from its inception in 1970 to the day of his untimely passing. What should we do? This question virtually monopolized our lives at the time.

We could have just crossed out this problem on each of the 600 printed problem sheets. In addition, we could have selected a replacement problem, but we would have had to write it in chalk by hand in every examination room, as there would not have been time to print it. Both options were pretty embarrassing, desperate solutions for the jury of the National Olympiad, chaired by the great mathematician Andrej Nikolaevich Kolmogorov, who arrived the day before and was to take part in approving the problems. The best resolution, surely, would have been to solve the problem, especially because its statement was quite beautiful and also since we had no counter example to it.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_1

Even today, over half a century later, I can close my eyes and see how each of us, 30 judges, all fine problem solvers, worked on the problem. A few sat at the table as if posing for Rodin's *Thinker*. Some walked around as if measuring the room's dimensions. Andrei Suslin, who would later prove the famous Serre's conjecture[1], went out for a thinking hike. Someone was lying on a sofa with his eyes closed. You could hear a fly. The intense thinking seemed to stop the time inside the room. We were unable, however, to stop the time outside. Night fell, and with it fell our hopes for solving the problem in time.

Suddenly, the silence was interrupted as a victorious outcry "I got it!" echoed through the halls and the watch tower of the castle. It came from Aleksandr "Sasha" Livshits, an undergraduate student at Leningrad (St. Petersburg) State University and a former winner of the Soviet Union National Mathematical Olympiad and the International Mathematical Olympiad (IMO) (a perfect score of 42 at the 1967 IMO in Yugoslavia).[2] His number theoretic solution used the method of trigonometric sums. This, however, was, the least of our troubles: we immediately translated the solution into the elementary language of colored polygons.

Now, we had options. A consensus was reached to leave the problem in. The problem and its solution were too beautiful to be thrown away. We knew, though, that the chances of receiving a single solution from 600 bright high school Olympians were extremely slim. Indeed, nobody solved it.

---

[1] Daniel Quillen proved it independently and received the Fields Medal primarily for it.

[2] Andrei Suslin informs me that as of 1991, Sasha worked as a computer programmer in Leningrad. On June 26, 2019, I found Professor Grigory Rozenblioum, who like Sasha Livshits, Andrey Suslin, and I, was an undergraduate member of the 1970 Soviet Math. Olympiad's jury. He shared with me the tragic life story of Sasha Livshits. Even though Sasha graduated from the Leningrad State University with high honors and, in 1975, defended a most impressive PhD dissertation, he was unable to find a research job due to being labeled "Jewish" in the land of anti-Semitism. Following several years of working in provincial Syktyvkar, Sasha's thesis advisor Professor Anatoly Vershik helped the 40-year-old Sasha to get a university position. Soon, Sasha defended a Russian doctoral degree (which is qualitatively higher than a PhD). Yet, in 2000, a mental illness caught up with Sasha, and, in 2008, aged 58, he succumbed to cancer. His collected works were published posthumously in 2014: http://www.mathsoc.spb.ru/pers/livshits/ANLivshits_book.pdf.

Aleksandr Nakhimovich "Sasha" Livshits with a typical USSR Mathematical Olympiad first prize: a giant pile of mathematical books. (The photograph was first published in Livshits, A. N., *Dynamic systems, ergodic theory, formal languages*, 2014, edited by Dr. Andrei Lodkin, the copyright holder, whom I thank for giving me the kind permission to use this photograph in my book.)

1.2 The Problem of Colored Polygons

Here is the problem.

Problem 1.1 (N.B. Vasiliev; IV Soviet Union National Mathematical Olympiad, 1970). The vertices of a regular n-gon are colored in finitely many colors (each vertex in one color) in such a way that for each color all vertices of that color form themselves a regular polygon, which we will call a *monochromatic* polygon. Prove that among the monochromatic polygons, there are two polygons that are congruent. Moreover, the two congruent monochromatic polygons can always be found among the monochromatic polygons with the least number of vertices.

I first told this story and the problem in my 1994 Colorado Mathematical Olympiad book [Soi9]. It appeared in the chapter "Further Explorations," and as such, I left the pleasure of discovering a solution to my readers. It is time for me to share proof with you.

Solution of Problem 1.1 by Aleksandr Livshits (in "polygonal translation"): Let me divide the problem into three parts: preliminaries, tool, and proof.

Preliminaries Given a system S of vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ in the plane with a Cartesian coordinate system, all emanating from the origin O. We would call this system S *symmetric* if there is an integer k, $1 \le k < n$, such that the rotation of every vector of S about O through the angle $\frac{2\pi k}{n}$ transforms S into itself.

Of course, the sum $\sum \vec{v_i}$ of all vectors of a symmetric system is $\vec{0}$ because $\sum \vec{v_i}$ does not change under rotation through the angle $0 < \frac{2\pi k}{n} < 2\pi$.

Place a regular n-gon P_n in the plane so that its center coincides with the origin O. Then, the n vectors drawn from O to all the vertices of P_n form a symmetric system (Fig. 1.1).

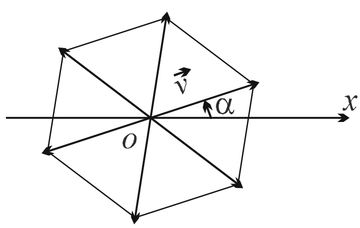


Fig. 1.1 The n vectors forming a symmetric system

Let $\vec{v}$ be a vector emanating from the origin O and making the angle α with the ray OX (Fig. 1.1). The symbol T^m will denote a transformation that maps $\vec{v}$ into the vector $T^m\vec{v}$ of the same length as $\vec{v}$ but making the angle $m\alpha$ with OX (Fig. 1.2).

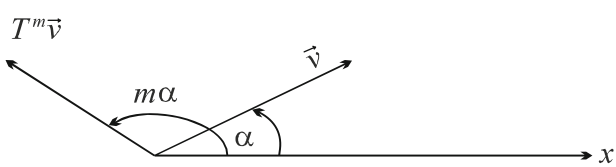


Fig. 1.2 Transformation T^m

To check your understanding of these concepts, please prove the following tool on your own.

Tool 1.2 Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ be a symmetric system S of vectors that transforms into itself under the rotation through the angle $0 < \frac{2\pi k}{n} < 2\pi$, $1 \leq k < n$, (you can think of $\frac{2\pi k}{n}$ as the angle between the two neighboring vectors of S). A transformation T^m applied to S produces the system $T^m S$ of vectors $T^m\vec{v_1}, T^m\vec{v_2}, \ldots, T^m\vec{v_n}$ that is symmetric if n does not divide km. If n divides km, then $T^m\vec{v_1} = T^m\vec{v_2} = \ldots = T^m\vec{v_n}$.

Solution of Problem 1.1 We will argue by contradiction. Assume that the vertices of a regular n-gon P_n are colored in r colors and we have, subsequently, r monochromatic polygons: n_1-gon P_{n_1}, n_2-gon P_{n_2}, ..., n_r-gon P_{n_r} such that no pair of congruent monochromatic polygons is created, i.e.,

$$n_1 < n_2 < \ldots < n_r.$$

We create a symmetric system S of n vectors going from the origin to all vertices of the given n-gon P_n. In view of tool 1.2, the transformation T^{n_1} applied to S produces a symmetric system $T^{n_1}S$. The sum of vectors in a symmetric system $T^{n_1}S$ is zero, of course.

On the other hand, we can first partition S in accordance with its coloring into r symmetric subsystems S_1, S_2, ..., S_r, obtain $T^{n_1}S$ by applying the transformation T^{n_1} to each system S_i separately, and then combine all $T^{n_1}S_i$. By tool 1.2, $T^{n_1}S_i$ is a symmetric system for $i = 2, \ldots, r$, but $T^{n_1}S_1$ consists of n_1 identical nonzero vectors. Therefore, the sum of all vectors of $T^{n_1}S$ is not zero. This contradiction proves that the monochromatic polygons cannot be all noncongruent. ■

Prove the last sentence of problem 1.1 on your own:

Problem 1.3 Prove that in the setting of problem 1.1, the two congruent monochromatic polynomials must exist among the monochromatic polynomials with the least number of vertices.

Readers familiar with complex numbers may have noticed that in the proof of problem 1.1, we can choose the given n-gon P_n to be inscribed in a unit circle and position P_n with respect to the axes so that the symmetric system S of vectors could be represented by complex numbers, which are precisely all nth degree roots of 1. Then, the transformation T^m would simply constitute raising these roots to the mth power.

1.3 Translation into the Language of Arithmetic Progressions

You might be wondering what this striking problem of colored polygons has in common with arithmetic progressions (APs), which are part of the section's title. Actually, everything! Problem 1.1 can be nicely translated into the language of infinite arithmetic progressions, or APs for short.[3]

[3] An infinite sequence $a_1, a_2, \ldots, a_n, \ldots$ is called an *arithmetic progression*, or AP, if for any integer $m > 1$, we have the equality $a_m = a_{m-1} + k$ for a fixed k, where k is a real number called *the constant difference* of the arithmetic progression.

Problem 1.4 In any coloring (partition) of a set of integers into finitely many infinite monochromatic APs, there are two APs with the same constant difference. Moreover, the largest constant difference necessarily repeats.

Equivalently,

Problem 1.5 Any partition of a set of integers into finitely many APs can be obtained *only* in the following way: N is partitioned into k APs, each with the same constant difference k (where k is a positive integer greater than 1); then, one of these APs is partitioned into finitely many APs with the same constant difference, and, then, one of these APs (at this stage, we have APs of two different constant differences) is partitioned into finitely many APs with the same constant difference, etc.

It was as delightful as it was valuable that our striking problem allowed two beautiful distinct formulations,: only because of that was I able to discover the prehistory of our problem.

1.4 Prehistory

Indeed, a year after I first published the story of this problem in 1994 [Soi9], I discovered that this unforgettable story actually had a prehistory! I became aware of it while watching a video recording of Ronald L. Graham's most elegant lecture "Arithmetic Progressions: From Hilbert to Shelah." To my surprise, Ron mentioned our problem in the language of partitions of integers into APs. Let me present the prehistory through the original e-mails so that you would discover the story the same way I did.

April 5, 1995; Soifer to Graham:

> In the beginning of your video "Arithmetic Progressions," you present a problem of partitioning integers into AP's. You refer to Mirsky–Newman. Can you give me a more specific reference to their paper? You also mention that their paper may not contain the result, but that it is credited to them. How come? When did they allegedly prove it?

April 5, 1995; Graham to Soifer:

> Regarding the Mirsky–Newman theorem, you should probably check with Erdős. I don't know that there ever was a paper by them on this result. Paul is in Israel at Tel Aviv University.

April 6, 1995; Soifer to Erdős:

> In the beginning of his video "Arithmetic Progressions," Ron Graham presents a problem of partitioning natural numbers into arithmetic progressions (with the conclusion that two progressions have the same constant difference). Ron refers to Mirsky–Newman. He gives no specific reference to their paper. He also mentions that their paper may not contain the result, but that it is credited to them ... Ron suggested that I ask you, which is what I am doing.

I have good reasons to find this out, as in my previous book and in the one I am writing now, I credit Vasiliev (from Russia) with creating this problem before early 1970. He certainly did, which does not exclude others from discovering it independently, before or after Vasiliev.

April 8, 1995; Erdős to Soifer:

In 1950 I conjectured that there is no exact covering system in which all differences are distinct, and this was proved by Donald Newman and [Leon] Mirsky a few months later. They never published anything, but this is mentioned in some papers of mine in the 50s (maybe in the Summa Brasil. Math. 11(1950), 113–123 [E50.07], but I am not sure).

April 8, 1995; Erdős to Soifer:

Regarding that Newman's proof, look at P. Erdős, *On a problem concerning covering systems*, Mat. Lapok 3(1952), 122–128 [E52.03].

I am looking at these early Erdős' articles. In his 1950 paper, he introduces covering systems of (linear) congruences. Since each linear congruence $x \equiv a(\mathrm{mod}\, n)$ defines an AP, we can talk about a covering system of APs and define it as a set of finitely many infinite APs, all with distinct constant differences, such that every integer belongs to at least one of the APs of the system. In his 1952 paper [E52.03], Paul introduces the problem for the first time in print (in Hungarian!)[4]:

I conjectured that if system [of k AP's with constant differences n_i respectively] is covering, then

$$\sum_{i=1}^{k} \frac{1}{n_i} > 1, \qquad (8)$$

that is the system does not uniquely cover every integer. This, however, I could not prove. For (8) Mirsky and Newmann [Newman] gave the following witty proof (the same proof was found later by Davenport and Rado as well).

Wow, Leon Mirsky, Donald Newman, Harold Davenport, and Richard Rado – quite a company of distinguished mathematicians, who worked on this bagatelle! Erdős then proceeds [E52.03] with presenting this company's proof of his conjecture, which uses infinite series and limits.

When viewing old video recordings of Paul Erdős' lectures at my University of Colorado at Colorado Springs, I found a curious historical detail that Paul mentioned in his March 16, 1989, lecture: he created this conjecture in 1950 while traveling by car from Los Angeles to New York!

---

[4] In English, this result was briefly mentioned, without proof, much later, in 1973 [E73.21] and 1980 [EG].

1.5 Completing the Go-Round

In 1959, Paul Erdős and János Surányi published a book on the theory of numbers. In the 2003 English translation [ESu2] of its 1996 second Hungarian edition, Erdős and Surányi present the result from Erdős' 1952 paper:

> In a covering system of congruences [of all distinct moduli], the sum of the reciprocals of the moduli is larger than 1.

Erdős and Surányi then repeat Mirsky–Newman–Davenport–Rado proof from Erdős' Hungarian 1952 paper [52.03] and call it Theorem 3. Then, comes a surprise:

> A. Lifsic [sic] gave an elementary solution to a contest problem that turned out to be equivalent to Theorem 3.
>
> Based again on exercises 9 and 10, it is sufficient to prove that it is not possible to cover the integers by finitely many arithmetic progressions having distinct differences in such a way that no two of them share a common element.

Erdős and Surányi then repeat the trick first discovered by us, the judges of the Soviet Union National Mathematical Olympiad, in May 1970, of converting the calculus problem into an elementary Olympiad problem about colored polygons! Here is how it goes:

> Wind the number line around a circle of circumference d. On this circle, the integers represent the vertices of a regular d-sided polygon ... The arithmetic progressions form the vertices of disjoint regular polygons that together cover all vertices of the d-sided polygon.

Erdős and Surányi continue by repeating, with credit, Sasha Livshits' solution of Kolya Vasiliev's problem of colored polygons that we have seen at the start of this chapter.[5] We have thus come a full circle, a merry-go-round from the Soviet Union Mathematical Olympiad to Paul Erdős and back to the same Olympiad. I hope you enjoyed the ride!

[5] Erdős and Surányi obtained the translation of the problem into the language of polygons and the polygonal proof from the 1988 Russian book [VE] by Nikolai Vasiliev and Andrei Egorov, which they credit for it. In this book, Vasiliev gives credit for the solution to Sasha Livshits – and, in a sign of extreme modesty, does not credit himself with creating this remarkable colored polygon problem independently from Erdős and in a different form.

Now, looking at the original 1996 Hungarian 2nd edition [ESu1] of the Erdős–Surányi book, I realize, with sadness, that Paul Erdős did not see the beauties of Sasha Livshits's proof – it did not appear in the Hungarian edition of 1996, the year Paul passed away. Clearly, Surányi alone added Livshits's proof to the 2003 English translation [ESu2] of the book.

Part II
Colored Plane

Chapter 2
Chromatic Number of the Plane: The Problem

A great advantage of geometry lies in the fact that in it the senses can come to the aid of thought and help find the path to follow.
– Henry Poincaré [Poi]

[I] can't offer money for nice problems of other people because then I will really go broke…
It is a very nice problem. If it were mine, I would offer $250 for it.
– Paul Erdős
Boca Raton, February 6, 1992

The most widely known problem in Euclidean Ramsey Theory is probably that of determining the chromatic number of the plane, $\chi(E^2)$.
– Ronald L. Graham and Eric Tressler [GT]

The unit distance graph in the plane … is simple enough to describe to a nonmathematician, and so enigmatic that finding its chromatic number is a new four-color map problem for graph theorists.
– Ronald L. Graham and Eric Tressler (Ibid.)

If Problem 8 [the chromatic number of the plane] takes that long to settle [as the Four-Color problem], we should know the answer by the year 2084.
– Victor Klee and Stan Wagon [KW]

Our good ole Euclidean plane, don't we know all about it? What else can there be after Pythagoras and Steiner, Euclid, and Hilbert? In this chapter, we will look at an open problem that exemplifies what is best in mathematics: Anyone can understand this problem; yet, no one has been able to conquer it in 73 years.

In August 1987, I attended an inspiring talk by Paul Halmos at Chapman College in Orange, California. It was entitled "Some problems you can solve, and some you cannot." This problem is an example of a problem that "you cannot solve."

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_2

"A fascinating problem... that combines ideas from set theory, combinatorics, measure theory, and distance geometry," write Hallard T. Croft, Kenneth J. Falconer, and Richard K. Guy in their book *Unsolved Problems in Geometry* [CFG].

"If Problem 8 takes that long to settle [as the celebrated Four-Color Conjecture], we should know the answer by the year 2084," write Victor Klee and Stan Wagon in their book *New and Old Unsolved Problems in Plane Geometry* [KW].

Are you ready? Here it is:

What is the smallest number of colors sufficient for coloring the plane in such a way that no two points of the same color are at a unit distance apart?

This number is called *the chromatic number of the plane* and is denoted by $\chi(E^2)$ or simply χ.

We will use R to denote the set of real numbers and the real line. The line equipped with the usual Euclidean distance, we will denote by E^1. Generalizing the line E^1, we get the Euclidean plane E^2 and the Euclidean space E^3, and we define the *n-dimensional* space R^n for any positive integer n as the set of all n-tuples $(x_1, x_2,..., x_n)$, where $x_1, x_2,..., x_n$ are real numbers. When the distance between two points $(x_1, x_2,..., x_n)$ and $(y_1, y_2,..., y_n)$ of R^n is defined by the equality

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2} \qquad (*)$$

we get the *Euclidean n-dimensional* space E^n. In other words, E^n is just the set R^n together with the distance d defined by (*).

To color the plane means to assign one color to every point of the plane. Please note that, here, we color without any restrictions and are not limited to "nice" tiling-like or map-like colorings. Given a positive integer n, we say that the plane is *n-colored*, if every point of the plane is assigned one of the given n colors.

Here, a *segment* will stand for just a two-point set (which are end points in a conventional treatment of a segment). Similarly, a *polygon* will stand for a finite set of points. A *monochromatic* set is a set, whose all elements are assigned the same color. In this terminology, we can formulate the chromatic number of the plane (CNP) problem as follows: What is the smallest number of colors sufficient for coloring the plane in a way that forbids monochromatic unit segments?

I do not know who first noticed the following result. Perhaps, Adam? Or Eve? To be a bit more serious, I do not think that ancient Greek geometers, for example, knew this nice fact, for they simply did not ask these kinds of questions!

Problem 2.1 (Adam and Eve). No matter how the plane is two-colored, it contains a monochromatic segment of length 1, i.e.,

$$\chi \geq 3.$$

Proof Toss on the two-colored plane an equilateral triangle T of side 1 (Fig. 2.1). We have only two colors, while T has three vertices (I trust you have not forgotten the Pigeonhole principle). Two of the vertices must be of the same color. They *are* at a distance 1 apart. ■

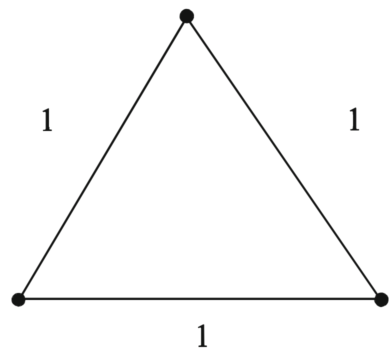


Fig. 2.1 At least 3 colors are necessary

We can do better than Adam and Eve:

Problem 2.2 No matter how the plane is three-colored, it contains a monochromatic segment of length 1, i.e.,

$$\chi \geq 4.$$

Proof by the Canadian Geometers, Brothers Leo and William Moser (1961, [MM]). Toss on the three-colored plane what we now call *the Mosers Spindle* (Fig. 2.2). Every edge in the spindle has the length 1.

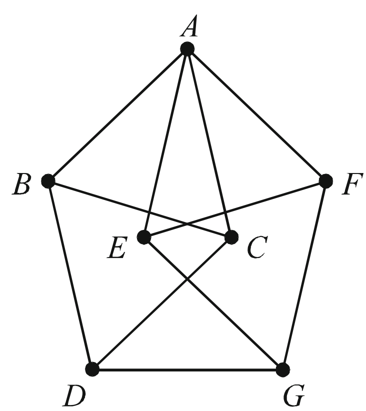


Fig. 2.2 The Mosers Spindle

Assume that the seven vertices of the spindle do not contain a monochromatic unit segment. Call the colors used in coloring the plane red, white, and blue. The solution now will faithfully follow the children's alphabet song "A B C D E F G .... ".

Let the point A be red, then B and C must be one white and one blue, respectively, and therefore, D must be red. Similarly, E and F must be one white and one blue, respectively, and therefore, G must be red. We have a monochromatic unit segment DG in contradiction to our assumption. ■

Observe The Mosers Spindle has worked for us in solving problem 2.2 precisely because *any* three vertices of the spindle contain two vertices that are at a distance 1 apart. This implies that *in* *a Mosers Spindle that forbids a monochromatic unit segment, at most two points can be of the same color.* Let us record this observation as a tool, which we will need later in Chapters 4 and 40.

Mosers' Tool 2.3 *Any* three vertices of the Mosers Spindle contain a unit segment. Consequently, in a Mosers Spindle that forbids a monochromatic unit segment, at most two vertices can be of the same color.

When I presented the Mosers' solution to high school mathematicians, everyone agreed that it was beautiful and simple. "But how do you come up with a thing like the spindle?", I was asked. As a reply, I presented a less elegant but a more naturally found solution. In fact, I would call it a second version of the same solution. Here, we touch on a curious aspect of mathematics. In mathematical texts, we often see the terms "second solution" and "third solution." However, which two solutions ought to be called distinct? We do not know. It is not defined and is thus a judgment call. Distinct solutions for one person could be viewed as versions of the same for another. It is interesting to notice that both versions were published in the same year, 1961, one in Canada and the other in Switzerland.

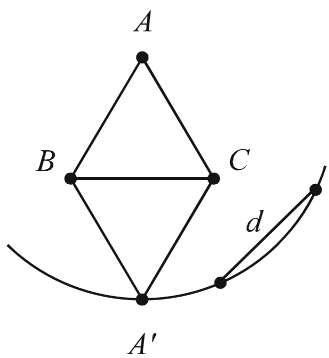


Fig. 2.3 At least 4 colors are necessary

Second Version of the Proof (Hugo Hadwiger, 1961, [Had4]). Assume that a three-colored red–white–blue plane does not contain a monochromatic unit segment. Then an equilateral triangle ABC of side 1 will have one vertex of each color (Fig. 2.3). Let A be red, then B and C must be one white and one blue, respectively. The vertex A' symmetric to A with respect to the side BC must be red as well. As we rotate our rhombus $ABA'C$ through *any* angle about A, the vertex A' will have to remain red due to the above argument. Thus, we get a whole red circle of radius AA'. Surely, it contains a cord d of length 1, both end points of which are red, in contradiction to our assumption. ■

Does an upper bound exist for χ? It is not immediately obvious. Can you find one? Think of tiling the plane with square tiles.

Problem 2.4 There is a 9-coloring of the plane that contains no monochromatic segments of length 1, i.e.,

$$\chi \leq 9.$$

Proof Tile the plane with unit squares. Now, we color one square in color 1 and its eight neighbors in colors 2, 3, ..., 9 (Fig. 2.4). The union of these 9 unit squares is a 3×3 square S, shown in bold. Translates of S (i.e., images of S under translations) tile the plane and determine how we color it in nine colors.

You can easily verify (do) that no distance d in the range $\sqrt{2} < d < 2$ is realized monochromatically in the plane. Thus, by shrinking all linear sizes by the factor of, say, 1.5, we get a 9-coloring that contains no monochromatic segments of length 1. (Observe: due to the above inequality, we have enough cushion so that it does not matter in which of the two adjacent colors we color the boundaries of the unit squares.) ■

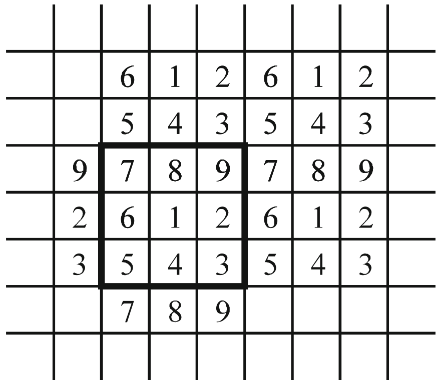


Fig. 2.4 9 colors suffice

Now that a tiling has helped us solve the above problem, it is natural to ask whether another tiling can help us improve the upper bound. One can indeed.

Problem 2.5 There is a 7-coloring of the plane that contains no monochromatic unit segments, i.e.,

$$\chi \leq 7.$$

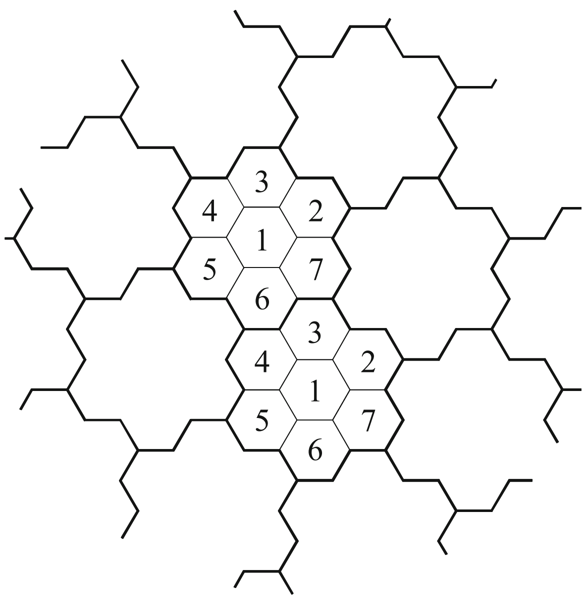


Fig. 2.5 A 7-coloring using a hexagonal tiling

Proof [Had3]. We can tile the plane by regular hexagons of side 1. Now, we color one hexagon in color 1 and its six neighbors in colors 2, 3, ..., 7 (Fig. 2.5). The union of these seven hexagons forms a "flower" P, a highly symmetric polygon P of 18 sides. Translates of P tile the plane and determine how we color the plane in seven colors. It is easy to compute (please do) that each color does not have monochromatic segments of any length d, where $2 < d < \sqrt{7}$. Thus, if we shrink all linear sizes by a factor of, say, 2.1, we will get a 7-coloring of the plane that has no monochromatic segments of length 1. (Observe: due to the above inequality, we have enough cushion so that it does not matter in which of the two adjacent colors we color the boundaries of the hexagons.) ■

This is the way the upper bound is proved in every book I know ([CFG] and [KW], for example). Yet, in 1982, the Hungarian mathematician László A. Székely found a clever way to prove the upper bound 7 without using hexagonal tiling.

Problem 2.6 (L. A. Székely, [Sze1]). Prove the upper bound $\chi \leq 7$ by tiling the plane with squares again.

Proof This is László Székely's proof from [Sze1]. His original picture needs a small correction in his Fig. 1, and boundary coloring needs to be addressed, which I am doing here. We start with a row of squares of diagonal 1, with cyclically alternating colors of the squares 1, 2, . . ., 7 (Fig. 2.6). We then obtain consecutive rows of colored squares by shifting the preceding row to the right through 2.5 square sides.

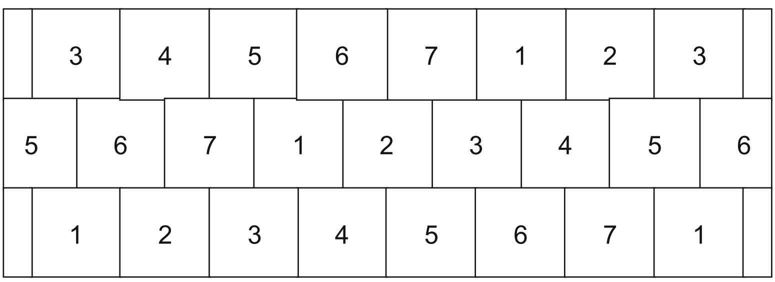


Fig. 2.6 A 7-coloring using square tiling

The upper and right boundaries are included in the color of each square, except the square's upper left and lower right corners. ■

In 1995, my former student and now a well-known puzzlist Edward Pegg, Jr. sent me two distinct 7-colorings of the plane. In the one I am sharing with you (Fig. 2.7), Ed uses 7-gons for six of the colors and tiny squares for the seventh color. In fact, the seventh color occupies only about one-third of 1% of the plane.

In Fig. 2.7, all thick black bars have a unit length. A unit of the tiling uses a heptagon and half a square.

The area of each square is 0.0041222051899307168162...

The area of each heptagon is 0.62265127164647629646...

Thus, the area ratio is 302.09620480194552853007836272658 28...

If one-third of 1% of the plane is removed, then the remainder can be six-colored with this tiling!

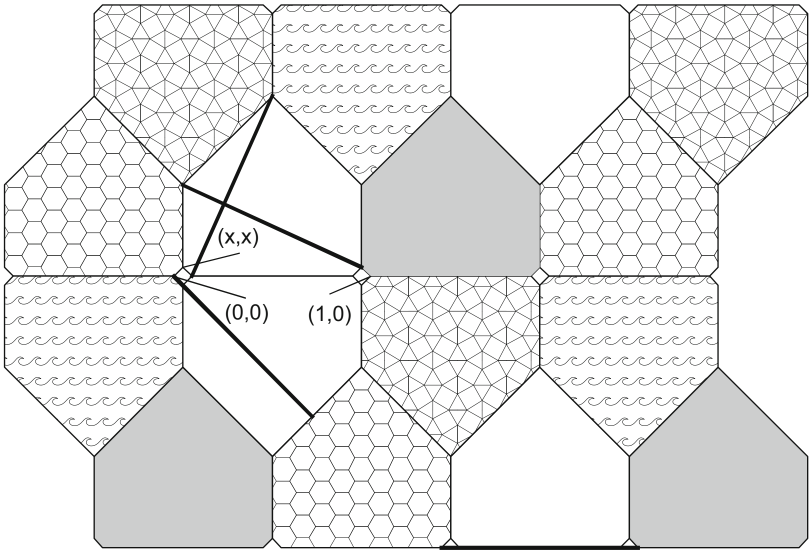


Fig. 2.7 Ed Pegg's 7-coloring with a small use of color 7

The lower bound for the chromatic number of the plane (problem 2.2) also has proofs that are fundamentally different from using the Mosers Spindle. In the early 1990s, I received from my colleague and friend Klaus Fischer of George Mason University a finite configuration of the chromatic number 4, different from the Mosers Spindle. Klaus had no idea who created it, so I commenced backtracking this construction. Klaus got it from our friend and colleague Heiko Harborth of Braunschweig Technical University, Germany, who, in turn, referred me to his source, Solomon W. Golomb of the University of Southern California, the famous inventor of polyomino. Solomon invented this graph as well and described it in the September 10, 1991, letter to me [Gol1]:

> The example you sketched of a 4-chromatic unit-distance graph with ten vertices is original with me. I originally thought of it as a 3-dimensional structure (the regular hexagon below, the equilateral triangle above it in a plane parallel to it), and all connected by unit-length toothpicks. The structure is then allowed to collapse down into the plane, to form the final Figure (Fig. 2.8). I have shown it to a number of people, including the late Leo Moser, Martin Gardner, and Paul Erdős, as well as Heiko Harborth. It is possible that Martin Gardner may have used it in one of his columns, but I don't remember. Besides my example and Mosers' original example (which I'm reasonably sure I have seen in Gardner's column), I have not seen any other "fundamental" examples. I believe what I had suggested to Dr. Harborth in Calgary was the possibility of finding a 5-chromatic unit-distance graph, having a much larger number of edges and vertices.

"The possibility of finding a 5-chromatic unit-distance graph" was on the minds of most of us, who worked on this problem. Does it exist? You will find a definitive answer later in this book.

In the consequent September 25, 1991, letter [Gol2], Sol Golomb informed me that he likely found this example, which I will naturally call *the Golomb Graph*, in the period 1960–1965.

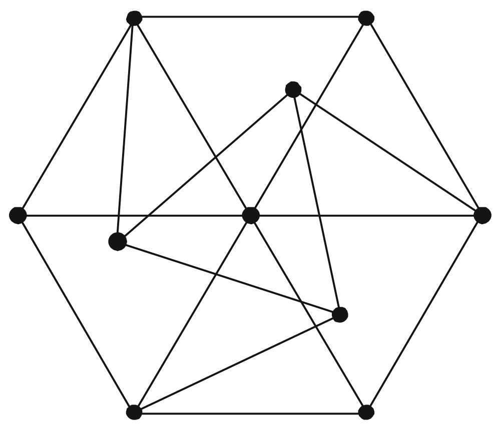


Fig. 2.8 The Golomb graph

Second Solution of Problem 2.2 Just toss the Golomb graph with all edges of unit length on a three-colored (red, white, and blue) plane (Fig. 2.8). Assume that in the graph, there are no adjacent vertices of the same color. Let the center vertex be colored red, then, since it is connected by unit edges to all vertices of the regular hexagon H, the vertices of H must be colored white and blue in an alternating manner. All vertices of the central equilateral triangle T are connected by unit edges to the three vertices of H of the same color, say, white. However, then, white cannot be used in coloring T, and, thus, T is colored red and blue. However, this implies that two of the vertices of T are assigned the same color. This contradiction proves that 3 colors are not enough to properly color the 10 vertices of the Golomb graph, let alone the whole plane. ■

It is amazing that the pretty easy solutions of problems 2.2 and 2.4 provided us with the best bounds known to mathematics prior to 2018 for the chromatic number of the plane χ *in the general case*. They were published more than 60 years ago (in fact, they are older than that: see the next chapter for an intriguing historical account). Still, all we knew at the time of the first edition of this book was

$$\chi = 4, \text{or } 5, \text{or } 6, \text{or } 7.$$

A very broad spread! Which do you think is the exact value of χ? The legendary Paul Erdős believed that it was $\chi \geq 5$.

The renown American geometer Victor Klee of the University of Washington shared with me in 1991 a highly intriguing story. In 1980, he lectured in Zürich, Switzerland. The celebrated 77-year-old mathematician Bartel L. van der Waerden (whom we will frequently meet later in this book) was in attendance. When Vic presented the state of this problem, Van der Waerden became very interested. Right there and then, during Vic's lecture, Bartel started working on the problem. He tried to prove that $\chi = 7$.

For many years, I believed that $\chi = 7$ (you will find my thoughts on the matter in *Predicting the Future*, later in this book). Paul Erdős used to say that

> God has a transfinite Book, which contains all theorems and their best proofs, and if He is well intentioned toward those, He shows them the Book for a moment.

If I ever deserved the honor and had a choice, I would have asked to peek at the page with the chromatic number of the plane problem. Wouldn't you?

Chapter 3
Chromatic Number of the Plane: A Historical Essay

[I] cannot trace the origin of this problem.
– Paul Erdős, 1961

[This is] a long-standing open problem of Erdős.
– Hallard T. Croft, 1967

It is often easier to be precise about Ancient Egyptian history than about what happened among our contemporaries.
– Nicolaas Govert de Bruijn
Eindhoven, July 5, 1995
e-mail to A. Soifer

It happened a long time ago and is not true.
– An old Russian joke

It is natural for one to inquire into the authorship of one's favorite problem. So, in 1991, I turned to countless articles and books. Some of the information that I found appears here in Table 3.1 and Diagram 3.1 – take a look. Are you confused? I was too!

As you can see, Douglas R. Woodall credits Martin Gardner, who, in turn, refers to Leo Moser. Hallard T. Croft calls it "a long-standing open problem of Erdős", Gustavus J. Simmons credits "Erdős, [Frank] Harary, and [William Thomas] Tutte," while Paul Erdős himself "cannot trace the origin of this problem"! Later Erdős credits "Hadwiger and Nelson," while Victor Klee and Stan Wagon state that the problem was "posed in 1960–61 by M. Gardner and Hadwiger." Croft comes again, this time with Kenneth J. Falconer and Richard K. Guy, to cautiously suggest that the problem is "apparently due to E. Nelson" [CFG]. Yet, Richard Guy did not know who "E. Nelson" was and why Guy and his coauthors "apparently" attributed the problem to him (my conversation with Richard Guy on the back seat of a taxi in Keszthely, Hungary, when we both attended Paul Erdős' 80th birthday conference in August of 1993).

Thus, at least *seven* (!) mathematicians were credited with creating this problem: Paul Erdős, Martin Gardner, Hugo Hadwiger, Frank Harary, Leo Moser, Edward Nelson, and William T. Tutte – a great group of mathematicians to be sure. But it was hard for me to believe that they all created the problem, be it independently or all seven together.

© Alexander Soifer 2024
A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_3

Table 3.1 Who created the chromatic number of the plane problem?

Publication	Year	Author(s)	Problem creator(s) or source named
[Gar2]	1960	Gardner	"**Leo Moser** ...writes..."
[Had4]	1961	Hadwiger (after Klee)	**Nelson**
[E61.22]	1961	Erdős	"I cannot trace the origin of this problem"
[Cro]	1967	Croft	"A long[18]-standing open problem of **Erdős**"
[Woo1]	1973	Woodall	**Gardner**
[Sim]	1976	Simmons	**Erdős**, **Harary**, and **Tutte**
[E80.38] [E81.23] [E81.26]	1980–1981	Erdős	**Hadwiger** and **Nelson**
[CFG]	1991	Croft, Falconer, and Guy	"Apparently due to **E. Nelson**"
[KW]	1991	Klee and Wagon	"Posed in 1960–61 by **M. Gardner** and **Hadwiger**"

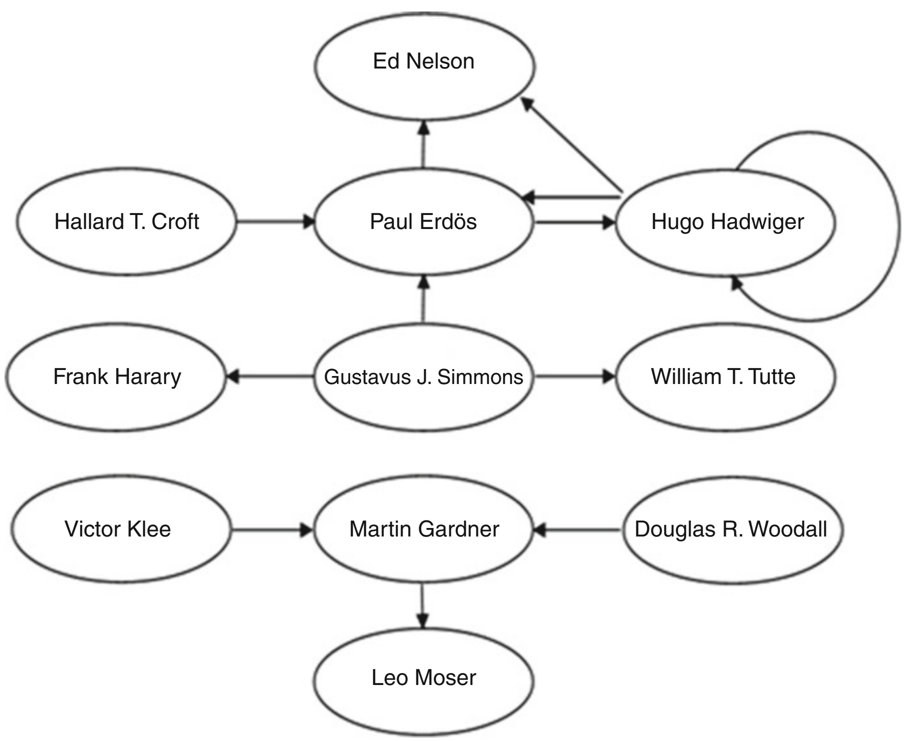


Diagram 3.1 Who created the chromatic number of the plane problem?

I felt an urge, akin to that of a private investigator, a Sherlock Holmes, to untangle the web of conflicting accounts. It took 6 months to solve this historical puzzle. A good number of mathematicians, through conversations and e-mails, contributed their insights: Branko Grünbaum, Peter D. Johnson, Tony Hilton, Ron Graham, and Klaus Fischer first come to mind. I am especially grateful to Paul Erdős, Victor Klee, Martin Gardner, Edward Nelson, and John Isbell for contributing critically important pieces of the puzzle. Only their accounts, recollections, and congeniality made these findings possible.

What follows is my 1991 investigation into the history of the problem. Today, I still stand by this research. It gives me great sadness to see that the players and informants of this *Story of Creation* are no longer with us. As my homage to them, I would like to list them here, alphabetically:

Paul Erdős, 26 March 1913–20 September 1996
Martin Gardner, 21 October 1914–22 May 2010
Ronald L. Graham, 31 October 1935–6 July 2020
Branko Grünbaum, 2 October 1929–14 September 2018
Hugo Hadwiger, 23 December 1908–29 October 1981
John R. Isbell, 27 October 1930–6 August 2005
Victor L. Klee, Jr., 18 September 1925–17 August 2007
Leo Moser, 11 April 1921–9 February 1970
William "Willy" O.J. Moser, 5 September 1927–28 January 2009
Edward Nelson, 4 May 1932–10 September 2014

I commenced my investigation on July 6, 1991, by mailing a page-long handwritten letter to Paul Erdős [Soi91/7/6ltr]. I open it by sharing my plans and then ask several questions. The first question is of our interest here:

> Dear Paul,
>
> I am writing a book, "*Mathematical Coloring Book*."
>
> It will have two parts: one about properties of colored objects (n-colored plane and chromatic number of the plane, colored numbers and Schur Theorem, colored polygons, etc.); the other part about coloring (coloring as a mean of solving tiling problems, coloring a map, etc.)
>
> I welcome your advice, problems for inclusion, including open problems (with their history if possible).
>
> In particular, I have a few questions for you.
>
> I am trying to reconstruct the history of the problem asking for the chromatic number of the plane $\chi(E^2)$ (minimal number of colors that color the plane without monochromatic segments of length 1.
>
> Folklore has it that this is your problem. Gustavus Simmons says that in an article. Martin Gardner mentions it in 1960 and says that he heard it from Leo Moser. Victor Klee and Stan Wagon, in their almost-published book, say that this problem was born in 1960. Woodall starts the story of the problem from Gardner's October 1960 mention of Leo Moser.
>
> I think the problem is older than 1960, and it is your problem. Please, let me know the true history, as I wish to give credit where credit is due, especially since this is my

favorite problem in all of mathematics. Did you originally conjecture that $\chi(E^2) = 4$? What do you think now: 4, 5, 6, or 7?

This seems to be my first mention of an aspiration to write a book and the first time I gave this book her title, *Mathematical Coloring Book*. I never thought it would become an 18-year-long obsession that resulted, as you know, in the first edition of this book.

In the July 12, 1991, letter, Paul writes [E91/7/12ltr]:

I first heard of the chromatic number of the plane problem from Leo Moser in 1958–1962.

On August 10, 1991, Paul shares his appreciation of the problem, for which he could not claim the authorship [E91/8/10ltr]:

The problem about the chromatic number of the plane is unfortunately not mine.

In a series of letters of July 12, 1991; July 16, 1991; August 10, 1991; and August 14, 1991, Paul formulates for me a good number of problems related to the chromatic number of the plane that he did create. We will look at some of Erdős' problems in the following chapters.

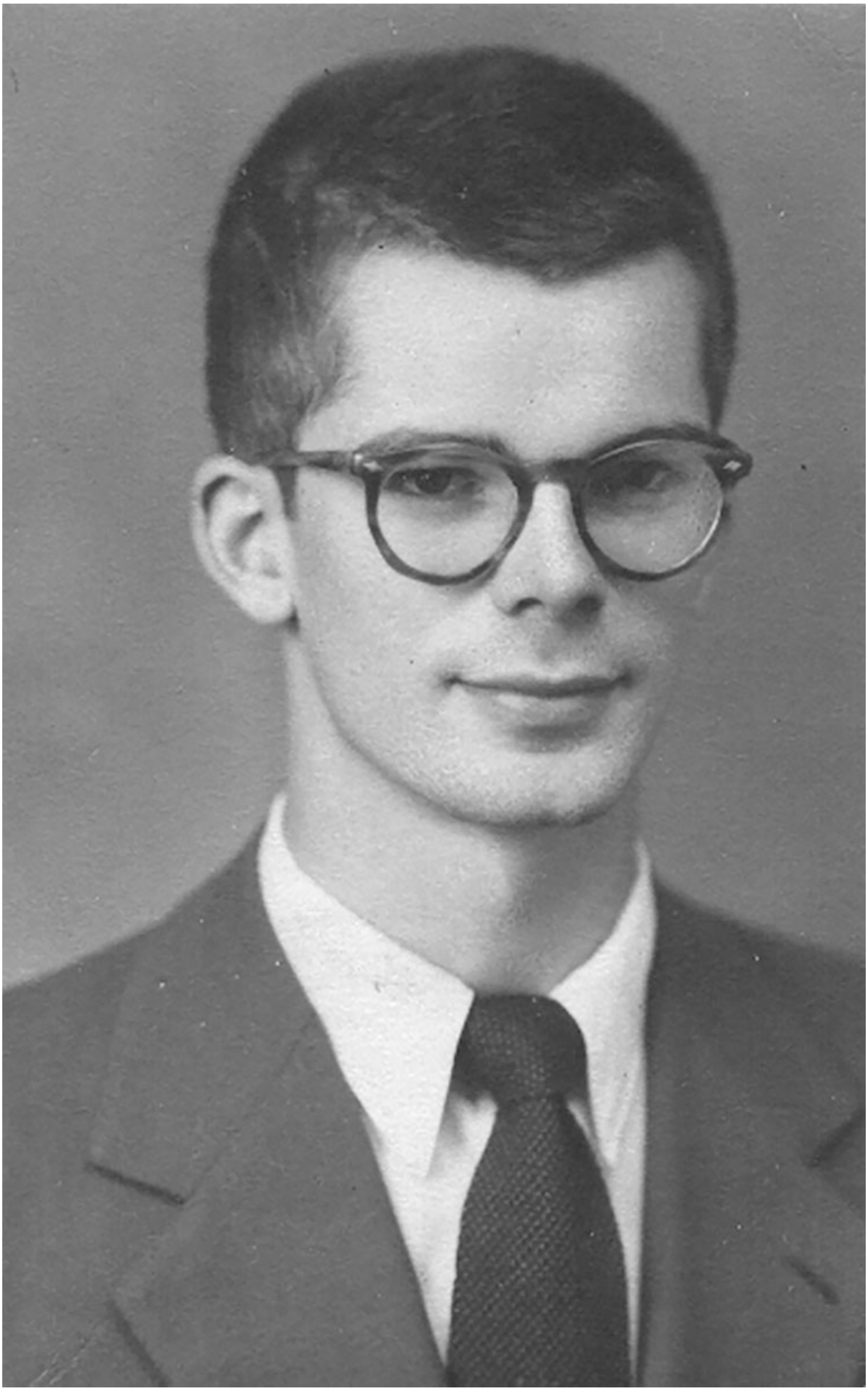
Having established that the author was not Paul Erdős, I moved down the list of the "candidates," and, on August 8, 1991, and again on August 30, 1991, I wrote to Victor Klee, Edward Nelson, and John Isbell. I shared with them my Table 3.1 and asked them to tell me what they knew about the birth of the problem. I also interviewed Princeton Professor Nelson over the phone on September 18, 1991.

Edward Nelson created what he named "a second four-color problem" (the first being the famous four-color problem of map coloring), which we will discuss in Part 4). In his October 5, 1991, letter [Nel2], he conveys the *Story of Creation*:

Dear Professor Soifer:

In the autumn of 1950, I was a student at the University of Chicago and among other things was interested in the Four-Color problem, the problem of coloring graphs topologically embedded in the plane. These graphs are visualizable as nodes connected by wires. I asked myself whether a sufficiently rich class of such graphs might possibly be subgraphs of one big graph whose coloring could be established once and for all, for example, the graph of all points in the plane with the relation of being unit distance apart (so that the wires become rigid, straight, of the same length, but may cross). The idea did not hold up, but the other problem was interesting in its own right and I mentioned it to several people.

I asked Professor Nelson for his photograph from the time when he created this problem. He referred me to his photograph published in *Time Magazine* from December 1949. I found the article about Eddie Nelson's successes with his photograph and called *Time Magazine*. They informed me that *if* they find that photograph in their archive, which was not given, they will charge me $300 for a one-time right to reproduce it in my book. Nelson was clearly unhappy with the *Time*'s reply. He found a photograph of himself ca. 1950 and generously sent it to me as a gift, which I share with you here.



Eddie Nelson, c. 1950. (Courtesy of Edward Nelson)

One of the people Eddie Nelson mentioned the problem to was John Isbell. Almost half a century later, Isbell still remembers the story very vividly when on August 26, 1991, he shares it with me [Isb1]:

> Dear Professor Soifer,
>
> I should certainly like to receive any future *Geombinatorici*, and I might contribute. There is an annoying problem I talked with Erdős and Pach about a few (5–6) years back; I sort of promised Pach a preprint, but didn't get enough results to publish. Maybe I could write it up as a problem or two and put in the bits of result I did get as a background.
>
> Authorship of plane chromatic no. problem: Of course I can't comment on Croft's attribution to Erdős "long ago" except that it is a vague reference (but you cite Croft et al. 1991 against early Croft, maybe because he learned [from Klee?] about the following.
>
> Ed Nelson told me the problem and $\chi \geq 4$ in November 1950, unless it was October – we met in October. I said what upper bound have you, he said none, and I worked out 7. I was a senior at the time (B.S., 1951). I think Ed had just entered U. Chicago as a nominal sophomore and taken placement exams which placed him a bit ahead of me, say a beginning graduate student with a gap or two in his background. I certainly mentioned the problem to other people between 1950 and 1957; Hugh Spencer Everett III, the author of the many-worlds interpretation of quantum mechanics, would certainly be one, and Elmer Julian Brody who did a doctorate under Fox and has long been at the Chinese University of Hong Kong and is said to be into classical Chinese literature would be another. I mentioned it to Vic Klee in 1958 $\pm$ 1 ...
>
> I don't see that Woodall's attribution to Gardner, who attributes it elsewhere, is worth a plugged nickel. If you said Erdős, Moser, and Nelson independently, you would probably be accurate in the eyes of the Recording Angel.

In September 1991, I had a most enjoyable phone conversation with the distinguished geometer Victor Klee. He too remembered hearing the problem from John Isbell in 1957–1958. In fact, it took place before September 1958 when Professor Klee left for Europe. There, *Klee passed this problem to Hugo Hadwiger*, who was collecting problems for the book *Open Problems in Intuitive Geometry,* to be written jointly by Erdős, Fejes Toth, Hadwiger, and Klee. To my great regret, this great book-to-be has never materialized.

Gustavus J. Simmons [Sim], in giving credit for the problem to "Erdős, Harary, and Tutte," no doubt had in mind their joint 1965 paper in which the three famous authors defined the dimension of a graph (see Chapter 13 on this). The year of 1965 was way too late for our problem's creation, and, besides, the three authors have never made any claims to such a discovery.

What were the roles of Paul Erdős, Martin Gardner, and Leo Moser in the *Story of Creation*? I am prepared to answer these questions, all except one: I am leaving to others to research Leo Moser's archive (it used to be maintained by his brother Willy Moser at McGill University in Montreal) and find out when and from whom Leo Moser came by the problem. What is important to me is that he did not create it independently from Edward Nelson, as Paul Erdős informed me in his July 16, 1991, letter [E91/7/16]:

I do not remember whether Moser in 1958 [possibly on June 16, 1958, the date from which we are lucky to have a photo record] told me how he heard the problem on the chromatic number of the plane, I only remember that it was not his problem.

Paul Erdős (left) and Leo Moser, June 16, 1958. (Courtesy of Paul Erdős)

Yet, Leo Moser made a valuable contribution to the survival of the problem. He gave it to both Paul Erdős and the wonderful mathematics expositor Martin Gardner. Gardner, due to his fine taste, recognized the value of this problem and included it in his October 1960 *Mathematical Games* column in *Scientific American* ([Gar2]), with the acknowledgment that he received it from Leo Moser of the University of Alberta. Thus, the credit for the first *publication* of the problem goes to Martin Gardner. It is beyond me why so many authors of articles and books, as late as 1973 ([Woo1], for example), gave credit for the *creation* of the problem to Martin Gardner, something he himself had never claimed. In my 1991 telephone conversation with him, Martin told me for a fact that the problem was not his, and he promptly listed Leo Moser as his source, both in print and in his archive, which he checked as I was waiting on the line.

Moreover, some authors ([KW], for example) who knew Edward Nelson's authorship, still credited Martin Gardner and Hugo Hadwiger as late as in 1991 because, it seems, only written, preferably published, word was acceptable to them. Following this logic, the creation of the celebrated four-color map coloring problem (4CP) must be attributed to Augustus De Morgan, who first *wrote* about it in his October 23, 1852, letter to William Rowan Hamilton, or better yet to Arthur Cayley, whose 1878 abstract included the *first non-anonymous publication* of the problem.[1] Yet, we all seem to agree that the 20-year-old Francis Guthrie created 4CP, even though he did not publish or even write a word about it! (See Part IV for more on this.)

Of course, a lone self-serving statement would be too weak a foundation for a historical claim. On the other hand, independent disinterested testimonies corroborating each other comprise as solid a foundation for the attribution of the credit as any publication. This is precisely what my inquiry has produced. Here is just one example of Nelson and Isbell's selflessness. Edward Nelson writes to me on August 23, 1991 [Nel1]:

> I proved nothing at all about the problem.

John Isbell corrects Nelson in his September 3, 1991, letter [Isb2]:

> Ed Nelson's statement which you quote, "I proved nothing at all about the problem," can come only from a failure of memory. He proved to me that the number we are talking about is ≥ 4, by precisely the argument in Hadwiger 1961. Hadwiger's attribution (on Klee's authority) of that inequality to me can only be Hadwiger's or Klee's mistake.

This brings us to the issue of the authorship of the bounds for χ:

$$4 \leq \chi \leq 7.$$

Once again, the entire literature is off the mark by giving credit for the first proofs to Hadwiger and the Mosers. Yes, in 1961, the famous Swiss geometer Hugo Hadwiger published [Had4] the chromatic number of the plane problem together with proofs of both bounds. But he writes there (*and nobody reads!!*):

> We thank Mr. V. L. Klee (Seattle, USA) for the following information. The problem is due to E. Nelson; the inequalities are due to J. Isbell.

Hadwiger does go on to say:

> Some years ago the author [i.e., Hadwiger] discussed with P. Erdős questions of this kind.

Did Hadwiger insinuate that he created the problem independently from Nelson? We will never know for sure, but I have my doubts about Hadwiger's (co)authorship. Hadwiger jointly with Hans Debrunner published an excellent, long problem paper in 1955 [HD1] that was extended to their wonderful, famous book in 1959 [HD2]; see also its 1964 English translation [HDK] with Victor Klee and the 1965 Russian translation [HD3] edited by the famous Russian geometer and expositor Isaac Moiseevich Yaglom. All these books (and Hadwiger's other papers) included a number of "questions of this kind," but did not once

[1] First publication could be attributed to De Morgan, who mentioned the problem in his 1860 book review in *Athenaeum* [DeM4], albeit anonymously – see more on this in Section 18.

include the chromatic number of the plane problem. Moreover, it seems to me that the problem in question is somewhat out of Hadwiger's "character": in all problems "of this kind," he preferred to consider closed sets rather than arbitrary sets, in order to take advantage of topological tools.

I shared with Paul Erdős these twofold doubts about Hadwiger independently creating the problem. It was especially important because Hadwiger in the above-quoted text mentioned Erdős as his witness or coauthor of sorts. Paul replied to me in his July 16, 1991, letter [E91/7/16] as follows:

> I met Hadwiger only after 1950, thus I think Nelson has priority (Hadwiger died a few years ago, thus I cannot ask him, but I think the evidence is convincing).

During his talk at the 25th Southeastern International Conference on Combinatorics, Graph Theory, and Computing at Florida Atlantic University, Boca Raton, Florida, at 9:30–10:30 a.m. on Thursday, March 10, 1994, Paul Erdős summarized the results of my historical research in the characteristically Erdősian style ([E94.60])[2]:

> There is a mathematician called Nelson who in 1950 when he was an epsilon, that is he was 18, discovered the following question. Suppose you join two points in the plane whose distance is 1. It is an infinite graph. What is [the] chromatic number of this graph?
>
> Now, de Bruijn and I showed that if an infinite graph, which is chromatic number k, it always has a finite subgraph, which is chromatic number k. So, this problem is really [a] finite problem and not an infinite problem. And it was not difficult to prove that the chromatic number of the plane is between 4 and 7. I would bet it is bigger than 4, but I am not sure. And the problem is still open.
>
> If it would be my problem, I would certainly offer money for it. You know, I can't offer money for every nice problem because I would go broke immediately. I was asked once what would happen if all your problems would be solved, could you pay? Perhaps not, but it doesn't matter. What would happen to the strongest bank if all the people who have money there would ask for money back? Or what would happen to the strongest country if they suddenly ask for money? Even Japan or Switzerland would go broke. You see, Hungary would collapse instantly. Even the United States would go broke immediately . . .
>
> Actually, it was often attributed to me, this problem. It is certain that I had nothing to do with the problem. I first learned the problem, the chromatic number of the plane, in 1958, in the winter, when I was visiting [Leo] Moser. He did not tell me from where this nor the other problems came from. It was also attributed to Hadwiger, but Soifer's careful research showed that the problem is really due to Nelson.

The leader of Ramsey Theory, Ronald L. Graham, also endorses the results of my historical investigation in his important 2004 problem paper [Gra6] in *Geombinatorics*:

> It is certainly not necessary to point out to readers of this journal any facts concerning the history and current status of this problem (which [is] due to Nelson in 1950) since

[2]Thanks to Professor Fred Hoffman, the tireless organizer of this annual conference, I have a videotape of this memorable Paul Erdős' talk and thus have transcribed Paul's words exactly.

the Editor Alexander Soifer has written a scholarly treatment of this subject in this journal [Soi18], [Soi19], [SS2].

Ron confirmed the validity of my historical research in his January 26, 2007, email:

Hi Sasha,
... I think it is clear by your historical research that Nelson gets credit for the chromatic number of the plane problem.
Best regards,
Ron

The results of my historical research are summarized in Diagram 3.2, where arrows show passing of the problem from one mathematician to another. In the end, Paul Erdős shares the problem with the world in numerous talks and articles.

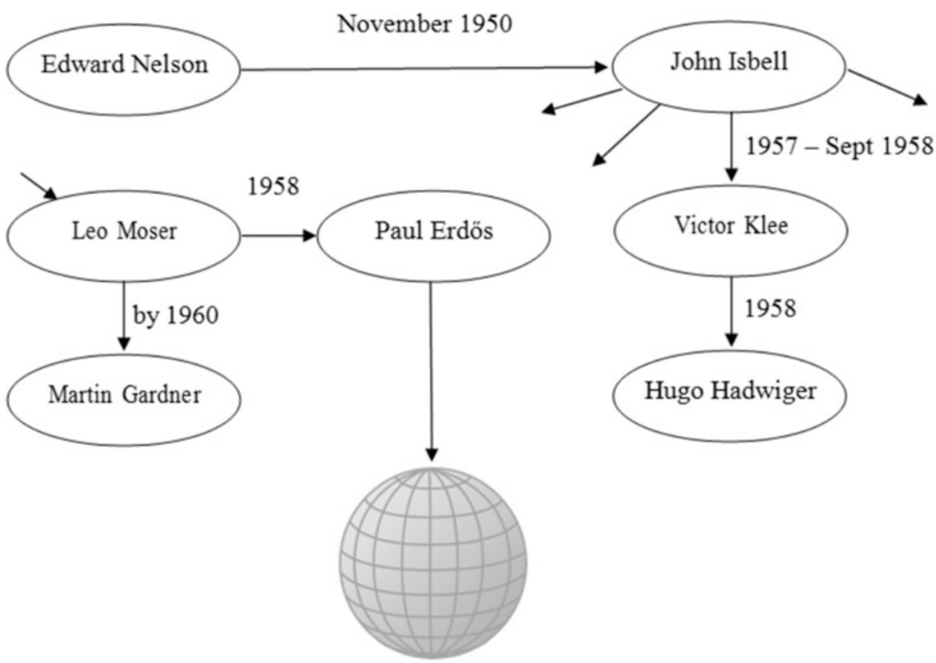


Diagram 3.2 Passing the baton of the chromatic number of the plane problem

Paul Erdős' and Ron Graham's acceptance of my research on the history of this fascinating problem has had a significant effect: most researchers and expositors now give credit to Edward Nelson for the chromatic number of the plane problem. There are, however, exceptions. László Lovász and K. Vesztergombi, for example, state in 2002 [LV] that

in 1944 Hadwiger and Nelson raised the question of finding the chromatic number of the plane.

Of course, the problem did not exist in 1944, in Hadwiger's cited paper or anywhere else. Moreover, Eddie Nelson was just an 11–12-year-old boy in 1944! In the same book, dedicated to the memory of Paul Erdős, one of the leading researchers of the problem László

Székely (who had in 1992 already attended my talk on the history of the problem in Boca Raton, where I presented the proof of Nelson's authorship), goes even further than Lovász and Vesztergombi in creating a myth [Sze3]:

> E. Nelson and J. R. Isbell, and independently Erdős and H. Hadwiger, posed the following problem . . .

The fine Russian researcher of this and related problems Andrei M. Raigorodskii copies Székely in his 2003 book [Raig6, p. 3], despite citing (and thus presumably knowing) my historical investigation in his survey [Raig3]:

> There were several authors. First of all, already in the early 1940s the problem was posed by remarkable mathematicians Hugo Hadwiger and Paul Erdős; secondly, E. Nelson and J. P. Isbell worked on the problem independently from Erdős and Hadwiger.[3]

Raigorodskii then "discovers" a nonexistent connection between World War II (!) and the popularity of the chromatic number of the plane problem[4]:

> In the 1940s there was W.W.II, and this circumstance is responsible for the fact that at first chromatic numbers [sic] did not raise too thunderous an interest.

In 2019, [BRa] L.I. Bogolyubsky and A.M. Raigorodskii drop John Isbell and call it "the Nelson–Erdős–Hadwiger problem."

I have won many battles in my life, for example, the change of the "Rolf Nevanlinna Prize" to "IMU Abacus Medal." However, I am giving up trying to do justice here and get the only correct name, **The Edward Nelson Problem** in this case. Even though Hugo Hadwiger admitted in print that he was not the author of the problem, the name "Hadwiger–Nelson" got stuck to the problem, just as Cardano did not author the Cardano formula and the Pythagoras theorem was known a millennium before the great Greek was born. Such is life with credits in mathematics. Most mathematicians view history as Cinderella that does not merit the respect they hold toward mathematics. History requires and deserves rigor and respect, gentlemen.

Not only Hadwiger but also the two famous Canadian problem people, the brothers Leo and William Moser, published in 1961 [MM] the proof of the lower bound $4 \leq \chi$ while solving a different problem. Although, in my opinion, their proof is not distinct from those by Nelson and by Hadwiger, the Mosers' emphasis on a finite set and their invention of the seven-point configuration, now called *The Mosers Spindle* (plural, "Mosers," for we have here two brothers) proved to be highly productive.

Now, we can finally give due credit to Edward Nelson for being the first in 1950 to prove the lower bound $4 \leq \chi$. Because of the bound 4, John Isbell recalls in his letter [Isb1] that Nelson "liked calling it a second four-color problem!" Nelson shared with me that he thought the chromatic number of the plane to be 4.

In phone interviews with Edward Nelson on September 18 and 30, 1991, I learned some information about the problem creator.

[3] My translation from Russian.

[4] Ibid.

Joseph Edward Nelson was born on May 4, 1932 (an easy number to remember: 5/4/32), in Decatur, Georgia, near Atlanta. The son of the secretary of the Italian YMCA,[5] Ed Nelson had studied at a *liceo* (Italian prep school) in Rome. In 1949, Eddie returned to the United States and entered the University of Chicago. The visionary president of the university, Robert Hutchins,[6] allowed students to avoid "doing time" at the university by passing lengthy placement exams instead. Ed Nelson had done so well on so many exams that he was allowed to go right on to graduate school without working on his bachelor's degree.

Time Magazine reported young Nelson's fine achievements in 14 exams on December 26, 1949 [Time], next to the report on the completion of the last war crime trials of World War II (Field Marshal Fritz Erich von Manstein was sentenced to 18 years in prison), assurances by General Dwight D. Eisenhower that he would *not* be a candidate in the 1952 presidential election (he certainly was – and won it), and promise to announce *Time*'s "A Man of the Half-Century" in the next issue (subsequently, the *Time*'s choice was Winston Churchill).

Upon obtaining his doctorate from the University of Chicago in 1955, Edward Nelson became the National Science Foundation's postdoctoral fellow at Princeton's Institute for Advanced Study in 1956. Three years later, he became a professor of mathematics at Princeton University. His main areas of interest were analysis and logic. In 1975, Edward Nelson was elected to the American Academy of Arts and Sciences and in 1997 to the National Academy of Sciences. Ed shared with me, with excitement, that he gave an invited mathematical talk in Vatican. During my 2002–2004 and 2006–2007 work at Princeton, I had the pleasure of interacting with Professor Nelson almost daily. We became friends and shared lunches in the established company of senior Princetonians. Ed had a wonderful contagious smile. He enjoyed smoking his pipe and sipping wine. Ed attended my talk on the chromatic number of the plane problem at Princeton's Discrete Mathematics Seminar that I dedicated "To Edward Nelson, who created this celebrated problem for us all." He passed away on September 10, 2014, in Princeton.

John Rolfe Isbell (October 27, 1930–August 6, 2005) was the first in 1950 to prove the upper bound $\chi \leq 7$. He used the same hexagonal 7-coloring of the plane that Hadwiger published in 1961 [Had4]. Please note that Hadwiger first used this coloring of the plane in 1945 [Had3] but for a different problem: His goal was to show that there are seven congruent closed sets that cover the plane (he also proved there that no five congruent closed sets cover the plane). Professor Isbell, PhD Princeton University 1954 under Albert Tucker, had been for decades on the faculty of mathematics at the State University of New York at Buffalo, where he later became professor emeritus. John Isbell passed away on 6 August 2005.

Presently, in expanding the book for this new edition, I simply ought to pose a Hadwigerian open problem:

Plane Covering Problem 3.0 Is there a closed set S such that there is a distance not realizable between any pair of points of S and the plane can be covered by six sets congruent to S?

---

[5] The Young Men's Christian Association (YMCA) is one of the oldest and largest not-for-profit community service organizations in the world.

[6] Robert Maynard Hutchins (1899–1977) was president (1929–1945) and chancellor (1945–1951) of the University of Chicago.

Paul Erdős' contribution to the survival and popularity of this problem is twofold. First of all, Paul kept the flaming torch of the problem brightly lit. He made the chromatic number of the plane problem well-known by posing it in his countless problem talks and many publications, for example, we see it in [E61.22], [E63.21], [E75.24], [E75.25], [E76.49], [E78.50], [E79.04], [ESi], [E80.38], [E80.41], [E81.23], [E81.26], [E85.01], [E91.60], [E92.19], [E92.60], and [E94.60].

Second, Paul Erdős created a good number of fabulous, related problems. We will discuss one of them in the next chapter.

In February 1992 at the 23rd Southeastern International Conference on Combinatorics, Graph Theory, and Computing in Boca Raton, during his traditional Thursday morning talk, I asked Paul Erdős how much he would offer for the first solution of the chromatic number of the plane problem. Paul replied (while I was jotting down on paper his words):

> I can't offer money for nice problems of other people because then I will really go broke.

I then transformed my question into the realm of mathematics and asked Paul "*Assume* this is *your* problem, Paul; how much would you initially offer for its first solution?" Paul answered:

> It is a very nice problem. If it were mine, I would offer $250 for it.

Later, the price went up for the first solution of just the lower bound part of the chromatic number of the plane problem. On Saturday, May 4, 2002, which by the way was precisely Edward Nelson's 70th birthday, Ronald L. Graham gave a talk on Ramsey theory at the Massachusetts Institute of Technology to about 200 high school participants in the USA Mathematical Olympiad. During the talk, he offered $1000 for the first proof or disproof of what he called, after Nelson, "Another Four-Color Conjecture." The talk commenced at 10:30 a.m. (as a member of the USA Mathematical Olympiad Subcommittee, I was in attendance and took notes).

Another Four-Color $1000 Problem 3.1 (R.L. Graham, May 4, 2002). Is it possible to four-color the plane to forbid a monochromatic distance 1?

In August 2003, in his talk "What is Ramsey Theory?" at Berkeley [Gra4], Graham asked for much more work for $1000:

$1000 Open Problem 3.2 (R.L. Graham, August 2003). Determine the value of the chromatic number χ of the plane.

It seems that Graham believed that the chromatic number of the plane takes on an intermediate value, between of its known boundaries, for in his two surveys [Gra7], [Gra8], he offered the following open problems:

$100 Open Problem 3.3 (R.L. Graham [Gra7], [Gra8]). Show that $\chi \geq 5$.[7]

$250 Open Problem 3.4 (R.L. Graham [Gra7], [Gra8]). Show that $\chi \leq 6$.

This prompted me to look at all published Erdős' predictions of the chromatic number of the plane. Let me summarize them here for you. First Erdős believed – and communicated it

[7] Ron Graham cites the O'Donnell theorem 50.4 (see it later in this book) as "perhaps, the evidence that χ is at least 5."

in 1961 [E61.22] and in 1975 [E75.24] – that the problem creator Nelson conjectured that the chromatic number was 4; Paul enters no prediction of his own. In 1976 [E76.49], Erdős asks:

Is this graph 4-chromatic?

In 1979 [E79.04], Erdős becomes more assertive:

It seems likely that the chromatic number is greater than four. By a theorem of de Bruijn and myself this would imply that there are n points $x_1, \ldots, x_n$ in the plane so that if we join any two of them whose distance is 1, then the resulting graph $G(x_1, \ldots, x_n)$ has chromatic number > 4. I believe such an n exists but its value may be very large.

A certainty comes in 1980 [E80.38] and [E80.41]:

I am sure that [the chromatic number of the plane] $\alpha_2 > 4$ but cannot prove it.

In 1981 [E81.23] and [E81.26], we read Erdős, respectively:

It has been conjectured [by E. Nelson] that $\alpha_2 = 4$, but now it is generally believed that $\alpha_2 > 4$.

It seems likely that $\chi(E^2) > 4$.

In 1985 [E85.01], Paul Erdős writes:

I am almost sure that $h(2) > 4$.

Once – just once – Erdős expresses mid-value expectations. It happened on Thursday, March 10, 1994 at the 25th Southeastern International Conference on Combinatorics, Graph Theory, and Computing in Boca Raton. Following Erdős' plenary talk (9:30–10:30 a.m.), I was giving my talk at 10:50 a.m., when suddenly Paul Erdős said (and I jotted it down):

Excuse me for interrupting, I am almost sure that the chromatic number of the plane is greater than 4. It is not a proof, but any measurable set without distance 1 in a very large circle has measure less than ¼. I also do not think that it is 7.

It is time for me to speak on the record and predict the chromatic number of the plane. In 2002, I was leaning toward predicting 7 or else 4 – somewhat disjointly from Graham and Erdős' apparent expectations. Limiting myself to just one value, still in 2002, I conjectured:

Chromatic Number of the Plane Conjecture 3.5 (A. Soifer, 2002)[8]. $\chi = 7$.

On January 26, 2007, in a personal e-mail to me, Ron Graham clarified the terms of awarding his prizes:

I always assume that we are working in ZFC (for the chromatic number of the plane!). My monetary awards can vary depending on which audience I am talking to. I always give the maximum of whatever I have announced (and not the sum!).

[8] See more on the predictions in Chapter 62.

Chapter 4
Polychromatic Number of the Plane and Results Near the Lower Bound

When a great problem withstands all assaults, mathematicians create many related problems. It gives them something to solve, plus sometimes there is an extra gain in this process, when an insight into a related problem brings new ways to see and conquer the original one. Numerous problems have been posed around the chromatic number of the plane. I would like to share with you my favorite among them.

It is convenient to say that a colored set S *realizes distance* d if S contains a monochromatic segment of length d; otherwise, we say that S *forbids distance* d.

Our knowledge about this problem starts with the celebrated 1959 book by Hugo Hadwiger and Hans Debrunner ([HD2] and, subsequently, its enhanced translations into Russian by Isaak M. Yaglom [HD3] and into English by Victor Klee [HDK]). Hadwiger reports in the book the contents of the September 9, 1958, letter he received from the young (at the time) Hungarian mathematician Aladár Heppes:

> Following an initiative by P. Erdős he [i.e., Heppes] considers decompositions of the space into disjoint sets rather than closed sets. For example, we can ask whether proposition 59 remains true in the case where the plane is decomposed into three disjoint subsets. As we know, this is still unresolved.

In other words, Paul Erdős asked whether it was true that if the plane is partitioned (colored) into three disjoint subsets, then one of the subsets must realize all distances. Soon, the problem took on its current "appearance." Here it is:

Erdős' Open Problem 4.1 What is the smallest number of colors needed for coloring the plane in such a way that no color realizes all distances?[1]

This number had to have a name, and, so, in 1992 [Soi5], I named it *the polychromatic number of the plane* and denoted it by χ_p. The name and the notation seemed so natural that, by now, it has become the standard and has (without credit) appeared in such encyclopedic books as [JT] and [GO].

[1] The authors of the fine problem book [BMP] incorrectly credit Hadwiger as the "first" to study this problem (p. 235). Hadwiger, quite typically for him, limited his study to partitions into *closed* sets.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_4

Since I considered this to be a very important open problem, I asked Paul Erdős to verify his authorship, alleged in passing by Hadwiger. As always, Paul was very modest in his July 16, 1991, letter to me [E91/7/16ltr]:

> I am not even quite sure that I created the problem: Find the smallest number of colors for the plane, so that no color realizes all distances, but if there is no evidence contradicting it we can assume it for the moment.

My notes show that during his unusually long 2-week visit of me in December 1991–January 1992 (we were working together on the book of Paul's open problems, entitled *Problems of pgom Erdős*), Paul confirmed his authorship of this problem. In the chromatic number problem, we were looking for colorings of the plane such that each color forbids distance 1. In the polychromatic number problem, we are coloring the plane in such a way that each color i forbids a distance d_i. For distinct colors i and j, the corresponding forbidden distances d_i and d_j may (but do not have to) be distinct. Of course, $\chi_p \le \chi$; therefore,

$$\chi_p \le 7.$$

Nothing else had been discovered during the first 12 years of this problem's life. Then, in 1970, Dmitry E. Raiskii, a student of the Moscow High School for Working Youth[2] 105, published [Rai] the lower and upper bounds for χ_p. Here, we will look at the lower bound, leaving the upper bound to Chapter 6.

Raiskii's Theorem 4.2 (D. E. Raiskii [Rai], 1970): $4 \le \chi_p$.

Three years after Raiskii's publication, in 1973, the British mathematician Douglas R. Woodall from the University of Robin Hood (I mean Nottingham), published a paper [Woo1] on problems related to the chromatic number of the plane. Among other things, he provided his own proof of the lower bound. As I showed in [Soi17], Woodall's proof stemmed from a triple application of two simple ideas of Hugo Hadwiger ([HDK], Problems 54 and 59).

In 2003, the Russian-turned-Israeli mathematician Alexei Kanel-Belov communicated to me an incredibly beautiful short proof of this lower bound by the new generation of young Russian mathematicians, all his students. The proof was found in 1997 by Alexei Merkov, a 10th grader from the Moscow High School 91, and communicated by Alexei Roginsky and Daniil Dimenstein at the Moscow Pioneer Palace [Poisk]. Following is Merkov's proof with my gentle modifications. It is truly "proof from the book," if you are familiar with Paul Erdős' famous metaphor.

Proof of D.E. Raiskii's Lower Bound Theorem by A. Merkov, 1997: Assume that the plane is colored in three colors, red, green, and blue, and that each color forbids a distance r, g, and b, respectively. Equip the three-colored plane with the Cartesian coordinates with the origin O, and, construct in the plane, three seven-point sets S_r, S_g, and S_b, with each being the Mosers Spindle (Fig. 4.1), such that all three spindles share O as one of their seven vertices and have edges all equal to r, g, and b, respectively.

[2] Students in such high schools hold regular jobs during the day and attend classes at night.

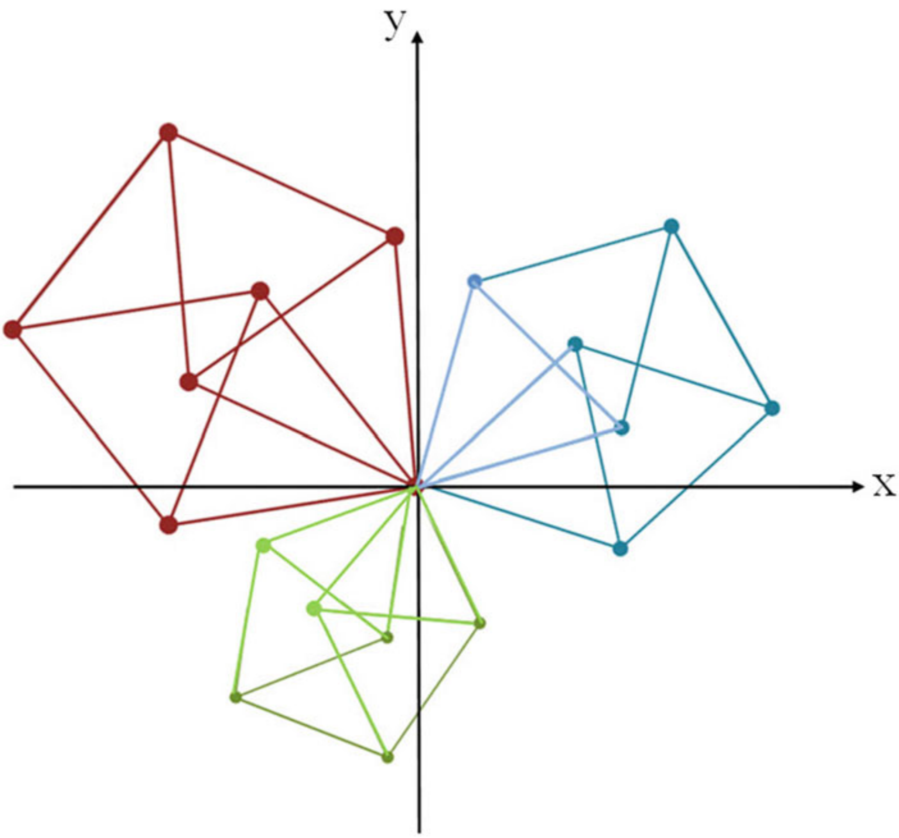


Fig. 4.1 The Cartesian plane with three Mosers Spindles

This construction defines 6 "red" vectors $v_1, \ldots, v_6$ from the origin O to each vertex of S_r; 6 "green" vectors $v_7, \ldots, v_{12}$ from O to the vertices of S_g; and 6 "blue" vectors $v_{13}, \ldots, v_{18}$ from O to the vertices of S_b , i.e., 18 vectors in all.

Now, introduce the 18-dimensional Euclidean space E^{18} and a function M from E^{18} to the plane E^2 naturally defined as follows:

$$M\colon (a_1, a_2, \ldots, a_{18}) \mapsto a_1 v_1 + a_2 v_2 + \ldots + a_{18} v_{18}.$$

This function induces a 3-coloring of E^{18} by assigning a point of E^{18} the color of the corresponding point of the plane. We will call the first six axes of E^{18} "red," the next six axes "green," and the last sic axes "blue."

Define by W the subset in E^{18} of all points whose coordinates include at most 1 coordinate equal to 1 for each of the three colors of the axes and the rest (15 or more) coordinates 0. It is easy to verify that W consists of 7^3 points. Let us fix an array of allowable W coordinates on the green and blue axes and vary allowable coordinates on the red axes. We get the seven-element set A of points in W. The image $M(A)$ of the set A under the map M forms in the plane, a translate of the original seven-point set S_r. If we fix another array of green and blue coordinates, we get another seven-element set in E^{18}, whose image under M forms in the plane that is another translate of S_r. Thus, the set W gets partitioned into 7^2 subsets, each of which maps by M into a translate of S_r.

Now, recall Mosers' Tool 2.3. It implies here that any translate of the Mosers Spindle S_r contains at most two red points out of its seven points. Since the set W has been *partitioned* into the translates of S_r, at most 2/7 of the points of W are red.

We can start all over again and, in a similar way, show that at most 2/7 of the points of W are white and that at most 2/7 of the points of W are blue. But 2/7 + 2/7 + 2/7 does not add up to 1! This contradiction implies that at least one of the colors realizes all distances, as required. ■

At the International Congress on Mathematical Education in 1992 in Quebec City, I spent much time with Nikolai Nikolaevich "Kolya" Konstantinov, whose mathematical circle at the Old Building of Moscow State University I attended on Saturday afternoons during the 1962–63 academic year, when I was an eighth grader. To my amazement, I learned that the hero of this chapter, Dmitrii "Dima" Raiskii, was Konstantinov's student as well, just 2 years my junior! It took me many years to get "the full story" out of Kolya Konstantinov, but it was worth waiting for his February 23, 2007, e-mail, which I am translating here for you from the Russian original:

> Dima Raiskii entered school Nr. 7 in 1965.[3] He was a part of a very strong group of students, from which several professional mathematicians came out, including Lena Nekhludova, who won gold medal of the International Mathematical Olympiad, Andrej Grjuntal, now chair of a department in the Institute of System Research, Vasilii Kozlov, now professor in the department of statistics of the Mechanics-Mathematical Faculty of the Moscow State University, and several well-known applied mathematicians.
>
> Teachers of main mathematical courses were also very strong, including Joseph Bernstein, Viktor Zhurkin, formerly a graduate of this school and now a well-known biochemist, working in the USA.
>
> The teaching method was based on students proving theorems of a course on their own, and on solving a large number of meaningful problems, which required creative abilities ...
>
> Dima performed well in mathematics, but was missing classes, and he had difficulties in other disciplines, in which teachers did not want to pass him because of small amount of earned credits. However, the main problem was at home. Dima's father thought his son was inept and insisted that Dima master a profession of a shoemaker, so that he could somehow feed himself. When I got to know Dima's family, I did not see his father, probably because by then he had already left the family, but I did not feel I had the right to ask about it.
>
> Without any help, on his own Dima had read Hadwiger and Debrunner's book on combinatorial geometry.[4] He told me that he solved a problem from that book and wanted to show it to me. His presentation of the proof was in a "hall style" – very careless and informal, and l did not understand it right away – I felt, nevertheless, that the proof seemed plausible.

[3]This was one of the Soviet Union's greatest high schools with emphasis on mathematics, where special courses were offered by some of the top Moscow State University professors.

[4][HD3].

Dima then wrote down his solution. Here I made sure that everything was correct. However, Dima did not have an experience of writing articles, and so I undertook the "combing" of the text and gave it the usual for a publication look – I introduced several notations and terms. My work was purely technical; the published text did not contain my single idea. There was, however, an example, inserted by the Editor of *Математические Заметки* [*Mathematical Notes*][5] Stechkin.[6] Then a funny episode happened. The inserted paragraph Stechkin ended with the phraise "the author thanks Stechkin for this example." Dima, however, thought that the word "author" refers to Stechkin in this case, and could not understand how Stechkin could thank himself.

Meanwhile clouds were thickening over Dima's head. The school wanted to expel him for absences, and he got into a children section of a psychiatric hospital. I visited him there. I saw lads of a school age behaving themselves quite freely. The counselors looked upon it nonchalantly – what can one ask of the sick ones? One boy, for example, asked, what would happen if to throw Brezhnev[7] into a toilette bowl and flush the toilette? And other silliness of the same kind.

After the release from the hospital, Dima [was expelled from the mathematical school number 7 and] transferred to the school [number 105] for working youth. There his affairs got even worse. He was finishing his senior 11^th^ year, and the teachers' council had to decide whether to graduate the student, who missed countless classes and had almost no grades. At that time, the school received a letter from England. The thing is, at the end of Dima's published article there was the school's number, where he studied at the time of the article's publication. The letter was written by the professor [could it be Douglas Woodall?] who worked on the same problem but did not succeed. He informed Raiskii that he was sending him all the materials because he would no longer work on this problem but hoped that Raiskii would be interested in acquainting himself with this unfinished work. This was not just a letter, but a thick packet, and the letter opened with "Dear Professor Raiskii." The lady principal looked very gloomily during the teachers' council meeting dedicated to the question of Raiskii's graduation. She opened the meeting by acquainting the teachers with the content of this letter. She then said, "Let us graduate him."

In conclusion, let me add that Raiskii's family difficulties continued. Of course, Dima's psyche was not fully normal, but I think that his mother's psyche played a more negative role in his life than his own psyche. Here is one of her tricks. After Dima was released from the hospital, she wrote a letter to the Minister of Education complaining about me and P. S. Alexandrov.[8] The school [number 7] principal Volkov showed me this letter (which the Ministry forwarded to the school). Dima's mother claimed in this letter that Alexandrov and Konstantinov politically corrupt the child and inoculate the child with the anti-Soviet views. The letter went on further to claim that Konstantinov

[5]The journal in which this article appeared.

[6]Sergei Borisovich Stechkin, a noted Russian mathematician – see his example and more about this story in Chapter 6.

[7]Head of the Soviet Union at the time.

[8]Pavel Sergeevich Alexandrov, a topologist and a member of the Soviet Academy of Sciences, one of Russia's great mathematicians.

established the power over all Moscow psychiatrists and they all dance to his tune. The principal read this letter to me seriously, without any smile, until the last phraise when he finally allowed himself to laugh. I do not think it would be interesting to describe other tricks of Dima's mother.

While a high school student, Dima tried to solve mathematical problems many times. In particular, while participating in the Moscow Mathematical Olympiad, he worked not at all on the problems of the Olympiad, but on his own problems. He then got involved in the Eastern games of the mind – but I am not an expert in them, and do not remember their names. After that, I think, you know more about Dima than I do.

I wish you success [with the book].

Kolya

On Christmas Day, December 25, 2003, the hero of this section, Dima Raiskii, told me how he came across the polychromatic number of the plane problem:

I learned about our coloring problem while reading the book *Combinatorial Geometry of the Plane* by Hadwiger and Debrunner [HD3]. This book was a part of the 3rd prize that I received at the Moscow Mathematical Olympiad of the 8th graders.

In my phone conversation with Dima Raiskii, I expressed my regret that he left mathematics after such a brilliant first paper. "Mathematicians appeared boring to me," Dima replied and added: "They were constantly suffering from a feeling of guilt toward each other or tried to make others repent. I felt much more at ease with *Go* players." So, Dima worked as a computer programmer and spent his free time playing *Go*. Then he gave up the city life, as he informed me on February 6, 2003:

I now settled in a remote village, where there is neither post nor computer. However, when I come to the city, I visit an internet-salon. What is new with your studies of African cultures? Are there meditative practices in Africa?

In his e-mails sent on the go from Internet cafés, Dima described his involvement in *Go*, meditation, and writing books to aid others with meditation and spirituality. On March 17, 2003, I read:

In the latter years I played *Go*. This is the only game richer than chess; it is popular in China, Japan, Korea, etc. One of my students later became the Russian Champion for players up to the age of 10. According to the tradition, many *Go* players do meditative exercises in the style of Zen because this game equally uses both sides of the brain. In a close circle, I taught Zen meditation. In the East, however, many Buddhist authorities use Christian texts for teaching meditation. I am now preparing a small book of exercises for people raised in the Christian culture . . .

P.S.: *Go* (brought to Europe by [the legendary world chess champion Emanuel] Lasker) is a most interesting object for computer modeling – in this regard, *Go* is richer than chess. One of my acquaintances is the European Champion in *Go* programming. Are people at Princeton involved in it?[9]

[9] At that time, I worked at Princeton University.

Dima asked me several times to publish his results as a joint work of his and Nikolai Nikolaevich "Kolya" Konstantinov, his and my mathematics teacher, who – for better or for worse – influenced my choice of mathematics as a profession. Dima insisted on sharing credit with Kolya, and Kolya categorically refused his share, as in his opinion, all the ideas belonged to Dima.

Dima does not communicate with many people. He even sends his regards to his Moscow teacher, Nikolai Konstantinov, via me in the United States. His e-mails to me are always inquisitive and warm. In his November 23, 2006, e-mail he expressed an appreciation for our correspondence:

> News from you always improve my mood. Give my regards to Nikolai Nikolaevich [Konstantinov].

In his December 19, 2007, e-mail, Dima wrote:

> I was always interested in the Eastern culture and studies of the Eastern religions. In the old times, however, I could not have publications [on these subjects], and instead had a lot of troubles. It seems likely that something will be published in the nearest time. This will start my public "biography." Will you be interested in my article? ... Yours always, Dima.

Dear Dima Raiskii, through the years of our correspondence, we became not only pen pals but also friends. The societal pressure altered his life not unlike the change in the life of Grigory Perelman, who abandoned mathematics at the peak of his creative powers, after conquering the celebrated Poincare and geometrization conjectures. Their unprotected moral purity and extreme sensitivity made it difficult for them to deal with the ills of society in general and the mathematical community in particular. Our friendship has provided Dima with an outlet for his thoughts and communication. I hope someone offered the same to Grisha Perelman.

PS: After the first edition of this book went into production, I informed Dima that his theorem and biography will appear in it, as will Van der Waerden's theorem and biography. On May 3, 2008, Dima replied:

> Sasha, thank you very much! My biography and biography of Van der Waerden – not a bad combination. I will be telling my fellow villagers: "Once upon a time I am sitting with Vanya, this Vanya, you know, which is der Waerden, who ... Will definitely read your book. Happy [W.W.II] Victory Day!"

■

Paul Erdős proposed yet another related problem (e.g., see [E85.01]). For a given finite set S of r positive numbers, a *set of forbidden distances* if you will, we define the graph $G_S(E^2)$, whose vertices are points in the plane, and a pair of points is adjacent if and only if the distance between them belongs to S. Denote

$$\chi_r = \max_S \chi\left(G_S\left(E^2\right)\right).$$

"It is easy to see that $\lim\limits_{r\to\infty} \frac{\chi_r}{r} = \infty$," Erdős writes, and poses a question:

Erdős' Problem 4.3 Does χ_r grow polynomially?

It is natural to call the chromatic number $\chi_S(E^2)$ of the graph $G_S(E^2)$ the *S-chromatic number of the plane*. One can pose a more general and hard problem, and in fact, it is an old problem of Paul Erdős ("I asked long ago," Paul says in [E94.60]):

Erdős' Open Problem 4.4 Given S, find the S-chromatic number $\chi_S(E^2)$ of the plane.

How difficult this problem is judge for yourselves: For an one-element set S, this is the chromatic number of the plane problem!

Chapter 5
De Bruijn–Erdős Reduction to Finite Sets and Results Near the Lower Bound

We can expand the notion of the chromatic number to any subset S of the plane. *The chromatic number* $\chi(S)$ of S is the smallest number of colors sufficient for coloring the points of S in such a way that forbids monochromatic unit segments.

In 1951, Nicolaas Govert de Bruijn and Paul Erdős published a highly powerful tool [BE2] that will help us with this and other problems. We will formulate and prove it in Part V. In our setting here, it implies the following.

Compactness Theorem 5.1[1] (N.G. de Bruijn, P. Erdős). The chromatic number of the plane is equal to the maximum chromatic number of its *finite* subsets.

Thus, as Paul Erdős used to say, the problem of finding the chromatic number of the plane is a problem about finite sets in the plane.[2]

There are easy questions about finite sets in the plane. Solve the following two problems on your own.

Problem 5.2 Find the smallest number δ_3 of points in a plane set whose chromatic number is equal to 3.

Problem 5.3 (L. Moser and W. Moser, [MM]). Find the smallest number δ_4 of points in a plane set whose chromatic number is 4. (Answer: $\delta_4 = 7$).

Victor Klee and Stan Wagon posed the following open problem in [KW]:

Open Problem 5.4 When k is 5, 6, or 7, what is the smallest number δ_k of points in a plane set whose chromatic number is equal to k?

Of course, problem 5.4 makes sense only if $\chi > 4$. In the latter case, this problem suggests a way to attack the chromatic number of the plane problem by constructing new "spindles."

When you worked on problems 5.2 and 5.3, you probably remembered our problems 2.1 and 2.2. Indeed, those problems provide optimal configurations (Figs. 2.1 and 2.2) for

[1]The axiom of choice is assumed in this result.

[2]Or so we all thought. Because of that, I choose to leave this chapter as it was written in the early 1990s. See Part XII of this book for axiomatic developments.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_5

problems 5.2 and 5.3. Both optimal configurations were built of equilateral triangles of side 1. Can we manage without them?

Problem 5.5 Find the smallest number σ_3 of points in a plane set without equilateral triangles of side 1 whose chromatic number is equal to 3.

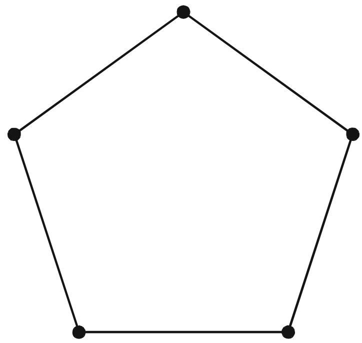


Fig. 5.1 An equilateral pentagon of side 1

Solution $\sigma_3 = 5$. The regular pentagon of side 1 (Fig. 5.1) delivers a minimal configuration of chromatic number 3.

It is easy to 2-color any four-point set A, B, C, D without equilateral triangles of side 1. Just color A red. All points at a distance 1 from A, color blue; these are second-generation points. All uncolored points at a distance 1 from any point of the second generation, we color red, and these are third-generation points. All uncolored points at a distance 1 from the points of the third generation, we color blue. If we did not color all four points, then we start this process all over again by coloring any uncolored point red. If this algorithm were not to define the color of any point uniquely, we would have an odd-sided n-gon with all sides 1, i.e., an equilateral triangle (since $n \leq 4$), which cannot be present, and thus would provide the desired contradiction. ■

For four colors, this question for a while was an open problem first posed by Paul Erdős in July 1975 (and published in 1976), who, as was usual for him, offered to "buy" the first solution – for $25.

Paul Erdős' $25 Problem 5.6 [E76.49]. Let S be a subset of the plane, which contains no equilateral triangles of size 1. Join two points of S if their distance is 1. Does this graph have chromatic number 3?

If the answer is no, assume that the graph defined by S contains no C_l (cycles of length l) for $3 \leq l \leq t$ and ask the same question.

It appears that Paul Erdős was not sure of the outcome, which was rare for him. Moreover, from the next publication of this problem in 1979 [E79.04], it is clear that Paul expected that triangle-free unit distance graphs had chromatic number at most 3 or else chromatic number 3 can be forced by prohibiting all small cycles up to C_k for a sufficiently large k:

Paul Erdős' $25 Problem 5.6' [E79.04]. "Let our n points [in the plane] be such that they do not contain an equilateral triangle of side 1. Then their chromatic number is probably at most 3, but I do not see how to prove this. If the conjecture would unexpectedly [sic] turn out to be false, the situation can perhaps be saved by the following new conjecture:

There is a k so that if the girth of $G(x_1,\ldots,x_n)$ is greater than k, then its chromatic number is at most three – in fact, it will probably suffice to assume that $G(x_1,\ldots,x_n)$ has no odd circuit of length $\leq k$."[3]

Erdős' first surprise arrived in 1979 from down under: Nicholas Wormald, then of the University of Newcastle, Australia, disproved the first, easier, triangle-free conjecture. Erdős paid \$25 reward for the surprise and promptly reported it in his next 1978 talk (published 3 years later [E81.23]):

> Wormald in a recent paper (which is not yet published) disproved my original conjecture – he found a [set] S for which [the unit distance graph] $G_1(S)$ has girth 5 and chromatic number 4. Wormald's construction uses elaborate computations and is fairly complicated.

In his paper [Wor], Wormald proved the existence of a set S of 6448 (!) points without triangles and quadrilaterals with all sides 1, whose chromatic number was 4. He was aided by a computer. I would like to give you a taste of the initial Wormald construction or, more precisely, the Blanche Descartes construction that Wormald was able to embed in the plane, but it is a better fit in Chapter 12 – so, see it there.

The size of Wormald's example, of course, did not appear to be anywhere near optimal. Surely, it must have been possible to do the job with less than 6448 points! In my March–1992 talk at the Southeastern International Conference on Combinatorics, Graph Theory, and Computing at Florida Atlantic University, I shared Paul Erdős' old question, but I put it in a form of competition:

A graph is called *unit-distance* if its two vertices are connected by an edge if and only if they are at distance 1 apart.

Open Problem 5.7 Find the smallest (in the number of vertices) unit-distance graph in the plane without equilateral triangles, whose chromatic number is 4. Construct such a graph.

The result exceeded my wildest dreams. A number of young mathematicians, including graduate students, were inspired by this talk and entered the race I proposed. Coincidentally, during that academic year, with the participation of the celebrated geometer Branko Grünbaum, and of Paul Erdős, whose problem papers set the style, I started a new and unique quarterly *Geombinatorics*, dedicated to problem-posing essays on discrete and combinatorial geometry and related areas. *Geombinatorics* is still alive and well now, 32 years later. The aspirations of the journal were clear from my 1991 Editor's Page in Issue 3 of Volume I:

> In a regular journal, a paper appears 1 to 2 (or more) years after the research is completed. By then even the author may not be excited any more about his results. In *Geombinatorics* we can exchange open problems, conjectures, aspirations, work-in-progress that is still exciting to the author, and therefore inspiring to the reader.

A true World Series played out on the pages of *Geombinatorics* around problem 5.7. The graphs obtained by the record setters were as mathematically significant as they were beautiful. I have to show them to you – see them discussed in detail in Chapters 14 and 15.

---

[3] The symbol $G(x_1,\ldots,x_n)$ denotes the graph on the listed inside parentheses n vertices, with two vertices adjacent if and only if they are a unit distance apart.

Many attempts to increase the lower bound of the chromatic number of the plane had not achieved their goal. Rutgers University's PhD student Rob Hochberg believed that the chromatic number of the plane was 4, while his roommate and fellow PhD student Paul O'Donnell was of the opposite opinion. They managed to get along despite this disagreement of the mathematical kind. On January 7, 1994, Rob sent me an e-mail to that effect:

> Alex, hello. Rob Hochberg here. (The one who's gonna prove $\chi(E^2) = 4$.) . . . It seems that Paul O'Donnell is determined to do his Ph. D. thesis by constructing a 5-chromatic unit-distance graph in the plane. He's got several interesting 4-chromatic graphs and great plans. We still get along.

Two months later, Paul O'Donnell's abstract in the *Abstracts* book of the Southeastern International Conference on Combinatorics, Graph Theory, and Computing in Boca Raton, Florida, included the following announcement:

> The chromatic number of the plane is between four and seven. A five-chromatic subgraph would raise the lower bound. If I discover such a subgraph, I will present it.

We all came to his talk of course (it was easy for me, as I spoke immediately before Paul in the same room). At the start of his talk, however, Paul simply said, "not yet," and went on to show his impressive 4-chromatic graph of girth 4. Five years later, on May 25, 1999, Paul O'Donnell defended his doctorate at Rutgers University.

Much was learned about 4-chromatic unit distance graphs. The best of these results, in my opinion, was contained in O'Donnell's dissertation. He completely solved Paul Erdős' problem 5.6 and delivered to Paul Erdős an ultimate surprise by negatively answering Erdős' general conjecture:

O'Donnell's Theorem 5.8 [Odo3, Odo4, Odo5]. There exist 4-chromatic unit distance graphs of arbitrary finite girth.

I choose to divide the proof of this result between Parts III and IX. See you there!

Chapter 6
Polychromatic Number of the Plane and Results Near the Upper Bound

6.1 Stechkin's 6-Coloring

In Chapter 4, we discussed the polychromatic number χ_p of the plane and looked at the 1970 paper [Rai] by Dmitry E. Raiskii, in which he was the first to prove that 4 is the lower bound of χ_p. The paper also contained the upper bound:

$$\chi_p \leq 6.$$

The example proving this upper bound was found by Sergei B. Stechkin and published with his permission by D.E. Raiskii in [Rai]. Stechkin has never gotten credit in the West for his example. Numerous articles and books credited Raiskii (except for Raiskii himself!). How did this happen? As everyone else, I read the English translation of Raiskii's paper [Rai]. It says (the words in italics are mine):

S.B. Stechkin noted that the plane can be decomposed into six sets such that all distances are not realized in any one of them. A corresponding example is presented here with the *author's solution*.

The author of what? – I was wondering. The author of the paper (as everyone decided)? But there is very little need for a "solution" once the example is found. I put Sherlock Holmes's cloak on and ordered a copy of the original Russian text. I read it in disbelief:

A corresponding example is presented here with the *author's permission*.

Stechkin *permitted* Raiskii to publish Stechkin's example! The translator mixed up two somewhat similar-looking Russian words and "innocently" created a myth (see Table 6.1):

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_6

Table 6.1 Translator's folly

Russian word	English translation
Решение	Solution
Разрешение	Permission

This is a great example in support of the expression "lost in translation." In reality, Sergei B. Stechkin was the editor of *Matematicheskie Zametki* (*Mathematical Notes*); he received Raiskii's manuscript, came up with the example, and inserted it in the manuscript with, I am sure, the agreement of Raiskii. Let us roll back to the mathematics of this example.

Problem 6.1 (S.B. Stechkin, [Rai]). $\chi_p \leq 6$, i.e., there is a 6-coloring of the plane such that no color realizes all distances.

Solution by S.B. Stechkin [Rai]. The "unit of the construction" is a parallelogram that consists of four regular hexagons and eight equilateral triangles, all of side lengths 1 (Fig. 6.1). We color the hexagons in colors 1, 2, 3, and 4. We partition the triangles of the titling into two types: We assign color 5 to the triangles with a vertex below their horizontal base and color 6 to the triangles with a vertex above their horizontal base. While coloring, we include with every hexagon its entire boundary, except its one rightmost and two lowest vertices; and every triangle does not include any of its boundary points.

We can now tile the entire plane with translates of the unit of the construction. ■

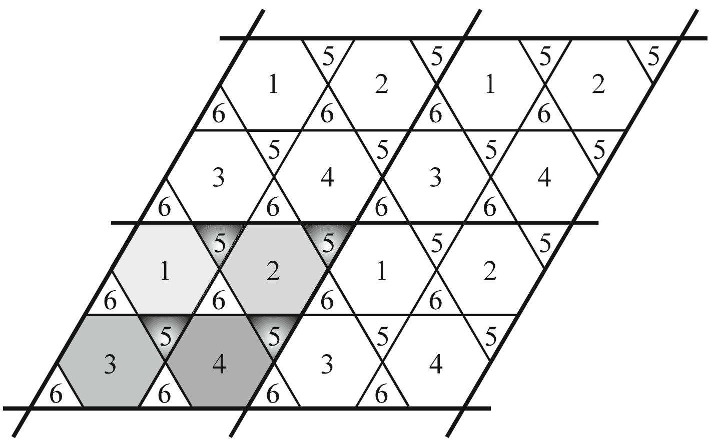


Fig. 6.1 The Stechkin 6-coloring of the plane

An easy construction solved problem 6.1 – easy to see after someone showed it to you. The trick was to find it, and Sergej Borisovich Stechkin found it first. Christopher Columbus too "just ran into" America! I got hooked.

6.2 The Best 6-Coloring of the Plane

I felt that if our ultimate goal was to find the chromatic number χ of the plane or to at least improve its known bounds ($4 \le \chi \le 7$), it may be worthwhile to somehow measure how close a given coloring of the plane is to achieving this goal. In 1992, I introduced such a measurement and named it *coloring type*.

Definition 6.2 (A. Soifer [Soi5], [Soi6], 1992). Given an n-coloring of the plane such that the color i does not realize the distance d_i ($1 \le i \le n$). Then we would say that this coloring is of *type* $(d_1, d_2, \ldots, d_n)$.

This new notion of type was so natural and helpful that it received the ultimate compliment of becoming a part of the mathematical folklore: it appeared everywhere without a credit to its inventor (look, for example, p. 14 of the fundamental 991-page-long monograph [GO]).

It would have been a great improvement in our search for the chromatic number of the plane if we were to find a 6-coloring of type (1, 1, 1, 1, 1, 1) or to show that one does not exist. With the appropriate choice of a unit, we can make the 1970 Stechkin coloring to have type $(1, 1, 1, 1, \frac{1}{2}, \frac{1}{2})$. Three years later, in 1973, Douglas R. Woodall [Woo1] found the second 6-coloring of the plane with all distances not realized in any color. Woodall's coloring had a special property that the author desired for his purposes: each of the six monochromatic sets was closed. His example, however, had three distinct "missing distances": It had type $(1, 1, 1, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{2\sqrt{3}})$. Woodall unsuccessfully tried to reduce the number of distinct distances, for he wrote "I have not managed to make two of the three 'missing distances' equal in this way" ([Woo1], p. 193).

In 1991, in search of a "good" coloring, I looked at the tiling with regular octagons and squares that tiled floors in many Russian public places (Fig. 6.2).

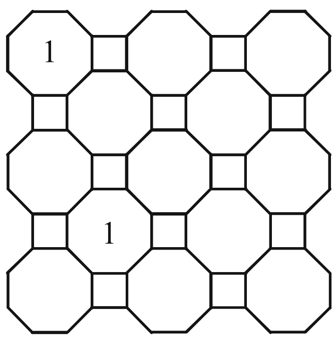


Fig. 6.2 Tiling used in many public places

But the "Russian public tiling" did not work! See it for yourself:

Problem 6.3 Prove that the set of all squares in the tiling of Fig. 6.2 (even without their boundaries) realizes all distances.

I then decided to shrink the squares until their diagonal became equal to the distance between the two closest squares. Simultaneously (!), the diagonal of the now nonregular octagon became equal to the distance between the two octagons marked with 1 in Fig. 6.2. I was in business!

Problem 6.4 (A. Soifer [Soi3], 1991). There is a 6-coloring of the plane of type (1, 1, 1, 1, 1, $\frac{1}{\sqrt{5}}$).

Solution We start with two squares, one of side 2 and the other of diagonal 1 (Fig. 6.3). We can use them to create the tiling of the plane with squares and (nonregular) octagons (Fig. 6.5). Colors 1, ..., 5 will consist of octagons; we will color all squares in color 6. With each octagon and each square, we include half of its boundary (bold lines in Fig. 6.4) without the end points of that half. It is easy to verify (please do) that the distance $\sqrt{5}$ is not realized in any of the colors 1, ..., 5 and the distance 1 is not realized in the color 6. By shrinking all linear sizes by a factor of $\sqrt{5}$, we get the 6-coloring of type (1,1,1,1,1,$\frac{1}{\sqrt{5}}$).

To simplify a verification, observe that the unit of my construction is bounded by the bold line in Fig. 6.5; its translates tile the plane. ■

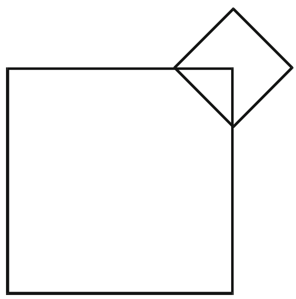


Fig. 6.3 Foundation squares

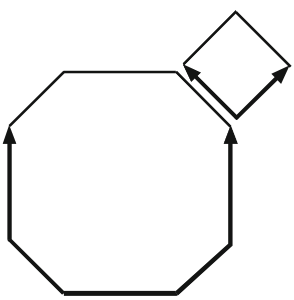


Fig. 6.4 Coloring of the boundaries

I had mixed feelings when I obtained the result of problem 6.4 in early August 1991. On the one hand, I knew the result was "close but no cigar": after all, a 6-coloring of type (1, 1, 1, 1, 1, 1) was not found. On the other hand, I thought that the latter 6-coloring may not exist, and, if so, my 6-coloring would be the best possible. There was another consideration as well. While in a PhD program in Moscow, I hoped to produce the longest paper that would still be refereed in by a major journal (and I got one published in 1973 that in manuscript was 56 pages long). This time, I was interested in a "dual record": how short can a paper be and still contain enough "stuff" to be refereed in and published? The paper [Soi6] solving problem 6.4 was two pages long, *including* three pictures. It was received on August 8, 1991, and

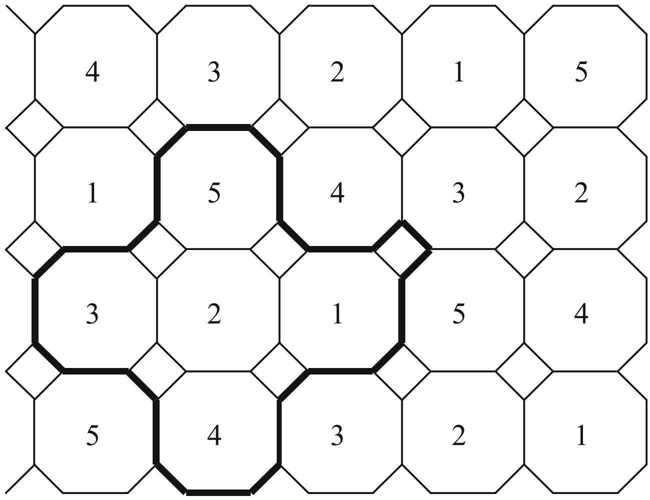


Fig. 6.5 The Soifer 6-coloring of the plane

accepted the next day by the *Journal of Combinatorial Theory, Series A* by the managing editor Bruce Rothschild of UCLA (University of California, Los Angeles). As nearly all journal editors have nearly always, Professor Rothschild insisted on objectivity. Referring to the chromatic number of the plane problem (CNP), I wrote that it was "my favorite open problem." Rothschild changed it in pencil to "an old problem." I accepted the edit as it was a condition for publication. In my book, I can finally declare that CNP is still my favorite open problem in all of mathematics.

This short paper also gave birth to a new definition and an open problem.

Definition 6.5 [HS1]. *An almost chromatic number* χ_a of the plane is the minimum number of colors that are required for coloring the plane so that almost all (i.e., all but one) colors forbid a unit distance and the remaining color forbids a distance.

We have the following inequalities for χ_a:

$$4 \leq \chi_a \leq 6.$$

The lower bound follows from Dmitry Raiskii's [Rai]. I proved the upper bound in 6.4 above [Soi6]. This naturally gave birth to a new problem, which is still open:

Open Problem 6.6 [HS1]. Find χ_a.

6.3 The Age of Tiling

Hadwiger's, Stechkin's, and my ornaments (Figs. 2.4, 6.2, and 6.6, respectively) delivered new mathematical results. They were also aesthetically pleasing. Have we contributed something, however little, to the arts? Not really. Nothing is new in the world of arts. We can find Henry Moore's aesthetics in pre-Columbian art and Picasso's cubistic geometrization

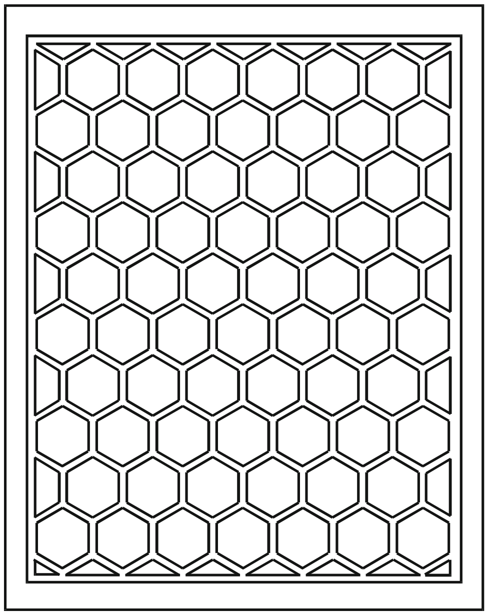


Fig. 6.6 Chinese lattice 1

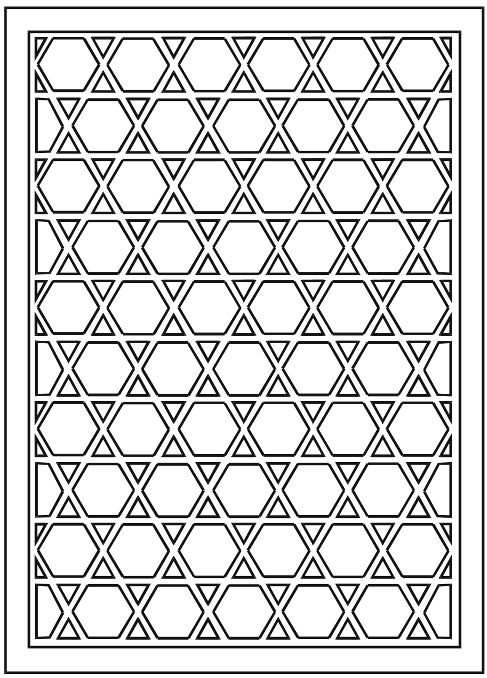


Fig. 6.7 Chinese lattice 2

of form in the art of sub-Saharan Africa. Our ornaments too were known for over 1000 years to the artists of China, India, Persia, Turkey, and Europe. Figures 6.6, 6.7, and 6.8, reproduced with kind permission from the Harvard-Yenching Institute from the wonderful 1937 book *A*

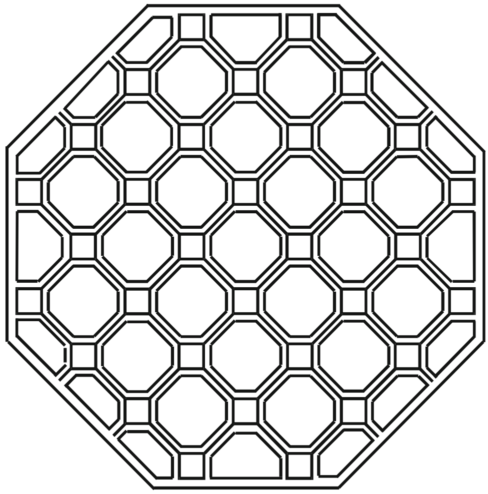


Fig. 6.8 Chinese lattice 3

Grammar of Chinese Lattice by Daniel Sheets Dye [Dye], show how those ornaments were implemented in old Chinese lattices.

If it is any consolation, I can point out that the Chinese ancestors did not invent the beauty and strength of honeycombs either: Bees were here first!

Chapter 7
Continuum of 6-Colorings of the Plane

In 1993, another 6-coloring was found by Ilya Hoffman and me ([HS1], [HS2]). Its type was $\left(1, 1, 1, 1, 1, \sqrt{2} - 1\right)$. The story of this discovery is noteworthy. In the summer of 1993, I was visiting my cousin in Moscow, a well-known New Vienna School composer, Leonid Hoffman. His 15-year-old son Ilya was studying violin at the Gnessin Music High School. Ilya set out to learn what I was doing in mathematics and did not accept any general answers. He wanted particulars. I showed him my 6-coloring of the plane (Problem 6.4), and the teenage musician got busy. The very next day he showed me . . . the Stechkin coloring (Fig. 6.2) that he discovered on his own! "Great," I replied, "but you are 23 years late." A few days later, he came up with a new idea of using a two-square tiling. Ilya had an intuition of a virtuoso fiddler and no mathematical culture – and so I calculated the sizes the squares had to have for the 6-coloring to do the job we needed. I wanted Ilya to be the sole author, but he insisted on our joint credit. And the joint work of the unusual mathematician–musician team was born. Ilya went on to graduate from the graduate school of Moscow Conservatory in the class of the celebrated violist and conductor Yuri Bashmet and is now one of Russia's hottest violinists and violists and the winner of several international competitions.

Problem 7.1 (I. Hoffman and A. Soifer [HS1], [HS2]). There is a 6-coloring of the plane of type $\left(1, 1, 1, 1, 1, \sqrt{2} - 1\right)$.

Solution Tile the plane with squares of diagonals 1 and $\sqrt{2} - 1$ (Fig. 7.1). We use colors 1, ..., 5 for larger squares and color 6 for all smaller squares. With each square, we include half of its boundary, its left and lower sides, without the end points of this half (Fig. 7.2).

To easily verify that this coloring does the job, observe the unit of the construction that is bounded by the bold line in Fig. 7.1. Its translates tile the plane and thus define its coloring. ■

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_7

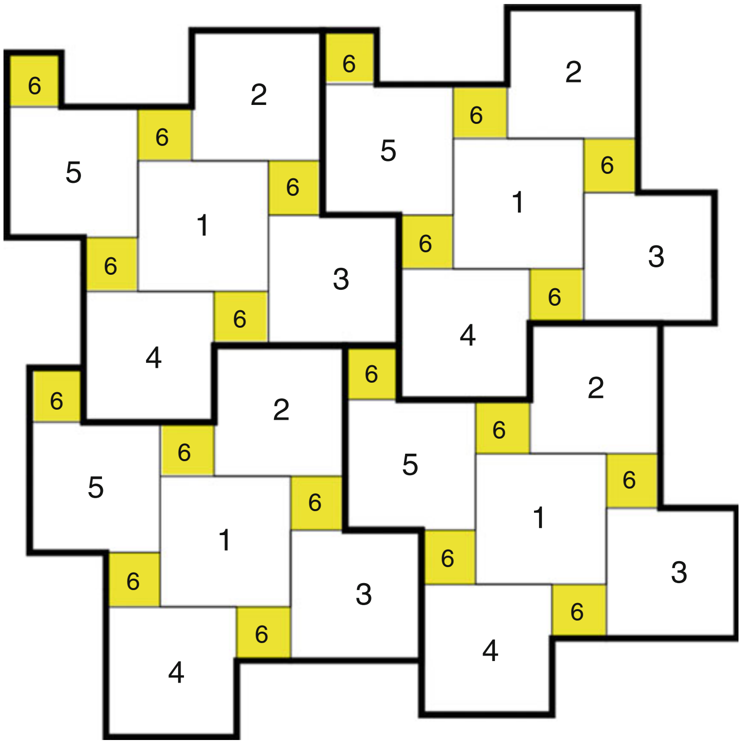


Fig. 7.1 The Hoffman–Soifer 6-coloring of the plane

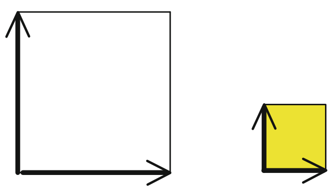


Fig. 7.2 Coloring of the boundaries

The two examples, found in the solutions of problems 6.4 and 7.1, prompted me in 1993 to introduce a new terminology for this problem and to translate the results and problems into this new language.

Open Problem 7.2 (A. Soifer [Soi7], [Soi8]). Find the *6-realizable set* X_6 of all positive numbers α such that there exists a 6-coloring of the plane of type $(1, 1, 1, 1, 1, \alpha)$.

In this new language, the results of problems 6.4 and 7.1 can be written as follows:

$$\frac{1}{\sqrt{5}}, \sqrt{2} - 1 \in X_6.$$

We now have two examples of "working" 6-colorings. But what do they have in common? It is not obvious, is it? One uses octagons, while the other does not. After a while, I realized that they were two extreme examples of a general case and, in fact, a much better result was possible, describing a whole continuum of working 6-colorings!

Theorem 7.3 (A. Soifer [Soi7], [Soi8]).

$$\left[\sqrt{2}-1, \frac{1}{\sqrt{5}}\right] \subseteq X_6,$$

i.e., for every $\alpha \in \left[\sqrt{2}-1, \frac{1}{\sqrt{5}}\right]$, there is a 6-coloring of type $(1, 1, 1, 1, 1, \alpha)$.[1]

Proof Let a unit square be partly covered by a smaller square, which cuts off the unit square into vertical and horizontal segments of lengths x and y, respectively, and forms with it an angle ω (Fig. 7.3). These squares induce the tiling of the plane that consists of nonregular octagons and "small" squares that are congruent to each other (Fig. 7.4).

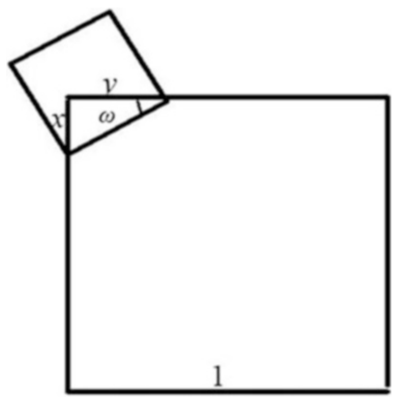


Fig. 7.3 The foundation squares

[1] Symbol $[a,b]$, $a < b$, as usual, stands for the line segment, including its end points a and b.

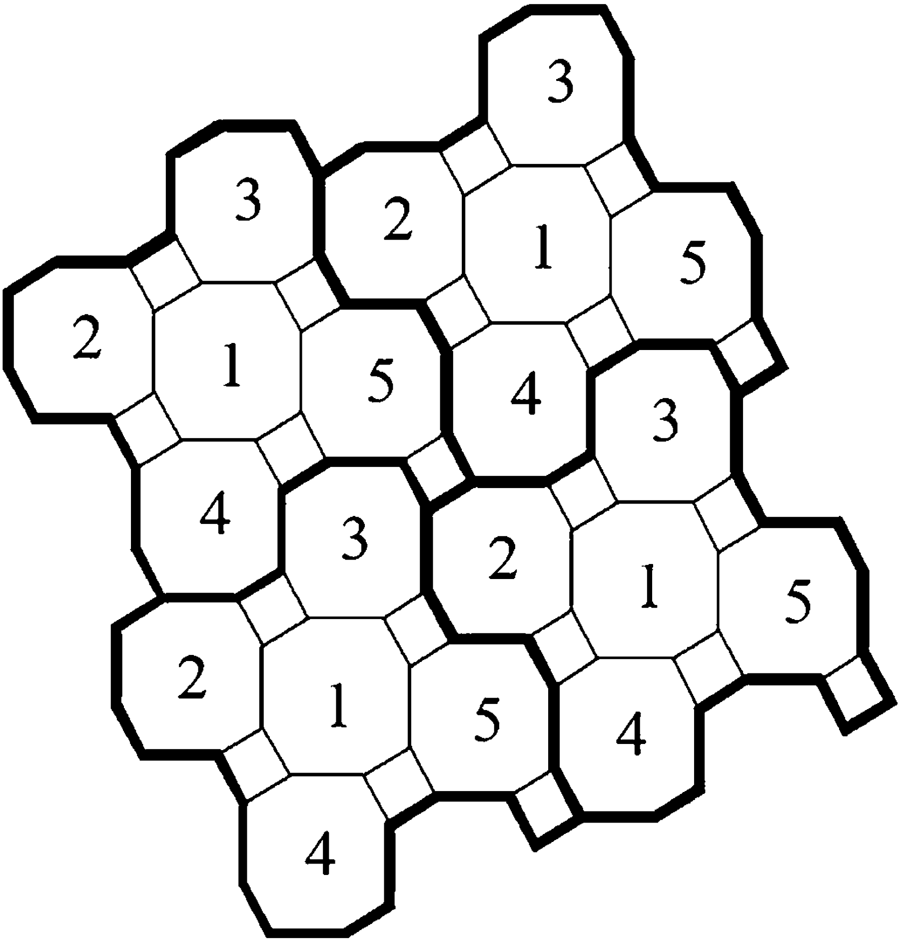


Fig. 7.4 The Soifer continuum of 6-colorings of the plane

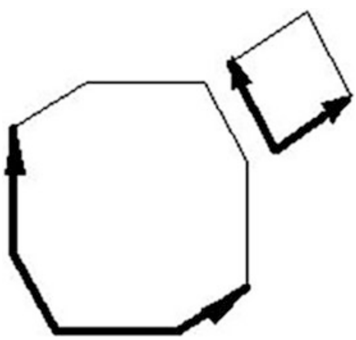


Fig. 7.5 Coloring boundaries

Now we are ready to color this tiling in 6 colors. Denote by *F* the unit of our construction, bounded by a bold line (Fig. 7.4) and consisting of five octagons and four small squares. Use colors 1 through 5 for the octagons inside *F* and color 6 for all small squares. Include in the colors of octagons and small squares the parts of their boundaries that are shown in bold in Fig. 7.5. Translates of F tile the plane and thus define the 6-coloring of the plane. We now wish to select the parameters to guarantee that each color forbids a distance.

At first, the complexity of computations appeared unassailable to me. However, a true Math Olympiad approach (i.e., good choices of variables, clever substitutions, and nice optimal properties of the chosen tiling) allowed for a successful sailing.

Let $x \leq y$ (Fig. 7.3). It is easy to see (Figs. 7.6 and 7.7) that we can split each small square into four congruent right triangles with sides x and y and a square of side $y - x$.

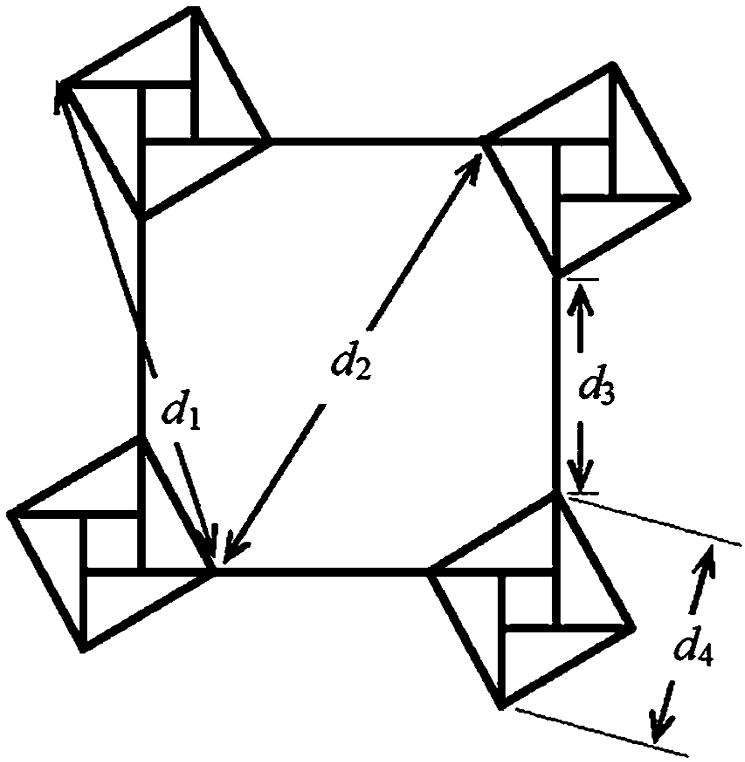


Fig. 7.6 A closer look at the tiling's foundation

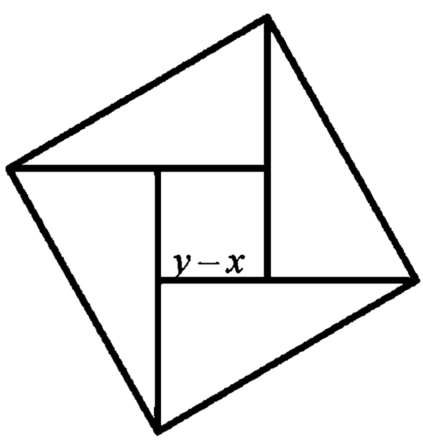


Fig. 7.7 A foundation close-up

The requirement for each color to forbid a distance produces the following system of two inequalities (see Fig. 7.6):

$$\begin{cases} d_1 \geq d_2 \\ d_3 \geq d_4 \end{cases}$$

Figures 7.6 and 7.7 allow for an easy representation of all d_i $(i = 1, 2, 3, 4)$ in terms of x and y. As a result, we get the following system of inequalities:

$$\left.\begin{array}{r} \sqrt{(1+y-x)^2+(2x)^2} \geq \sqrt{1+(1-2x)^2} \\ 1-x-y \geq \sqrt{2(x^2+y^2)} \end{array}\right\} \tag{7.2}$$

Solving for x in each of the two inequalities in (7.2) separately, we unexpectedly get the following system:

$$x^2+2(1-y)x+\left(y^2+2y-1\right) \geq 0$$
$$x^2+2(1-y)x+\left(y^2+2y-1\right) \leq 0.$$

Therefore, we get the equation (!) in x and y:

$$x^2+2(1-y)x+\left(y^2+2y-1\right) = 0.$$

Treating this as the equation in variable x, we obtain a *unique* (!) solution for x as a function of y that satisfies the system (7.2) of inequalities:

$$x = \sqrt{2-4y}+y-1, \text{where } 0 \leq y \leq 0.5. \tag{7.3}$$

Since $0 \leq x \leq y$, we get even narrower bounds for y: $0.25 \leq y \leq \sqrt{2}-1$. For any value of y within these bounds, x is uniquely determined by (7.3) and is accompanied by the *equalities* (!) $d_1 = d_2$ and $d_3 = d_4$.

Thus, we showed that for every $y \in \left[0.25, \sqrt{2}-1\right]$, there is a 6-coloring of type $(1, 1, 1, 1, 1, \alpha)$. But what values can α take on? Surely,

$$\alpha = \frac{d_4}{d_2}. \tag{7.4}$$

Let us introduce a new variable $Y = \sqrt{2-4y}$, where $Y \in \left[2-\sqrt{2}, 1\right]$ and figure out x and y from (7.3) as functions of Y:

$$4y = -Y^2+2$$
$$4x = -Y^2+4Y-2 \tag{7.5}$$

Now substituting from (7.1) and (7.2) the expressions for d_4 and d_2 into (7.4), and using the two equalities (7.5) to get rid of x and y everywhere, we get a "nice" expression for α^2 as a function of Y (do verify my algebraic manipulations on your own):

$$\alpha^2 = \frac{Y^4 - 4Y^3 + 8Y^2 - 8Y + 4}{Y^4 - 8Y^3 + 24Y^2 - 32Y + 20}.$$

By substituting $Z = Y - 2$, where $Z \in \left[-\sqrt{2}, -1\right]$, we get a simpler function α^2 of Z:

$$\alpha^2 = 1 + \frac{4Z\left(Z^2 + 2Z + 2\right)}{Z^4 + 4}.$$

To observe the behavior of the function α^2, we compute its derivative:

$$\left(\alpha^2\right)' = -\frac{4}{\left(Z^4 + 4\right)^2}\left(Z^6 + 4Z^5 + 6Z^4 - 12Z^2 - 16Z - 8\right).$$

Normally, there is nothing promising about finding the exact roots of an algebraic polynomial of a degree greater than 4. But we are positively lucky here, for this sixth-degree polynomial can be nicely decomposed into factors:

$$\left(\alpha^2\right)' = -\frac{4}{\left(Z^4 + 4\right)^2}\left(Z^2 - 2\right)\left[(Z+1)^2 + 1\right]^2.$$

Hence, the derivative has only two zeros. In fact, in the segment of our interest, $Z \in \left[-\sqrt{2}, -1\right]$, the only extremum of α^2 occurs when $Z = -\sqrt{2}$. Going back from Z to Y to y, we see that on the segment $y \in \left[0.25, \sqrt{2} - 1\right]$, the function $\alpha = \alpha(y)$ decreases from $\alpha = \frac{1}{\sqrt{5}} \approx 0.44721360$ (i.e., a 6-coloring of problem 6.4) to $\alpha = \sqrt{2} - 1 \approx 0.41421356$ (i.e., a 6-coloring of problem 7.1). Since the function $\alpha = \alpha(y)$ is continuous and increasing on $\left[0.25, \sqrt{2} - 1\right]$, it takes on *each* intermediate value from the segment $\left[\sqrt{2} - 1, \frac{1}{\sqrt{5}}\right]$ and only *once*.

We have proved the required result and much more:

For every angle ω between the small and the large squares (see Fig. 7.3)*, there are unique sizes of the two squares (and unique square intersection of parameters x and y) such that the constructed 6-coloring has type* $(1, 1, 1, 1, 1, \alpha)$ *for a uniquely determined α.*

This is a remarkable fact: the working solutions barely exist – they form something of a curve in a three-dimensional space formed by the angle ω and two linear variables x and y! We thus found a continuum of permissible values for α and a continuum of working 6-colorings of the plane. ■

Remark The problem of finding the 6-realizable set X_6 has a close relationship with the problem of finding the chromatic number χ of the plane. Its solution would shed light – if not solve – the chromatic number of the plane problem:

If $1 \notin X_6$, then $\chi = 7$;
If $1 \in X_6$, then $\chi \leq 6$.

Open Problem 7.4 (A. Soifer [Soi5]). Find X_6.

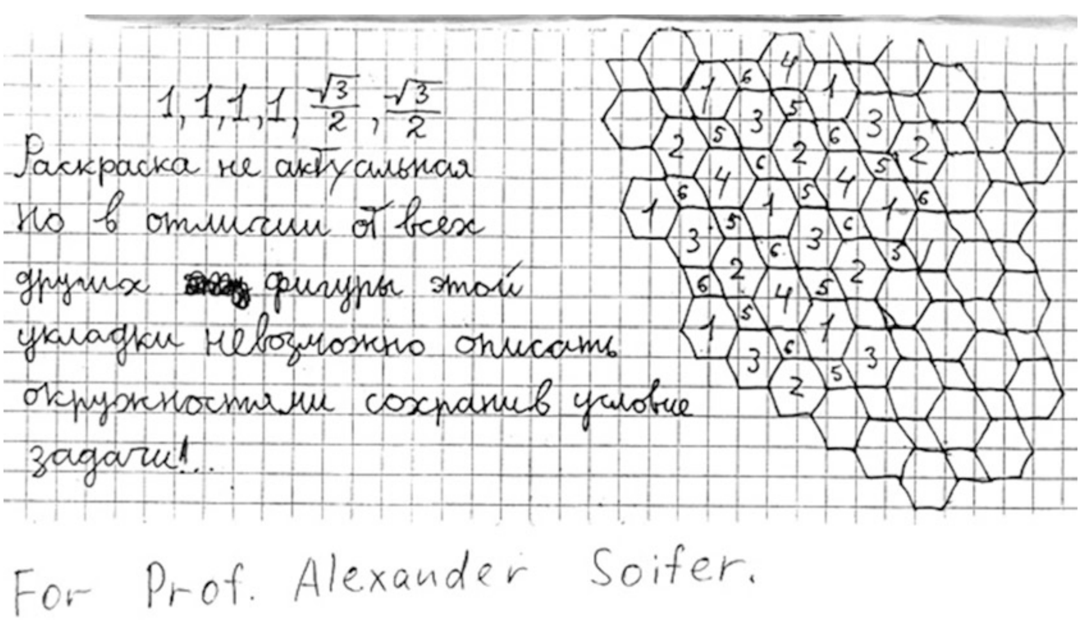
I am sure you understand that this problem, formulated in just two words, is extremely difficult.

In 1999, the Russian authorities accused my young coauthor and fine young violinist Ilya Hoffman of computer hacking (even though he did not pocket any money) and imprisoned him before the trial as a "danger to the society." I flew to Moscow, met with the presiding judge, a middle-aged pretty lady, in a black gown, of course. We were alone in her office. I asked, "What danger to the society does my nephew-violinist present?" The judge replied that she was not at liberty to do what she thought was right. I understood: she could have lost her job for that – or worse.

I met with Valery Vasilyevich Borshchev, a member of the Russian Parliament "Duma" and a human rights supporter. I also met with the vice president of the Russian Academy of Sciences and the head of the Judicial Division of the Academy, Vladimir Nikolaevich Kudryavtsev, who listened to me and generously volunteered to write a "Friend of the Court" opinion if the case were to reach the level of the City of Moscow Court or higher. Permit me to tell you a few words about the celebrated Jurist Kudryavtsev (10 April 1923–5 October 2007).

In 1951, Stalin's prosecutor general, Vyshinsky, announced a new legal doctrine: "One is guilty whom the court finds guilty." He called the presumption of innocence "bourgeois superstition." A young senior lieutenant rose to speak against the new Stalin's doctrine announced by Vyshinsky. This extraordinary hero was Vladimir N. Kudryavtsev. It was unforgettable to meet this brave man and get his full understanding and support.

When the trial finally took place, Ilya was released home from the courtroom. While in prison, he was not allowed to play his viola and violin, so Ilya wrote music and mathematics. The following page he sent to me from his prison cell:



Ilya discovered a new 6-coloring of the plane. Four colors consist of regular hexagons of diameter 1 and two colors occupy rhombuses. By carefully assigning colors to the boundaries, we get a 6-coloring of type $(1, 1, 1, 1, \frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2})$.

When my writings require an English translation of brilliant Russian poetry, I connect with Ilya for a joint translation work. “Always invite me to play linguistic combinatorics – I’m very pleased,” wrote Ilya to me on New Year’s Day, January 1, 2023.

Chapter 8
Chromatic Number of the Plane in Special Circumstances

As you know from Chapters 4 and 6, 3 years after Dmitry E. Raiskii, in 1973, Douglas R. Woodall published the paper [Woo1] on problems related to the chromatic number of the plane. In it, he provided his own proofs of Raiskii's inequalities of problems 4.1 and 6.1. In the same paper, Woodall also formulated and attempted to prove a lower bound for the chromatic number of the plane for the special case of map-type coloring of the plane. This was the main result of [Woo1]. However, in 1979, the mathematician from the University of Aberdeen Stephen Phillip Townsend found an error in Woodall's proof and constructed a counterexample, demonstrating that one essential idea of Woodall's proof was false. Townsend had also found a proof of Woodall's statement, which was very much more elaborate than Woodall's unsuccessful attempt.

The intriguing history of this discovery and Townsend's wonderful proof are a better fit in Chapter 25, as a part of our discussion of map coloring – do not overlook them. Here I will formulate an important corollary of Townsend's proof.

Chromatic Number of a Map-Type Colored Plane 8.1 The chromatic number of the plane under a map-type coloring is 6 or 7.

Woodall showed that this result implies one more meritorious statement:

Closed Chromatic Number of the Plane 8.2 [Woo1]. The chromatic number of the plane under coloring with closed monochromatic sets is 6 or 7.

I prefer not to use the Greek word "lemma" since there is an appropriate English word "tool". Moreover, I would like to offer you the following tool from topology to prove on your own. We will use this tool in the proof that follows.

Tool 8.3 If a bounded closed set S does not realize a distance d, then there is $\varepsilon > 0$ such that S does not realize any distance from the segment $[d - \varepsilon, d + \varepsilon]$.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_8

Proof of Result 8.2 [Woo1] Assume that the union of closed sets $A_1, A_2, \ldots, A_n$ covers the plane and that for each i the set A_i does not realize a distance d_i. Place onto the plane a unit square lattice L and choose an arbitrary closed unit square U of L. Choose also i from the set $\{1, 2, \ldots, n\}$. Denote by $C(U)_i$ the closed set that contains all points of the plane that are at most at the distance d_i from a point in U. The set $A_i \cap C(U)_i$ is closed and bounded; thus, by tool 8.3, there is $\varepsilon_i(U)$ such that no two points of A_i, at least one of which lies in U, realize any distance from the segment.

$$[d_i - \varepsilon_i(U), d_i + \varepsilon_i(U)]. \tag{8.1}$$

Denote by $\varepsilon(U)$ the minimum of $\varepsilon_i(U)$ over all $i = 1, 2, \ldots, n$.

Now, for the square U, we choose a positive integer $m(U)$ such that

$$\frac{1}{2^{m(U)}}\sqrt{2} < \frac{1}{2}\varepsilon(U). \tag{8.2}$$

On the unit square U, we place a square lattice L' of little closed squares u of side $\frac{1}{2^{m(U)}}$. The inequality (8.2) guarantees that the diagonal of u is shorter than half of $\varepsilon(U)$.

For each little square u contained in each unit square U of the entire plane, we determine $f(u) = \min\{i : u \cap A_i \neq \varnothing\}$ and, then, for each $i = 1, 2, \ldots, n$ define the monochromatic color set of our new *n-coloring* of the plane as follows:

$$B_i = \bigcup_{f(u)=i} u \tag{8.3}$$

As unions of closed squares u, each B_i is closed, and all B_i together cover the plane. The interiors of these n sets B_i are obviously disjoint. All there is left to prove is that the set B_i does not realize the distance d_i. Indeed, assume that the points b, c of B_i are at the distance d_i apart. The points b, c belong to little squares u_1, u_2, respectively, with each little square of side $\frac{1}{2^{m(U)}}$. Due to the definition (8.3) of B_i, the squares u_1, u_2 contain points a_1, a_2 from A_i, respectively. With vertical bars denoting the distance between two points and by utilizing the inequality (8.2), we get:

$$|b, c| - \varepsilon(U) < |a_1, a_2| < |b, c| + \varepsilon(U),$$

i.e.,

$$d_i - \varepsilon(U) < |a_1, a_2| < d_i + \varepsilon(U),$$

which contradicts (8.1).

Thus, the chromatic number under the conditions of result 8.2 is not smaller than that under the conditions of result 8.1. ■

During 1993–1994, a group of three young undergraduate students, Nathanial Brown, Nathan Dunfield, and Greg Perry, in a series of three essays, their first publications, proved on the pages of *Geombinatorics* [BDP1, BDP2, BDP3][1] that a similar result is true for coloring the plane with open monochromatic sets. As was easy to predict, the youngsters became professors of mathematics – Nathan at the University of Illinois at Urbana-Champaign and Nathanial at Pennsylvania State University.

Open Chromatic Number of the Plane 8.4 (Brown–Dunfield–Perry). The chromatic number of the plane under coloring with open monochromatic sets is 6 or 7.

[1] The important problem book by Brass–Moser–Pach [BMP] mistakenly cites only one of these series of three papers. It also incorrectly states that the authors proved only the lower bound 5, whereas they raised the lower bound to 6.

Chapter 9
Measurable Chromatic Number of the Plane

Check for updates

9.1 Definitions

As you know, the *length* of a segment $[a, b]$, $a < b$, on the line E^1 is defined as $b - a$. The *area* A of a rectangle $[a_1, b_1] \times [a_2, b_2]$, $a_i < b_i$, in the plane E^2 is defined as $A = (b_1 - a_1)(b_2 - a_2)$. The French mathematician Henri Léon Lebesgue (1875–1941) generalized the notion of area to a vast class of plane sets. In place of area, he used the term measure. For a set S in the plane, we define its measure $\mu^*(S)$ as follows:

$$\mu^*(S) = \inf \sum_i A(R_i) \tag{9.1}$$

with the infimum taken over all coverings of S by a countable sequence $\{R_i\}$ of rectangles. When the infimum exists, S is said to be *Lebesgue-measurable* or – since we consider here no other measures – a *measurable* set – if for any set B in the plane $\mu^*(B) = \mu^*(B \cap S) + \mu^*(B \backslash S)$. For a measurable set S, its measure is defined by $\mu(S) = \mu^*(S)$.

Any rectangle is measurable, and its measure coincides with its area. It is shown in every measure theory text that all closed sets and all open sets are measurable. Giuseppe Vitali (1875–1932) was the first to show that in the standard system of axioms **ZFC** for set theory (Zermelo–Fraenkel system plus the axiom of choice), there are nonmeasurable subsets of the set R of real numbers.

We will use the same definition (9.1) for Lebesgue measure on the line E^1, when the infimum is naturally taken over all countable covering sequences $\{R_i\}$ of segments. For a measure of S on the line, we will use the symbol $l(S)$. Generalization of the notion of measure to n-dimensional Euclidean space E^n is straightforward; here we will use the symbol $\mu_n(S)$. In particular, for $n = 2$, we will omit the subscript and simply write $\mu(S)$.

9.2 Bounds for the Measurable Chromatic Number of the Plane

While a graduate student in Great Britain, Kenneth J. Falconer proved the following important result [Fal1]:

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_9

Falconer's Theorem 9.1 Let $E^2 = \bigcup_{i=1}^{4} A_i$ be a covering of the plane by four disjoint measurable sets. Then at least one of the sets A_i realizes distance 1.

I found his 1981 publication [Fal1] to be too concise and not self-contained for the result that I viewed as very important. Accordingly, I asked Kenneth Falconer, professor and dean at the University of St. Andrews in Scotland, for a more detailed self-contained exposition. In February 2005, I received Kenneth's manuscript [Fal2], handwritten especially for this book, which I am delighted to share with you.

Before we prove his result, we need to arm ourselves with some basic definitions and tools of the measure theory.

A non-empty collection $\beth$ of subsets of E^2 is called *σ-field*, if $\beth$ is closed under taking complements and countable unions, i.e.,

(*) If $A \in \beth$, then $E^2 \backslash A \in \beth$ and

(**) If $A_1, A_2, \ldots, A_n, \ldots \in \beth$, then $\bigcup_{i=1}^{\infty} A_i \in \beth$.

Exercise 9.2 Show that any σ-field $\beth$ is closed under a countable intersection and set difference. Also, show that $\beth$ contains the empty set $\varnothing$ and the whole space E^2.

It is shown in all measure theory textbooks that the collection of all measurable sets is a σ-field. The intersection of all σ-fields containing the closed sets is a σ-field containing the closed sets and is the minimal such σ-field with respect to inclusion. Its elements are called *Borel sets*. Since closed sets are measurable and the collection of all measurable sets is a σ-field, it follows that all Borel sets are measurable.

(Observe that in place of the plane E^2, we can consider the line E^1 or an n-dimensional Euclidean space E^n and define their Borel sets.)

The following notations will be helpful:

$C(x, r)$ – A circle with center at x and radius r

$B(x, r)$ – A circular disk (or ball) with center at x and radius r

For a measurable set S and a point x, we define the *Lebesgue density*, or simply *density*, of S at x as follows:

$$D(S,x) = \lim_{x \to 0} \frac{\mu(S \cap B(x,r))}{\mu(B(x,r))},$$

where $\mu(B(x, r))$ is, of course, equal to πr^2.

Lebesgue Density Theorem (LDT) 9.3 For a measurable set $S \subset E^2$, the density $D(S, x)$ exists and equals 1 if $x \in S$ and 0 if $x \in E^2 \backslash S$, except for a set of points x of measure 0.

For a measurable set A, denote

$$\tilde{A} = \{x \in A : D(A,x) = 1\}.$$

Then, due to LDT, we get $\mu(\tilde{A}\triangle A) = 0$, i.e., $\tilde{A}$ is "almost the same" as A.[1] Observe also that $\mu(S \cap B(x, r))$ is a continuous function of x for $r > 0$; therefore, $\tilde{A}$ is a Borel set.

We will define the *density boundary* of a set A as follows:

$$\partial A = \{x : D(A, x) \neq 0, 1 \text{ or does not exist}\}.$$

By LDT,

$$\mu(\partial A) = 0.$$

You can prove on your own or read in [Cro] the proof of the following tool:

Tool 9.4 For a measurable set $A \subset E^2$, such that both $\mu(A) > 0$ and $\mu(E^2 \backslash A) > 0$, we have $\partial A \neq \varnothing$.

Tool 9.5 If $E^2 = \bigcup_{i=1}^{4} A_i$ is a covering of the plane by four disjoint measurable sets, then $E^2 = \bigcup_{i=1}^{4} \tilde{A}_i$ is a disjoint union with the complement $\mathfrak{M} \equiv \bigcup_{i=1}^{4} \partial A_i$.

Proof follows from tool 9.4 and the observation that if $x \in \partial A_i$ then also $x \in \partial A_j$ for some $j \neq i$. ■

The next tool claims the existence of two concentric circles with the common center in $\mathfrak{M}$, which intersects $\mathfrak{M}$ in length 0.

Tool 9.6 Let $\mathfrak{M}$ be as defined in tool 9.5; there exists $x \in \mathfrak{M}$ such that

$$l(C(x, 1) \cap \mathfrak{M}) = l\left(C\left(x, \sqrt{3}\right) \cap \mathfrak{M}\right) = 0.$$

I will omit the proof but include Falconer's insight: "The point of this lemma is that if we place the 'double equilateral triangle' [see Fig. 9.1] of side 1 in almost all orientations with a vertex at x, the point x essentially has 'two colours' in any colouring of the plane, and other points just one colour. (Note $|xw| = \sqrt{3}$.)"

Tool 9.7 Let $E^2 = \bigcup_{i=1}^{4} A_i$ be a covering of the plane by four disjoint measurable sets, none of which realizes distance 1. Let $x \in \mathfrak{M}$ as defined in tool 9.6, and without loss of generality, let $x \in \partial A_1$ and $x \in \partial A_2$. Then $l\left(C\left(x, \sqrt{3}\right) \backslash \left(\tilde{A}_1 \cup \tilde{A}_2\right)\right) = 0$.

Proof Since $x \in \partial A_1$ and $x \in \partial A_2$, there exists $\varepsilon > 0$ such that

(1) $\varepsilon < \frac{\mu(A_1 \cap B(x, r))}{\pi r^2} < 1 - \varepsilon$ for some arbitrarily small r and
(2) $\varepsilon < \frac{\mu(A_2 \cap B(x, r))}{\pi r^2} < 1 - \varepsilon$ for some arbitrarily small r

[1] Here $A \Delta B$ stands for the symmetric difference of these two sets, i.e., $A \Delta B = (A \backslash B) \cup (B \backslash A)$.

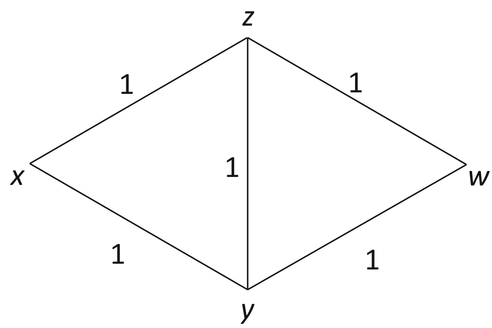


Fig. 9.1 A unit diamond

Consider the diamond (Fig. 9.1) consisting of two unit equilateral triangles xyz and yzw, where x is the point fixed in the statement of this tool, and $y, z, w \notin \mathfrak{M}$ (this happens for almost all orientations of the diamond, due to tool 9.6). Thus, suppose $y \in \tilde{A}_{i(y)}$, $z \in \tilde{A}_{i(z)}$, and $w \in \tilde{A}_{i(w)}$, where $i(y), i(z), i(w) \in \{1, 2, 3, 4\}$. For a sufficiently small r, say $r < r_0$, we get:

(3) $1 - \frac{\varepsilon}{4} < \frac{\mu\left(A_{i(y)} \cap B(y, r)\right)}{\pi r^2} \leq 1$;

(4) $1 - \frac{\varepsilon}{4} < \frac{\mu\left(A_{i(z)} \cap B(z, r)\right)}{\pi r^2} \leq 1$;

(5) $1 - \frac{\varepsilon}{4} < \frac{\mu\left(A_{i(w)} \cap B(w, r)\right)}{\pi r^2} \leq 1$.

We can now choose $r < r_0$ such that (1) holds (as well as (3), (4), and (5)). Let v be a vector going from the origin to a point in $B(0, r)$ and consider translation of the diamond x, y, z, w through v, i.e., to the diamond $x + v, y + v, z + v, w + v$. Now (1), (3), (4), and (5) imply that

$$\frac{1}{\pi r^2}\mu\left(\left\{B(0, r) : x + v \in A_1, y + v \in A_{i(y)}, z + v \in A_{i(z)}, w + v \in A_{i(w)}\right\}\right) > \varepsilon - \frac{\varepsilon}{4} - \frac{\varepsilon}{4} - \frac{\varepsilon}{4} > 0.$$

Thus, we can choose $v \in B(0, r)$ such that $x + v \in A_1, y + v \in A_{i(y)}, z + v \in A_{i(z)}, w + v \in A_{i(w)}$. Since by our assumption none of the sets A_i, $i = 1, 2, 3, 4$ realizes distance 1, we conclude (by looking at the translated diamond) that $1 \neq i(y), 1 \neq i(z)$, $i(y) \neq i(z)$, $i(z) \neq i(w)$, and $i(w) \neq i(y)$.

The same argument, using (2), (3), (4), and (5), produces $2 \neq i(y)$, $2 \neq i(z)$, $i(y) \neq i(z)$, $i(z) \neq i(w)$, and $i(w) \neq i(y)$. Therefore, $i(y)$, $i(z)$ are 3 and 4, respectively, in some order, and thus, $i(w) = 1$ or 2, i.e., $w \in \tilde{A}_1$ or $w \in \tilde{A}_2$.

By tool 9.6, this holds for almost every orientation of the diamond. Since $|xw| = \sqrt{3}$, we conclude that for almost all $w \in C\left(x, \sqrt{3}\right)$, we get $w \in \tilde{A}_1$ or $w \in \tilde{A}_2$. Thus, $l\left(C\left(x, \sqrt{3}\right) \setminus \left(\tilde{A}_1 \cup \tilde{A}_2\right)\right) = 0$, as required. ■

Tool 9.8 Let C be a circle of radius $r > \frac{1}{2}$ and let E_1, E_2 be disjoint measurable subsets of C such that $l(C \setminus (E_1 \cup E_2) = 0$. Then if $\phi = 2 \sin^{-1}\left(\frac{1}{2r}\right)$ is an irrational multiple of π, either E_1 or E_2 contains a pair of points distance 1 apart.

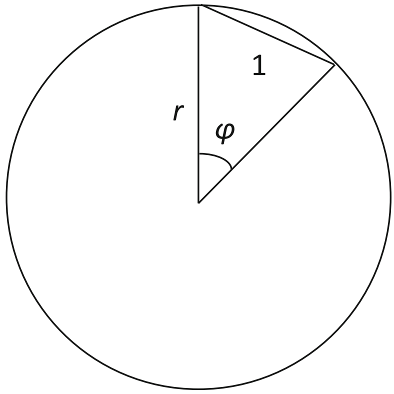


Fig. 9.2 Parameterizing C by angle θ (mod 2^{π})

Proof Assume that neither E_1 nor E_2 contains a pair of points distance 1 apart. Parameterize C (Fig. 9.2) by angle θ (mod 2π).

Let $l(E_1) > 0$, then by LDT, there is θ and $\varepsilon > 0$ such that

$$l(E_1 \cap (\theta - \varepsilon, \theta + \varepsilon)) > \frac{3}{4} 2\varepsilon.$$

Let θ_1 be an angle. Since ϕ is an irrational multiple of π, there is a positive integer n such that

$$|\theta_1 - (2n\phi + \theta)| < \frac{1}{4}\varepsilon \text{ (mod } 2\pi).$$

Since neither E_1 nor E_2 contain a pair of points distance 1 apart, we get (with angles counted mod 2π):
$l(E_1 \cap (\theta + k\phi - \varepsilon, \theta + k\phi + \varepsilon)) = l(E_1 \cap (\theta - \varepsilon, \theta + \varepsilon))$ for even k and
$l(E_1 \cap (\theta + k\phi - \varepsilon, \theta + k\phi + \varepsilon)) = 2\varepsilon - l(E_1 \cap (\theta - \varepsilon, \theta + \varepsilon))$ for odd k.
In particular, $l(E_1 \cap (\theta + 2n\phi - \varepsilon, \theta + 2n\phi + \varepsilon)) > \frac{3}{4} 2\varepsilon$, thus

$$l(E_1 \cap (\theta_1 - \varepsilon, \theta + \varepsilon)) > \frac{3}{4} 2\varepsilon - \frac{\varepsilon}{4} - \frac{\varepsilon}{4} = \varepsilon.$$

Hence, for all θ_1,

$$\frac{l(E_1 \cap (\theta_1 - \varepsilon, \theta + \varepsilon))}{2\varepsilon} \geq \frac{1}{2},$$

and by LDT $l(C \backslash E_1) = 0$. This means that E_1 includes almost all of C and, therefore, contains a pair of points until a distance apart, a contradiction. ■

Surprisingly, we need a tool from abstract algebra or number theory.

Tool 9.9 For any positive integer m, $\left(1 - i\sqrt{11}\right)^{2m} \neq (-12)^m$.

Proof It suffices to note that $Q\left(\sqrt{-11}\right)$ is a Euclidean quadratic field; therefore, its integer ring $Z\left(\sqrt{-11}\right)$ (with units +1/−1) has a unique factorization (see Chapters 7 and 8 in the standard abstract algebra textbook [DF] for a proof). ■

Note I believe that an alternative proof is possible: it should be not hard to show that the left side cannot be an integer for any m. ■

Now we are ready to prove Falconer's Theorem 9.1.

Proof of Falconer's Theorem Let $E^2 = \bigcup_{i=1}^{4} A_i$ be a covering of the plane by four disjoint measurable sets, none of which realizes distance 1. Due to tool 9.6, there is $x \in \mathfrak{M}$ such that $l\left(C\left(x, \sqrt{3}\right) \setminus \left(\tilde{A}_1 \cup \tilde{A}_2\right)\right) = 0$. Taking $E_1 = \tilde{A}_1, E_2 = \tilde{A}_2$, and $r = \sqrt{3}$, we get the desired result by tool 9.8 – if only we can prove that $\phi = \sin^{-1}\left(\frac{1}{2\sqrt{3}}\right)$ is an irrational multiple of π.

Assume that $m\theta$ is an integer multiple of 2π for some integral $2m$. Since $\sin\theta = \frac{1}{2\sqrt{3}}$; $\cos\theta = \frac{\sqrt{11}}{2\sqrt{3}}$, we get

$$\left(\frac{\sqrt{11}}{2\sqrt{3}} + i\frac{1}{2\sqrt{3}}\right)^{2m} = 1,$$

i.e.,

$$\left(1 - i\sqrt{11}\right)^{2m} = (-12)^m.$$

We are done, as the last equality contradicts tool 9.9. ■

9.3 Kenneth J. Falconer

I am always interested in learning about the life and personality of the author of a result that impressed me, aren't you? Accordingly, I asked Kenneth to tell me about himself and his life. The following account comes from his September 30, 2005, e-mail to me.

> I was born on 25th January 1952 at Hampton Court on the outskirts of London (at a maternity hospital some 100 metres from the gates of the famous Palace). This was two weeks before Queen Elizabeth II came to the throne and when food rationing was still in place. My father had served in India for 6 years during the war while my mother brought up my brother, 12 years my senior, during the London blitz. My parents were both school teachers, specialising in English, my brother studied history before becoming a Church of England minister, and I was very much the 'black sheep' of the family, having a passionate interest in mathematics and science from an early age . . .

I gained a scholarship to Corpus Christi College, Cambridge to read mathematics and after doing well in the Mathematical Tripos I continued in Cambridge as a research student, supervised by Hallard Croft. I worked mainly on problems in Euclidean geometry, particularly on convexity and of tomography (the mathematics of the brain scanner) and obtained my Ph.D. in 1977.

I had the good fortune to obtain a Research Fellowship at Corpus Christi College, where I continued to study geometrical problems, including the fascinating problem of the chromatic number of the plane, showing in particular that the chromatic number of a measurable colouring of the plane was at least 5. Also around this time I worked on generalisations of the Kakeya problem (the construction of plane sets of zero area containing a line segment in every direction). Thus I encountered Besicovitch's beautiful idea of thinking of such sets as duals of what are now termed 'fractals', with directional and area properties corresponding to certain projections of the fractals. This led to my 'digital sundial' construction – a subset of R^3 with prescribed projections in (almost) all directions . . .

In 1980 I moved to Bristol University as a Lecturer, where the presence of theoretical physicist Michael Berry, and analyst John Marstrand were great stimuli. Here I started to work on geometric measure theory, or fractal geometry, in particular looking at properties of Hausdorff measures and dimensions, and projections and intersections of fractals . . .

It became clear to me that much of the classical work of Besicovitch and his School on the geometry of sets and measures had been forgotten, and in 1985 I published my first book 'The Geometry of Fractal Sets' to provide a more up to date and accessible treatment. This was around the time that fractals were taking the world by storm, following Mandelbrot's conceptually foundational work publicised in his book 'The Fractal Geometry of Nature' which unified the mathematics and the scientific applications of fractals. My book led to requests for another at a level more suited to postgraduate and advanced undergraduate students and in 1990 I published 'Fractal geometry – Mathematical Foundations and Applications' which has been widely used in courses and by researchers, and has been referred to at conferences as 'the book from which we all learnt our fractal mathematics'. A sequel 'Techniques in Fractal Geometry' followed in 1998. In collaboration with Hallard Croft and Richard Guy, I also authored 'Unsolved Problems in Geometry,' a collection of easy to state unsolved geometrical problems. Happily (also sadly!) many of the problems in the book are no longer unsolved! . . .

In 1993 I was appointed Professor at the University of St Andrews in Scotland, where I have been ever since. Although St Andrews is a small town famous largely for its golf, the University has a thriving mathematics department, in particular for analysis and combinatorial algebra, to say nothing of its renowned History of Mathematics web site. I became Head of the School of Mathematics and Statistics in 2001, with the inevitable detrimental effect on research time. I was elected a Fellow of the Royal Society of Edinburgh in 1998, and to the Council of the London Mathematical Society in 2000 ...

My main leisure activity is long distance walking and hillwalking. I have climbed all 543 mountains in Britain over 2500 feet high. I am a keen member of the Long Distance Walkers Association, having been Editor of their magazine 'Strider' from 1986–91 and Chairman from 2000–03. I have completed the last 21 of the LDWA's annual hundred mile non-stop cross-country walks in times ranging from 26 to 32 hours.

* * *

Let me add the latest honor bestowed on Professor Falconer in early 2018. He was appointed Regius Chair of Mathematics by the Queen of England. The Regius Chair of Mathematics is one of the oldest and most prestigious positions in England, founded by King Charles II in 1668.

Chapter 10
Coloring in Space

For myself, I like a universe that includes much that is unknown and, at the same time, much that is knowable.
– Carl Sagan

When in 1958, Paul Erdős learned about the chromatic number of the plane problem, he created a number of related problems, some of which we have discussed in the preceding chapters. Paul also generalized the problem to the n-dimensional Euclidean spaces E^n. On October 2, 1991, I received a letter from him, which contained a historic remark [E91/10/2ltr][1]:

I certainly asked for the chromatic number of $E^{(n)}$ long ago (30 years).

Paul was interested in both the asymptotic behavior as n increased and in the exact values of the chromatic number for small n, first of all for $n = 3$.

As we have already discussed in Chapter 4, in 1970, Dmitry E. Raiskii [Rai] proved the lower bound for n-dimensional Euclidean spaces.

Lower Bound 10.1 (D. Raiskii, 1970). $n + 2 \leq \chi(E^n)$.

For $n = 3$, this, of course, gives $5 \leq \chi(E^3)$. This lower bound for the three-dimensional space had withstood 30 years, until in January 2000, Oren Nechushtan of Tel Aviv University improved it (and published 2 years later [Nec]):

Best Lower Bound 10.2 (O. Nechushtan): $6 \leq \chi(E^3)$.

At the end of his paper, Nechushtan remarks:

The proof given above can be used to obtain an explicit unit distance graph of the space with chromatic number 5. This graph has less than 400 vertices. The author believes that it is possible to reduce the size of such a graph.

Nechushtan's paper continues to inspire still today. The reduction in size was achieved not long ago by others, who started with Nechushtan's steps and then diverged from them. In 2020, Aubrey de Grey constructed a 6-chromatic unit distance graph in 3-dimensional space

[1] Curiously, Paul wrote an improbable date on this letter: "1977 VII 25": at that time, I still lived in Moscow.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_10

on just 59 vertices [G3]; he also reported that Exoo and Ismailescu independently constructed such a graph on 79 vertices.

The obvious upper bound of 27 for the chromatic number of three-dimensional space was reduced to 21; it was proved by David Coulson of Melbourne University in [Cou1]. Curiously, the credit for the bound of 21 Coulson seems to give to me:

> The upper bound of 21 was found [*Mathematical Coloring Book* (to appear in 1996–1997)] using a colouring based on the lattice $D_3 \cong A_3$, which is the face-centered cubic lattice.

First of all, no one had a copy of *The Mathematical Coloring Book* in 1996–1997 and that includes David Coulson and me – the book was finished and published a decade+ later, in November 2008. Moreover, I had nothing to do with proving this upper bound. David Coulson then goes on to further reduce the upper bound to 18 [Cou1]. Please, pay attention to the dates, as it seems Coulson's papers were slow to appear in print. The upper bound of 18 was first submitted in 1993 to the *Transaction of the American Mathematical Society* (on September 27, 1993, and I received an e-mail from Coulson to that effect). Then (I assume due to the lack of interest in the *Transactions* for combinatorial mathematics), Coulson submitted it to *Discrete Mathematics* on April 24, 1995; he revised the paper on August 30, 1996, and it was finally published in 1997 [Cou1], 4 years after the initial submission.

Coulson then achieved an amazing improvement: He obtained the upper bound of 15 using a face-centered cubic lattice (see Conway and Sloane [CS] for more about three-dimensional lattices). The upper bound of 15 also took 4 years to appear in print. It was submitted to *Discrete Mathematics* on December 9, 1998. A month later, I received this manuscript to referee under the number DM 9298. Amazingly, a copy of my February 27, 1999, report survives. I suggested five stylistic improvements and wrote:

> I found the main result to be of high importance to the field. Indeed, Coulson has dramatically improved his own previous bound of 18 by proving that 15 is an upper bound of the chromatic number of the 3-space. He conjectures that 15 is the best possible upper bound if one uses a lattice-based coloring. His argument in favor of this conjecture is good, and we would encourage the author to pursue the proof . . .
>
> The author hints that his methods may produce similar results in other dimensions. Again, the referee would encourage the pursuit of these results.

I am at a loss to explain why the revised manuscript was received by the editors only about 2 years later, on December 11, 2000. While writing these lines, I am looking at the uncorrected proofs that I received from the author – they are dated 2001. The paper was published much later yet, in 2002 [Cou2].

In August 2002, David Coulson and I played a very unusual role at the Congress of the World Federation of National Mathematics Competitions in Melbourne, Australia: we were co-presenters of the 80-minute-long plenary talk, entitled *50 Years of Chromatic Number of the Plane* (we did not sing in a duet but rather presented one person at a time. I spoke about the problem, its history, and results in the plane. On his part, David described his results on the upper bounds of the chromatic number of the Euclidean 3-space. After the talk, I invited David to submit a version of his part of the talk to *Geombinatorics*, where it appeared very quickly, in the January issue of 2003 [Cou3].

Best Upper Bound of 3-Space 10.3 (D. Coulson, 1998–2002). $\chi(E^3) \le 15$.

The curiosity surrounding this result did not end with its publication. It was published again in 2003 by a pair of authors, Rados Radoičić and Géza Tóth [RT]. By the time they received the proofs, the authors saw Coulson's *publication*. They added it to the bibliography and chose to publish their proof based on the same tiling of the 3-space. In a copy of this paper downloaded from an author's homepage, I read:

> Very recently, Coulson [C02] also [sic] proved our [sic] Theorem, moreover, he found essentially the same coloring.

The comment in the published journal version was fairer toward Coulson:

> **Added in proof.** Very recently, Coulson [C02] has independently found a very similar 15-coloring of 3-space.

I respectfully disagree with the characterization "very recently," for Coulson first submitted his paper years earlier than Radoičić and Tóth, on December 9, 1998. It is quite possible that Radoičić and Tóth found their proof independently and before reading Coulson's proof. Yet, I am assigning credit to Coulson alone for the following reasons:

1. Radoičić and Tóth saw Coulson's *publication* before they received their proofs.
2. Their proof is not essentially different from Coulson's.
3. Coulson first submitted his paper many years prior, in 1998.
4. Coulson has circulated his preprint fairly widely ever since 1998 (I was one of the recipients).

In the same paper [RT], Rados Radoičić and Géza Tóth contributed an important upper bound for the chromatic number of the four-dimensional Euclidean space:

Upper Bound of 4-Space 10.4 (Radoičić and Tóth [RT]).

$$\chi\left(E^4\right) \le 54.$$

As I mentioned in my referee report, Coulson conjectured that the upper bound of 15 is best possible *for lattice coloring*. I conjectured much more: I thought 15 was the exact value for the 3-space as likely as 7 was for the plane:

Chromatic Number of 3-Space Conjecture 10.5 (A. Soifer, 2002).

$$\chi\left(E^3\right) = 15.$$

Life in four and five dimensions was studied by Kent Cantwell in his 1996 work [Can1], who found the following lower bounds:

Lower Bounds for E^4 and E^5 10.6 (Cantwell, 1996).

$$\chi\left(E^4\right) \geq 7;$$
$$\chi\left(E^5\right) \geq 9.$$

Now, what do you think about Fig. 10.1?

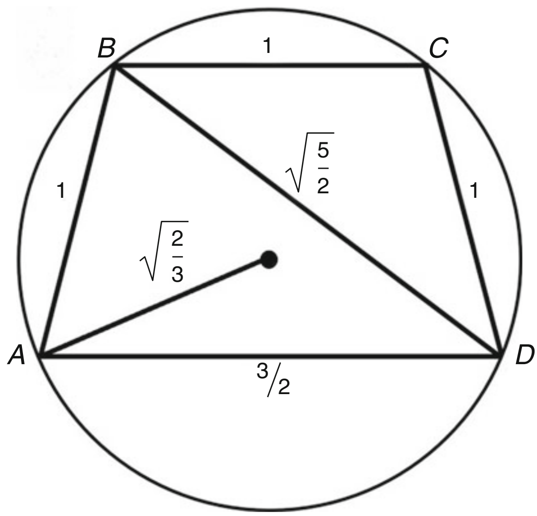


Fig. 10.1 A useful trapezoid

I hear you replying, "a trapezoid" and wondering: "what has this middle-school-like geometry drawing got to do with chromatic numbers of n-spaces?" In fact, Fig. 10.1 is the key in Leonid L. Ivanov's proof [Iva] of the first Cantwell's lower bound. Ivanov writes: "Lower bound [6] was proved by D.E. Raiskii. In 4-dimensional space we made one more step forward." This is incorrect, for Ivanov's proof came a whole 10 years after the 1996 Cantwell publication, and, Moscow State University, with a fine group working on these problems, should have done a better publication search. In 2006, Leonid Ivanov must have been an undergraduate student, for only in 2011 did he defend his PhD dissertation "Investigation of optimal configurations in problems about chromatic number of spaces and in the Borsuk problem" that included this proof. Thus, Ivanov does not get credit in this book for discovering the inequality $\chi(E^4) \geq 7$. However, his proof is shorter than Cantwell's, just one page. Would you like to see it? All right. There are no illustrations in Ivanov's proof; I am adding Fig. 10.1 for you to easily "see" through the proof. I am also shrinking the proof presentation to ca. 3/4 of a page.

Proof Assume that the Euclidean space E^4 is six-colored without a monochromatic unit segment. Toss on the plane $<x_1, x_2>$ spanned by the first two coordinate axes a circle O_1 of radius $\sqrt{2/3}$ with the center at the origin. Let O_2 be the circle in the plane $<x_3, x_4>$ of radius $\sqrt{1/3}$ with the center at the same origin. Inscribe in O_2 a unit equilateral triangle. Its vertices has to be colored in three colors. Since the distance between every point of O_1 and every point of O_2 is 1, this leaves at most three colors for the circle O_1.

Now choose on O_1 a sequence of points A, B, C, D with unit distances between neighbors (see Fig. 10.1). Let vertical bars denote the length of a segment; a simple calculation gives $|AD| = 3/2$; $|BD| = \sqrt{5/2}$. As we can see, any triangle with sides like in ABC can be extended to an isosceles trapezoid $ABCD$ (Fig. 10.1).

Let AA' be a unit segment and D at distance 3/2 from A and A'. Let A and D be of different colors. Let S be the set of all points B that are at a distance 1 from A and $\sqrt{5/2}$ from D. S is a 2-dimensional sphere in the 3-space orthogonal to AD. The radius of S is $\sqrt{15}/4$ (it is equal to the altitude from B in a triangle like ABD (Fig. 10.1). Choose on S a unit equilateral triangle $B_1 B_2 B_3$ (it exists because $\sqrt{15}/4 > \sqrt{3}/2$). Among its vertices, two must differ in color from D; let it be B_1, B_2. Extend each of the triangles AB_1D, AB_2D to isosceles trapezoids AB_1C_1D; AB_2C_2D, respectively. These trapezoids are inscribable in a circle of radius $\sqrt{2/3}$; therefore, by the argument of the first paragraph, the vertices of trapezoids must utilize at most three colors. This leaves for C_1 only the same color as A and so does for C_2 in its trapezoid. Finally, observe that $|B_1C_1| = |B_2C_2| = |B_1B_2| = |C_1C_2|$. ■

It took almost two decades and three coauthors for the first Cantwell's lower bound to be improved. Geoffrey Exoo, Dan Ismailescu, and Michael Lim achieved it in 2014. They constructed a 65-vertex unit distance graph of chromatic number 9 and embedded it into the Euclidean 4-space [EIL].

Lower Bound for 4-Space 10.7 (Exoo, Ismailescu, and Lim, 2014).

$$\chi\left(E^4\right) \geq 9.$$

Thus, the best-known bounds for the 4-space are $9 \leq \chi(E^4) \leq 54$, which, in my opinion, are about equally distant from the exact value.

Conjecture for 4-Space 10.8 (Soifer, 2002).

$$\chi\left(E^4\right) = 31.$$

On March 31, 2008, I received an impressive submission [Cib] to *Geombinatorics* from Josef Cibulka of Charles University in Prague. His main result offered the new lower bound for the chromatic number of E^6:

Lower Bounds for E^6 10.9 (Cibulka, 2008).

$$\chi\left(E^6\right) \geq 11.$$

In reply to my inquiry, Josef answered on April 1, 2008:

> I am first year graduate student; actually, most results of the submitted paper are from my diploma thesis.

In 2014, Geoffrey Exoo and Dan Ismailescu improved Cibulka's result and established the lower bounds of chromatic numbers of Euclidean spaces E^n for $n = 6, 7, 10, 11, 12, 13$, and 14. Let us record their achievement:

Lower Bounds for Euclidean Spaces E^n for $n = 6,7,10,11,12,13,14$; 10.10 ([EI8], Exoo–Ismailescu, 2014).

$$\chi(E^6) \geq 12;$$
$$\chi(E^7) \geq 16;$$
$$\chi(E^{10}) \geq 26;$$
$$\chi(E^{11}) \geq 26;$$
$$\chi(E^{12}) \geq 36;$$
$$\chi(E^{13}) \geq 36;$$
$$\chi(E^{14}) \geq 36.$$

In January 2015, *Geombinatorics* carried a paper [KaT] by Matthew Kahle and Bira Taha with new lower bounds of chromatic numbers of Euclidean spaces E^n for $n = 8,9,10,11,12$.

Lower Bounds for Euclidean Spaces E^n for $n = 8,\ldots,12$; 10.11 ([KT], Kahle–Taha, 2015).

$$\chi(E^8) \geq 19;$$
$$\chi(E^9) \geq 21;$$
$$\chi(E^{10}) \geq 26 \ (\text{Exoo–Ismailescu had it for } E^{10} \text{ a year before})$$
$$\chi(E^{11}) \geq 32;$$
$$\chi(E^{12}) \geq 32.$$

In 2018, the Russian group of Danila Cherkashin, Anatoly Kulikov, and Andrei Raigorodskii improved several lower bounds of chromatic numbers of Euclidean spaces E^n for $n = 9, \ldots, 12$.

Lower Bounds for Euclidean Spaces E^n for $n = 8,\ldots,12$; 10.12 ([CKR], Cherkashin–Kulikov–Raigorodskii, 2015).

$$\chi(E^9) \geq 22;$$
$$\chi(E^{10}) \geq 30;$$
$$\chi(E^{11}) \geq 35;$$
$$\chi(E^{12}) \geq 37.$$

There are papers, [Sze3] and [BRa], presenting lower bounds of chromatic numbers of E^n for $n > 12$. I refer you to these two papers for the results.

On February 12, 2022, Andrii Arman, Andriy V. Bondarenko, Andriy Prymak, and Danylo Radchenko uploaded to arXiv the second version of their paper, containing a good number of upper bound results [ABPR]. The authors point out that they obtained these results using explicit constructions of colorings of E^n based on sublattice coloring schemes.

Upper Bounds for Euclidean Spaces E^n 10.13 [ABPR].

$$\chi\left(E^4\right) \leq 49;$$

$$\chi\left(E^5\right) \leq 140;$$

$$\chi(E^n) \leq 7^n/2 \text{ for } n \in \{6,8,24\};$$

$$\chi\left(E^7\right) \leq 1372;$$

$$\chi\left(E^9\right) \leq 17253;$$

$$\chi(E^n) \leq 3^n \text{ for all } n \leq 38 \text{ and } n = 48, 49.$$

A long time ago, Paul Erdős conjectured, and often mentioned in his problem talks, for example, [E75.24], [E75.25], [E79.04], [E80.38], [E81.23], and [E81.26], that the chromatic number $\chi(E^n)$ of E^n grows exponentially in n.

Erdős' Conjecture on the Asymptotic Behavior of the Chromatic Number of R^n 10.14 $\chi(E^n)$ tends to infinity exponentially.

This conjecture was settled in the positive by a set of two results: the 1972 exponential upper bound, found by D.G. Larman and C.A. Rogers, and the 1981 exponential lower bound, established by P. Frankl and R.M. Wilson:

Frankl–Wilson's Asymptotic Lower Bound 10.15 (Frankl and Wilson, 1981, [FW]).

$$(1 + o(1))1.2^n \leq \chi(E^n).$$

Larman–Rogers' Asymptotic Upper Bound 10.16 (1972, [LR]).

$$\chi(E^n) \leq (3 + o(1))^n.$$

Asymptotically, Larman–Rogers' upper bound remains the best possible still today. Frankl–Wilson's asymptotic lower bound has been improved by Andrei Raigorodskii:

Raigorodskii's Asymptotic Lower Bound 10.17 (2000, [Raig2]).

$$(1.239\ldots + o(1))^n \leq \chi(E^n).$$

In Chapter 4, you came across the *polychromatic number* χ_p of the plane and in Chapters 4, 6, and 7 the results of it. This notion naturally generalizes to the *polychromatic number* $\chi_p(E^n)$ of the Euclidean n-dimensional space E^n. Dmitry E. Raiskii was the first to publish a relevant result [Rai]:

Raiskii's Lower Bound 10.18 $n + 2 \leq \chi_p(E^n).$

Larman–Rogers' upper bound 10.15 implies the same asymptotic upper bound for the polychromatic number:

Larman–Rogers Upper Bound 10.19

$$\chi_p(E^n) \leq (3 + o(1))^n.$$

They also conjectured that $\chi_p(E^n)$ grows exponentially in n. The positive proof of the conjecture, started by Larman and Rogers, was completed by Frankl and Wilson [FW]:

Frankl–Wilson Lower Bound 10.20

$$(1 + o(1))1.2^n \leq \chi_p(E^n).$$

Asymptotically, Larman–Rogers' upper bound remains the best possible still today. Frankl–Wilson's asymptotic lower bound was improved by Andrei M. Raigorodskii:

Raigorodskii's Asymptotic Lower Bound 10.20 (2000, [Raig2]).

$$(1.239\ldots + o\ (1))^n \leq \chi_p(E^n).$$

Obviously, there is a gap between the lower and upper bounds, and it would be very desirable to narrow it down.

Problem 4.4 can be considered in an n-dimensional Euclidean space too. For a given finite set S of positive numbers, called *a set of forbidden distances*, we define the graph $G_S(E^n)$, whose vertices are points of the Euclidean n-space E^n, and a pair of points is adjacent if the distance between them belongs to S. We will naturally call the chromatic number $\chi_S(E^n)$ of the graph $G_S(E^n)$ the *S-chromatic number of n*-space E^n. The following problem is as general as it is extremely hard:

Erdős' Open Problem 10.21 Given S, find the S-chromatic number $\chi_S(E^n)$ of the space E^n.

By de Bruijn–Erdős compactness theorem, which we mentioned in Chapter 5, the problem of investigating the S-chromatic number of E^n is a problem about finite graphs.[2]

* * *

As you can see, there are now a good number of results presenting lower and upper bounds of chromatic numbers of Euclidean spaces. We all understand that these bounds are far from the actual values of chromatic numbers. In my opinion, there is a simple formula for these chromatic numbers. See my 2002 general conjecture for the chromatic number $\chi(E^n)$ of E^n for any $n > 1$ near the end of this book.

[2] De Bruijn–Erdős Theorem assumes the axiom of choice – see later chapters of this book, analyzing the influence of axioms we choose for set theory on combinatorial results.

Chapter 11
Rational Coloring

I would like to mention here one more direction of assault on the chromatic number of the plane. By placing Cartesian coordinates on the plane E^2, we get an algebraic representation of the plane as the set of all ordered pairs (x, y) with coordinates x and y from the set R of real numbers and the Euclidean distance function:

$$E^2 = \{(x, y) : x, y \in R\} \tag{11.1}$$

Due to De Bruijn–Erdős Theorem 5.1, it suffices to deal with finite subsets of E^2; thus, we can restrict the coordinates in (11.1) to some subset of R. The problem is, which one should be chosen?

A set A is called *countable* if there is a one-to-one correspondence between A and the set of positive integers N.

For any set C, we define C^2 as the set of all ordered pairs (c_1, c_2), where c_1 and c_2 are elements of C:

$$C^2 = \{(c_1, c_2) : c_1, c_2 \in C\}$$

Open Problem 11.1 Find a countable subset C of the set of real numbers R such that the chromatic number $\chi(C^2)$ is equal to the chromatic number $\chi(E^2)$ of the plane.

The set of all algebraic numbers would work, but it would not advance our search for the chromatic number of the plane. The set Q of all rational numbers would not work, as Douglas R. Woodall showed in 1973.

Result 11.2 (D.R. Woodall, [Woo1]). $\chi(Q^2) = 2$.

Proof by D.R. Woodall [Woo1]. We need to color the points of the rational plane Q^2, i.e., the set of ordered pairs (r_1, r_2), where r_1 and r_2 are rational numbers. We partition Q^2 into disjoint classes as follows: we put two pairs (r_1, r_2) and (q_1, q_2) into the same class if and only if both $r_1 - q_1$ and $r_2 - q_2$ have odd denominators when written in their lowest terms (an integer n is written in its lowest terms as $n/1$).

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_11

This partition of Q^2 into subsets has an important property: if the distance between two points of Q^2 is 1, then both points belong to the same subset of the partition. Indeed, let the distance between (r_1, r_2) and (q_1, q_2) be equal to 1. This precisely means that

$$\sqrt{(r_1 - q_1)^2 + (r_2 - q_2)^2} = 1,$$

i.e.,

$$(r_1 - q_1)^2 + (r_2 - q_2)^2 = 1.$$

Let $r_1 - q_1 = a/b$ and $r_2 - q_2 = c/d$ be these differences written in their lowest terms. We have

$$\left(\frac{a}{b}\right)^2 + \left(\frac{c}{d}\right)^2 = 1$$

i.e.,

$$a^2 d^2 + b^2 c^2 = b^2 d^2.$$

Therefore, b and d must both be odd (can you see why?), i.e., by our definition above (r_1, r_2) and (q_1, q_2) must belong to the same subset.

Since any class of our partition can be obtained from any other class of the partition by a translation (can you prove this?), it suffices for us to color just one class and extend the coloring to the whole Q^2 by translations. Let us color the class that contains the point (0,0). This class consists of the points (r_1, r_2), where, in their lowest terms, the denominators of both r_1 and r_2 are odd (can you see why?). We color red the points of the form $\left(\frac{o}{o}, \frac{o}{o}\right)$ and $\left(\frac{e}{o}, \frac{e}{o}\right)$ and color blue the points of the form $\left(\frac{o}{o}, \frac{e}{o}\right)$ and $\left(\frac{e}{o}, \frac{o}{o}\right)$, where o stands for an odd number and e for an even number. In this coloring, two points of the same color may not be at a distance 1 apart (prove this on your own). ■

Then there came a "legendary unpublished manuscript," as Peter D. Johnson, Jr. referred [Joh8] to the paper by Miro Benda, then of the University of Washington, and Micha Perles, then of the Hebrew University, Jerusalem. The widely circulated and admired manuscript was called *Coloring of Metric Spaces.* Peter Johnson tells its story on the pages of *Geombinatorics* [Joh8]:

> The original manuscript of "Colorings…," from which some copies were made and circulated (and then copies were made of the copies, etc.), was typed in Brazil in 1976. I might have gotten my first or second generation copy in 1977 … The paper was a veritable treasure trove of ideas, approaches, and results, marvelously informative and inspiring.
>
> During the early and mid-1980s "Colorings…" was mentioned at a steady rhythm, in my experience, at conferences and during visits. I don't remember who said what about it, or when (except for a clear memory of Joseph Zaks mentioning it, at the University of

Waterloo, probably in 1987), but it must surely win all-time prize for name recognition in the "unpublished manuscript" category.

Johnson's story served as an introduction and homage to the conversion of the unpublished manuscript into the Benda–Perles publication [BP] in *Geombinatorics*' January 2000 issue.

This paper, dreamed up over a series of lunches the two authors shared in Seattle in the fall of 1975, created a new, algebraic approach to the chromatic number of the plane problem. Moreover, it formulated a number of open problems, not directly connected to the chromatic number of the plane, problems that gave algebraic chromatic investigations their own identity. Let us take a look at but a few of their results and problems. First of all, Benda and Perles prove (independently; apparently, they did not know about Woodall's paper) Woodall's result 11.2 about the chromatic number of the rational plane. They are a few years too late to coauthor the result, but their analysis allows an insight into the algebraic structure that we do not find in Woodall's paper. They then use this insight to establish more sophisticated results and the structure of the rational spaces they study.

Result 11.3 (Benda and Perles [BP]). $\chi(Q^3) = 2$.

Result 11.4 (Benda and Perles [BP]). $\chi(Q^4) = 4$.

Benda and Perles then pose important problems.

Open Problem 11.5 (Benda and Perles [BP]). Find $\chi(Q^5)$ and, in general, $\chi(Q^n)$.

Open Problem 11.6 (Benda and Perles [BP]). Find the chromatic number of $Q^2(\sqrt{2})$ and, in general, of any algebraic extension of Q^2.

This direction was developed by Peter D. Johnson, Jr. from Auburn University [Joh1], [Joh2], [Joh3], [Joh4], [Joh5], and [Joh6]; Joseph Zaks from the University of Haifa, Israel [Zak1], [Zak2], [Zak4], [Zak6], and [Zak7]; Klaus Fischer from George Mason University [Fis1] and [Fis2]; Kiran B. Chilakamarri [Chi1], [Chi2], and [Chi4]; Michael Reid, Douglas Jungreis, and David Witte ([JRW]); and Timothy Chow [Cho]). In fact, Peter Johnson published in 2006 in *Geombinatorics* "A Tentative History and Compendium" of this direction of research inquiry [Joh9]. I refer you to this survey and works cited there for many exciting results of this algebraic direction.

In recent years, Matthias Mann from Germany entered the scene and discovered partial solutions to problem 11.6, which he published in *Geombinatorics* [Man1].

Result 11.8 (Mann [Man1]). $\chi(Q^5) \geq 7$.

This jump from $\chi(Q^4) = 4$ explains the difficulty in finding $\chi(Q^5)$, the exact value of which remains open. Mann then found a few more important lower bounds [Man2].

Results 11.9 (Mann [Man2]).

$$\chi\left(Q^6\right) \geq 10;$$

$$\chi\left(Q^7\right) \geq 13;$$

$$\chi(Q^8) \geq 16.$$

In reply to my request, Matthias Mann wrote about himself on January 4, 2007:

> Since I have not spent too much time on unit-distance-graphs since the last article in Geombinatorics 2003, I do not have any news concerning this topic. To summarize, I found the following [lower bounds of] chromatic numbers:
>
> $$Q^5 \geq 7$$
>
> $$Q^6 \geq 10$$
>
> $$Q^7 \geq 13$$
>
> $$Q^8 \geq 16$$
>
> The result for Q^8 improved the lower bounds for the dimensions 9–13.
>
> For the Q^7 I think that I found a graph with chromatic number 14, but up to now I cannot prove this result because I do not trust the results of the computer in this case.
>
> Now something about me: I was born on May 12th, 1972, and studied mathematics at the University of Bielefeld, Germany from 1995 – 2000. I wrote my Diploma-thesis (the "Dipl.-Math." is the old German equivalent to Master' degree) in 2000. It was supervised by Eckhard Steffen, who has worked on edge-colorings. I had the opportunity to choose the topic of my thesis freely, so I read the book "Graph Coloring Problems" by Tommy Jensen and Bjarne Toft (Wiley Interscience 1995) and was very interested in the article about the Hadwiger–Nelson Problem, and found the restriction to rational spaces even more interesting. After reading articles of Zaks and Chilakamarri (a lot of them in *Geombinatorics*), I started to work on the problem with algorithms.
>
> Unfortunately, I had no opportunity to write a Ph.D. thesis about unit-distance graphs, so I started work as an information technology consultant in 2000.

In the previous chapter, you have already met Josef Cibulka, a first-year graduate student at Charles University in Prague, who improved some of Mann's lower bounds for the chromatic numbers of rational spaces:

Lower Bounds for Q^5 and Q^7 11.10 (Cibulka, [Cib]).

$$\chi(Q^5) \geq 8;$$

$$\chi(Q^7) \geq 15.$$

As I mentioned in the previous section, in 2018, Danila Cherkashin, Anatoly Kulikov, and Andrei Raigorodskii improved several lower bounds of chromatic numbers of Euclidean spaces E^n for $n = 9, \ldots, 12$. In fact, the same lower bounds for rational spaces are even more impressive.

Lower Bounds for Rational Euclidean Spaces Q^n for $n = 8,\ldots,12$; 11.11 ([CKR], Cherkashin–Kulikov–Andrei Raigorodskii, 2015).

$$\chi(Q^9) \geq 22;$$
$$\chi(Q^{10}) \geq 30;$$
$$\chi(Q^{11}) \geq 35;$$
$$\chi(Q^{12}) \geq 37.$$

Woodall's 2-coloring of the rational plane (result 11.2) has been used by the Australian student Michael Payne to construct a wonderful example of a unit distance graph – see it near the end of this book.

Part III
Coloring Graphs

Chapter 12
Chromatic Number of a Graph

12.1 The Basics

The notion of a graph is so basic, so unrestrictive, that graphs appear in all fields of mathematics, and indeed in all fields of scientific inquiry.

A *graph* G is just a non-empty set $V(G)$ of *vertices* and a set $E(G)$ of unordered pairs $\{v_1, v_2\}$ of vertices called *edges*. If $e = \{v_1, v_2\}$ is an edge, we say that e and v_1 are *incident* as are e and v_2; we also say that v_1 and v_2 are *adjacent* or are *neighbors.* Simple, don't you think?

By all standards of book writing, I am now supposed to give you an example of a graph. Why not create your own example instead? As the set V of vertices, take the set of all cities you have ever visited. Call two cities a and b from V adjacent if you have ever traveled from one of them to the other. Let us denote your travel graph by $T(Y)$.

We can certainly represent $T(Y)$ in the plane. Just take a map of the world, plot a dot for each city you've been in, and draw the lines (edges) of all your travels (the shape of edges does not matter, but do not connect two adjacent vertices a and b of $T(Y)$ by more than one edge even if you have traveled various routes between a and b).

We often represent graphs in the plane, as we have just done for $T(Y)$. When we do that, the only things that matter are the set of vertices (but not their positions) and which vertices are adjacent (but not the shape of edges; we presume that edges have no points in common, except vertices of the graph incident with them).

In fact, you can think of a graph as a set of pins, some of which are connected by rubber bands. So, we consider the graph unchanged if we reposition the pins and stretch the rubber bands. Thus, we call two graphs *isomorphic* if "pins" of one of them can be repositioned and its "rubber bands" stretched so that this graph becomes graphically identical to the other graph.

More formally, two graphs, G and G_1, are called *isomorphic* if there is one-to-one correspondence f: $V \to V_1$ of their vertex sets that preserves adjacency, i.e., vertices v_1 and v_2 of G are adjacent if and only if the vertices $f(v_1)$ and $f(v_2)$, respectively, of G_1 are adjacent.

For example, the two graphs in Fig. 12.1 are isomorphic, whereas the two graphs in Fig. 12.2 are not (prove both facts on your own or see the proof, for example, in [BS] pp. 102–105).

I would like to get to our main interest, coloring, as soon as possible. Thus, I will stop my introduction to graphs here and refer you to [BS] for a little more about graphs; you will find

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_12

much more in the books dedicated exclusively to graphs, such as [BLW], [Har0], [BCL], and a great number of other books. In fact, graph theory is lucky: It has inspired more enjoyable books than most other relatively new fields.

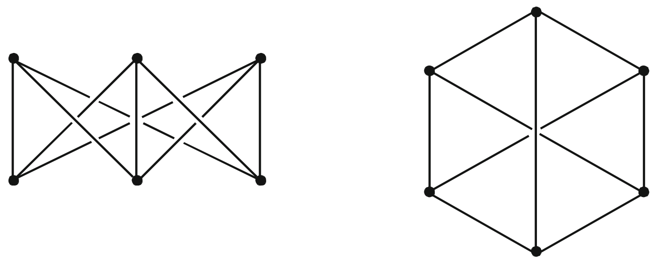


Fig. 12.1 A pair of isomorphic graphs

Fig. 12.2 A pair of non-isomorphic graphs

The notion of the chromatic number of the plane (Chapter 2) was motivated by a much older notion of the chromatic number of a graph. As Paul Erdős put it in his 1991 letter to me [E91/10/2ltr]:

Chromatic number of a graph is ancient.

The *chromatic number* $\chi(G)$ of a graph G is the minimum number n of colors with which we can color the vertices of G in such a way that no edge of G is *monochromatic* (i.e., no edge ab has both vertices a and b identically colored). In this case, we also say that G is a *n-chromatic graph*.

A graph G is called *n-colorable* if it can be colored in n colors without monochromatic edges. In this case, of course, $\chi(G) \leq n$.

Let us determine the chromatic numbers of some popular (and important) graphs.

A *n-path* P_n from x to y is a graph consisting of n distinct vertices $v_1, v_2, \ldots, v_n$ and edges $v_1v_2, \ldots, v_{n-1}v_n$, where $x = v_1$, $y = v_n$. If $n \geq 3$ and we add the edge $v_n v_1$, then we obtain a *n-cycle* C_n.

Problem 12.1 Prove that

$$\chi(C_n) = \begin{cases} 2, \text{if } n \text{ is even} \\ 3, \text{if } n \text{ is odd} \end{cases}$$

Problem 12.2 For a graph G, $\chi(G) \leq 2$ if and only if G contains no n-cycles for any odd n.

Such a graph has a special name: *bipartite graph*.

In particular, the *complete bipartite graph* $K_{m,n}$ has m vertices of one color and n vertices of the other, and two vertices are adjacent if and only if they have different colors. In Fig. 12.3, you can find examples of complete bipartite graphs.

A *complete graph* K_n consists of n vertices, every two of which are adjacent. In Fig. 12.4, you will find complete graphs K_n for small values of n.

Problem 12.3 Is there an upper limit to the chromatic numbers of graphs?

Solution Since every two vertices of K_n are adjacent, they must all be assigned distinct colors. Thus, $\chi(K_n) = n$, and there is no upper limit to the chromatic numbers of graphs. ■

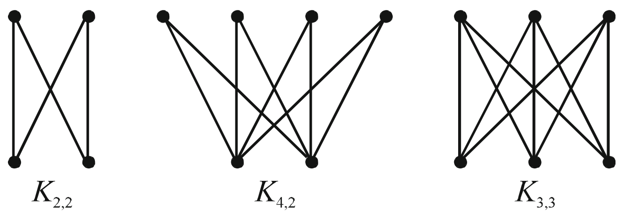


Fig. 12.3 Complete bipartite graphs

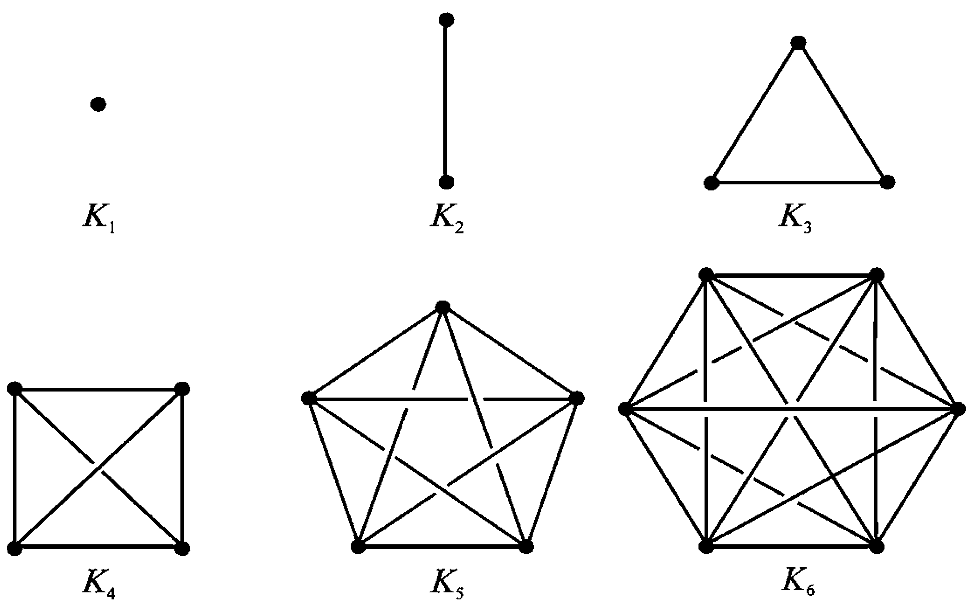


Fig. 12.4 Complete graphs

The number of edges incident to a vertex v of the graph G is called the *degree* of v and is denoted by $\deg_G v$. The maximal degree of a vertex in G is denoted by $\Delta(G)$.

If v is a vertex of a graph G, then $G - v$ denotes a new graph obtained from G by deleting v and its incident edges.

Problem 12.4 For any graph G with finitely many vertices,

$$\chi(G) \leq \Delta(G) + 1.$$

Proof Let G be a graph of chromatic number $\chi(G) = n$. If there is a vertex v in G, such that $\chi(G - v) = n$, we replace G by $G - v$. We can continue this process of deleting one vertex at a time with its incident edges until we get a graph G_1 such that $\chi(G_1) = n$ but $\chi(G_1 - v) \leq n - 1$ for any vertex v of G_1.

Let v_1 be the vertex of maximum degree in G_1, then

$$\Delta(G) \geq \Delta(G_1) = \deg_{G_1} v_1.$$

If we can prove that $\deg_{G1} v_1 \geq n - 1$, then coupled with the above inequality, we would get $\Delta(G) \geq n - 1$, which is exactly the desired inequality.

Assume the opposite, i.e., $\deg_{G1} v_1 \leq n - 2$. Since $\chi(G_1 - v_1) \leq n - 1$, we color the graph $G_1 - v_1$ in $n - 1$ colors. In order to get a $(n-1)$-coloring of G_1, we have to just color the vertex v_1. We can do it because $\deg_{G1} v_1 \leq n - 2$, i.e., v_1 is adjacent to at most $n - 2$ other vertices of G_1 (Fig. 12.5); thus, at least one of the $n - 1$ colors is unused around v_1. We use it on v_1. Thus, $\chi(G_1) \leq n - 1$, in contradiction to $\chi(G_1) = n$. ■

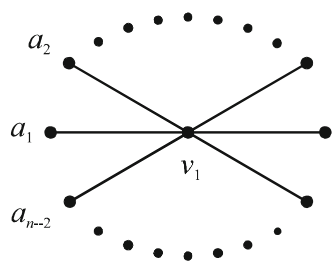


Fig. 12.5 Vertices adjacent to v_1

R. Leonard Brooks of Trinity College, Cambridge, in his now classic theorem, reduced the above upper bound by 1 (for most graphs), to the best possible general bound. His result was communicated by William T. Tutte on November 15, 1940, and published the following year [Bro].

Brooks' Theorem 12.5 ([Bro]). If $\Delta(G) = n > 2$ and the graph G has no component K_n, then

$$\chi(G) \leq \Delta(G).$$

12.2 Chromatic Number and Girth

W. T. Tutte, R. L. Brooks and Company pulled off the *Blanche Descartes* stint not unlike the better-known Nicolas Bourbaki. Arthur M. Hobbs and James G. Oxley convey the story of Blanche Descartes in the memorial article "William T. Tutte 1917–2002" [HOx]:

> Not long after he started his undergraduate studies at Cambridge, Tutte was introduced by his chess-playing friend R. Leonard Brooks to two of Brooks's fellow mathematics students, Cedric A. B. Smith, and Arthur Stone. The four became fast friends and Tutte came to refer to the group as "the Gang of Four," or "the Four" [Tut]. The Four joined the Trinity Mathematical Society and devoted many hours to studying unsolved mathematics problems together.
>
> They were most interested in the problem of squaring a rectangle or square, that is, of finding squares of integer side lengths that exactly cover, without overlaps, a rectangle or square of integer side lengths. If the squares are all of different sizes, the squaring is called perfect. While still undergraduates at Cambridge, the Four found an ingenious solution involving currents in the wires of an electrical network . . .

The Gang of Four were typical lively undergraduates. They decided to create a very special mathematician, Blanche Descartes, a mathematical poetess. She published at least three papers, a number of problems and solutions, and several poems. Each member of the Four could add to Blanche's works at any time, but it is believed that Tutte was her most prolific contributor.

The Four carefully refused to admit that Blanche was their creation. Visiting Tutte's office in 1968, Hobbs had the following conversation with him:

> **Hobbs**: "Sir, I notice you have two copies of that proceedings. I wonder if I could buy your extra copy?"
> **Tutte**: "Oh, no, I couldn't sell that. It belongs to Blanche Descartes."

And yet, I found at least one occasion when Tutte allowed to use his name in place of Blanche Descartes. Paul Erdős narrates [E87.12]:

> Tutte sometimes published his results under the pseudonym Blanche Descartes, and in one of my papers quoting this result I referred to Tutte. Smith wrote me a letter saying that Blanche Descartes will be annoyed that I attributed her results to Tutte (he clearly was joking since he knew that I know the facts), but Richard [Rado] was very precise and when in our paper I wanted to refer to Tutte, Richard only agreed after I got a letter from Smith stating that my interpretation was correct.

You may wonder, what paper by Blanche Descartes does Paul Erdős refer to? Our story commences with the problem [Des1] Blanche Descartes published in April 1947. To simplify the original language used by Descartes, let me introduce here a notion of the *girth of a graph* G as the smallest number of edges in a cycle in G.

Descartes' Problem 12.6 ([Des1], 1947). Find a 4-chromatic graph of girth 6.

Descartes' solution appeared in 1948 [Des2]. This was the start of an exciting train of mathematical thought. In 1949, the first Russian graph theorist Alexander A. Zykov produced the next result [Zyk1]. He limited the restriction to just triangles but asked in return for an arbitrarily large chromatic number:

Zykov's Result 12.7 ([Zyk1], 1949). There exists a triangle-free graph of an arbitrarily large chromatic number.

Zykov's 1949 comprehensive publication [Zyk1] contained a construction, proving his result. The cold war and the consequent limited exchange of information apparently made Zykov's advance unknown in the West. Four years later, in 1953, Peter Ungar formulated the same problem in the *American Mathematical Monthly* [Ung], which attracted much of attention and results. *The Monthly* chose not to publish the proposer's solution (which was supposedly similar to that of Zykov). Instead, in 1954, *The Monthly* published a solution by Blanche Descartes [Des3], which both generalized Descartes' own 1948 result and solved Zykov–Ungar's problem: Descartes constructed graphs of an arbitrarily large chromatic number, which contained no cycles of less than six lines. This [Des3] was the Blanche Descartes' paper that Paul Erdős referred to in the above quote, and it was written by William T. Tutte alone.

John B. Kelly and Le Roy M. Kelly obtained a very similar construction in the same year [KK]. Finally, Jan Mycielski, originally from Poland, professor emeritus at the University of Colorado at Boulder, published his original construction [Myc] in 1955. (Following Paul Erdős' stays with me, I usually drove him to Boulder to stay with Jan.)

Let us look at the mathematics of this explosion of constructions. We will start with an exercise showing how to increase the chromatic number by attaching 3-cycles.

Problem 12.8 Let T be a 3-cycle with its vertices labeled 1, 2, and 3, and R a set of 7 vertices labeled 1, 2, ..., 7 (Fig. 12.6). For *each* three-element subset V of R, we construct a copy T_V of T and attach it to R by joining vertex 1 of T_V with the lowest numbered vertex of V, vertex 2 of T_V with the middle-numbered vertex of V, and vertex 3 of T_V with the highest numbered vertex of V. In Fig. 12.7, for example, this connection is drawn for $V = \{2, 3, 6\}$.

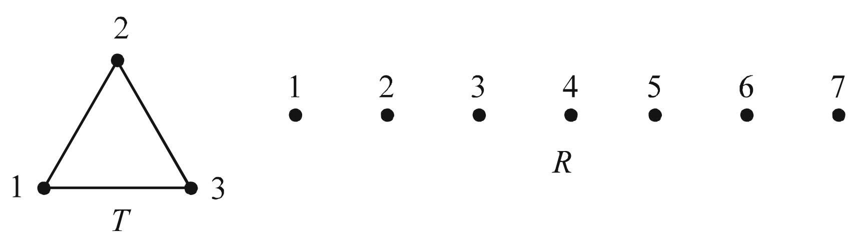


Fig. 12.6 A3-cycle and the foundation set R

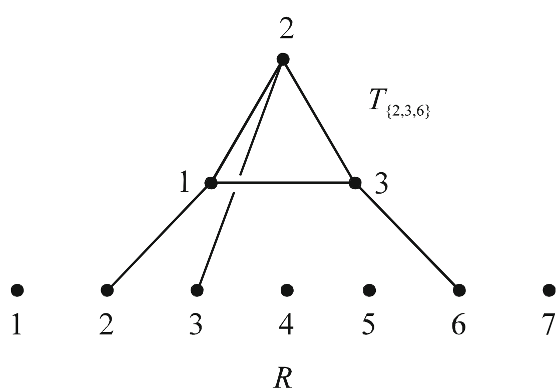


Fig. 12.7 A 3-cycle attached to the foundation set R

Since the number of three-element subsets of the seven-element set R is equal to $\binom{7}{3} = 35$, the resultant graph G has $7 + 3 \cdot 35 = 112$ vertices. Prove that $\chi(G) = 4$.

Proof Four colors suffice to color G, since each T_v in G can be colored with the first three colors and all the vertices of R with the fourth color. Thus, $\chi(G) \leq 4$.

Assume now that the graph G is three-colored. Then by the pigeonhole principle, among the seven vertices of R, there are three, say vertices 2, 3, and 6, which are colored in the same color, say color A. Then (see Fig. 12.7), color A is not present in the coloring of $T_{\{2,3,6\}}$; thus, $T_{\{2,3,6\}}$ is two-colored. However, this is a contradiction since by problem 12.1, a 3-cycle cannot be two-colored. Hence, $\chi(G) = 4$. ■

Problem 12.9 Use the construction of problem 12.8 with the Mosers Spindle (Fig. 2.2) in place of T and a 25-point set R. What is the chromatic number of the resulting graph G? How many vertices does G have?

The answer should serve as a *hint*: a 5-chromatic graph on

$$25 + 7\binom{25}{7} = 3{,}364{,}925 \text{ vertices.}■$$

In his monograph [Har0], Frank Harary discloses the secret authorship of one result: "This so-called lady [Descartes] is actually a non-empty subset of {Brooks, Smith, Stone, Tutte}; in this [Des3] case {Tutte}." Let us take a look at Blanche Descartes' (Tutte's) construction.

Blanche Descartes's Construction 12.10 [Des3]. For any integer $n > 1$, there exists a n-chromatic graph G_n of girth 6.

Proof For the case $n = 2$, we just pick a 6-cycle: $G_2 = C_6$. For $n \geq 3$, we define a sequence of graphs $G_3, G_4, \ldots, G_n, \ldots$ by induction. Let G_3 be a 7-cycle: $G_3 = C_7$.

Assume that the graph G_k is defined and has M_k vertices. We need to construct G_{k+1}. The construction is the same as in problem 12.8. Let R be a set of $k(M_k - 1) + 1$ vertices. For *each* M_k-element subset U of R, we construct a copy G_k^U of G_k, then pick a one-to-one correspondence f^U between the vertices of U and G_k^U (two M_k-element sets surely have one), and, finally, connect by edges the corresponding vertices of U and G_k^U. The resulting graph is G_{k+1}.

Thus, we constructed the graphs $G_3, G_4, \ldots, G_n, \ldots$ No graph G_n has a cycle of less than six edges (can you prove it?).

By induction, we can prove that $\chi(G_n) \geq n$ for every $n \geq 3$. Indeed, G_3 is 3-chromatic as an odd cycle. Assume that $\chi(G_k) \geq k$ for some $k \geq 3$. We need to prove that

$$\chi(G_{k+1}) \geq k + 1.$$

If, on the contrary, $\chi(G_{k+1}) \leq k$, then by the pigeonhole principle, out of $k(M_k - 1) + 1$ vertices of the set R, there will have to be an M_k-element subset U of vertices all colored in the same color, say color A. But then color A is not present in the copy G_k^U of G_k, i.e., the graph G_k can be $(k-1)$-colored in contradiction to the inductive assumption. The induction is complete.

Please note that we proved the inequality $\chi(G_n) \geq n$. Since we want to have the equality, we may have to delete (one at a time) some vertices of G_n and their incident edges until we end up with G'_n such that $\chi(G'_n) = n$. ■

Mycielski's Construction 12.11 [Myc]. For any integer $n > 1$, there exists a triangle-free n-chromatic graph.

Proof You may ask, why should we bother to prove the result that is weaker than Descartes' and Kelly and Kelly's result 12.10? Simply because this is a different construction, and it will work best for us in Chapter 15 and other places.

Start with a triangle-free (k–1)-chromatic graph G. For each vertex v_i of G, add a new vertex w_i adjacent to all neighbors of v_i. Next, add a new vertex z adjacent to all new vertices w_i. The chromatic number of this new graph is k, and it is still triangle-free. ■

Observe that if we were to start with a 5-cycle, then the graph generated by Mycielski's construction is the unique smallest triangle-free 4-chromatic graph (Fig. 12.8). Three years later, in 1958, this graph was independently found for different purposes by Herbert Grötzsch [Gro], and, thus, it makes sense to call it the *Mycielski–Grötzsch graph.* We will discuss Grötzsch's reasons for constructing this graph in Chapter 20.

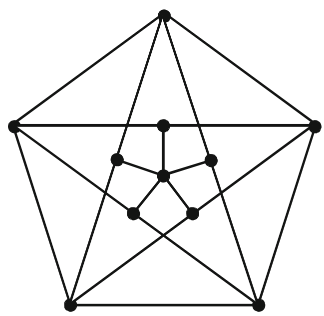


Fig. 12.8 The Mycielski–Grötzsch graph

The next major advance in our train of thought took place in 1959, when Paul Erdős, using probabilistic methods, dramatically strengthened result 12.10:

Erdős' Theorem 12.12 (P. Erdős, [E59.06]). For any two integers $m, n \geq 2$, there exists a n-chromatic graph of girth m.

An alternative, non-probabilistic proof of this result was obtained in 1968 by the Hungarian mathematician László Lovász [Lov1].

The greatest result was still to come 30+ years after Lovász. Paul O'Donnell proved the existence of 4-chromatic unit distance graphs of arbitrary girth. We will look at this remarkable piece of work later in this book.

Paul Erdős posed numerous exciting open problems related to the chromatic number of a graph. Let me share with you one such still open problem that I found in Paul's 1994 problem paper [E94.26].

Erdős' Open Problem on 4-Chromatic Graphs 12.13 Let G be a 4-chromatic graph with lengths of the cycles $m_1 < m_2 < \ldots$ Can $\min(m_{i+1} - m_i)$ be arbitrarily large? Can this happen if the girth of G is large?

12.3 Wormald's Application

In Chapter 5, I described Paul Erdős' problem partially conquered by Nicholas C. Wormald [Wor]. Wormald's first step was to construct what I will call the *Wormald graph*; he then embedded it in the plane. In his construction, Wormald used the Blanche Descartes construction of problem 12.10. In problem 12.8, I showed how one can use this construction. Analogously, Wormald uses a 5-cycle in place of T and a 13-point foundation set R. For *each* five-element subset V of R, he constructs a copy T_V of T, fixes a one-to-one correspondence of the vertices of V and T_V, and attaches T_V to V by connecting the corresponding vertices.

He ends up with the graph G on $13 + 5\binom{13}{5} = 6448$ vertices. Wormald uses 5-cycles because his goal is to construct a 4-chromatic graph of girth 5. I leave the pleasure of proving these facts to you:

Problem 12.14 (N.C. Wormald, [Wor]). Prove that the Wormald graph G is indeed a 4-chromatic, girth 5 graph.

So, what is so special about Nicholas Wormald's paper [Wor] of 1979? Even though independently discovered (I think), didn't he use the construction that was published 25 years earlier by Blanche Descartes [Des2]? The real Wormald's achievement was to *embed* his huge 6448-vertex graph in the plane, i.e., to draw his graph in the plane with all adjacent vertices, and only them, distance 1 apart. In my talk at the conference dedicated to Paul Erdős' 80th birthday in Keszthely, Hungary, in July 1993, I presented Wormald's graph as a picture frame without a picture inside it, to indicate that Wormald proved the existence and did not actually draw his graph. Nick Wormald accepted my challenge and shortly after the conference, on September 8, 1993, mailed to me a drawing of the actual plane embedding of his graph. I am happy to share his drawing with you here. Ladies and gentlemen, the Wormald Graph! (Fig. 12.9).

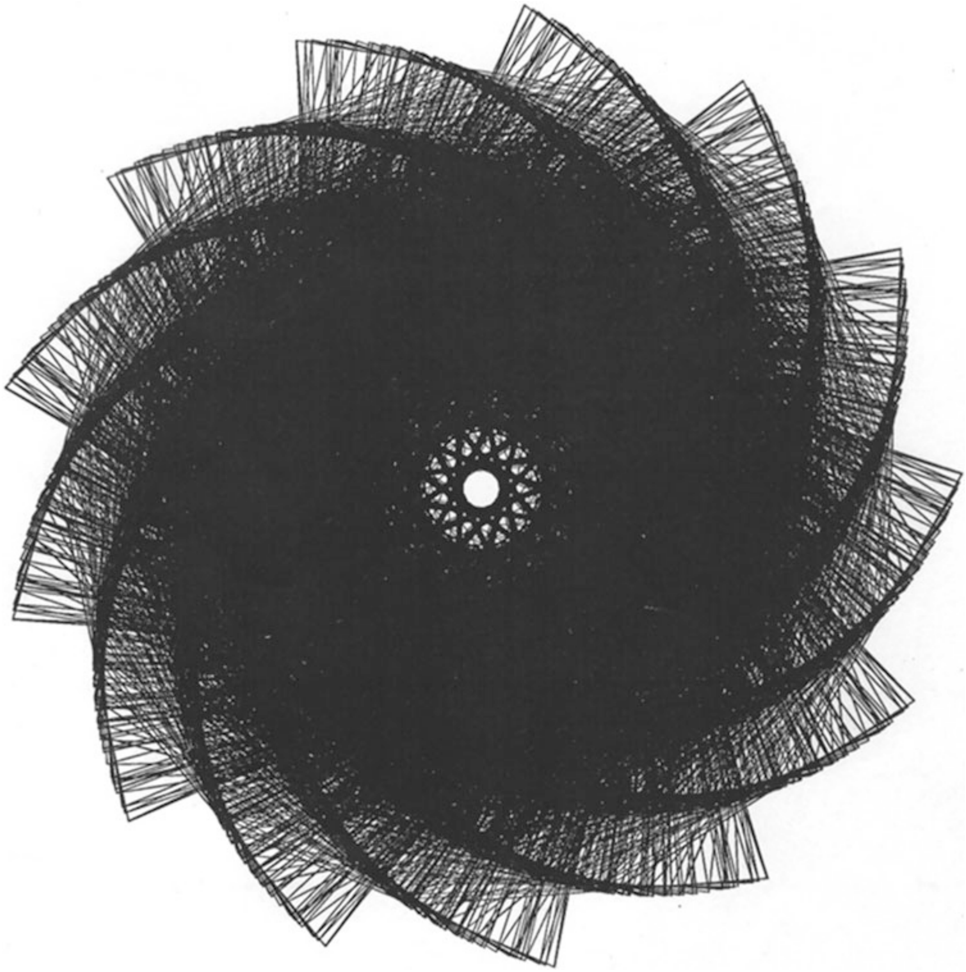


Fig. 12.9 The 6448-vertex Wormald Graph embedded in the plane

In his doctoral dissertation (May 25, 1999, Rutgers University), Paul O'Donnell offered a much simpler embedding than Wormald's – see it in Chapter 14, where I present Paul's machinery for embedding unit distance graphs in the plane.

The following problem, however, is open.

Open Problem 12.15 Find the smallest number λ_4 of vertices in a 4-chromatic graph without 3- and 4-cycles.

We know, of course, [Wor] that $\lambda_4 \leq 6448$. Chapters 14, 15, and 16 will be dedicated to major improvements in this direction.

Chapter 13
Dimension of a Graph

13.1 Dimension of a Graph

In 1965, a distinguished group of mathematicians consisting of Paul Erdős, Frank Harary, and William Thomas Tutte created a notion of the dimension of a graph [EHT].

They defined the *dimension of a graph* G, denoted by $dimG$, as the minimum number n such that G can be embedded in an *n-dimensional* Euclidean space E^n with every edge of G being a unit segment. Here, we will call such an embedding *1-embedding*.

Dimensions of some popular graphs can be easily found.

Problem 13.1 [EHT]. Prove the following equalities for complete graphs:

$$dimK_3 = 2,$$

$$dimK_4 = 3,$$

$$dimK_n = n - 1.$$

The symbol $K_n - x$ denotes the graph obtained from the complete graph K_n by deleting one edge x; due to the symmetry of all edges, this graph is well-defined.

Problem 13.2 [EHT]. Prove that

$$dim(K_3 - x) = 1,$$

$$dim(K_4 - x) = 2.$$

In general,

$$dim(K_n - x) = n - 2.$$

Now let us take a look at complete bipartite graphs.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_13

Problem 13.3 [EHT].[1] Prove that for $m \geq 1$,

$$dimK_{m,1} = 2,$$

except for $m = 1, 2$ when $dimK_{m,1} = 1$.

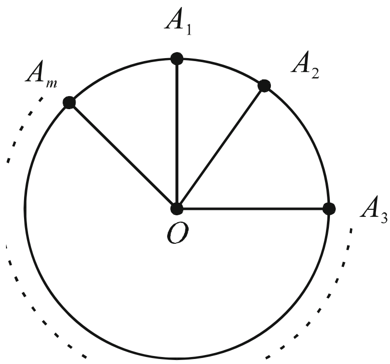


Fig. 13.1 Embedding of $K_{m,1}$ in the plane

Proof Let S be a circle of radius 1 with center at O. By connecting arbitrary m points A_1, A_2, ..., A_m of S with O, we get a desired embedding of $K_{m,1}$ in the plane (Fig. 13.1).

The graphs $K_{1,1}$ and $K_{2,1}$ can obviously be 1-embedded in the line E^1; thus, $dimK_{2,1} = 1$. ■

Problem 13.4 [EHT]. Prove that for $m \geq 2$,

$$dimK_{m,2} = 3,$$

except for $m = 2$ when $dimK_{2,2} = 2$.

Proof Let ABC be an isosceles triangle with $|AB| = |BC| = 1$. As we rotate ABC about AC, point B orbits a circle S (Fig. 13.2). By connecting arbitrary m points A_1, A_2, ..., A_m of S with both A and C, we get a desired embedding of $K_{m,2}$ in the 3-space E^3.

[1] The article [EHT] contains a minor oversight: it says "Obviously, for every $n > 1$, $dimK_{m,1} = 2$."

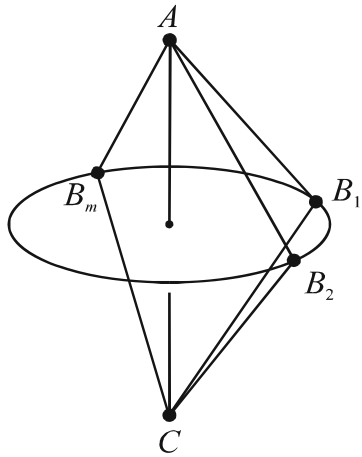


Fig. 13.2 Rotating *ABC* about *AC* creates a circle *S*

Since $m \geq 2$, the graph $K_{m,2}$ can be 1-embedded in the plane E^2 if and only if $m = 2$. Prove the last statement on your own. ■

Problem 13.5 [EHT]. Prove that for $m \geq n \geq 3$,

$$dimK_{m,n} = 4.$$

Proof In the solution of problem 13.4 (see Fig. 13.2), we had points of a one-dimensional "circle" (i.e., the two points *A* and *C*) at a distance 1 from the points of a circle *S*. Similarly, in the Euclidean four-dimensional space E^4, we can find two circles S_1 and S_2 such that any point of S_1 is at a distance 1 from any point of S_2. We pick the circle S_1 in the plane through the coordinate axes *X* and *Y* and the circle S_2 in the plane through the coordinate axes *Z* and *W*. Both S_1 and S_2 have the center at the origin $O = (0, 0, 0, 0)$ and radius $\frac{1}{\sqrt{2}}$. We then just pick *m* points on S_1 and *n* points on S_2.

This solution was obtained by Lenz in 1955 according to Paul Erdős. Formally (i.e., algebraically) it goes as follows. Let $\{u_i\}$ be the *m* vertices of the first color, and let $\{v_j\}$ be the *n* vertices of the second color (remember, we are constructing a complete bipartite graph). We pick coordinates in E^4 for $u_i = (x_i, y_i, 0, 0)$ and for $v_j = (0, 0, z_j, w_j)$ in such a way that they lie on our circles S_1 and S_2, respectively, i.e., $x_i^2 + y_i^2 = \frac{1}{2}$ and $z_j^2 + w_j^2 = \frac{1}{2}$. Then, the distance between every pair u_i, v_j will be equal to 1 (verify it using the definition of the distance in E^4).

It is not difficult to show (do) that for $m \geq n \geq 3$, the graph $K_{m,n}$ cannot be 1-embedded in the 3-space E^3 . ■

Problem 13.6 [EHT]. Find the dimension of the Petersen graph shown in Fig. 13.3.

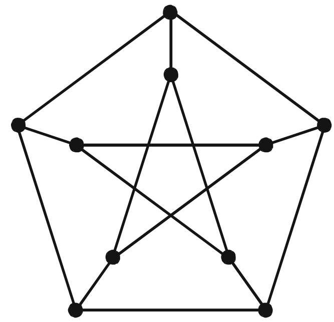


Fig. 13.3 The Petersen graph

Solution I enjoyed the style of the article [EHT]. I quote this solution in its entirety in order to show you what I mean:

It is easy to see (especially after seeing it) that the answer is 2; see Fig. 13.4. ■

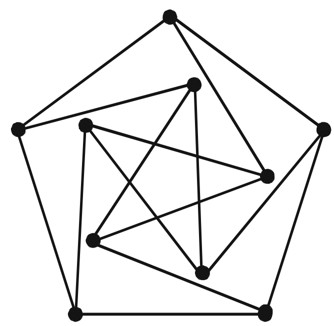
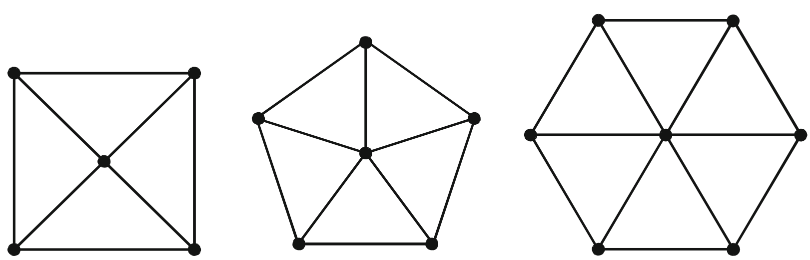


Fig. 13.4 The 1-embedding of the Petersen graph in the plane

Paul Erdős told me that it was Frank Harary's joke, as the latter wrote this solution for their joint article.

By connecting all vertices of a *n-gon* ($n \geq 3$) with one other vertex, we get the graph W_n called the *wheel with n spokes*. Figure 13.5 shows some popular wheels.



W_4 W_5 W_6

Fig. 13.5 Wheel graphs

Problem 13.7 [EHT]. The edges and vertices of a cube form a graph Q^3. Find its dimension.

Solution $dimQ^3 = 2$. Just think of Fig. 13.6 as drawn in the plane! ■

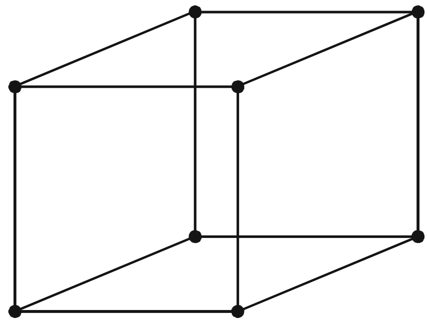


Fig. 13.6 The 1-embedding of the cube graph in the plane

The following two problems are for your own enjoyment.

Problem 13.8 [EHT]. Prove that

$$dimW_n = 3,$$

except for the "odd" number $n = 6$ when $dimW_6 = 2$.

Problem 13.9 [EHT]. A *cactus* is a graph in which no edge is on more than one cycle. Prove that for any cactus C,

$$dimC \leq 2.$$

I hope that you have enjoyed finding dimensions of graphs. There is no known systematic method for determining it. However, look, it has its good side. As the authors of [EHT] write, "the calculation of the dimension of a given graph is at present in the nature of mathematical recreation."

There is, however, one general inequality in [EHT] that connects the dimension and the chromatic number of a graph.

Problem 13.10 [EHT]. For any graph G,

$$dimG \leq 2\chi(G).$$

I totally agree with the authors of [EHT] that "the proof of this theorem is a simple generalization of the argument" used in problem 13.5. However, for the benefit of young readers not too fluent with n-dimensional spaces, I am presenting here both a geometric ideology of the solution and a formal algebraic proof.

Geometric Idea Let $\chi(G) = n$. In the Euclidean $2n$-dimensional space E^{2n}, we can find n circles $S_1, S_2, \ldots, S_n$ such that the distance between any two points from distinct circles is equal to 1. We pick the circle S_1 in the plane through the coordinate axes X_1 and X_2; the circle

S_2 in the plane through the coordinate axes X_3 and X_4; ...; the circle S_n in the plane through the coordinate aces X_{2n-1} and X_{2n}. All n circles have radius $\frac{1}{\sqrt{2}}$ and their centers at the origin.

Finally, when we color G in n colors (it can be done since $\chi(G) = n$), we get, say, k_1 points of color 1, k_2 points of color 2, ..., k_n points of color n. Accordingly, we pick arbitrary k_1 points on S_1, k_2 points on S_2, ..., k_n points on S_n for the desired 1-embedding of G.

Algebraic Solution Let $\{u_i^1\}$ be the k_1 vertices of color 1, $\{u_i^2\}$ the k_2 points of color 2, ..., $\{u_i^n\}$ the k_n vertices of color n. We pick coordinates in E^{2n} for these vertices as follows:

$$u_i^1 = (x_i^1, x_i^2, 0, 0, \ldots, 0)$$
$$u_i^2 = (0, 0, x_i^3, x_i^4, 0, \ldots, 0)$$
$$\ldots\ldots\ldots\ldots\ldots$$
$$u_i^n = (0, 0, 0, 0, \ldots, x_i^{2n-1}, x_i^{2n})$$

in such a way that they lie on our circles S_1, S_2, ..., S_n, respectively (see "Geometric Idea" above), i.e.,

$$(x_i^1)^2 + (x_i^2)^2 = \frac{1}{2}$$
$$(x_i^3)^2 + (x_i^4)^2 = \frac{1}{2}$$
$$\ldots\ldots\ldots\ldots$$
$$(x_i^{2n-1})^2 + (x_i^{2n})^2 = \frac{1}{2}$$

Then the distance between every pair of points that belong to different circles is equal to 1 (can you see why?). Thus, the distance between any two points of different colors of the graph G in this embedding in E^{2n} is equal to 1. (We do not have to care at all about the distances between two points of G of the same color: they are not adjacent in G.) We got a 1-embedding of G in the space E^{2n}; therefore, $dimG \leq 2n$. ■

If it so happened that every vertex of graph G_1 is also a vertex of graph G, *and* every edge of G_1 is also an edge of G, then graph G_1 is called a *subgraph* of graph G.

Prove on your own the following property of subgraphs.

Problem 13.11 For every subgraph G_1 of a graph G,

$$dimG_1 \leq dimG.$$

During his December 1991–January 1992 2-week visit of me in Colorado Springs, Paul Erdős posed the following (quite solvable, I think) problem:

Erdős' Open Problem 13.12 What is the smallest number of edges in a graph G such that $dimG = 4$?

13.2 Euclidean Dimension of a Graph

I enjoyed the Erdős–Harary–Tutte paper [EHT] very much. There was, however, one more thing I expected from the notions of 1-embedding and dimension but did not get it from [EHT]. I hoped that they would unite *the chromatic number of a plane set* (Chapter 5) and *the chromatic number of a graph* (Chapter 12). Here is what I meant. I wanted to consider such embeddings of a graph G in the plane E^2 (and more generally in the n-dimensional space E^n) that *the chromatic number of a plane set V of vertices of the embedded graph G is equal to the chromatic number of G.*

It was certainly not the case with the 1-embeddings discussed above. The chromatic number of the 1-embedded set V of vertices of a graph G may be not uniquely defined. Take, for example, the cycle C_4. We can 1-embed it in the plane so that its vertex set V has the chromatic number 2 (just think of a square), but we can also 1-embed C_4 so that V has the chromatic number 3 (think of a rhombus with a $\pi/3$ angle).

The notions of the chromatic number of the plane and the chromatic number of a plane set were generalized by Paul Erdős to Euclidean n-spaces over half a century ago:

> Let S be a subset of the n-dimensional Euclidean space E^n (S may coincide with E^n). The *chromatic number* $\chi(S)$ of S is the smallest number of colors with which we can color the points of S in such a way that no color contains a monochromatic segment of length 1.[2]

Thus, if we adjoin two points a,b of S with an edge if and only if the distance $|ab| = 1$, we will get the graph G with the chromatic number equal to the chromatic number of its vertex set S:

$$\chi(G) = \chi(S). \qquad (*)$$

Two new definitions, as well as most of the problems below, occurred to me on September 9, 1991. I remember this day very well: I was present in the delivery room when my daughter Isabelle Soulay Soifer was born on this very day at 6 in the evening.

On September 12, 1991, I sent the news to Paul Erdős:

> On the Jewish New Year, 9/9/1991 the baby girl Isabelle Soifer was born.

In my September 15, 1991, letter, I shared with Paul the mathematical thoughts that visited me as I was waiting in the delivery room for Isabelle's arrival in this world:

> I enjoyed Erdős–Harary–Tutte 1965 article where *dimension* of a graph was introduced. (Apparently Harary and Tutte did not particularly like it: dimension of a graph did not appear in the books on graph theory.)
>
> In my book though I am going to introduce a more precise notion. An embedding of a graph G into E^n we call *Euclidean* if two vertices v,w of G are adjacent if and only if in E^n the segment vw has length 1. *Euclidean dimension of a graph G* is the minimum n such that there is a Euclidean embedding of G in E^n (notation *EdimG*). Of course, $dimG \le EdimG$. But a strict inequality is possible: let W_6 be the wheel with 6 spokes, and W_6*'* [a wheel] without 1 spoke [my drawings in the letter are the Figs. 13.7 and 13.8

[2] Victor Klee was the first to prove (unpublished) that $\chi(E^n)$ is finite for any positive integer n.

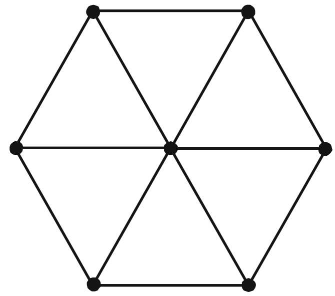


Fig. 13.7 Wheel W_6

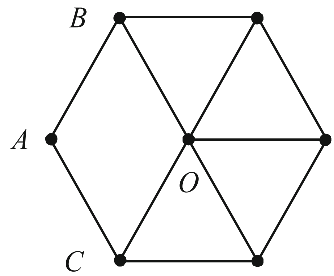


Fig. 13.8 Wheel W_6 with one spoke removed

below]. Then $dimW_6' = 2 < 3 = EdimW_6'$. Also, there is a graph G and its subgraph G_1 such that $EdimG_1 > EdimG$. Just take $W_6' \subseteq W_6$.

This Euclidean dimension (rather than dimension) of a graph connects precisely chromatic numbers of a graph and [of] a plane set:

If a graph G is *Euclideanly embedded* in E^n, then $\chi(G) = \chi(V)$, where $\chi(G)$ is the chromatic number of the graph, and $\chi(V)$ is the chromatic number of the vertex set V of G (i.e., subset of E^n).

What do you think?

Paul Erdős' reply arrived on October 2, 1991. Following my family affairs, "I am very sorry to hear about your father's death [Yuri Soifer, June 20, 1907 – June 17, 1991], but congratulations for the birth of Isabelle." Paul expressed his approval of the new notion of the Euclidean dimension of a graph and commenced posing problems about it [E91/10/2ltr]:

"Can *EdimG* – *dimG* be arbitrarily large?"

Little did I know at the time that, in fact, Paul Erdős himself with Miklós Simonovits invented the Euclidean dimension before me – in 1980 – they called it *faithful dimension* [ESi], and Paul did not remember his own baby definition when he discussed it with me. Of course, the credit for the discovery goes to Erdős and Simonovits. In my opinion, however, the term "Euclidean dimension" more faithfully names the essence of the notion, and, so, I will keep this term here.

Let us summarize the definitions and the early knowledge that we have.

We call a one-to-one mapping of the vertex set V of a graph G into a Euclidean space E^n the *Euclidean embedding of G into* E^n if two vertices v,w of G are adjacent if and only if the distance between $f(v)$ and $f(w)$ is equal to 1.

In other words, to obtain a Euclidean embedding of G into E^n, we need to draw G in E^n with every edge of G being a unit segment and the distance between two nonadjacent vertices being not equal to 1.

We define the *Euclidean dimension* of a graph G, denoted by $EdimG$, as the minimum number n such that G has a Euclidean embedding into E^n.

Now, we do get the desired connection:

Problem 13.13 The chromatic number of a graph G is equal to the chromatic number of its vertex set V when G is *Euclideanly embedded* in E^n for some n.

The two dimensions are connected by the following inequality:

Problem 13.14 Prove that for any graph G,

$$dimG \le EdimG.$$

For some popular graphs, we have the equality:

Problem 13.15 For any complete graph K_n,

$$dimK_n = EdimK_n$$

i.e., $EdimK_n = n - 1$.

Problem 13.16 For any complete bipartite graph $K_{m,n}$,

$$dimK_{m,n} = EdimK_{m,n}$$

Problem 13.17 For any wheel W_n,

$$dimW_n = EdimW_n$$

i.e., $EdimW_n = 3$, except for the "odd" number $n = 6$ when $EdimW_6 = 2$.

Problem 13.19 For any graph G,

$$dimG \le EdimG \le 2\chi(G).$$

The new notion makes sense only if there is a graph G for which $dimG \ne EdimG$. And it does exist:

Problem 13.20 Find a graph G such that

$$dimG < EdimG.$$

The inequality $dimG_1 \leq dimG$ that is trivially true for any subgraph G_1 of a graph G may not be true at all for the Euclidean dimension:

Problem 13.21 Construct an example of a graph G and its subgraph G_1 such that

$$EdimG_1 > EdimG.$$

Solutions to Problems 13.20 and 13.21 Take the wheel W_6 with six spokes (Fig. 13.7) and knock out one spoke (Fig. 13.8). Let us prove that the resulting graph W'_6 has the Euclidean dimension 3, even though $EdimW_6 = 2$.

Indeed, when we draw the graph W'_6 in the plane so that its every edge is a segment of length 1, the rigid construction of W'_6 leaves no options for the distance OA. It is equal to 1 even though the spoke is missing! Thus, there is no Euclidean embedding of W'_6 in the plane. It is easy to Euclideanly embed W'_6 in 3-space E^3: start with the plane W'_6 depicted in Fig. 13.8 and rotate BAC in the space about the axis BC until the distance OA is not 1.

We proved that $EdimW'_6 > EdimW_6$. Thus, problem 13.21 is solved. Problem 13.20 is solved at the same time because $dimW'_6 = 2$, and, therefore,

$$EdimW'_6 > dimW'_6.\blacksquare$$

The question that Paul Erdős posed to me, "Can *Edim*G – *dim*G be arbitrarily large?" – was answered in the positive by him and Simonovits 11 years earlier:

Problem 13.22 [ESi]. For any positive n, there is a graph G such that $dimG \leq 4$ while $n–2 \leq EdimG \leq n–1$.

Hint In problem 13.5, we saw that for $n \geq 3$, $dimK_{n,n} = 4$. Let G be the graph obtained from $K_{n,n}$ by removing a 1-factor, i.e., G is a graph on $2n$ vertices $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ with edges $x_i\, y_j$ for all $i \neq j$. Clearly, $dimG \leq dimK_{n,n} = 4$. Show that G cannot be Euclideanly embedded in the space E^{n-3} but can be Euclideanly embedded in E^{n-1}. ■

Erdős and Simonovits also found an upper bound for the Euclidean dimension of a graph. Their results showed that the Euclidean dimension of a graph G is related to its maximal vertex degree $\Delta(G)$ and not to its chromatic number $\chi(G)$:

Problem 13.23 [ES]. For any graph G, $EdimG \leq 2\Delta(G)+ 1$.

Nine years later, this bound was slightly improved by László Lovász, Michael Saks, and Alexander Schrijver:

Problem 13.24 [LSS]. For any graph G, $EdimG \leq 2\Delta(G)$.

Surprisingly, this bound still seems to be the best-known.

We are now ready to continue our discussion of Nicholas Wormald's paper [Wor], started in Chapter 5 and continued in Chapter 12. The big deal was not to construct his 4-chromatic graph G without 3- and 4-cycles. The real Wormald's trick was to Euclideanly embed his huge 6448-vertex graph G in the plane. And he did it with the use of his ingenuity and a computer. Read more about his embedding in his paper [Wor]. Here I would like to discuss one approach to the chromatic number of the plane problem. The rest of this chapter, which appeared in the first 2009 edition of this *Mathematical Coloring Book*, became dated when Aubrey de Grey came onstage; however, I would like to keep it here as a historical artifact, showing the thinking at that time of most colleagues working on this problem.

If you believe that the chromatic number of the plane χ is at least 5, here is what you can do to prove it. You can create a 5-chromatic graph G and then Euclideanly embed G in the plane. "Easier said than done," you say? Sure, but let us discuss it, then who knows? You just might succeed!

The 3,364,925-vertex graph G, which we constructed in problem 12.9, is surely 5-chromatic. But we constructed it with the use of the Mosers Spindle; thus, G has a lot of triangles. It may be too rigid to have a Euclidean embedding in the plane.

We can replace the Mosers Spindle with, say, the Mycielski–Grötzsch graph (Fig. 12.8) and use the same construction as we did in problems 12.8, 12.9, and 12.10. We would get a 5-chromatic graph G with $41 + 11\binom{41}{11} = 34{,}754{,}081{,}689$ vertices. This graph G has no triangles. But does it have a Euclidean embedding in the plane? To begin with, I do not think (check it out) the Grötzsch graph itself has a Euclidean embedding in the plane.

It seems that we need to start with a highly "flexible" graph having a Euclidean embedding in the plane. Let us start with the Wormald graph G (see Section 12.3) and the foundation set R of $6447 \times 4 + 1 = 25{,}789$ points. For every 6448-element subset V of R, we attach a copy G_V of G. We get a 5-chromatic graph G_1 without triangles, with $25{,}789 + 6448\binom{25798}{6448}$ vertices. Does G_1 have an Euclidean embedding? Computers are much better today than in 1978, when Nicholas Wormald completed his paper [Wor]. Are computers good enough for this task? Are we good enough to break through these computational walls?

Chapter 14
Embedding 4-Chromatic Graphs in the Plane

14.1 A Brief Overture

In Chapters 1 and 2, we got acquainted with examples of 4-chromatic unit distance graphs, the Mosers Spindle, and the Golomb graph. In Chapters 5 and 12, we encountered Paul Erdős' $25 Problem 5.6 and its partial solution by Nicholas Wormald, who used Blanche Descartes' construction of a 4-chromatic graph and his own embedding of that graph in the plane. Wormald's result was improved time and again on the pages of *Geombinatorics* by Paul O'Donnell, Rob Hochberg, and Kiran Chilacamari. Upon constructing a promising graph G, the authors of the new 4-chromatic unit distance examples used a two-part approach to complete their task:

1. *Graph-Theoretic Part*: Show that the chromatic number of a graph G is 4 and that the graph has no short cycles.
2. *Geometric Part*: Embed G in the plane in such a way that every pair of adjacent vertices is distance 1 apart *and* nonadjacent vertices are not 1 apart (like in the previous chapter dealing with the Euclidean dimension).

In this chapter, we will concentrate on the essentials of Part 2 – tools for embedding in the plane, as developed and presented by Paul O'Donnell [Odo3], [Odo4], and [Odo5]. In the next chapter, we will look at the world records in the new sport of embedding. Do use a pencil as you read this chapter.

We say that a k-vertex graph G with vertices $V = \{u_1, u_2, \ldots, u_k\}$ is *attached* to a set of vertices $V^* = \left\{u_1^*, u_2^*, \ldots, u_k^*\right\}$ if the vertices of G are connected via a *matching* to V (i.e., via a one-to-one correspondence of V to V^* and connection of the corresponding vertices by new edges).

The *shadow of* G, denoted by G^*, is the set to which G is attached. We often choose the graph G to be an odd cycle. Odd cycles are often attached to subsets of a large *independent* (i.e., no pair of vertices is adjacent) set of size n. The n independent vertices are called the *foundation vertices.*

If the vertices of G are placed at points of the plane so that the adjacent vertices are exactly a unit distance apart, we say that this is a *unit distance embedding* of G. Thus, in the plane, if the odd cycle $\{u_1, u_2, \ldots, u_k\}$ is attached to $\left\{u_1^*, u_2^*, \ldots, u_k^*\right\}$, then the vertices $u_1, u_2, \ldots, u_k$, $u_1^*, u_2^*, \ldots, u_k^*$ are fixed points in the plane such that for some permutation σ, u_i is a unit

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_14

distance from $u_{\sigma(i)}$ and from u_{i-1} and u_{i+1} (indices are added modulo k) for $1 < i < k$. Since the vertices can be relabeled, we assume that u_i is adjacent to u_i^* in the attachment. Usually, we do not want distinct vertices to be placed at the same point in the plane. If the vertices of G are placed at distinct points of the plane so that the adjacent vertices are exactly a unit distance apart, we say that this is *a proper unit distance embedding* of G. A graph with a proper unit distance embedding is called a *unit distance graph in the plane.* In this chapter, higher-dimensional analogues are not considered, and so, a *unit distance graph* will mean a unit distance graph in the plane. In our geometric contexts, the terms point and vertex will be used interchangeably, while the term edge will mean a unit length edge. The following continuity argument of the attachment procedure is important, and Paul O'Donnell uses it in most of the results of this chapter:

Continuity Argument 14.1 Given fixed points in the plane,

$$u_1^* = (x_1, y_1), u_2^* = (x_2, y_2), \dots, u_n^* = (x_n, y_n)$$

and a point u_1 on the unit circle centered at u_1^*. Let u_i be a point distance 1 from both u_{i-1} and u_1^* for $2 \le i \le k$. (In the following examples, the distance between u_{i-1} and u_1^* is less than 2, so there are two points satisfying the distance restrictions; let u_i be the one closer to the corresponding point in the attachment.) For an appropriately chosen arc along the unit circle centered at u_1^*, u_i is a continuous function of u_1. If there exist two points u_1^{short} and u_1^{long} such that the distance between u_1^{short} and the corresponding u_k^{short} is less than 1, while the distance between u_1^{long} and the corresponding u_k^{long} is greater than 1, then due to continuity, there must be a point, u_1^{unit}, such that the distance between u_1^{unit} and the corresponding u_k^{unit} is exactly 1. In other words, the set of points $\{u_1^*, u_2^*, \dots, u_k^*\}$ has a k-cycle attached, namely, $\{u_1^{unit}, u_2^{unit}, \dots, u_k^{unit}\}$. ■

The foundation points are distributed among four regions. They are placed inside the δ-balls centered at the following four points:

$C_1 = (0, 0)$
$C_2 = (0, 0.9)$
$C_3 = (0.9, 0.9)$
$C_4 = (0.9, 0)$

Since δ is very close to zero, it is impossible to attach a cycle to k points if they are all inside the same δ-ball. The partitioning of the foundation points is designed to prevent such an occurrence. Can a k-cycle be attached if the points are distributed among at least two of the δ-balls? Yes, they can. First, k-cycles are attached to k foundation points placed exactly at some or all of C_1, C_2, C_3, or C_4. Next, the points are moved slightly so that the k-cycles are attached to k distinct points, each placed inside the appropriate δ-ball surrounding C_1, C_2, C_3, or C_4. This prevents the foundation vertices from coinciding. Then, some of the vertices are moved slightly to eliminate all coincidences.

In *Geombinatorics* [Odo5] (but not in his dissertation [Odo3]), O'Donnell introduces a useful notion of type:

A set of foundation vertices has *type* $(a_1, a_2, a_3, a_4)_\delta$ if it consists of a_i vertices placed inside the δ-ball around C_i, $1 \le i \le 4$.

14.2 Attaching a 3-Cycle to Foundation Points in Three Balls

Only δ-balls around points C_1, C_2, and C_3 are dealt with for the basic argument. To distinguish between the preliminary and final situations, the foundation vertices coincident with C_1, C_2, and C_3 are denoted; the paths or cycles attached to them are denoted by $v_1, v_2, \ldots, v_k$, while the foundation vertices inside the δ-balls around C_1, C_2, and C_3 are denoted by $\{u_1^*, u_2^*, \ldots, u_k^*\}$ and the paths or cycles, attached to them, are denoted by $u_1, u_2, \ldots, u_k$.

O'Donnell starts by attaching a triangle.

Tool 14.2 A 3-*cycle* can be attached to the set of foundation points $\{C_1, C_2, C_3\}$

Proof Using the points listed in the *Appendix* at the end of this chapter (rounded to five decimal places), two three-vertex unit distance paths are attached to C_1, C_2, and C_3. In the first path, T_1^{short}, T_2^{short}, T_3^{short}, the distance from T_1^{short} to T_3^{short} is less than 0.99. In the second path, T_1^{long}, T_2^{long}, T_3^{long}, the distance from T_1^{long} to T_3^{long} is greater than 1.01 (see Fig. 14.1).

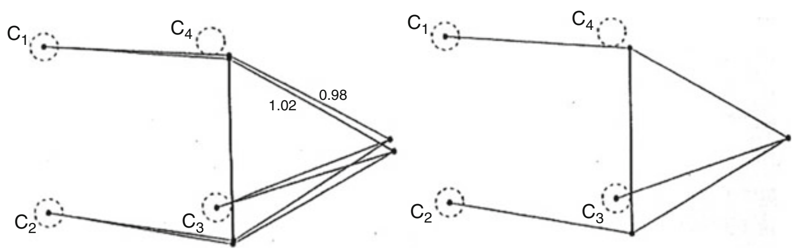


Fig. 14.1 A "too short" attachment and a "too long attachment" are shown together on the left. The "just right" attachment is on the right (all unlabeled edges have unit length)

Since one path is obtained from the other by continuously sliding the starting vertex, by continuity argument, there must be a path for which the distance between the first and last vertices is exactly one. This is a required, attached 3-cycle. ■

Now, we will relax the condition of Tool 14.2 and allow the three foundation vertices to be anywhere inside δ-balls and not just at their centers. Given $\delta > 0$, let u_1^*, u_2^*, and u_3^* be the foundation vertices placed anywhere inside δ-balls, centered at C_1, C_2, and C_3, respectively. If δ is small enough, we will show that a cycle can be attached to the foundation set $\{u_1^*, u_2^*, u_3^*\}$, which is very close to the cycle attached to the foundation set $\{C_1, C_2, C_3\}$ (which we have already accomplished in Tool 14.2).

Tool 14.3 There exists $\delta > 0$ such that a 3-cycle can be attached to any foundation vertex set of type $(1, 1, 1, 0)_\delta$.

Proof Given $\varepsilon > 0$, choose δ so that we can find "too short" and "too long" paths whose vertices are less than ε from the corresponding "too short" and "too long" paths attached to $\{C_1, C_2, C_3\}$. This is possible due to the continuity argument. As in Tool 14.2, we get a "just right" path, which is the required cycle. ■

14.3 Attaching a *k*-Cycle to a Foundation Set of Type $(a_1, a_2, a_3, 0)_\delta$

To generalize the above construction to *k*-cycles, where $k > 3$ is odd, other special points are needed. The three *triangle points*, denoted by T_1, T_2, and T_3, are the points of the 3-cycle attached to C_1, C_2, and C_3, respectively. The three *spoke points*, denoted by S_1, S_2, and S_3, respectively, are the points such that S_i is a unit distance from T_i and C_i, for $1 \leq i \leq 3$. We define "triangle" points T_i^{short} and T_i^{long} and "spoke points" S_i^{short} and S_i^{long} analogously for $1 \leq i \leq 3$. At first, cycles or paths are attached, which coincide with these triangles and spoke points. The shadows of these cycles coincide with the center points C_1, C_2, and C_3. We then use the continuity argument to show the existence of cycles very close to these.

Tool 14.4 Let $k \geq 3$ be an odd number. For all positive integers a_1, a_2, a_3 such that $a_1 + a_2 + a_3 = k$, a *k*-cycle consisting only of edges from T_i to T_{i+1} (addition modulo 3), and T_i to S_i, for $1 \leq i \leq 3$, can be attached to the union of a_i points at C_i, $1 \leq i \leq 3$.

Proof The work involved in attaching 5-cycles contains all the details of the general case. Suppose, for example, we want to attach a 5-cycle to the set $\{u_1^*, u_{1a}^*, u_{1b}^*, u_2^*, u_3^*\}$ (where the number in the subscript indicates the δ-ball containing the vertex).

By Tool 14.3, for δ small enough, we can attach a 3-cycle u_1, u_2, u_3 to the foundation vertices $\{u_1^*, u_2^*, u_3^*\}$. We just need to insert a "detour" into this cycle. Instead of going from u_1 to u_2, we go from u_1 to u_{1a} to u_{1b}, which is arbitrarily close to u_1. We then continue to u_2 to u_3 and finally back to u_1. Of course, we cannot actually construct the 5-cycle directly from the 3-cycle. Instead, we construct "too short" and "too long" 5-paths with corresponding vertices within ε of the vertices of the "too short" and "too long" 3-paths used to construct the 3-cycle (see Tools 14.2 and 14.3). Given ε, we choose δ such that this is possible. By the continuity argument, we get a "just right" 5-path. This is an attached 5-cycle (see Fig. 14.2).

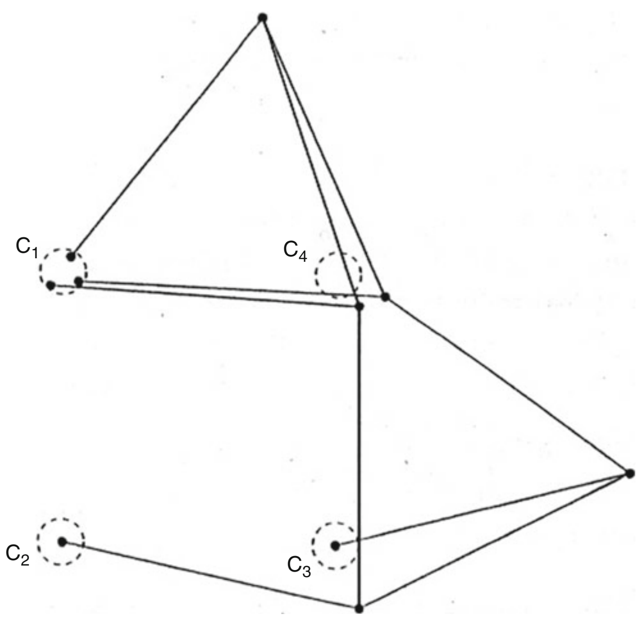


Fig. 14.2 A 5-cycle attached to a set of type (3, 1, 1, 0) δ

Of course, it did not matter that three foundation vertices were in the same δ-ball. Only two were necessary for the argument to work. The basic idea is to take a 3-cycle u_1, u_2, u_3 and construct a 5-cycle u_1, z, u_1, u_2, u_3. It does not matter where the foundation vertex z is so long as it's close enough to u_1 so that the unit length edges can be connected to z (i.e., z should be less than 2 units away from u_1).

For example, suppose we want to attach a 5-cycle to $\{u_1^*, u_{1a}^*, u_2^*, u_3^*, u_{3a}^*\}$. We find "too short" and "too long" 5-paths that are arbitrarily close to the corresponding "too short" and "too long" 3-paths. By the continuity argument, we get a 5-cycle (see Fig. 14.3).

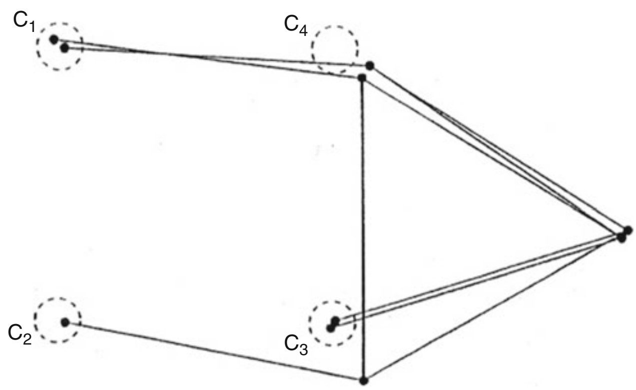


Fig. 14.3 A 5-cycle attached to a set of type (2, 1, 2, 0)δ

Similarly (using induction and considering two cases as discussed above), k-cycles can be attached to k points by first looking at a (k–2)-cycle attached to k–2 points and, then, performing the insertion procedure described above. The cycle will look like a triangle with a few spokes coming off some of the vertices. ■

By symmetry, we can now attach k-cycles to sets of types $(a_1, 0, a_3, a_4)_\delta$ and $(0, a_2, a_3, a_4)_\delta$. What if we need to place the foundation vertices inside all four of the δ-balls? In fact, for our purposes, we need only the case when the partitioning of the foundation vertices puts just one foundation vertex in the δ-ball around C_4, and so, only this case needs to be considered.

14.4 Attaching a k-Cycle to a Foundation Set of Type $(a_1, a_2, a_3, 1)_\delta$

Tool 14.5 Let $k \geq 5$ be an odd number. For all positive integers a_1, a_2, a_3, a_4 such that $a_1 + a_2 + a_3 + a_4 = k$ and $a_4 = 1$, there exists $\delta > 0$ such that a k-cycle can be attached to any foundation set of type $(a_1, a_2, a_3, 1)_\delta$.

Proof The argument of the previous chapter applies here as well. At least one of a_1, a_2, a_3 is greater than 1, say a_1. We first find a (k–2)-cycle attached to a set of type $(a_l - 1, a_2, a_3, 0)_\delta$ and then replace the vertex u_1 in the cycle by a path u_1, u_4, u_{1a}. This produces an attached k-cycle. Like before, we really do all the work on the "too short" and "too long" paths and use the continuity argument to prove the existence of the desired "just right" cycle (see Fig. 14.4). ■

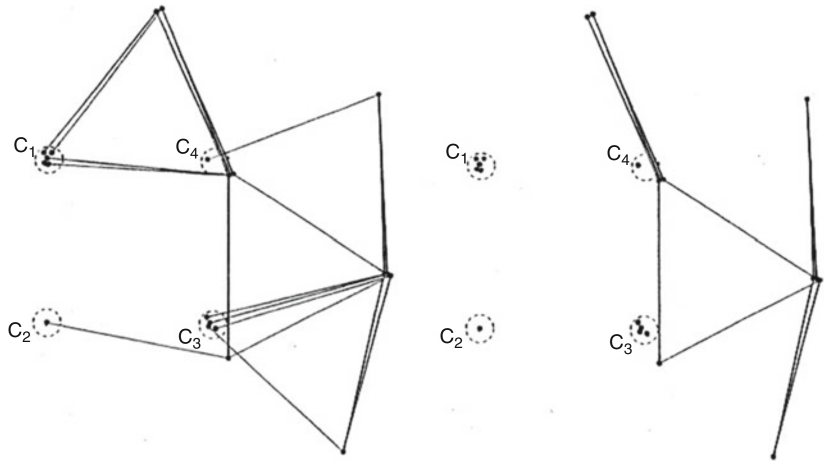


Fig. 14.4 An 11-cycle attached to a set of type (5, 1, 4, 1) δ. The cycle and the attaching edges are shown on the left. The cycle alone is shown on the right

14.5 Attaching a k-Cycle to the Foundation Sets of Types $(a_1, a_2, 0, 0)_\delta$ and $(a_1, 0, a_3, 0)_\delta$

We have shown that an odd cycle can be attached to k points placed inside δ-balls around any three or all four of C_1, C_2, C_3, and C_4. But what if the points are distributed between δ-balls around just two of the center points? The crucial step is still attaching a triangle. Once it is shown that a triangle can be attached to the center points, the previous arguments show that a k-cycle can be attached for any odd $k > 3$. We simply think of one of the δ-balls as two overlapping δ-balls (so, now, we have three balls).

Tool 14.6 Let $k \geq 3$ be an odd positive integer. For all positive integers a_1, a_2 such that $a_1 + a_2 = k$, there exists δ such that a k-cycle can be attached to any foundation set of type $(a_1, a_2, 0, 0)_\delta$.

Proof Without loss of generality, assume $a_1 \geq 2$. We attach a 3-cycle to two vertices at C_1 and one vertex at C_2, using the same notation as before for the triangle points; only here, the triangle points with subscripts 1 or 2 correspond to C_1, while those with subscript 3 correspond to C_2 (see the left drawing of Fig. 14.5).

Using the points listed in the Appendix (rounded to five decimal places), two 3-vertex paths are attached to C_1, C_1, and C_2. In the first path, T_1^{short}, T_2^{short}, T_3^{short}, the distance from T_1^{short} to T_3^{short} is less than 1. In the second path, T_1^{long}, T_2^{long}, T_3^{long}, the distance from T_1^{long} to T_3^{long} is greater than 1. Since one path is obtained from the other by continuously sliding the starting vertex, by the continuity argument, there must be a path for which the distance between the first and the last vertices is exactly 1. This is a desired, attached 3-cycle.

Now, we attach the k-cycle. Let a_1' and a_1'' be positive integers such that $a_1' + a_1'' = a_1$. We treat C_1 as if it were two separate vertices C_1' and C_1'' and use the machinery from the previous section to find δ such that any a_1' points in the δ-ball around C_1', a_1'' points in the δ-ball around C_1'', and a_2 points in the δ-ball around C_2, can have a k-cycle attached. In other words, any a_1

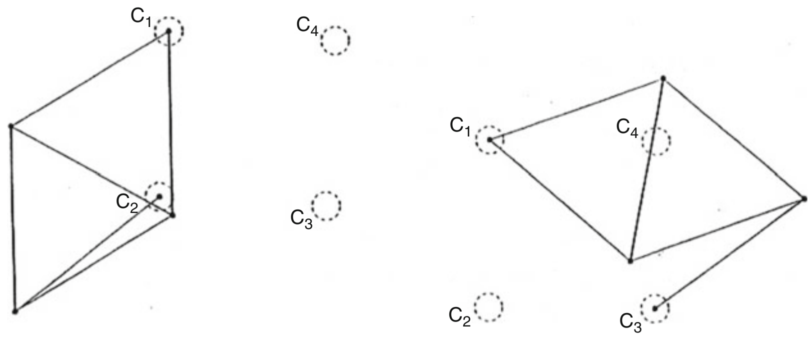


Fig. 14.5 Attaching a 3-cycle to $\{C_1, C_l, C_2\}$ on the left. Attaching a 3-cycle to $\{C_1, C_l, C_3\}$ on the right

points in the δ-ball around C_1 and a_2 points in the δ-ball around C_2 can have a k-cycle attached. ■

This allows the attachment of k-cycles if the center points are at a distance of 0.9 from each other, like C_l and C_2. The configuration consisting of C_l and C_3 can be handled similarly.

Tool 14.7 Let $k \geq 3$ be an odd positive integer. For all positive integers a_l, a_3 such that $a_l + a_3 = k$, there exists δ such that a k-cycle can be attached to any foundation set of type $(a_1, 0, a_3, 0)_\delta$.

Proof We just need to show that we can attach a 3-cycle to one vertex at the center of one δ-ball and two vertices at the center of the other. As is the proof of the previous tool, we use the "too short," "too long," and "just right" continuity argument (see the right drawing of Fig. 14.5). The Appendix at the end of this chapter contains the coordinates of the special points (rounded to five decimal places). ■

14.6 Removing Coincidences

If two vertices from a graph are placed at the same points in the plane, small cycles may inadvertently be created. We must ensure that no vertices coincide. For δ small enough, the regions containing the foundation vertices are disjoint from the regions containing cycle vertices. Furthermore, the foundation vertices can be placed anywhere in the δ-balls, so we choose distinct locations for all of them. It is possible, however, for cycle vertices to coincide. In small graphs, it can be verified computationally that this doesn't occur. For larger graphs, Paul O'Donnell has developed procedures to remove these coincidences.

If the vertices from two different attached cycles coincide, then one foundation vertex is moved slightly, causing all the vertices of the one attached cycle to move slightly, whereas no vertices of the other cycle move. "Slightly" means not enough to introduce any new coincidences. If the vertices from the same cycle coincide, a modification of this method is used to remedy it.

Tool 14.8 If there is an embedding of a unit distance graph G with $m \geq 1$ pairs of coincident vertices, then there is an embedding with fewer than m pairs of coincident vertices.

Proof Given an embedding of G with coincident vertices u and w, we shift some of the vertices of G, subject to several restrictions: no foundation vertex can move outside its δ-ball and no new coincidences may be introduced. Let ε_1 be the minimal distance between any foundation point and the boundary of the δ-ball containing it, and let ε_2 be the minimal distance between any two non-coincident vertices; we define

$$\varepsilon = \min\left\{\varepsilon_1, \frac{\varepsilon_2}{2}\right\}.$$

Given $\varepsilon > 0$, we choose δ', $0 < \delta' < \varepsilon$, such that if a foundation vertex is moved a distance less than δ', then no vertex moves a distance ε or greater. Since the foundation vertices are not moved more than ε_1, they remain inside their δ-balls; thus, all k-cycles can still be attached. Since all non-coincident pairs of vertices are at least ε_2 apart, the movement by less than $\varepsilon_2/2$ does not create new coincidences. Let us consider two cases.

Case 1 Assume that u and w are on different cycles: u is on the cycle $u = u_1, u_2, \ldots$ while w is on the cycle $w = w_1, w_2, \ldots$

Let u_j be a vertex such that no w_i is attached to the foundation vertex u_j^*. Moving u_{j+1} along the unit circle centered at u_{j+1}^* causes each vertex in the cycle

$$u_{j+1}, u_{j+2}, \ldots, u_k, u_1, \ldots, u_{j-1}$$

to move to maintain a unit distance from its foundation vertex and from the preceding cycle vertex. We move u_{j+1} so that no vertex has moved more than ε, and, thus, there is a point unit distance from u_{j-1} to u_{j+1} and distance less than δ' from u_j. This point is the new location of u_j. Now, we move u_j^* the same distance so that it is a unit distance from the new u_j. Of course, moving u_j^* may shift the vertices of cycles attached to it by distances less than ε, but no new coincidences are introduced. Since u_1 moves and w_1 does not, at least one coincidence is removed.

Case 2 Assume that u and w are on the same cycle; to reflect this, we call them u_1 and u_i. We choose a cycle vertex u_j different from the coincident vertices and apply the procedure described in case 1. The only foundation vertex that moves is u_j^*. The only point in the ε-ball around the coincident vertices, which is at a distance 1 from u_1^* and u_i^*, is the original location of those points (u_1 and u_i). Since u_1 and u_i move, while u_1^* and u_i^* do not, they no longer coincide. As before, no new coincidences are introduced. ■

This has been a display of Paul O'Donnell's embedding machinery and his presentation of it [Odo3], [Odo4], and [Odo5]. Can we get an immediate reward from his tool chest? As you know from Chapter 12, Wormald embedded his 6448-vertex graph in the plane. He started with 13 foundation points forming the vertices of a regular 13-gon, attached and embedded $\binom{13}{5}$ 5-cycles, and made sure that no coincidences occurred.

O'Donnell was able to do it much easier – let us take a look.

14.7 O'Donnell's Embeddings

Embedding the Wormald Graph Place four foundation vertices in each of the δ-balls centered at C_1, C_2, and C_3, plus one foundation vertex in the δ-ball centered at C_4. The embedding tools above allow the attachment of all 5-cycles and elimination of all coincidences that may occur. The unit distance embedding of the Wormald graph is thus accomplished! ■

Wormald hints that with considerable effort, he could probably embed a larger Blanche Descartes graph, which is constructed by attaching all 7-cycles to the foundation of 19 vertices. No wonder he does not actually deal with it: for one, this is a 352,735-vertex graph, and, thus, calculations would have grown dramatically. Moreover, Wormald admits that he does not see his approach going any further than a graph of girth 6.

The embedding of this 352,735-vertex graph too becomes trivial, compliments of O'Donnell's embedding tools.

Embedding the 352,735-Vertex Graph Indeed, just place six foundation vertices in each of the δ-balls centered at C_1, C_2, and C_3, plus one foundation vertex in the δ-ball centered at C_4. The embedding tools above allow the attachment of all 7-cycles and elimination of all coincidences that may occur. ■

The next chapter is dedicated to the World Records of Embedding set before the 2009 first edition of this book – join me for the exciting World Series!

Appendix

Vertices used to show a cycle can be attached to vertices at the points C_1, C_2, and C_3.

T_1^{short}	(0.99635, 0.08533)
T_2^{short}	(0.98269, 1.08524)
T_3^{short}	(1.84978, 0.58709)
T_4^{short}	(0.9980, 0.06319)
S_1^{short}	(0.57208, – 0.82020)
S_2^{short}	(0.65177, 0.14158)
S_3^{short}	(1.64588, 1.56608)
S_{14}^{short}	(1.60981, – 0.70439), distance 1 from C_4 and T_1^{short}
S_{24}^{short}	(1.77788, 0.47888), distance 1 from C_4 and T_2^{short}
S_{34}^{short}	(1.81111, – 0.41216), distance 1 from C_4 and T_3^{short}
T_1^{long}	(0.99541, 0.09567)
T_2^{long}	(0.98069, 1.09556)
T_3^{long}	(1.85956, 0.61850)
T_4^{long}	(0.99280, 0.11977)

(continued)

S_1^{long}	(0.58056, − 0.81422)
S_2^{long}	(0.65971, 0.14848)
S_3^{long}	(1.62357, 1.59025)
S_{14}^{long}	1.65414, − 0.65671), distance 1 from C_4 and T_1^{long}
S_{24}^{long}	(1.77374, 0.48640), distance 1 from C_4 and T_2^{long}
S_{34}^{long}	(1.82462, − 0.38089), distance 1 from C_4 and T_3^{long}
T_1	(0.99591, 0.09038)
T_2	(0.98173, 1.09028)
T_3	(1.85476, 0.60261)
S_1	(0.57623, − 0.81729)
S_2	(0.65565, 0.14494)
S_3	(1.63492, 1.57815)
S_{14}	(1.63230, − 0.68098), distance 1 from C_4 and T_1
S_{24}	(1.77587, 0.48255), distance 1 from C_4 and T_2
S_{34}	(1.81794, − 0.39671), distance 1 from C_4 and T_3

Vertices used to show a cycle can be attached to vertices at the points C_1, C_2:

T_1^{short}	(− 0.06194, 0.99808)
T_2^{short}	(0.83339, 0.55268)
T_3^{short}	(0.75995, 1.54998)
T_4^{short}	(0.83339, 0.55268)
S_1^{short}	(0.83339. 0.55268)
S_2^{short}	(− 0.06194, 0.99808)
S_3^{short}	(0.94288, 0.56685)
T_1^{long}	(− 0.08916, 0.99602)
T_2^{long}	(0.81800, 0.57522)
T_3^{long}	(0.74037, 1.57220)
T_4^{long}	(0.81800, 0.57522)
S_1^{long}	(0.81800, 0.57522)
S_2^{long}	(− 0.08916, 0.99602)
S_3^{long}	(0.95233, 0.59493)
T_1	(− 0.07551, 0.99715)
T_2	(0.82580, 0.56397)
T_3	(0.75029, 1.56111)
S_1	(0.82580, 0.56397)
S_2	(− 0.07551, 0.99715)
S_3	(0.94768, 0.58079)

Chapter 15
Embedding World Series

During December 1991–January 1992, Paul Erdős and I were working on the book *Problems of pgom Erdős* at my home in Colorado Springs.[1] Ron Graham called and invited me to meet him in person at a Florida Atlantic University's conference, which I did not know existed. I had just started publishing the problem-posing quarterly *Geombinatorics*, and, at that conference, I introduced it to my colleagues for the first time while giving a talk on the chromatic number of the plane problem. As a result, a group of brilliant young PhD students, including Paul O'Donnell and Rob Hochberg, got excited about the problem and the new journal. *Geombinatorics* has become the main home for related problems and results, outshining, in this regard, top journals on combinatorial theory and discrete geometry. One of the most exciting consequences was the competition for the smallest unit distance triangle-free graph, which I named *Embedding World Series.*

As you recall from Chapter 5, in 1975, Paul Erdős posed a problem to prove or disprove the existence of 4-chromatic unit distance graphs of girths 4, 5, and higher. Nicholas Wormald constructed a girth 5 graph on 6448 vertices (Chapter 12). In my talk, I asked for the smallest example, and the World Series began in earnest on the pages of *Geombinatorics*! New records were set by Paul O'Donnell, Rob Hochberg, and Kiran Chilakamarri; some new record graphs earned names, such as the Moth Graph and the Fish Graph, and appeared on the covers of *Geombinatorics*. In a book form, for the first time, these World Series were presented in the first 2009 edition of this book. You will see that, once a mathematical construction and proof were out of the way, the record holders went on to find a "beautiful" symmetric embedding of their graphs, the ones to which they – or else I – gave special names.

[1] When this second edition is finished, I will get back to finishing *Problems of pgom Erdős* and *Memory in Flashback: A Mathematician's Adventures on Both Sides of the Atlantic*, both under contracts with Springer; so, stay tuned.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_15

15.1 A 56-Vertex, Girth 4, 4-Chromatic Unit Distance Graph [Odo1]

As we have touched on in Section 12.2, in 1955, Jan Mycielski [Myc] invented a method of constructing triangle-free graphs of an arbitrary chromatic number k: Start with a triangle-free (k–1)-chromatic graph G. For each vertex $v_i \in V(G)$, add a vertex w_i adjacent to all vertices in the neighborhood of v_i. Next, add a vertex z adjacent to all of the new vertices. The chromatic number of this new graph is k, and it is still triangle-free. Let us call this graph *Mycielskian of G* and denote it as $M(G)$. Unfortunately, the resultant graph does not often embed in the plane. Notice that if a vertex of G has degree 3 or more, then the Mycielskian $M(G)$ of G contains a $K_{2,3}$ subgraph. The plane contains no unit distance $K_{2,3}$ subgraph, so the starting graph G must have a maximum degree at most 2 for the Mycielskian to be a unit distance graph. Thus, the only candidates for the unit distance version of the Mycielski construction are unions of paths and cycles. The Mycielskian of an odd cycle does not embed in the plane, however, so the Mycielski construction does not give a 4-chromatic unit distance graph. The Mycielskian of at least one even cycle does embed, though.

The 5-cycle u_1, u_2, u_3, u_4, u_5 is said to be *attached* to the set of vertices $\{v_1, v_2, v_3, v_4, v_5\}$ if v_i is adjacent to u_i for $1 \le i \le 5$ (see Fig. 15.1). Such an attachment is a useful operation because it can increase the chromatic number of a graph from 3 to 4 without introducing any 3-cycles.

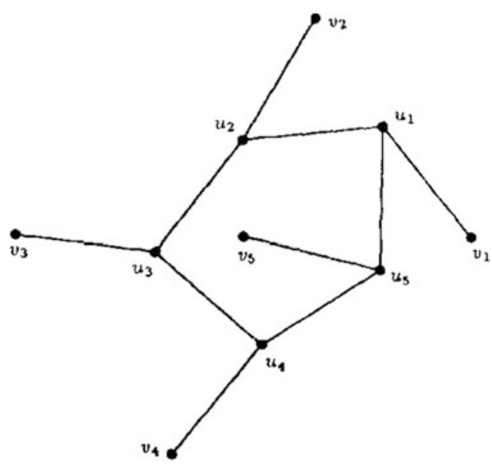


Fig. 15.1 The 5-cycle u_1, u_2, u_3, u_4, u_5 is *attached* to $\{v_1, v_2, v_3, v_4, v_5\}$

The graph H in Fig. 15.2 is the Mycielskian of the 10-cycle C_{10}. It can be shown with basic geometry and algebra that H can be embedded in the plane, but O'Donnell reports that Rob Hochberg pointed out a nicer proof, which shows *why* this is so. H is a subgraph of the projection of the 5-cube along a diagonal onto the plane. The coordinates of the vertices v_1, v_3, v_5, v_7, v_9 are the fifth roots of unity, while the edges are all of unit length since they are translations of these unit vectors. This graph is only 3-chromatic; thus, we will attach 5-cycles to make it 4-chromatic.

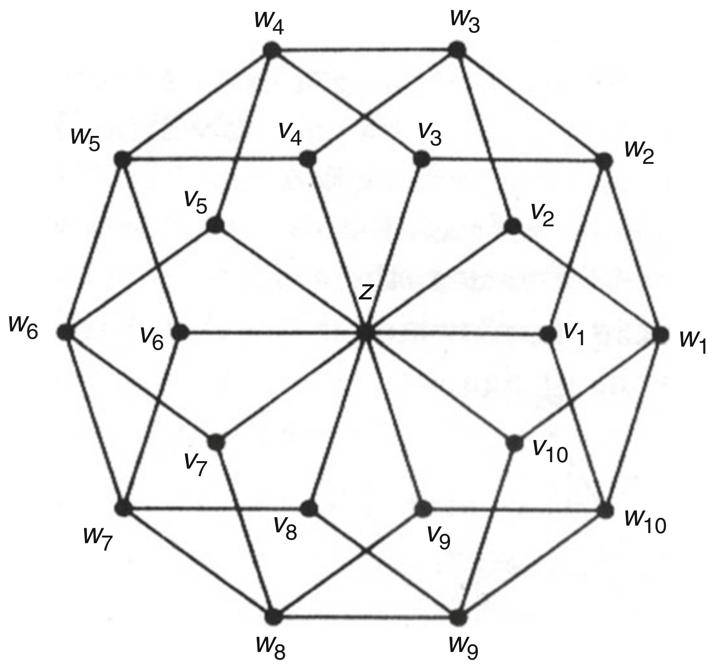


Fig. 15.2 $H = M(C_{10})$, the Mycielskian of C_{10}

Construction 15.1 A 5-cycle can be attached to the subgraph $R = \{v_1, v_3, v_5, v_7, v_9\}$ of the graph H depicted in Fig. 15.2.

Proof Center a regular pentagon of side length 1 at the origin and rotate it until the distance from one of its vertices to v_1 is 1 (see Fig. 15.3). Then the respective distances from the other vertices of the pentagon to the other vertices of the graph will all be 1.

Does this attachment remind you of the construction of the Golomb Graph (Fig. 2.8)? It should, for the Golomb Graph was O'Donnell's inspiration for this nice construction. As far as I am concerned, it reminds me of Fig. 13.4, obtained by Erdős–Harary–Tutte.

Moreover, you will see in Part X: *Ask What Your Computer Can Do for You* that this construction is extensively used today and has earned the name the *spindling method.* ■

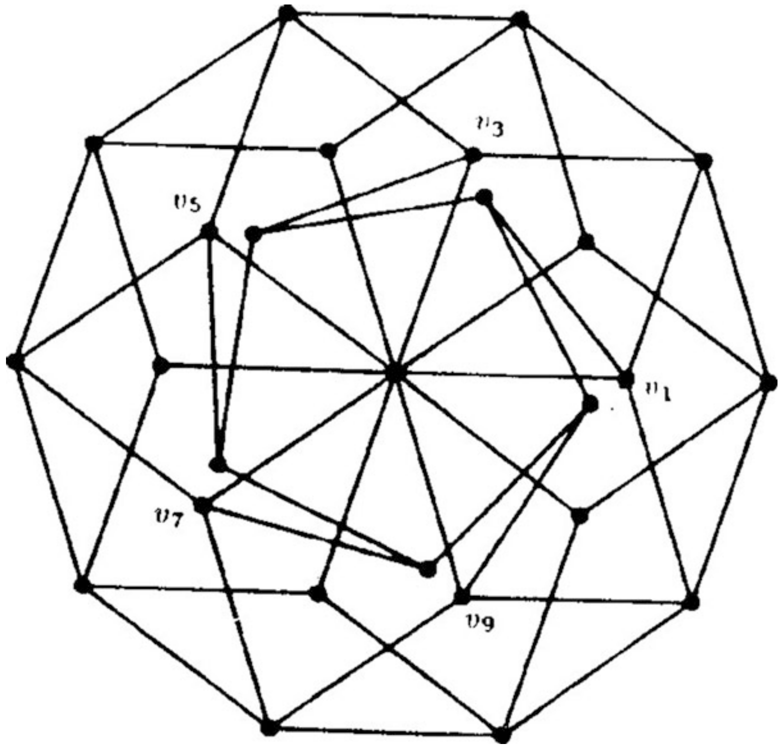


Fig. 15.3 H with one 5-cycle attached

Construction 15.2 A 5-cycle can be attached to $T = \{w_1, w_3, w_6, w_8, z\}$ (see Fig. 15.2).

Proof The proof relies on the intermediate value theorem and the continuity argument introduced in the beginning of Chapter 14. Described a little less formally, we try to attach a 5-cycle to the five vertices of T so that the cycle edges and the connecting edges are all of unit length. In fact, we try it twice. The problem with the attachments is that in the first one, one of the edges in the cycle is too short and in the second, it is too long. Since one configuration is obtained from the next by a continuous transformation, there exists an attachment where the same edge has unit length. Thus, T can have a 5-cycle attached (see Fig. 15.4).

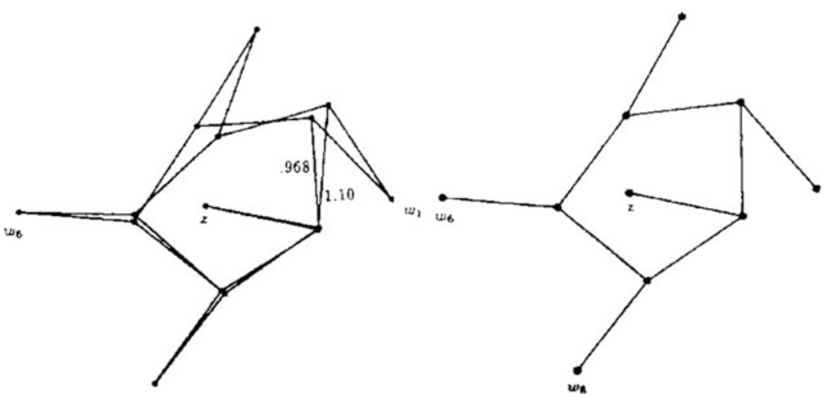


Fig. 15.4 The “short” attachment and the “long” attachment shown together on the left. The “just right” attachment is on the right. (All unlabeled edges are of unit length)

The most efficient way to verify these attachments is by a computer, although it is necessary to make sure that the error made by approximating the numbers does not affect any of the inequalities. The error in the numbers listed below is $<10^{-5}$, which does not affect the results.

Let C_i be the unit circle around the ith vertex in H. Let u_1, u_2, u_3, u_4, u_5 be a path with unit length edges and with u_i on C_i. This almost gives an attached 5-cycle. The attaching edges are all of a unit distance since each u_i is on the unit circle around some vertex in H, and the four-path edges are of a unit distance. This path can be slid back and forth in a continuous manner with each u_i tracing out an arc on C_i. One such path is approximately (0.95, 0.74413), (−0.04916, 0.70312), (−0.62463, −0.11470), (0.13436, −0.76580), (0.974661, −0.22369). The vertices of this path form a "too short" attachment where all distances are one, except from u_5 to u_1 where the distance is about 0.968.

A second path that can be obtained from the first one by the continuous sliding is (1.1, 0.85536), (0.13069, 0.60954), (−0.61938, −0.05183), (0.10423, −0.74206), (0.97027, −0.24204). The vertices of this path form a "too long" attachment where the u_5 to u_1 distance is 1.10. By continuity, there is a "just right" attachment where the edge from u_5 to u_1 is exactly one. The exact coordinates of this attachment are unknown, but, for claiming their existence, it suffices to show that a 5-cycle can be attached to our set (see Fig. 15.4). ■

Construction 15.1 allows us to attach a 5-cycle to $\{v_1, v_3, v_5, v_7, v_9\}$. Similarly, we can attach another 5-cycle to $\{v_2, v_4, v_6, v_8, v_{10}\}$. We get the new graph and call it H'. In a proper 3-coloring of H', the vertices in $\{v_1, v_3, v_5, v_7, v_9\}$ cannot get the same color since that leaves only two colors for the attached 5-cycle. The same holds for $\{v_2, v_4, v_6, v_8, v_{10}\}$. This is enough to rule out most of the 3-colorings of H'. In fact, except for the vertices of the attached 5-cycles, the coloring of H' is completely determined up to symmetries. This coloring is shown in Fig. 15.5 (the attached 5-cycles are not shown in the figure). Note that there are numerous ways to color the attached 5-cycles, but their attachment forces the rest of the graph to have a unique coloring up to a permutation of the colors and rotation of the graph.

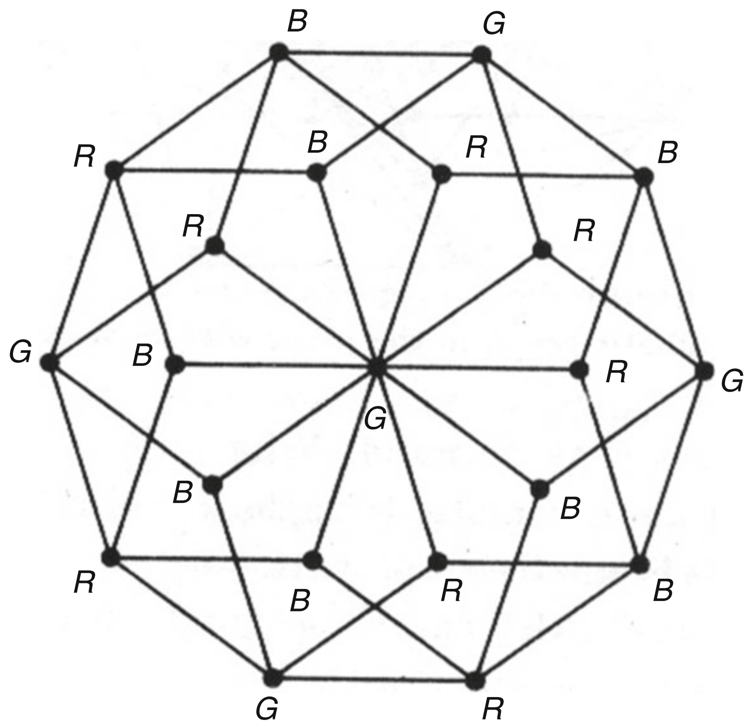


Fig. 15.5 The vertices of H must have this coloring up to symmetries when two 5-cycles are attached. (The attached 5-cycles are not shown)

In particular, in every 3-coloring of H', for some j, $1 \leq j \leq 5$, the set $\{w_j, w_{j+2}, w_{j+5}, w_{j+7}, z\}$ (addition modulo 10) is monochromatic, where z and w_i are as in Fig. 15.2. By attaching 5-cycles to all five of these sets, we exclude all 3-colorings. The result is a 4-chromatic graph. Moreover, since H is triangle-free, this new graph is also triangle-free. Approximation of the coordinates of the vertices ensures that there are no coincident vertices.

It is time to count the vertices of our construction: H has 21 vertices, then two 5-cycles are added, and then five 5-cycles more. The result is a triangle-free, 4-chromatic graph on 56 vertices (Fig. 15.6).

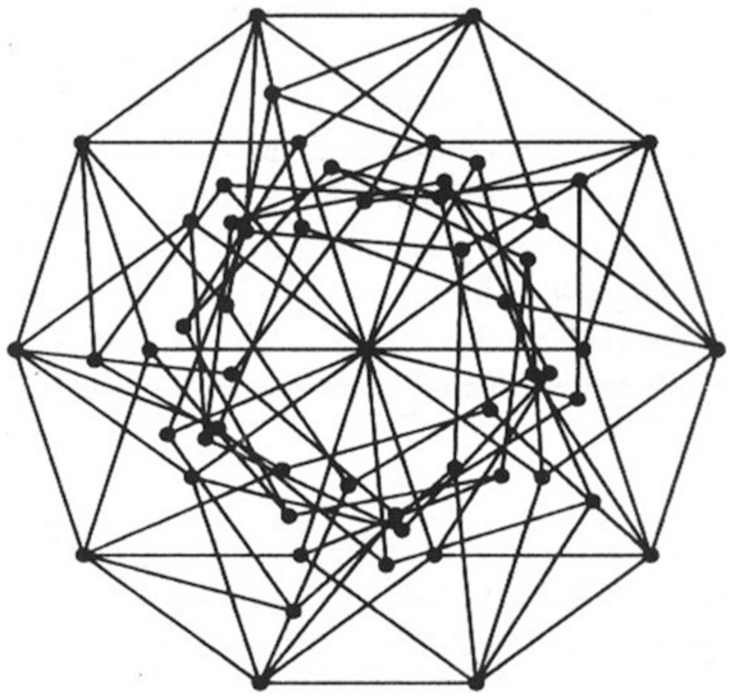


Fig. 15.6 A 56-vertex 4-chromatic graph in the plane with no 3-cycles

Beating the 6448-vertex Wormald graph with the new world record of a 56-vertex graph was a striking achievement.

In closing, Paul O'Donnell observes: One reason to search for triangle-free graphs is that they seem to be flexible. For example, *H* can be bent into a 4-chromatic graph, containing many Mosers Spindles (Fig. 15.7).

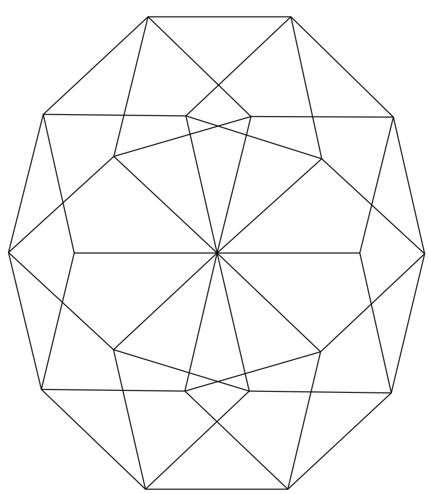


Fig. 15.7 *H* can be bent so that new edges (unit distances) are introduced. The chromatic number of this graph is 4

Paul ends [Odo1] with the ultimate goal (or ultimate musing):

Perhaps flexibility will prove useful in the construction of a 5-chromatic unit-distance graph in the plane!

15.2 A 47-Vertex, Girth 4, 4-Chromatic Unit Distance Graph [Chi]

Professor Kiran Chilakamarri, then of the Ohio State University (and later of Texas Southern University) was one of the early researchers of the chromatic number of the plane. Among other related things, he was much interested in constructing the smallest possible example of a 4-chromatic, triangle-free unit distance graph. I have little doubt that his work was well on the way when Paul O'Donnell published the first, 56-vertex breakthrough in these real (unlike baseball) World Series. In the fall of 1995, Kiran responded by beating Paul's world record with the 47-vertex Moth Graph of his own, which I promptly published on the cover of the January 1995 issue of *Geombinatorics*.

Chilakamarri constructs his example in stages, at each stage describing the properties shared by all possible colorings of the graph constructed. He begins with a graph on 12 vertices and 20 edges, which he calls the *core graph* shown in Fig. 15.8.

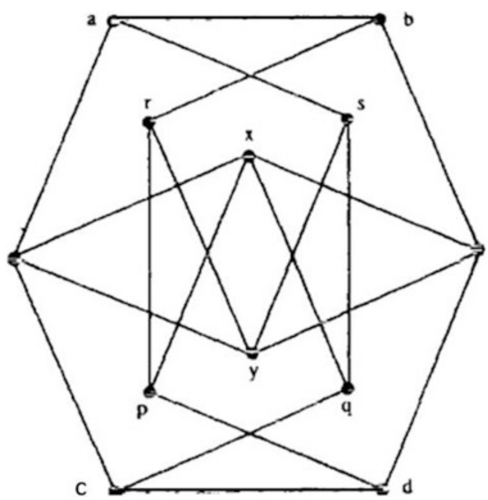


Fig. 15.8 A core graph

Chilakamarri then invents the *right-wing* graph (Fig. 15.9) on 10 vertices and 12 edges and symmetrically the *left-wing* graph.

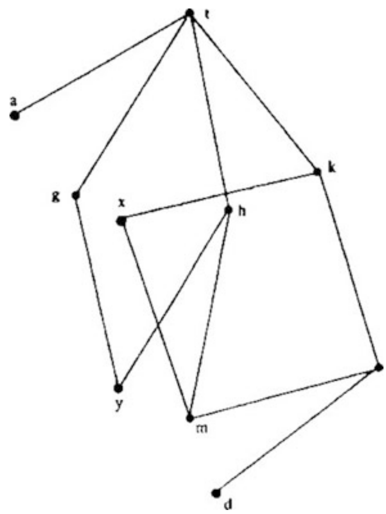


Fig. 15.9 The right-wing graph

He then attaches the wings to the core graph and gets the *Butterfly Graph* (Fig. 15.10).

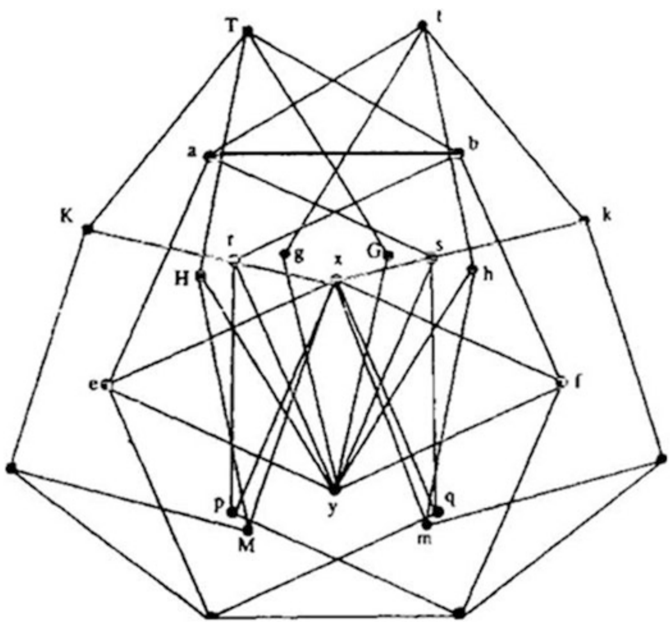


Fig. 15.10 The Chilakamarri butterfly graph

Finally, he joins two butterflies to produce the 47-vertex graph, which he proves to be 4-chromatic (Fig. 15.11).

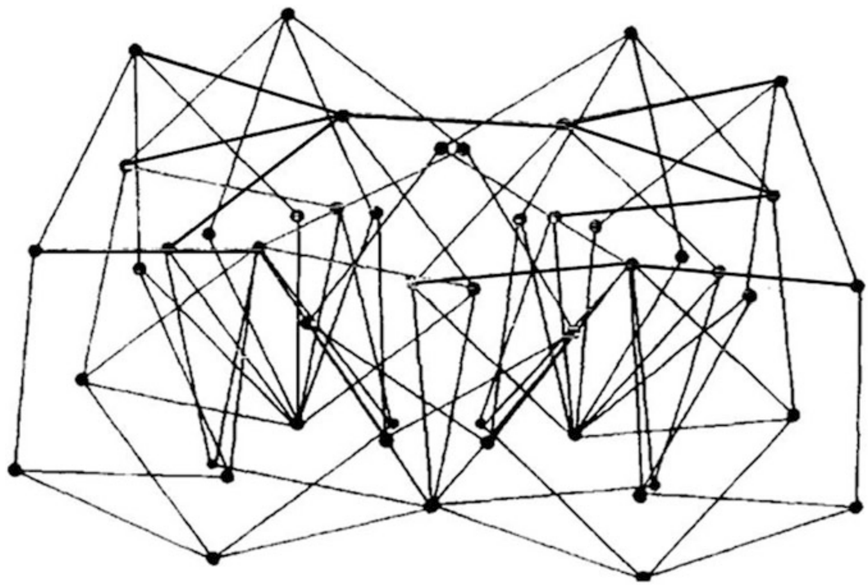


Fig. 15.11 The Chilakamarri Moth graph

Kiran then proves "the existence" (i.e., the existence of an embedding in the plane) of the Moth Graph by producing its coordinates. Finally, he proves that the Moth graph has girth 4· by checking that (a) the vertices of the core graph do not form an equilateral triangle, (b) the left wing has no equilateral triangle, (c) as we add the wings to the core, no new edges are created, and, finally, (d) as we join two butterfly graphs, no new edges are created other than the edge we have added. ■

As Kiran Chilakamarri set his new world record of 47, our World Series became so intense in mid-1995, that Chilakamarri in this July 1995 paper mentions in the endnotes "Paul O'Donnell tells me he is shrinking the size of the example (≤40?)..." Moreover, Robert Hochberg modified O'Donnell's 56-vertex construction to get a 46-vertex, unit distance, triangle-free 4-chromatic graph and thus beat Chilakamarri's world record of 47, but, Rob, to my regret, decided against publishing it because he too learned that O'Donnell was getting ready to roll out yet another new world record, the 40-vertex graph.

15.3 A 40-Vertex, Girth 4, 4-Chromatic Unit Distance Graph [Odo2]

Similarly, to [Odo1] approach, Paul O'Donnell starts with the Mycielskian of the 5-cycle C_5. This 11-vertex Mycielski–Grötzsch graph (we saw it in Fig. 12.8) is the smallest triangle-free 4-chromatic graph. Since it is not a unit distance graph, we modify it by taking out the "central" vertex adjacent to the 5 "new" vertices and replacing it with five vertices, each adjacent to a pair of "new" vertices as shown in Fig. 15.12. Call this graph *H*.

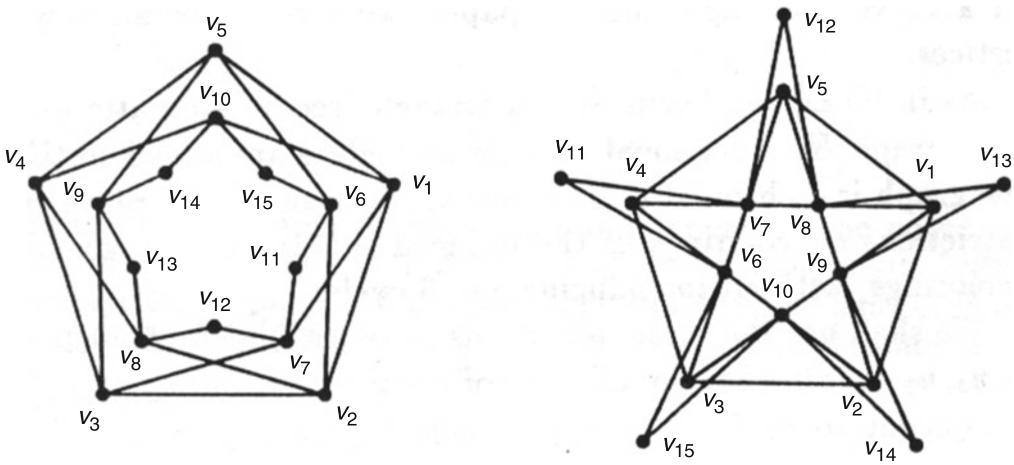


Fig. 15.12 An "instructive" drawing of H on the left. A unit distance embedding of H is on the right

H is 3-chromatic, but all the 3-colorings share a valuable property. In every 3-coloring, one of the sets $\{v_{1+i}, v_{6+i}, v_{11+(i+1)}, v_{11+(i+2)}, v_{11+(i+3)}\}$, for $0 \le i \le 4$ (where the parentheses indicate addition modulo 5), is monochromatic. By attaching 5-cycles, one of which is shown in Fig. 15.13, to all such sets, all 3-colorings get excluded. Thus, the resultant graph H' is 4-chromatic and still triangle-free. It remains to show that H' is a unit distance graph.

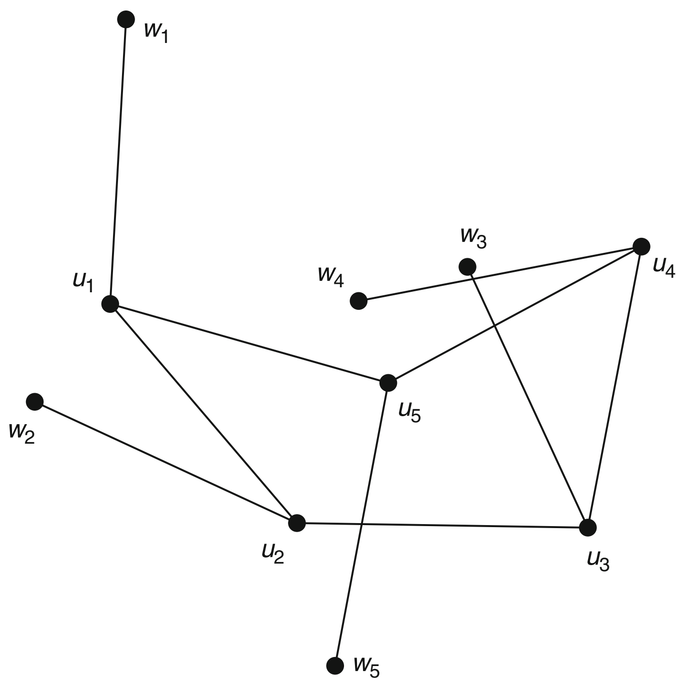


Fig. 15.13 The 5-cycle u_1, u_2, u_3, u_4, u_5 is attached to $\{w_1, w_2, w_3, w_4, w_5\}$

Construction 15.3 A 5-cycle can be attached to $T = \{v_1, v_6, v_{12}, v_{13}, v_{14}\}$ (see the right Fig. 15.12).

Proof We try to attach a 5-cycle w_1, w_2, w_3, w_4, w_5 so that the cycle edges and all the connecting edges are of unit length (see Fig. 15.14). It is fairly simple to attach a unit distance *path* w_1, w_2, w_3, w_4, w_5 to *T*. The hard part is getting w_5 and w_1 to be at a unit distance to complete the cycle.

Define a continuous function $f(\theta)$ to be the length of the edge $\{w_1, w_5\}$ when vertex w_1 is placed at an angle θ and at a unit distance from v_1 and each subsequent w_i is placed at a unit distance from both w_{i-1} and its corresponding vertex in *T*. Typically, there are two possible positions for w_i, so a precise description of $f(\theta)$ would include how all of the choices are made. It suffices to say that there exists $f(\theta)$ satisfying the above description and continuous on some interval [a, b] on which $f(a) < 1$ and $f(b) > 1$. By the intermediate value theorem, for some $\theta_0 \in [\mathrm{a}, b], f(\theta_0) = 1$. ■

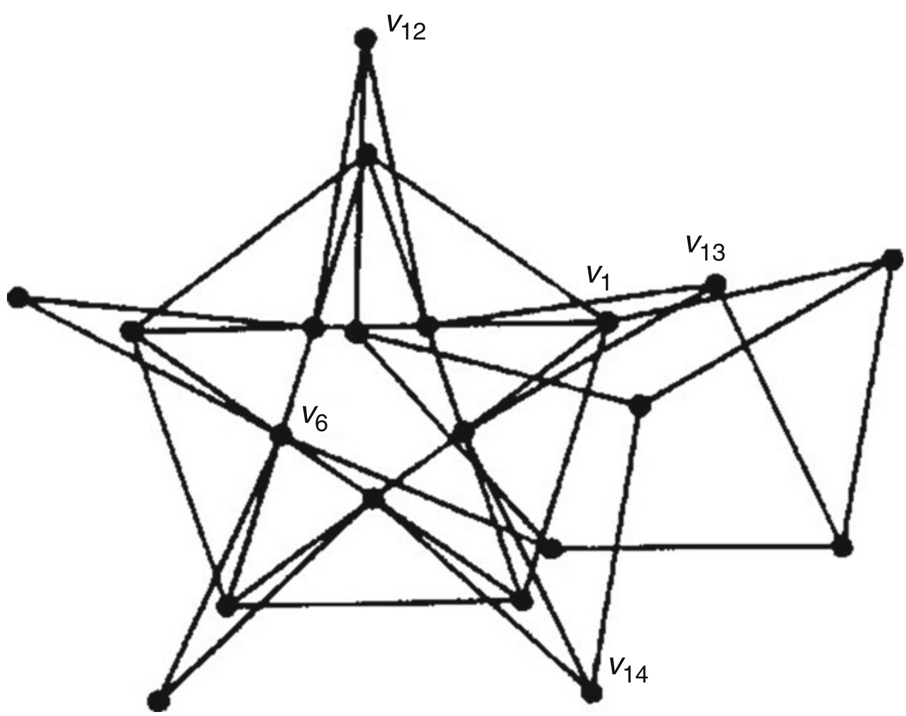


Fig. 15.14 *H* with a 5-cycle attached to *T*

By attaching 5-cycles to *T* to all five of its rotations, we obtain a graph with the desired properties. Since *H* had 15 vertices and we attached five 5-cycles, the result is a 4-chromatic, triangle-free unit distance graph on 40 vertices (Fig. 15.15).

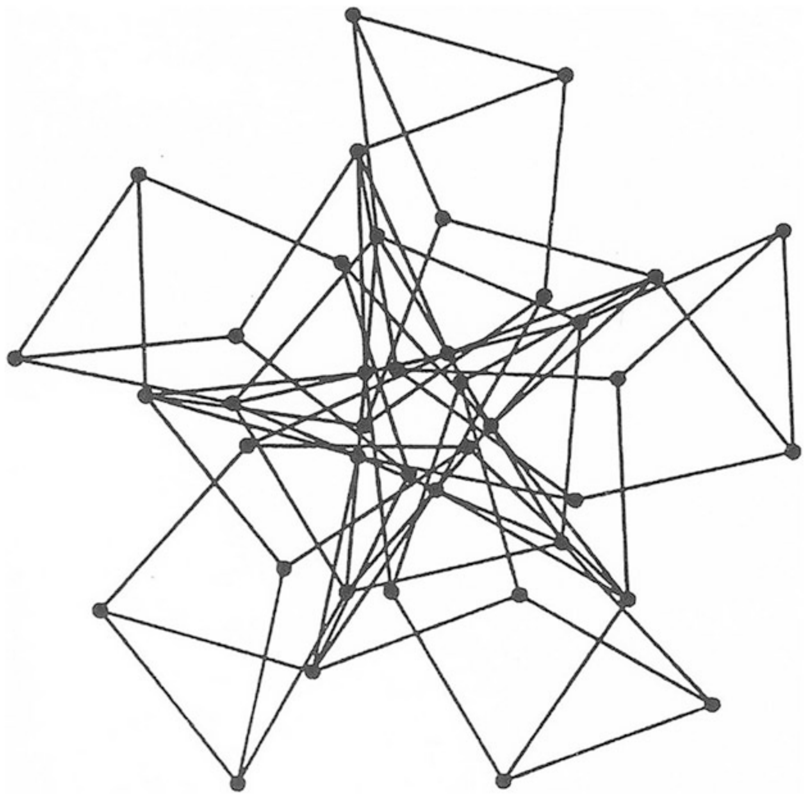


Fig. 15.15 The O'Donnell pentagonal graph

The new world record of 40 was fabulous; I gladly published it on the cover of the July 1995 issue of *Geombinatorics.* However, it was not the end of the Embedding World Series. Paul ended his essay [Odo2] with a promise of more things to come:

> Some related questions are still wide open. Given k, is there a 4-chromatic unit-distance graph with no $< k$-cycles? What is the smallest 4-chromatic triangle-free unit-distance graph? And of course, is there a 5-chromatic unit-distance graph in the plane? Stay tuned to *Geombinatorics* research quarterly for further developments.

15.4 A 23-Vertex, Girth 4, 4-Chromatic Unit Distance Graph

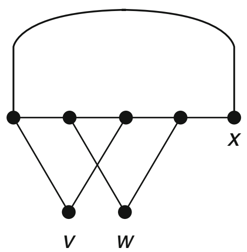
Indeed, more things did come. It remains a mystery to me why Paul O'Donnell did not include in his doctorate dissertation the two world records he has jointly set with Rob Hochberg. In the dissertation, Paul mentions this achievement briefly, as if in passing:

> In joint work with R. Hochberg [HO], the upper bounds on the sizes of the smallest 4-chromatic unit-distance graphs with girths 4 and 5 were lowered even more. A 23-vertex, girth 4, 4-chromatic unit-distance graph was found. The construction involved a generalized version of cycle attachment. A 45-vertex, girth 5, 4-chromatic

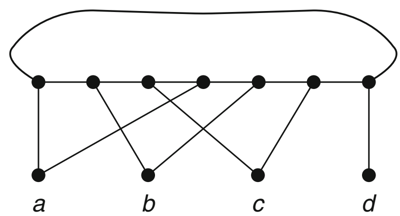
unit-distance graph was found. The construction involved a generalized version of cycle attachment.

That is all! Fortunately, I published their remarkable paper in *Geombinatorics* in April 1996 [HO], and, so, we are able to revisit it here.

In [Odo1] and [Odo2], Paul O'Donnell used the idea of attaching odd cycles to specified subsets of vertices of a starting independent set. Here, Rob Hochberg and Paul O'Donnell use a more complicated notion of *attaching*: a cycle might not have all of its vertices attached to the independent set, and some vertices in the independent set may have more than one vertex of the cycle attached to them. Figure 15.16 illustrates two applications of this idea.



Here, a 5-cycle is partially attached to the independent set $\{v, w\}$. In any 3-coloring of this graph, if v and w get the same color, then x must also get that color.



Here, a 7-cycle is attached to the independent set $\{a, b, c, d\}$. Any coloring of this graph that, makes the independent set monochromatic, must use at least 4 colors. Note that this graph has girth 5.

Fig. 15.16 Attaching odd cycles to independent sets

In Fig. 15.17a, the 5-cycle (u_1, u_2, u_3, u_4, u_5) is partially attached (by dashed lines) to $\{w, y\}$. Observe that in any 3-coloring, if w and y get the same color, then u_5 must also receive that color.

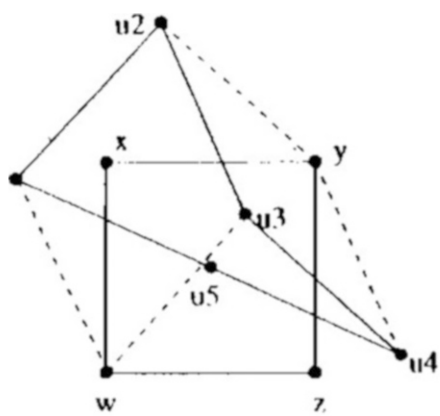


Fig. 15.17a Attaching a 5-cycle

To these three vertices $\{w, y, u_5\}$, we then attach the (bold) 5-cycle $(v_1, v_2, v_3, v_4, v_5)$, as shown in Fig. 15.17b.

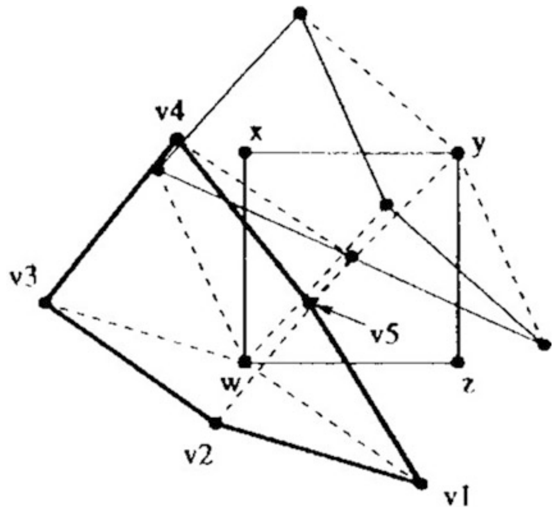


Fig. 15.17b Attaching a 5-cycle

Now, in any 3-coloring of this graph, if w and y (and hence u_5) receive the same color, then there are only two colors left, for the attached odd cycle, making such a 3-coloring impossible. But in any 3-coloring of the square $\{w, x, y, z\}$, one of the pairs $\{w, y\}$ or $\{x, z\}$ must be monochromatic. So, we take a copy of the two 5-cycles shown in Fig. 15.17b (flipped about a horizontal axis so that they are now attached to the pair $\{x, y\}$). With the coincidence at the center of the square, this adds only 9 new vertices (rather than 10 – every vertex counts when we set world records!), creating a 23-vertex graph with no 3-coloring. This graph is shown in Fig. 15.18. I named it the Hochberg–O'Donnell Fish Graph.

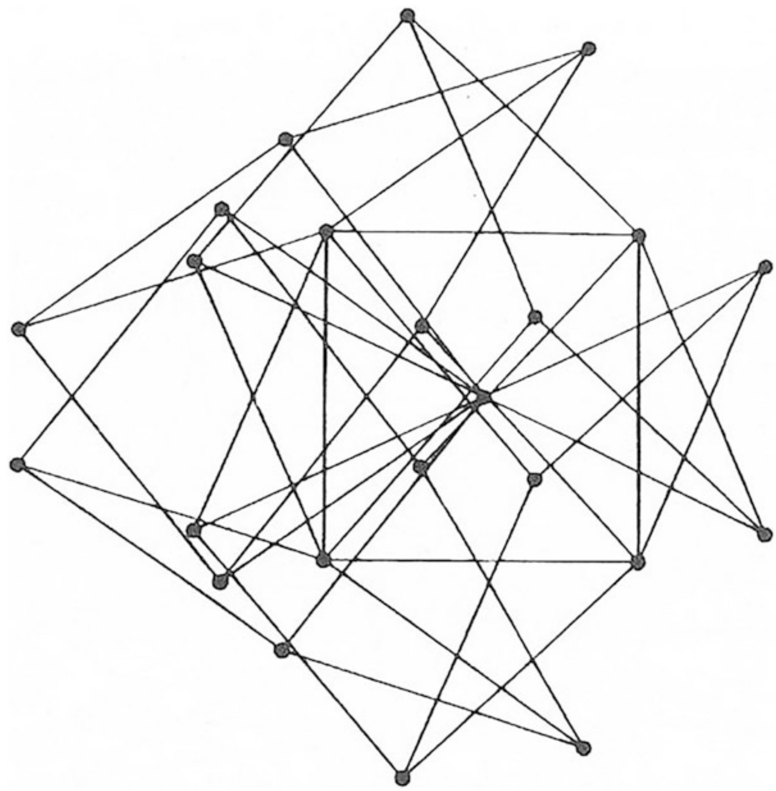


Fig. 15.18 The Hochberg–O'Donnell Fish Graph

It remains to be shown that the graph is indeed of a unit distance. Clearly, it suffices to show that the 5-cycles can be attached in the way we described. The proof relies on the intermediate value theorem and the continuity argument. We try to attach a cycle to a specified set of vertices so that the cycle edges and the connecting edges are all of unit length. In fact, we do it twice: in the first, one of the edges in the cycle will be too short and, in the second, it will be too long. Since one configuration can be obtained from the other by a continuous transformation (which does not alter the lengths of the unit length edges), there exists an attachment where the same edge has length 1. This works for all the attachments and partial attachments in these constructions. We have looked at this argument in greater detail earlier in this chapter where we discussed O'Donnell's 56- and 40-vertex record graphs. ■

The problem remained open:

Problem 15.4 What is the smallest number of vertices in a 4-chromatic unit distance graph of girth 4?

As you know, the smallest 4-chromatic, triangle-free graph is the Mycielski–Grötzsch Graph on 11 vertices. The Fish Graph satisfies all the Grötzsch conditions plus one extra: It is a unit distance graph. It was remarkable that Rob and Paul managed with merely 23 vertices. Was this the smallest possible number of vertices? I was not sure. I thought that it was fairly close to it.

In the course of 2 years, on the pages of *Geombinatorics*, we traveled from 6448 vertices all the way to 23, a fine achievement. This record survived for two decades, until fairly recently Geoffrey Exoo and Dan Ismailescu have completely settled my Problem 15.4. Chapter 16 is dedicated to their work.

15.5 A 45-Vertex, Girth 5, 4-Chromatic Unit Distance Graph

Recall the Petersen Graph (Fig. 13.3) and its unit distance embedding in the plane (Fig. 13.4), which was discovered by the distinguished triumvirate of mathematicians Erdős–Harary–Tutte in their 1965 article [EHT] (yes, that is where they famously observed, "It is easy to see (especially after seeing it)." Here Hochberg and O'Donnell pursue their second idea (see Fig. 15.16 on the right). Accordingly, in Fig. 15.19, a 7-cycle (shown in bold) is attached to a 4-vertex independent set of the Petersen graph.

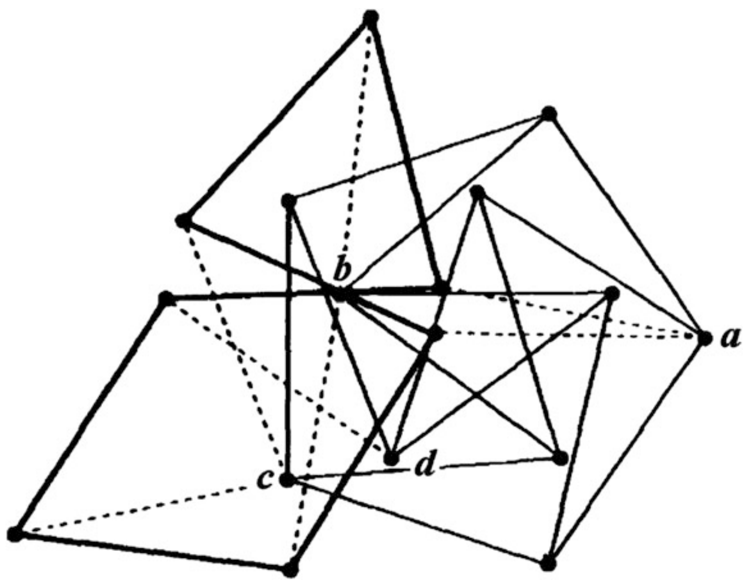


Fig. 15.19 The Petersen Graph with a 7-cycle attached (by dashed lines)

The authors then simply write: "By the pigeonhole principle, in any 3-coloring of the Petersen Graph, one of the five rotations of the set $\{a, b, c, d\}$ will be monochromatic." Can you figure out how the pigeons help here? Upon pondering for a few minutes, I have understood it (though not sure whether the authors had the same argument in mind): In a 3-coloring of the Petersen Graph, at least 4 out of its 10 vertices must appear in the same color (that is the Pigeonhole Principle). Now, which four vertices could that be (here the Pigeonhole Principle is of no help)? The answer is two vertices on the outer pentagon and two on the inner star. You can now verify (do) that the only pair of the outer monochromatic vertices that allows two inner vertices in the same color, up to a rotation is a, c (Fig. 15.19). It is then clear that a, c must be accompanied in the same color by the vertices b, d of the inside!

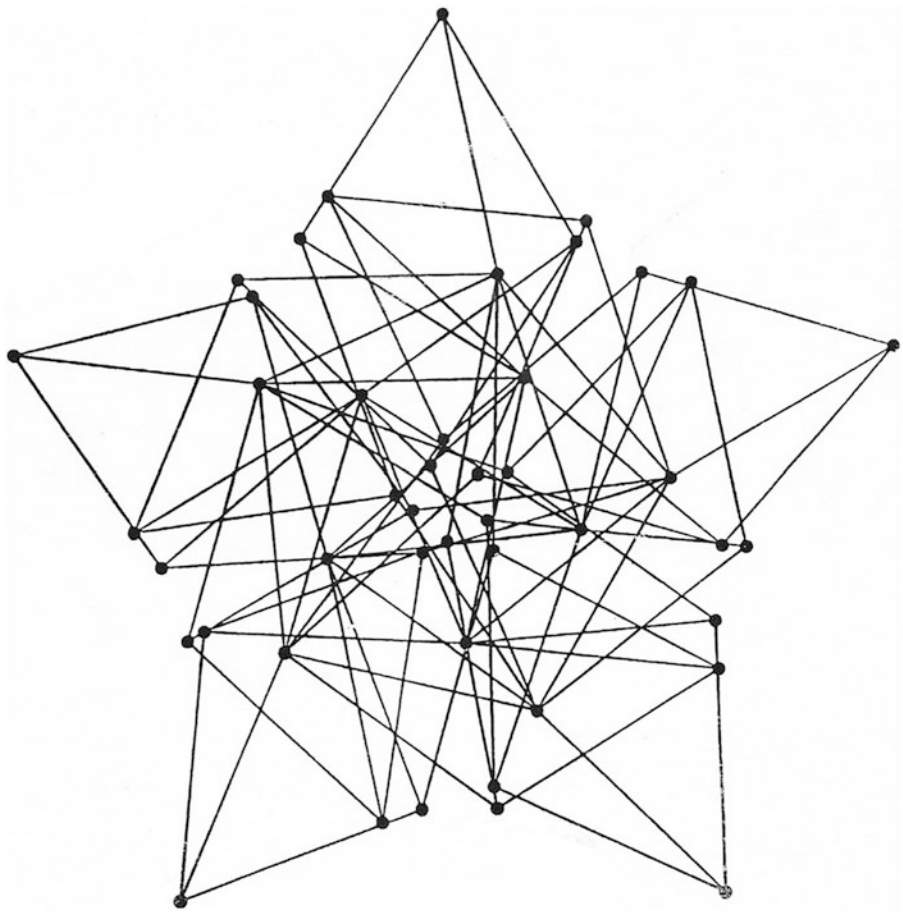


Fig. 15.20 The Hochberg–O'Donnell Star Graph

When 7-cycles are attached to all five rotations of $\{a, b, c, d\}$, the resulting graph will not be 3-colorable. This gives a 45-vertex 4-chromatic graph with no 3-cycles and 4-cycles. This beautiful graph is shown in Fig. 15.20. I gave it the name honoring its creators, the Hochberg–O'Donnell Star and published it on the front cover of the April 1996 issue of *Geombinatorics.*

Finally, we need to show that the Star Graph is indeed embeddable in the plane. It suffices to show that the 7-cycles can be attached in the way we described. The proof again relies on the intermediate value theorem and the continuity argument. We need to attach a 7-cycle to a specified set of vertices so that the cycle edges and the connecting edges are all of unit length. Instead, we do it twice: in the first one, one of the edges in the cycle will be too short and, in the second one, it will be too long. Since one configuration can be obtained from the other by a continuous transformation (which does not alter the lengths of the unit length edges), there exists an attachment where the same edge has length one. This works for all the attachments and partial attachments in these constructions. We looked at this argument in greater detail earlier in this chapter where we discussed O'Donnell's 56- and 40-vertex record graphs. ■

Open Problem 15.5 What is the smallest number of vertices in a 4-chromatic unit distance graph of girth 5?

I hope that you have enjoyed getting acquainted with the beautiful new graphs and the world records they represent. Tables 15.1 and 15.2 summarize the world records' history as it stood at the time of this book's first edition publication and underscore the role of *Geombinatorics* as the playing field of this World Series. In the next chapter, you will learn about the recent new achievements!

Table 15.1 World records up to 2008: smallest unit distance 4-chromatic graph of girth 4

Number of vertices	Author	Publication date	Journal
6448	N. Wormald	1979	[Wor]
56	P. O'Donnell	July 1994	*Geombinatorics* IV(1), 23–29
47	K. Chilakamarri	January 1995	*Geombinatorics* IV(3), 64–76
46	R. Hochberg	1995	Unpublished
40	P. O'Donnell	July 1995	*Geombinatorics* V(1), 31–34
23	R. Hochberg and P. O'Donnell	April 1996	*Geombinatorics* V(4), 137–141

Table 15.2 World records up to 2008: smallest unit distance 4-chromatic graph of girth 5

Number of vertices	Author	Publication date	Journal
6448	N. Wormald	1979	[Wor]
45	R. Hochberg and P. O'Donnell	April 1996	*Geombinatorics* V(4), 137–141

It is time to move on: We still have a lot of exciting colored and coloring mathematics to experience. Armed with great results on colored integers in Part VII, we will return to Paul O'Donnell's dissertation: Part IX will be dedicated to his main results.

Chapter 16
Exoo–Ismailescu: The Final Word on Problem 15.4

16.1 A Brief Story of Submission

On September 20, 2016, I received an e-mail from Dan Ismailescu:

> Can you please tell us if you had the chance to look over the paper we submitted a few months ago?

The e-mail below from Ismailescu, copied to his coauthor Geoffrey Exoo, was in fact dated a whole 8 months earlier, on January 26, 2016, and read as follows:

> Dear Professor Soifer,
>
> Attached you will find an article to be considered for publication in Geombinatorics.
>
> The paper deals with triangle-free 4-chromatic unit distance graphs, a problem which we know is very dear to you. We hope you will like it. We included the signed copyright form as well. Please let us know if there is anything else on our end that has to be done.
>
> Many thanks
>
> Dan Ismailescu

I sent my reply to the authors on September 25, 2016, after an extensive but futile search of my inbox:

> Dear Dan and Geoff,
>
> I searched my inbox several times – and found only your old submissions.
>
> I am terribly sorry, but your January submission must have fallen into a black hole.
>
> Now that I know that you sent your submission on 01/26, I will put your paper in the queue at that date. I do need you to resubmit the essay and the copyright form.
>
> I am glad you inquired, for this gives us an opportunity to fix the problem.
>
> And yes, I can't wait to read your paper – you are right, the topic *is* dear to my heart. Moreover, I am interested in your research!

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_16

Best wishes,

Alexander

P.S.: it would be great to receive your resubmission now, today–tomorrow, as I am selecting material for the October issue.

In 10 min, I had the essay in my hands and dove into it. It was a treasure. My response was immediate:

Wow! 17!

You came so close to 11 of the Grötzsch graph. Congratulations to you both. Write more for us. Your essay will appear in October or next, January issue.

Yours always,

Alexander

Geoffrey replied:

Thanks. I always loved Chapter 15 (in the coloring book)!

– Geoff

The nearly-lost essay was so impressive that I invited Geoffrey Exoo to join the editorial board of *Geombinatorics*. I am grateful to Geoff for his acceptance. Now that you know the *Story of Submission*, let us look at its content.

To make their essay [EI1] self-contained, the authors begin with a historical chapter concisely presenting the contents of Chapters 2, 3, and 15 of this book. They end the history with open Problem 15.4, which they solve completely:

We construct triangle-free 4-chromatic unit distance graphs with 21, 19, and 17 vertices, respectively. Moreover, we present evidence that the value 17 cannot be improved.

16.2 Constructing Triangle-Free, 4-Chromatic Unit Distance Graphs

Exoo and Ismailescu convey their result in the style of Hochberg–O'Donnell of the previous Chapter 15 so well that I pass the flaming pen to them for this chapter and only change references they use to those already present in my book.

We first describe the main idea behind our constructions. Our presentation is modeled after [HO].

Let G be a triangle-free, 3-chromatic unit distance graph with $n \geq 7$ vertices and e edges. For instance, the disjoint union of two unit 5-cycles is such a graph. Since $x(G) = 3$ and $n \geq 7$, it follows that $\alpha(G)$, the *independence number of* G, is at least $[n/3] \geq 3$.[1]

Let $I = \{1, 2, 3\}$ be a three-vertex independent set of G. We augment graph G by adding 5 new vertices, all different from the vertices of G, and 10 new edges in the following manner. We say that the 5-cycle $[a, b, c, d, e]$ is *attached to the independent set* I if a and c are adjacent

[1] The independence number $\alpha = \alpha(G)$ of a graph G is the cardinality of a maximum independent set of vertices.

to vertex 1, b and d are adjacent to vertex 2, and e is adjacent to vertex 3. Furthermore, suppose that this attachment can be done so that all these 10 edges have unit length – see Fig. 16.1 for an illustration.

We, thus, have 5 new vertices a, b, c, d, e, and 10 new unit edges: $\{a, b\}$, $\{b, c\}$, $\{c, d\}$, $\{d, e\}$, $\{e, a\}$ as the edges of the 5-cycle as well as $\{1, a\}$, $\{2, b\}$, $\{1, c\}$, $\{2, d\}$, $\{3, e\}$.

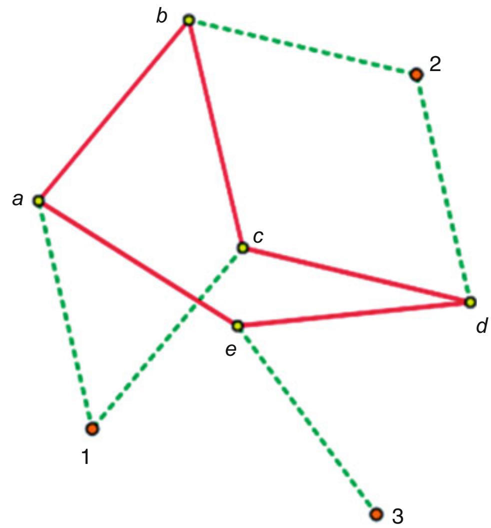


Fig. 16.1 Attaching a 5-cycle $[a, b, c, d, e]$ to the independent set $I = \{1, 2, 3\}$

The operation described above is very useful because it can increase the chromatic number of a graph from 3 to 4 without introducing any 3-cycles (triangles). We present the argument below.

Suppose that the initial graph G has k independent sets of size 3. To each one of these sets, we attach a 5-cycle as described above. We obtain a new unit distance graph H, which has $n + 5\,k$ vertices and $e + 10\,k$ edges. We claim that this graph is triangle-free and its chromatic number is at least 4.

The fact that H is triangle-free follows from the way that the 5-cycles are attached. For the sake of contradiction, assume that H can be properly colored with three colors. Since $n \geq 7$, it follows that there exist three vertices of G that are assigned the same color. These three vertices clearly form an independent set, say $I = \{1, 2, 3\}$. There is a 5-cycle $[a, b, c, d, e]$ attached to this independent set. Since vertices 1, 2, and 3 receive the same color, say red, none of the new vertices a, b, c, d, e can be colored red. It follows that each of a, b, c, d, e can only be assigned one of the remaining two colors. But this is impossible, since a 5-cycle cannot be properly colored with only two colors. So, H is 4-chromatic and triangle-free as desired.

The problem is that k may be large, and, therefore, H would have many vertices. But maybe we do not have to attach a 5-cycle to *every* independent set of size 3! The crucial idea of our approach is summarized in the two paragraphs below.

(1) Let G be a triangle-free, 3-chromatic unit distance graph. For a given proper 3-coloring of the vertices, and a given independent set I, we say that I is *monochromatic* if all vertices of I receive the same color.
(2) Let $\mathcal{I}$ be a collection of independent sets of size 3 such that for every proper 3-coloring of G, there exists a set $I \in \mathcal{I}$, which is monochromatic. It is then sufficient to attach 5-cycles *only* to the independent sets from $\mathcal{I}$, and the resulting graph will still be 4-chromatic.

16.3 A Triangle-Free, 4-Chromatic Unit Distance Graph on 21 Vertices

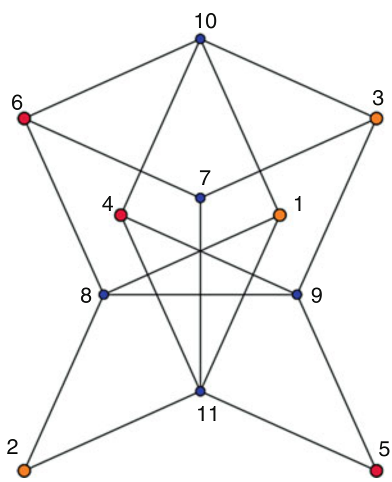
We are going to construct such a graph in two stages. First, we present a triangle-free, 3-chromatic unit distance graph G of order 11 and size 18, on the vertex set $\{1,2, \ldots, 11\}$, with the additional property that both $\{1, 2, 3\}$ and $\{4, 5, 6\}$ are independent sets in G and at least one of them is monochromatic under any proper 3-coloring of the vertices of G. Second, we will attach unit 5-cycles to the vertices of each of these independent sets, thus obtaining a graph of order 21 and size 38. This graph is still triangle-free and has chromatic number 4. Indeed, if it were 3-colorable, it would imply that the vertices of a unit 5-cycle could be properly 2-colored.

Lemma 16.1 Consider the graph G with vertex set $\{1, 2, 3, \ldots, 9, 10, 11\}$ and edge set:

$\{\{1, 8\}, \{1, 10\}, \{1, 11\}, \{2, 8\}, \{2, 11\}, \{3, 7\}, \{3, 9\}, \{3, 10\}, \{4, 9\}$ and
$\{4, 10\}, \{4, 11\}, \{5, 9\}, \{5, 11\}, \{6, 7\}, \{6, 8\}, \{6, 10\}, \{7, 11\}, \{8, 9\}\}$.

G is triangle-free and has chromatic number 3. Moreover, G has a faithful unit distance embedding, and, for every proper 3-coloring of the vertices, at least one of the independent sets $\{1, 2, 3\}$ or $\{4, 5, 6\}$ is monochromatic.

Proof The first two properties are straightforward to check. A faithful unit distance embedding is provided below in Fig. 16.2.



$$z_1 = \left((-1+\sqrt{7})/4, (+1+\sqrt{7})/4\right)$$
$$z_2 = \left((-1-\sqrt{7})/4, (+1-\sqrt{7})/4\right)$$
$$z_3 = \left((+1+\sqrt{7})/4, (+3+\sqrt{7})/4\right)$$
$$z_4 = \left((+1-\sqrt{7})/4, (+1+\sqrt{7})/4\right)$$
$$z_5 = \left((+1+\sqrt{7})/4, (+1-\sqrt{7})/4\right)$$
$$z_6 = \left((-1-\sqrt{7})/4, (+3+\sqrt{7})/4\right)$$
$$z_7 = (0,1),\ z_8 = (-1/2, 1/2),\ z_9 = (1/2, 1/2)$$
$$z_{10} = \left(0, (1+\sqrt{7})/2\right),\ z_{11} = (0,0)$$

Fig. 16.2 A unit distance embedding of the graph in Lemma 16.1.

A simple computer program shows that there are exactly 16 different proper 3-colorings.

{{**1**, **2**, **3**, 4, 5, 6}, {7, 8, 10}, {9, 11}}, {{**1**, **2**, **3**, 4, 5, 6}, {7, 8}, {9, 10, 11}}
{{**1**, **2**, **3**, 4, 5, 6}, {8, 10, 11}, {7, 9}}, {{**1**, **2**, **3**, 4, 5, 6}, {8, 11}, {7, 9, 10}}
{{**1**, **2**, **3**, 4, 5}, {7, 8, 10}, {6, 9, 11}}, {{**1**, **2**, **3**, 4, 6}, {5, 7, 8, 10}, {9, 11}}
{{**1**, **2**, **3**, 4, 6}, {5, 7, 8}, {9, 10, 11}}, {{**1**, **2**, **3**, 4}, {5, 7, 8, 10}, {6, 9, 11}}
{{**1**, **2**, **3**, 5, 6}, {4, 7, 8}, {9, 10, 11}}, {{**1**, **2**, **3**, 6}, {4, 5, 7, 8}, {9, 10, 11}}
{{1, 2, **4**, **5**, **6**}, {3, 8, 11}, {7, 9, 10}}, {{1, 2, 7, 9}, {8, 10, 11}, {3, **4**, **5**, **6**}}
{{1, 3, **4**, **5**, **6**}, {8, 10, 11}, {2, 7, 9}}, {{1, 3, **4**, **5**, **6**}, {8, 11}, {2, 7, 9, 10}}
{{1, **4**, **5**, **6**}, {3, 8, 11}, {2, 7, 9, 10}}, {{1, 7, 9}, {8, 10, 11}, {2, 3, **4**, **5**, **6**}}.

Note that for each of the 16 3-colorings above, either the vertices 1, 2, and 3 fall in the same color class or the vertices 4, 5, and 6 have this property. This proves Lemma 16.1. ■

Next, we show how to attach a unit 5-cycle to the independent set {1, 2, 3}. Note that since {4, 5, 6} is the reflection of {1, 2, 3} across the y-axis, the same construction is going to work for {4, 5, 6} as well.

Lemma 16.2 Attaching a unit 5-cycle to the independent set {1, 2, 3}. Consider the points

$$z_1 = \left(\left(-1+\sqrt{7}\right)/4,\ \left(1+\sqrt{7}\right)/4\right)$$
$$z_2 = \left(\left(-1-\sqrt{7}\right)/4,\ \left(1-\sqrt{7}\right)/4\right),\ \text{and}$$
$$z_3 = \left(\left(1+\sqrt{7}\right)/4,\ \left(3+\sqrt{7}\right)/4\right).$$

Then there exist five points which we label 12, 13, 14, 15, and 16 such that

$$\begin{aligned}\|z_1 - z_{12}\| &= \|z_2 - z_{13}\| = \|z_1 - z_{14}\| = \|z_2 - z_{15}\| = \|z_3 - z_{16}\| = \\ &= \|z_{12} - z_{13}\| = \|z_{13} - z_{14}\| = \|z_{14} - z_{15}\| = \|z_{15} - z_{16}\| = \|z_{16} - z_{12}\| = 1.\end{aligned}$$

Proof Since $\|z_1 - z_{12}\| = \|z_1 - z_{14}\| = 1$ we set $z_{12} = z_1 + \left(\frac{1-a^2}{1+a^2},\ \frac{2a}{1+a^2}\right)$ and $z_{14} = z_1 + \left(\frac{1-b^2}{1+b^2},\ \frac{2b}{1+b^2}\right)$ for some reals a, b to be found later.

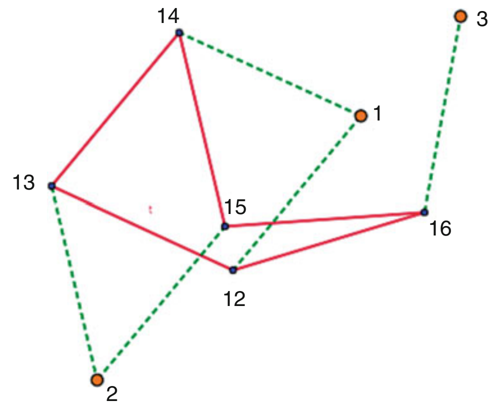


Fig. 16.3 Attaching the unit 5-cycle [12, 13, 14, 15, 16] (solid edges) to the monochromatic independent set {1, 2, 3}. The dashed segments are the edges joining the vertices of the independent set to the 5-cycle

Similarly, since $\|z_3 - z_{16}\| = 1$ set $z_{16} = z_3 + \left(\frac{1-c^2}{1+c^2}, \frac{2c}{1+c^2}\right)$ for some real c.

Since [1, 12, 13, 14] and [2, 13, 14, 15] are both rhombi, it follows that $z_{13} = z_{12} + z_{14} - z_1$ and $z_{15} = z_2 + z_{14} - z_{13} = z_1 + z_2 - z_{12}$.

We need to make sure that $\|z_{14} - z_{15}\| = \|z_{15} - z_{16}\| = \|z_{16} - z_{12}\| = 1$. Straightforward calculations show that these conditions are equivalent to the following system in the variables a, b, and c:

$$\begin{cases} (13-4\sqrt{7})a^2b^2 + 4\sqrt{7}(a+b)(ab+1) + 5(a^2+b^2) + 16ab + 13 + 4\sqrt{7} = 0 \\ (5-\sqrt{7})a^2c^2 + (2\sqrt{7}+2)(a+c)(1+ac) + (3+\sqrt{7})(a^2+c^2) + 8ac + 9 + 3\sqrt{7} = 0 \\ a^2c^2 - 4a^2c + 4ac^2 - 11a^2 + 16ac - 3c^2 + 4a - 4c + 1 = 0. \end{cases}$$

The last two equations depend on a and c only, and it can be checked that there are four real solutions. Using resultants, we can compute the minimal polynomials of a and c; they both have degree 6:

$$\begin{aligned} \text{minpoly } a = 16a^6 + \left(261\sqrt{7}+651\right)a^5 + \left(1233\sqrt{7}+3303\right)a^4 + \left(1538\sqrt{7}+4046\right)a^3 - \\ - \left(318\sqrt{7}+834\right)a^2 - \left(579\sqrt{7}+1533\right)a + \left(305\sqrt{7}+807\right) \\ \text{minpoly } c = 8c^6 + \left(27\sqrt{7}+75\right)c^5 + \left(159\sqrt{7}+399\right)c^4 + \left(414\sqrt{7}+1038\right)c^3 + \\ + \left(606\sqrt{7}+1638\right)c^2 + \left(579\sqrt{7}+1539\right)c + \left(255\sqrt{7}+671\right). \end{aligned}$$

Each real root a of the first polynomial above (and there are four of them) can then be used to solve the first equation of the system, which is a quadratic in b. The minimal polynomial of b can be computed as well, but, since it has degree 12 and rather large coefficients, we omit it. We obtain a total of eight real solutions (a, b, c) of the original system. In the end, we select

$$a = -2.1426706655\ldots, \quad b = 4.5753087820\ldots, \quad c = -1.2036751317\ldots,$$

which produces Fig. 16.3. This completes the proof of Lemma 16.2. ■

All that is left to do is to now attach the unit 5-cycles to the independent sets {1,2,3} and {4,5,6}. Recall that these sets are reflections of one another across the y-axis, and, hence, no additional work is needed. The result is a triangle-free, 4-chromatic unit distance graph of order 11 + 2 • 5 = 21 and size 18 + 2 • 10 = 38 whose embedding is shown in Fig. 16.4.

For the reader interested in checking our construction numerically, we provide the approximate coordinates of the vertices 12 through 16 below. Again, vertices 17 through 21 are obtained by reflecting these points with respect to the y-axis.

$$z_{12} = (-0.2308468159,\ 0.1449716273),\quad z_{13} = (-1.1396618925,\ 0.5621708083)$$
$$z_{14} = (-0.4973772488,\ 1.3286370087),\quad z_{15} = (-0.2691531840,\ 0.3550283726)$$
$$z_{16} = (+0.7281531072,\ 0.4283779561).$$

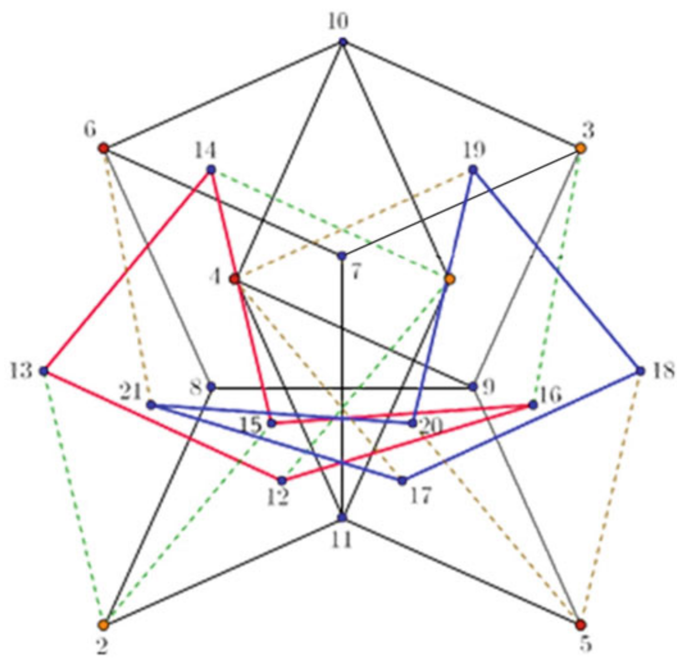


Fig. 16.4 A triangle-free, 4-chromatic unit distance graph of order 21. The edges of the base graph are the thin segments. The two unit 5-cycles are drawn in thick solid lines. The edges connecting the independent sets {1, 2, 3} and {4, 5, 6} to their respective cycles are shown in dashed segments

16.4 A Triangle-Free, 4-Chromatic Unit Distance Graph on 19 Vertices

The crucial property of the "core" graph given in Lemma 16.1 is the existence of the two independent sets, at least one of which is monochromatic under any proper 3-coloring.

Naturally, one may ask whether there exist small, triangle-free, 3-chromatic unit distance graphs that contain *one* independent set of three vertices, which is monochromatic under *any* proper 3-coloring. If such a graph were to exist, we would only need to attach one unit 5-cycle to the vertices of the independent set to create a triangle-free, 4-chromatic unit distance graph. It turns out that such graphs exist with as few as 14 vertices. It follows that we can further improve the record and obtain triangle-free, 4-chromatic unit distance graphs with only $14 + 5 = 19$ vertices. The details are presented below.

We are going to construct such a graph in two stages. We start with a triangle-free, 3-chromatic unit distance graph G of order 14 and size 25 with the additional property that $\{1,2,3\}$ is an independent set that is monochromatic under any proper 3-coloring of the vertices of G. Next, we will attach a unit 5-cycle to the vertices of this independent set, thus obtaining a graph of order 19 and size 35. This graph is still triangle-free and has chromatic number 4. Indeed, if it were 3-colorable, it would imply that the vertices of the unit 5-cycle can be properly 2-colored, a contradiction.

Lemma 16.5 Consider the graph G with vertex set $\{1, 2, 3, \ldots, 12, 13, 14\}$ and edge set:

$\{\{1, 4\}, \{1, 5\}, \{1, 10\}, \{2, 4\}, \{2, 5\}, \{3, 7\}, \{3, 8\}, \{3, 10\}, \{4, 7\},$
$\{4, 11\}, \{4, 12\}, \{5, 6\}, \{5, 8\}, \{5, 13\}, \{6, 9\}, \{6, 12\}, \{7, 9\}, \{7, 13\},$
$\{8, 11\}, \{8, 14\}, \{9, 14\}, \{10, 11\}, \{10, 13\}, \{12, 13\}, \{13, 14\}\}.$

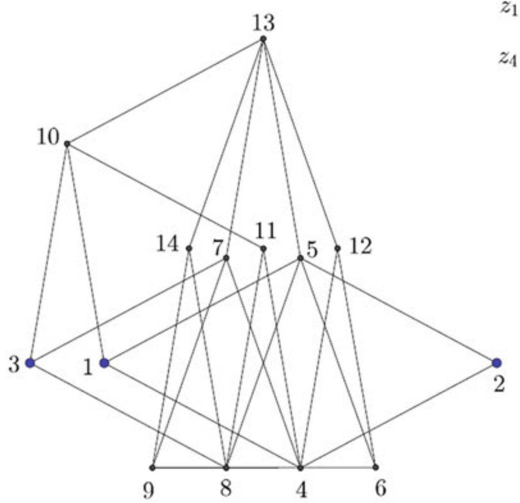
G is triangle-free and has chromatic number 3. Moreover, G has a faithful unit distance embedding, and, for every proper 3-coloring of the vertices, the independent set $\{1, 2, 3\}$ is monochromatic.

Proof A simple computer program shows that there are exactly 10 different proper 3-colorings.

$\{\{\mathbf{1}, \mathbf{2}, \mathbf{3}, 6, 11, 13\}, \{4, 8, 9, 10\}, \{5, 7, 12, 14\}\}$
$\{\{\mathbf{1}, \mathbf{2}, \mathbf{3}, 6, 13\}, \{4, 8, 9, 10\}, \{5, 7, 11, 12, 14\}\}$
$\{\{\mathbf{1}, \mathbf{2}, \mathbf{3}, 6, 11, 13\}, \{4, 8, 9\}, \{5, 7, 10, 12, 14\}\}$
$\{\{\mathbf{1}, \mathbf{2}, \mathbf{3}, 6, 11, 14\}, \{4, 8, 9, 13\}, \{5, 7, 10, 12\}\}$
$\{\{\mathbf{1}, \mathbf{2}, \mathbf{3}, 6, 11\}, \{4, 8, 9, 13\}, \{5, 7, 10, 12, 14\}\}$
$\{\{\mathbf{1}, \mathbf{2}, \mathbf{3}, 9, 11, 13\}, \{4, 6, 8, 10\}, \{5, 7, 12, 14\}\}$
$\{\{\mathbf{1}, \mathbf{2}, \mathbf{3}, 9, 13\}, \{4, 6, 8, 10\}, \{5, 7, 11, 12, 14\}\}$
$\{\{\mathbf{1}, \mathbf{2}, \mathbf{3}, 9, 11, 12\}, \{4, 6, 8, 13\}, \{5, 7, 10, 14\}\}$
$\{\{\mathbf{1}, \mathbf{2}, \mathbf{3}, 9, 11, 13\}, \{4, 6, 8\}, \{5, 7, 10, 12, 14\}\}$
$\{\{\mathbf{1}, \mathbf{2}, \mathbf{3}, 9, 11\}, \{4, 6, 8, 13\}, \{5, 7, 10, 12, 14\}\}.$

Note that the vertices 1, 2, and 3 fall in the same color class for any of the 10 3-colorings above.

It is easy to verify that G is triangle-free and 3-chromatic. A faithful unit distance embedding is provided in Fig. 16.5 below. Note that vertices 4, 6, 8, and 9 are collinear in this embedding and this is why the edge $\{6,9\}$ passes through vertices 4 and 8.



$$z_1=(0,0)\,,\ z_2=\left(2\sqrt{7}/3,0\right),\ z_3=(-1/3,0)$$
$$z_4=\left(\sqrt{7}/3,-\sqrt{2}/3\right),\ z_5=\left(\sqrt{7}/3,\sqrt{2}/3\right)$$
$$z_6=\left((+1+\sqrt{7})/3,-\sqrt{2}/3\right)$$
$$z_7=\left((-1+\sqrt{7})/3,+\sqrt{2}/3\right)$$
$$z_8=\left((-1+\sqrt{7})/3,-\sqrt{2}/3\right)$$
$$z_9=\left((-2+\sqrt{7})/3,-\sqrt{2}/3\right)$$
$$z_{10}=\left(-1/6,\sqrt{35}/6\right)$$
$$z_{11}=\left((-1+2\sqrt{7})/6,(\sqrt{35}-2\sqrt{2})/6\right)$$
$$z_{12}=\left((+1+2\sqrt{7})/6,(\sqrt{35}-2\sqrt{2})/6\right)$$
$$z_{13}=\left((-1+2\sqrt{7})/6,(\sqrt{35}+2\sqrt{2})/6\right)$$
$$z_{14}=\left((-3+2\sqrt{7})/6,(\sqrt{35}-2\sqrt{2})/6\right)$$

Fig. 16.5 A unit distance embedding of the graph in Lemma 16.5.

Next, we show how to attach a unit 5-cycle to the independent set {1, 2, 3} above.

Lemma 16.6 Attaching a unit 5-cycle to the independent set {1, 2, 3}.

Consider the points $z_1=(0,0)$, $z_2=\left(2\sqrt{7}/3,0\right)$, and $z_3=(-1/3,0)$. Then, there exist five points, which we denote as 15, 16, 17, 18, and 19 such that $\|z_1-z_{15}\|=\|z_2-z_{16}\|=\|z_1-z_{17}\|=\|z_2-z_{18}\|=\|z_3-z_{19}\|==\|z_{15}-z_{16}\|=\|z_{16}-z_{17}\|=\|z_{17}-z_{18}\|=\|z_{18}-z_{19}\|=\|z_{19}-z_{15}\|=1$.

Proof Since $\|z_1-z_{15}\|=\|z_1-z_{17}\|=1$ we set $z_{15}=z_1+\left(\frac{1-a^2}{1+a^2},\frac{2a}{1+a^2}\right)$

$$z_{17}=z_1+\left(\frac{1-b^2}{1+b^2},\frac{2b}{1+b^2}\right)$$

for some reals a, b to be found later.

Similarly, since $\|z_3-z_{19}\|=1$ set $z_{19}=z_3+\left(\frac{1-c^2}{1+c^2},\frac{2c}{1+c^2}\right)$ for some real c.

Since [1, 15, 16, 17] and [2, 16, 17, 18] are both rhombi, it follows that $z_{16}=z_{15}+z_{17}-z_1$ and $z_{18}=z_2+z_{17}-z_{16}=z_1+z_2-z_{15}$.

We need to make sure that $\|z_{17}-z_{18}\|=\|z_{18}-z_{19}\|=\|z_{19}-z_{15}\|=1$.

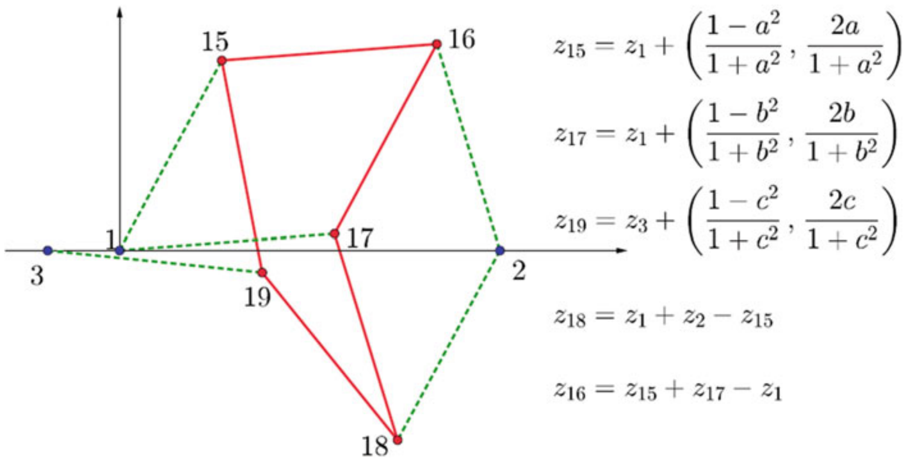


Fig. 16.6 Attaching the unit 5-cycle [15, 16, 17, 18, 19] (solid edges) to the monochromatic independent set {1, 2, 3}. The dashed segments are the edges joining the vertices of the independent set to the vertices of the 5-cycle

Straightforward calculations show that these conditions are equivalent to the following system in the variables a, b, and c:

$$\begin{cases} (55+24\sqrt{7})a^2b^2 + 19(a^2+b^2) + 72ab + 55 - 24\sqrt{7} = 0 \\ 3a^2c^2 + (\sqrt{7}-2)(a^2+c^2) + (7\sqrt{7}-17)ac + 9\sqrt{7} - 24 = 0 \\ a^2c^2 - 2a^2 + 9ac - 5c^2 + 1 = 0. \end{cases}$$

The last two equations depend on a and c only, and the real solutions are of the form $(a, c) = (\pm 0.5970459772\ldots, \pm 0.0511624986\ldots)$. Using resultants, it is easy to obtain the minimal polynomials of a and c; they are both of degree 6.

$$\text{minpoly } a = a^6 + \left(738\sqrt{7} - 1950\right)a^4 + \left(2523 - 954\sqrt{7}\right)a^2 + \left(16192 - 6120\sqrt{7}\right)$$
$$\text{minpoly } c = 18c^6 + \left(482\sqrt{7} - 1271\right)c^4 + \left(1805 - 683\sqrt{7}\right)c^2 + \left(844 - 319\sqrt{7}\right).$$

For each of these four solutions, we obtain two real values for b using the first equation of the system. In total, discounting symmetries, there are four different ways that the unit 5-cycle can be attached.

We used the values $a = 0.5970459772\ldots$, $b = 0.0380685953\ldots$, and $c = -0.0511624986\ldots$, which produced the result in Fig. 16.6. Now, overlaying Figs. 16.5 and 16.6, we obtain the desired triangle-free, 4-chromatic unit distance graph of order 19 – see Fig. 16.7 below.

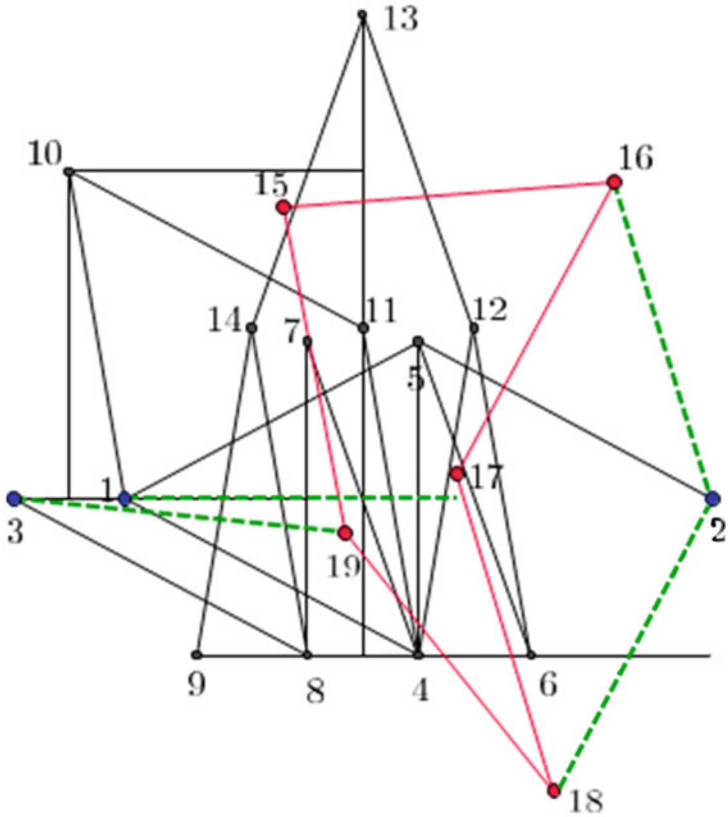


Fig. 16.7 A triangle-free, 4-chromatic unit distance graph of size 19

16.5 A Triangle-Free, 4-Chromatic Unit Distance Graph on 17 Vertices

At this point, it is natural to ask how close are we to finding the triangle-free, 4-chromatic unit distance graph of the smallest possible order? We already know that 19 vertices is optimal if we are to use the 5-cycle attaching approach. We searched for graphs of order n that satisfy the following properties:

- 4-Chromatic and *edge-critical*, that is, removal of any edge produces a graph, which is 3-colorable
- Triangle-free and contains no forbidden subgraph of order up to 7 inclusive – see [PP] for a list of such graphs.

We then look at these graphs and decide whether they are unit distance graphs or not. We found no such graphs with $n \leq 15$. For $n = 16$, there is exactly one unit distance graph, which unfortunately is not a *faithful* unit distance graph since every embedding forces two accidental edges and these additional edges lead to a graph, which contains triangles. However, we found one good graph of order 17.

Theorem 16.7 Consider the graph G with vertex set {1, 2, 3, ..., 16, 17} and edge set:

{{1, 8}, {1, 12}, {1, 15}, {2, 9}, {2, 10}, {2, 16}, {3, 9}, {3, 15}, {3, 16}, {3, 17}, {4, 10}, {4, 11}, {4, 16}, {5, 11}, {5, 13}, {5, 16}, {6, 12}, {6, 13}, {6, 14}, {7, 12}, {7, 14}, {7, 15}, {8, 13}, {8, 14}, {8, 17}, {9, 11}, {9, 13}, and {10, 15}, {10, 17}, {11, 14}, {12, 17}}.

G is a triangle-free, 4-chromatic, faithful unit distance graph.

Proof It can be verified that G is triangle-free, has chromatic number 4, and is edge-critical (removal of any edge results in a graph that is 3-colorable). We want to find a unit distance embedding of G.

Start by setting $z_{17} = [0, 0]$, $z_3 = [1, 0]$, and

$$z_8 = z_{17} + \left[\frac{a^2-1}{a^2+1}, \frac{2a}{a^2+1}\right], z_{10} = z_{17} + \left[\frac{b^2-1}{b^2+1}, \frac{2b}{b^2+1}\right], z_9 = z_3 + \left[\frac{c^2-1}{c^2+1}, \frac{2c}{c^2+1}\right],$$
$$z_{14} = z_8 + \left[\frac{d^2-1}{d^2+1}, \frac{2d}{d^2+1}\right], z_{16} = z_3 + \left[\frac{e^2-1}{e^2+1}, \frac{2e}{e^2+1}\right], z_{11} = z_9 + \left[\frac{f^2-1}{f^2+1}, \frac{2f}{f^2+1}\right],$$

where a, b, c, d, e, and f are the real parameters to be determined later.

The graph contains nine 4-cycles:

[1, 8, 17, 12], [1, 12, 7, 15], [2, 9, 3, 16], [2, 10, 4, 16], [3, 15, 10, 17], [4, 11, 5, 16], [5, 11, 9, 13], [6, 12, 7, 14], [6, 13, 8, 14];

Hence, $z_1 - z_8 + z_{17} + z_{12} = 0$, $z_1 - z_{12} + z_7 - z_{15} = 0, \ldots$

Solving for z_1, z_2, z_7, z_9, z_{11}, z_{13}, z_{14}, z_{15}, z_{16}, we obtain that

$$\begin{aligned}
z_1 &= z_8 + z_{12} - z_{17}, \\
z_2 &= -2z_3 + z_4 + 2z_8 - z_{10} + z_{12}, z_7 = z_3 - z_8 + z_{10}, z_9 = z_3 - z_4 + z_{10}, \\
z_{11} &= 2z_3 - z_4 + z_5 - 2z_8 + 2z_{10} - z_{12}, z_{13} = -z_3 + 2z_8 - z_{10} + z_{12}, \\
z_{14} &= z_3 + z_6 + z_{10} - z_8 - z_{12}, z_{15} = z_3 + z_{10} - z_{17}, \\
z_{16} &= -2z_3 + 2z_4 + 2z_8 - 2z_{10} + z_{12}.
\end{aligned}$$

At this point, all z_i, $1 \leq i \leq 17$ can be expressed in terms of a, b, c, d, e, and f. The conditions $\|z_i - z_j\| = 1$ for every $\{i, j\} \in E$ are equivalent to the following polynomial system of six equations and six unknowns:

$$\begin{cases} a^2be - 2ab^2e - 2abe^2 + 3b^2e^2 + a^2 - 2ab - 2ae + b^2 + be + e^2 = 0. \\ b^2d^2f^2 + 2b^2df + bd^2f + 2bdf^2 + b^2 + 2bd + bf + 2df + f^2 + 3 = 0. \\ a^2b^2c^2 - 3a^2b^2 + 8a^2bc - 3a^2c^2 - 8ab^2c + 8abc^2 - 3b^2c^2 + \\ +a^2 + 8ab - 8ac - 15b^2 + 8bc + c^2 - 3 = 0. \\ 3a^2c^2d^2 + 3a^2c^2 + 16a^2cd - a^2d^2 + 32ac^2d + 16acd^2 + 3c^2d^2 + \\ +15a^2 + 16ac + 32ad + 35c^2 + 16cd + 15d^2 + 63 = 0. \\ a^2d^2f^2 - a^2d^2ef + a^2de^2f - a^2def^2 + ad^2e^2f - ad^2ef^2 + ade^2f^2 - a^2de + \\ +a^2df + a^2e^2 - a^2ef - ad^2e + ad^2f + ade^2 + adf^2 + ae^2f - aef^2 + d^2e^2 - \\ -d^2ef + de^2f - def^2 + e^2f^2 + ad - ae + af - de + df + 3e^2 - ef + 1 = 0. \\ abcd(ab - ac - ad + 2bc + 2bd - 2cd) + a^2c^2d^2 + b^2c^2d^2 + a^2b^2 - a^2bc - a^2bd + \\ +a^2cd + 2ab^2c + 2ab^2d - 2abc^2 - 2abd^2 + 2ac^2d + 2acd^2 + 3b^2c^2 + b^2cd + \\ +3b^2d^2 - bc^2d - bcd^2 - 2ab + 2ac + 2ad + 6b^2 - bc - bd + c^2 + cd + d^2 + 3 = 0. \end{cases}$$

This system has 48 real solutions (a, b, c, d, e, f), which translates into 12 different embeddings discounting symmetries. The particular solution provided below generates the embedding in Fig. 16.8.

$$a = 2.431002435387093, b = 14.100395866873264, c = 5.1601979612829399,$$
$$d = -2.723692158669673, e = 1.478517551159578, f = 0.171178647198113.$$

The minimal polynomials of a, b, c, d, e, and f have degree 128, too large to be written here.

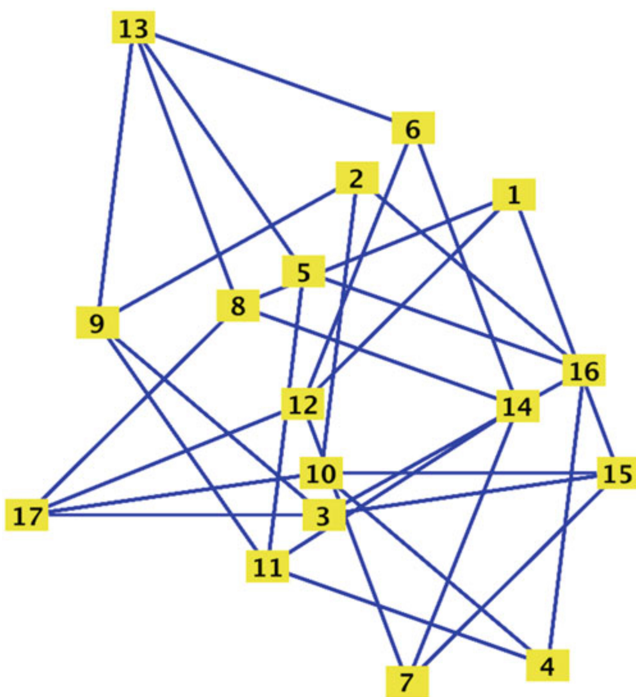


Fig. 16.8 The Exoo–Ismailescu record graph a triangle-free, 4-chromatic unit distance graph of order 17

The Exoo–Ismailescu graph decorates the front cover of the October 2016 issue of *Geombinatorics.* ■

On March 26, 2023, Geoffrey Exoo answered my request for an insight into their work.

Note on 21-19-17 Graphs

The overall structure of our search for triangle-free unit distance graphs works as follows. First we generate a (large) list of candidate graphs that we know are 4-chromatic and which do not contain any of the known forbidden subgraphs. Then we eliminate the candidates individually.

The first step in the process involves generating two lists of forbidden subgraphs. The first list contains those graphs that will be checked during the candidate generation phase of the process. These are graphs that can be checked very quickly. In this case only three graphs are available, K_3, $K_{2,3}$, and the Möbius ladder of order 8. K_3 is needed to guarantee the graphs are triangle free. The other two are the only triangle-free, edge-minimal, forbidden subgraphs of order less than 10[2].

The second list of forbidden subgraphs is created by running the procedure outlined here on smaller orders (up to 15), a procedure which found other triangle free 4-chromatic graphs that were not unit distance graphs. Checking whether these graphs are unit distance takes too much time to be used in the graph generation phase, but is useful later in the process to eliminate individual candidates.

The third step involves configuring a fast randomized graph coloring function that attempts to 3-color graphs very quickly. This function can be tuned to successfully 3-color well over 99 percent of graphs of order 16 (or less) that can be so colored, and it does so in approximately one percent of the time that an exhaustive search would take.

The fourth step is to generate all edge-minimal graphs of a given order (e.g., 16) that do not contain the forbidden subgraphs on the small list and which were not successfully 3-colored by the randomized coloring function. The graphs are generated in two different ways. First by a modified version of the graph generating program in Brendan McKay's nauty package. Second, by my own program, which is less efficient than nauty in general, but is designed specifically for the unit-distance problem. However, since there are so few forbidden subgraphs, nauty was faster for this problem. The lists generated by the two programs have been compared.

The fifth step is to run exhaustive searches for 3-colorings on each of the graphs output in the previous step. This eliminates a few of the candidates that the randomized coloring function missed.

The sixth step uses the large list of forbidden subgraphs (of orders up to 15) to eliminate a few more graphs.

I congratulate Geoffrey Exoo and Dan Ismailescu for their commendable achievement of completely solving my Problem 15.4 and thank them for sharing with us their very insightful presentation of the 21-19-17 progress. They will resurface later in this book.

[2] This can be checked by looking at the Globus–Parshall list, where the Möbius ladder is identified as graph $F(8, 12, 3)$. We independently verified the list, at least for triangle-free graphs, many years ago.

Chapter 17
Edge Chromatic Number of a Graph

17.1 Vizing's Edge Chromatic Number Theorem

We can assign a color to each edge of a graph instead of its vertices. This gives birth to the following notion:

A graph G is called *n-edge-colorable* if we can assign one of the n colors to each edge of G in such a way that the adjacent edges are colored differently.

The *edge chromatic number* $\chi_1(G)$ (also known as the *chromatic index*) of a graph G is the smallest number n of colors for which G is n-edge-colorable.

The following two statements follow straight from the definitions:

Problem 17.1[1] For any graph G,

$$\chi_1(G) \geq \Delta(G).$$

Problem 17.2 For any subgraph G_1 of a graph G,

$$\chi(G_1) \leq \chi(G).$$

In 1964, the Russian mathematician Vadim Georgievich Vizing published [Viz1] a wonderful result about the edge chromatic number of a graph. His proof is fairly long, but so nice that I am going to present it here in its entirety. *Do read it with pencil and paper!*

The Vizing Theorem 17.3 (V.G. Vizing, [Viz1]). If G is a non-empty graph, then

$$\chi_1(G) \leq \Delta(G) + 1 \qquad (^*)$$

i.e., the edge chromatic number $\chi_1(G)$ of a graph is always equal to Δ or $\Delta + 1$, where $\Delta = \Delta(G)$.

[1] $\Delta(G)$ is defined in Chapter 8.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_17

Proof I enjoyed a version of Vizing's proof in [BCL]. My presentation is based on it. I did make it more visual by including many illustrations, splitting one case into two, and adding a number of elucidations.

Part I. Preparation for the Assault We will argue by contradiction. Assume that the inequality (*) is not true. Then among the graphs for which (*) is not true, let G be a graph with the *smallest* number of edges. In other words, G is not $(\Delta + 1)$-edge-colorable; but the graph G' obtained from G by removing one edge e, is $[\Delta(G') + 1]$-edge-colorable. Since obviously $\Delta(G') \leq \Delta(G)$, the graph G' is $(\Delta + 1)$-edge-colorable.

Let G' be actually edge-colored in $\Delta + 1$ colors, i.e., every edge of the graph G, except $e = uv$ (this equality simply denotes that the edge e connects vertices u and v), is colored in one of the $\Delta + 1$ colors in such a way that the adjacent edges are colored differently. For each edge $e' = uv'$ of G that is incident with u (including e), we define its *dual color* as any one of the $\Delta + 1$ colors that is not used to color the edges incident with vertex v'. (Since the degree of any v' does not exceed Δ, we always have at least one color to choose as dual. It may so happen that distinct edges have the same dual color – it is all right.)

We are going to construct a sequence of distinct edges $e_0, e_1, \ldots, e_k$ that are all incident with u as follows (see Fig. 17.1). Let $e = e_0$ have the dual color α_1 (i.e., α_1 is not the color of any edge of G incident with v). There must be an edge, call it e_1, of color α_1 incident with u (for if not, then the edge e could be colored α_1, thus producing a $(\Delta + 1)$-edge coloring of G). Let α_2 be the dual color of e_1. If there is an edge of color α_2 incident with u and distinct from e_0 and e_1, we denote it by e_2 and its dual color by α_3, etc. We constructed a *maximal* (i.e., as long as possible) sequence $e_0, e_1, \ldots, e_k,$ $k \geq 1$, of *distinct* edges. The last edge e_k by construction is colored α_k and has the dual color α_{k+1}.

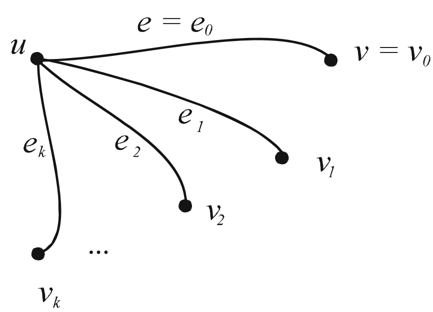


Fig. 17.1

If there were no edge of color α_{k+1} incident with u, then we would recolor each edge of our sequence $e_0, e_1, \ldots, e_k$ in its dual color and, thus, achieve a $(\Delta + 1)$-edge coloring of G (do verify that). This contradicts our initial assumption.

Therefore, there is an edge e_{k+1} of color α_{k+1} incident with u; but since we have constructed the longest sequence of distinct edges $e_0, e_1, \ldots, e_k$, the edge e_{k+1} must coincide with one of them: $e_{k+1} = e_i$ for some i, $1 \leq i \leq k$. Since the edges coincide, so do their colors: $\alpha_{k+1} = \alpha_i$. The color α_k of the edge e_k may not be the same as the dual color α_{k+1} of e_k: $\alpha_{k+1} \neq \alpha_k$. Thus, we get $\alpha_{k+1} = \alpha_i$ for some i, $1 \leq i < k$. Denote $t = i - 1$; then, the last equality can be written as follows:

$$a_{k+1} = a_{t+1}$$

for some t, $0 \leq t \leq k - 1$. Finally, this means that *the edges* e_k *and* e_t *have the same dual color*.

Now, the last preparatory remarks.

(a) For each color α among the $\Delta + 1$ colors, there is an edge of color α adjacent to the edge $e = uv$ (for if not, e could be colored α, thus producing $(\Delta + 1)$-edge coloring of G). But since there are at most Δ edges incident with the vertex u, there is a color, call it β, assigned to an edge incident with the vertex v that is not assigned to any edge incident with u.

(b) The color β must be assigned to at least one edge incident with the vertex v_i for each $i = 1, 2, \ldots, k$ (see Fig. 17.1). Indeed, if we assume that there is a vertex v_m, $1 \leq m \leq k$, such that no edge incident with v_m is colored β, then we can change the color of e_m to β and also change the color of each e_i, $0 \leq i \leq m$, to its dual color to obtain a $(\Delta + 1)$-edge coloring of G (verify this).

Part II. The Assault A sequence of edges $a_1, a_2, \ldots, a_n$ of a graph is called a *path of length n* if the consecutive edges of the sequence are adjacent (Fig. 17.2). You can trace a path with a pencil without taking it off the paper all the way from the *initial vertex of the path* v_0 to the *terminal vertex of the path* v_n. The edge a_1 is called the *initial edge*, while the edge a_n is the *terminal edge* of the path.

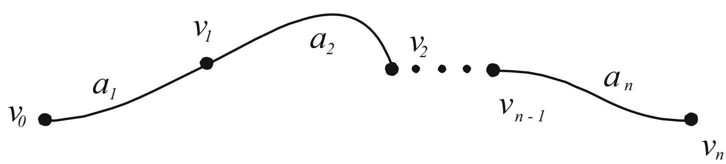


Fig. 17.2

Define two paths P and R as follows: Their initial vertices are v_k and v_t, respectively, and each of the paths has the *maximum possible length with edges alternately colored* β *and* $\alpha_{k+1} = \alpha_{t+1}$ (we established in Part I that colors α_{k+1} and α_{t+1} coincide). Denote the terminal vertices of the paths P and R by w and w', respectively, and consider five possibilities for w and w'.

Case 1 $w = v_m$ for some m, $0 \leq m \leq k - 1$ (Fig. 17.3).

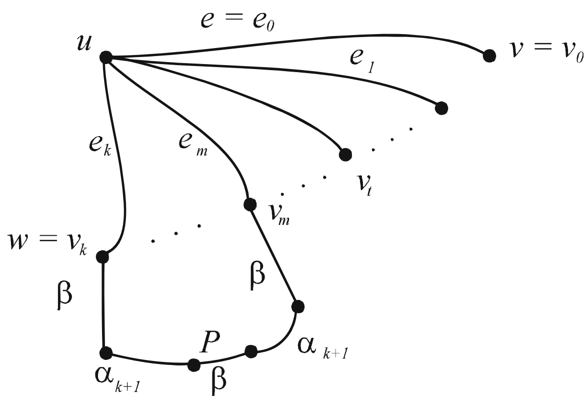


Fig. 17.3

Observe that the color $\alpha_{k\,+\,1}$ as the dual color of the edge e_k may not be adjacent to e_k; therefore, the initial edge of the path P must be colored β and $m \neq k$.

The terminal edge of P must be colored β as well. Indeed, if alternatively, the terminal edge of P were colored $\alpha_{k\,+\,1,}$ then we would be able to make P longer by adding one more edge incident with v_m and colored β (it exists as we noticed in (b) at the end of Part I of this proof).

Note that the vertex v_t is *not* on P unless $v_m = v_{t.}$ Indeed, assume that v_t is on P and $v_t \neq v_m$, then v_t is incident with the edges of P (Fig. 17.4). One of them must be colored $\alpha_{k\,+\,1}$ (and the other β), but the dual color of e_t is $\alpha_{t\,+\,1} = \alpha_{k\,+\,1}$; therefore, no edge of color $\alpha_{k\,+\,1}$ may be adjacent to e_t. This contradiction proves that v_t is not on P unless $v_t = v_m$.

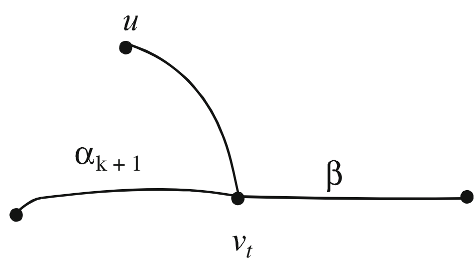


Fig. 17.4

We are ready to finish Case 1. Interchange the colors β and $\alpha_{k\,+\,1}$ on the edges of P. Please note (and prove) that as a result of this interchange, we do *not* alter the dual colors of edges e_i for any $i < m$ and end up with no edge of color β incident with v_m. Now to obtain a (Δ+1)-edge coloring of G, we just change the color of e_m to β and change the color of every e_i for $0 \leq i < m$ to its dual. (Do verify that we get a (Δ+1)-edge coloring of G.) We got a contradiction, for G is not (Δ+1)-edge-colorable.

Case 2 $w' = v_m$ for some m, $0 \leq m \leq k$ (Fig. 17.5).

Color $\alpha_{k+1} = \alpha_{t+1}$ as the dual color of the edge e_t may not be adjacent to e_t; therefore the initial edge of the path R must be colored β and $m \neq t$.

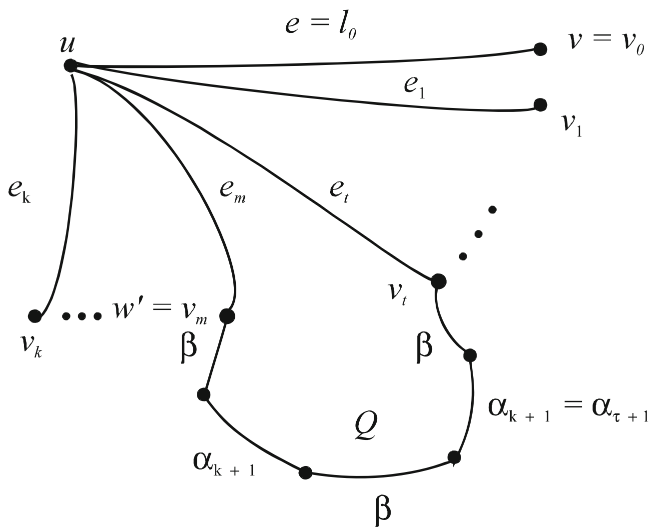


Fig. 17.5

The terminal edge of R must be colored β as well. Indeed, if alternatively, the terminal edge of R we colored α_{k+1}, then we would be able to make R longer by adding one more edge incident with v_m and colored β (it exists as we showed in (*b*) at the end of Part I of this proof).

The vertex v_k is *not* on R unless $v_m = v_k$ (the proof is identical to a relevant argument in Case 1 above). Now, we interchange the colors β and α_{k+1} of the edges of R. As a result of this interchange, we do *not* alter the dual colors of edges e_i for any $i \neq t$ and end up with no edge of color β incident with v_m.

If $m < t$, we finish as in Case 1. If $m > t$, then we change the color of e to β and change the color of every e_i, $0 \leq i < m$, to its dual. In either case, we get a (Δ+1)-edge coloring of G, which is a contradiction.

Case 3 $w \neq w_m$ for any m, $0 \leq m < k$ and $w \neq u$. As in Case 1, the initial edge of P must be colored β.

We interchange the colors β and α_{k+1} of the edges of P. As a result (just like in Case 1), we do not alter the dual colors of edges e_i for any $i < k$ and end up with no edge of color β incident with v_k. As in the previous cases, we can now obtain a (Δ+1)-edge coloring of G, a contradiction.

Case 4 $w' \neq v_m$ for any $m \neq t$ and $w' \neq u$. This case is similar to case 3 – consider it on your own.

Case 5 $w = w' = u$ (Figs. 17.6 and 17.7).

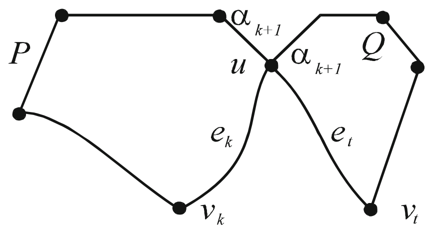


Fig. 17.6

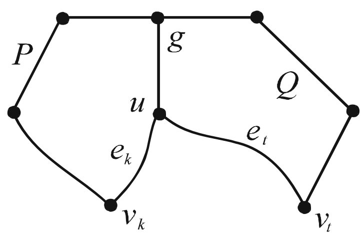


Fig. 17.7

Since by definition of β, u is incident with no edge-colored β, the terminal edge of both paths P and R is colored α_{k+1}.

If P and R have no edges in common (Fig. 17.6), then u is incident with two edges colored α_{k+1}, which cannot occur in edge coloring of a graph. But if P and R do have an edge in common, then there is a vertex (g in Fig. 17.7) incident with at least three edges of P and R. Since each of these three edges is colored β or α_{k+1}, two of them must be assigned the same color, which cannot occur with two adjacent edges of an edge-colored graph. In either case, we have obtained a contradiction. ■

This remarkable theorem partitions graphs into two classes: *class one*, when $\chi_1(G) = \Delta(G)$, and *class two*, when $\chi_1(G) = \Delta(G) + 1$.

Each class does contain a graph. The graph in Fig. 17.8 is of class one. The graph in Fig. 17.9 is of class 2. Can you prove it?

$$\chi_1(G) = 3 = D(G)$$

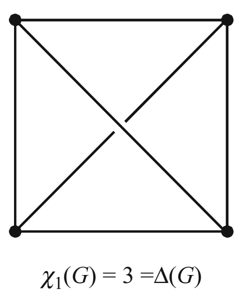
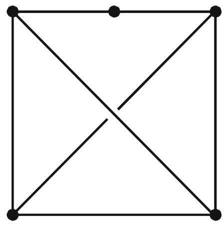


Fig. 17.8 A class one graph

$$\chi_1(G) = 4 = \Delta(G) + 1$$



$\chi_1(G) = 4 = \Delta(G) + 1$

Fig. 17.9 A class two graph

Problem 17.4 Prove that a n-cycle C_n ($n \geq 3$) is of class one if n is even and of class two if n is odd.

Problem 17.5 Prove that a complete graph K_n is of class one if n is even and of class two if n is odd.

Proof This problem does not sound exciting, does it? You are in for a nice surprise, a true mathematical recreation! In fact, do not read any further just yet, try to solve it on your own. Then read this solution, which comes from [BCL].

1. Assume the graph K_n is edge-colored in $\Delta(K_n) = n–1$ colors. Every vertex is incident with $n–1$ edges, which must be colored differently. Therefore, *every* vertex is incident with an *edge* of every color.

Now take color 1. Every vertex of K_n is incident with an edge of color 1, and the edges of color 1 are not adjacent. Therefore, the edges of color one partition the n vertices of K_n into disjoint pairs. Hence, n must be even.

We proved that if K_n is a graph of class one, then n is even.

2. Let us now prove that, conversely, the graph K_{2n} is of class one.

It is true for $n = 1$. Assume $n \geq 2$. Denote the vertices of K_{2n} by $v_0, v_1, \ldots, v_{2n–1}$. We arrange the vertices $v_0, v_1, \ldots, v_{2n–1}$ in a regular $(2n–1)$-gon and place v_0 in its center. Every two vertices we join by a straight line segment, thereby creating K_{2n}.

We are ready to color the edges of K_{2n} in $2n–1$ colors. We assign the color i ($i = 1, 2, \ldots, 2n–1$) to the edge v_0v_i and to all edges that are perpendicular to v_0v_i. We are done! All edges are colored: indeed, we assigned n edges to each color for the total of $n(2n–1)$ edges, which is the number of edges of K_{2n}. No two edges of the same color are adjacent: they clearly do not share a vertex. Figure 17.10 illustrates for you all edges of color 1 for K_8. Edge sets of other colors are obtained from this one by rotations about the center v_0 – this fact is true for the general case of K_{2n}.

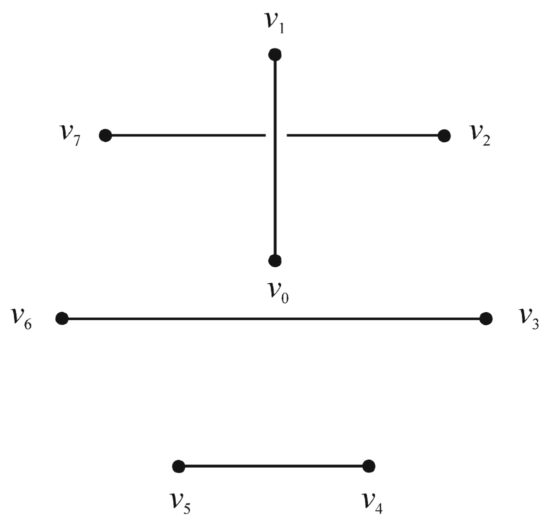


Fig. 17.10

Which class of graphs is "larger"? It does not appear at all obvious! Paul Erdős and Robin J. Wilson showed in 1977 ([EW]) that *almost all graphs are of class one*. "Almost all" is made precise by the authors of [EW]:

Problem 17.6 ([EW]) If U_n is the number of graphs with n vertices of class one and V_n is the total number of graphs with n vertices, then $\frac{U_n}{V_n}$ approaches 1 as n approaches infinity.

But how do we determine which graph belongs to which class? Nobody knows!

In 1973, Lowell W. Beineke and R.J. Wilson published [BW] the following simple sufficient condition for a graph to be of the second class:

The *edge independence number* $\beta_1(G)$ of a graph G is the maximum number of mutually nonadjacent edges of G ("Mutually nonadjacent edges" means that every two edges are nonadjacent.)

Problem 17.7 ([BW]). Let G be a graph with q edges. If

$$q > \Delta(G)\ \ \beta_1(G),$$

then G is of class two.

Proof Assume G is of class one, i.e., $\chi_1(G) = \Delta(G)$; hence, we can think of G as being $\Delta(G)$-edge-colored. How many edges of the same color can we have in G? At most $\beta_1(G)$ because the edges of the same color must be mutually nonadjacent. Therefore, the number of edges q of G is at most $\Delta(G)\ \beta_1(G)$, which contradicts the given inequality. G is of class two.

Problem 17.8 For any graph G with p vertices,

$$\beta_1(G) \leq \left\lfloor \frac{p}{2} \right\rfloor$$

where $\left\lfloor \frac{p}{2} \right\rfloor$ denotes the maximum integer not exceeding $\frac{p}{2}$.

Proof Assume that the graph G has $\beta_1(G)$ mutually nonadjacent edges. The p vertices of G are thereby partitioned into $\beta_1(G)$ two-vertex subsets plus perhaps one more subset (of vertices non-incident with any of the $\beta_1(G)$ edges). Therefore, $\beta_1(G) \leq \frac{p}{2}$, but as an integer, $\beta_1(G) \leq \left\lfloor \frac{p}{2} \right\rfloor$. ■

Problems 17.7 and 17.8 join in for an immediate corollary:

Problem 17.9 Let G be a graph with p vertices and q edges. If

$$q > \Delta(G) \cdot \left\lfloor \frac{p}{2} \right\rfloor,$$

then G is of class two.

The last problem shows that graphs with a relatively large ratio of their number of edges and number of vertices are "likely" to be of class two.

Yet, conditions of Problems 17.7 and 17.9 are far from being necessary. Can you think of a counter example? Here is one for you:

Problem 17.10 Show that the Peterson graph (Fig. 13.3) is of class two even though it does not satisfy the inequalities of problems 17.7 and 17.9.

Can we use another approach to shed light on this mysterious, relatively rare class two? We can gain an insight if we limit our consideration to planar graphs, i.e., those that can be embedded in the plane without intersection of edges. It is easy to find (do) class two planar graphs G with maximum degree $\Delta(G)$ equal to 2, 3, 4, and 5. We do not know whether maximum degree 6 or 7 can be realized in a class two planar graph. In 1965, Vadim Vizing [Viz2] proved that higher maximum degrees are impossible.

Problem 17.11 (Vizing, [Viz2, theorem 4]). If G is a planar graph with $\Delta(G) \geq 8$, then G belongs to class one.

The following problem is still awaiting its solution:

Open Problem 17.12 Find the criteria for a graph G to belong to class two.

17.2 Total Insanity Around Total Chromatic Number Conjecture

You are entitled to your own opinion,
but you are not entitled to your own facts.

— Daniel Patrick Moynihan, U.S.
Senator from New York

In February 1992, I gave my first talk at the International Southwestern Conference on Combinatorics, Graph Theory, and Computing at Florida Atlantic University, Boca Raton, Florida. The talk was about the problem of finding the chromatic number of the plane and my investigation into the authorship of the problem. My investigative skills must have looked good, for the British graph theorist Hugh Hind shared with me another controversy. In his manuscript on the total chromatic number conjecture, Hugh gave credit for the conjecture to Vizing and Behzad. As a condition of publication, the referee *demanded* (!) that the credit be given to Behzad alone. Even though Hind thought that both mathematicians authored the conjecture independently and deserved credit, he felt that he had no choice but to comply with the referee's demand. Hugh asked me to investigate the authorship of the total chromatic number conjecture.

I was shocked. The referee's ultimatum, backed by the editor (who sent the referee report to the author), seemed to be nothing short of a cold war on the mathematical front. What were the referee's and the editor's motives? Was it a retaliation for the Soviet anti-Semitism and violations of scientific norms? Was it a retaliation for the leading Soviet graph theorist Alexander A. Zykov's ridiculously giving in his book [Zyk3] credit for the Kuratowski planarity theorem to both Pontryagin and Kuratowski? Of course, Zykov's crediting Pontryagin was outrageous, and Pontryagin deserved no credit whatsoever. However, life is no math – it does not multiply two negatives to get a positive – two wrongs make no right. Surely, the referee and the editor of Hugh Hind's manuscript acted every bit as wrongly as the Soviet apparatchiks – unless the referee had historical factual grounds to deny Vizing credit, which they never disclosed. I accepted Hugh's call to investigate. What follows in this chapter is my investigative report.

The *total chromatic number* $\chi_{ve}(G)$ of a graph G is the minimum number of colors required for coloring the vertices *and* edges of G so that incident and adjacent elements are never assigned the same color.

Total Chromatic Number Conjecture 17.13 For any graph G,

$$x_{ve}(G) \leq \Delta(G) + 2.$$

I started my historical investigation right away, in Boca Raton during the same conference (February 1992). I contacted, in person, the well-known graph theorist Mark K. Goldberg, professor of computer science at Rensselaer Polytechnic Institute. This was a very lucky choice, for Mark was an eyewitness to the story. Goldberg told me that in December 1964, he arrived in Akademgorodok (Academy Town), located just outside Novosibirsk, a major city in Siberia, to apply for their PhD program in mathematics. During this trip, he interacted with the junior research staff member Vadim G. Vizing, who shared with Goldberg his edge chromatic number theorem and the total chromatic number conjecture.

Three years later, I was able to ask **Vadim Vizing** himself to share with me historical details of the total chromatic number conjecture. I learned from Bjarne Tofts, a professor at Odense University in Denmark, that Vadim G. Vizing was presently visiting him. On March 12, 1995, I asked Toft to pass my e-mail with numerous questions to Vizing. I asked biographical questions and, of course, questions about the conjecture. Two days later, on March 14, 1995, I received the following reply (my translation from Vizing's Russian):

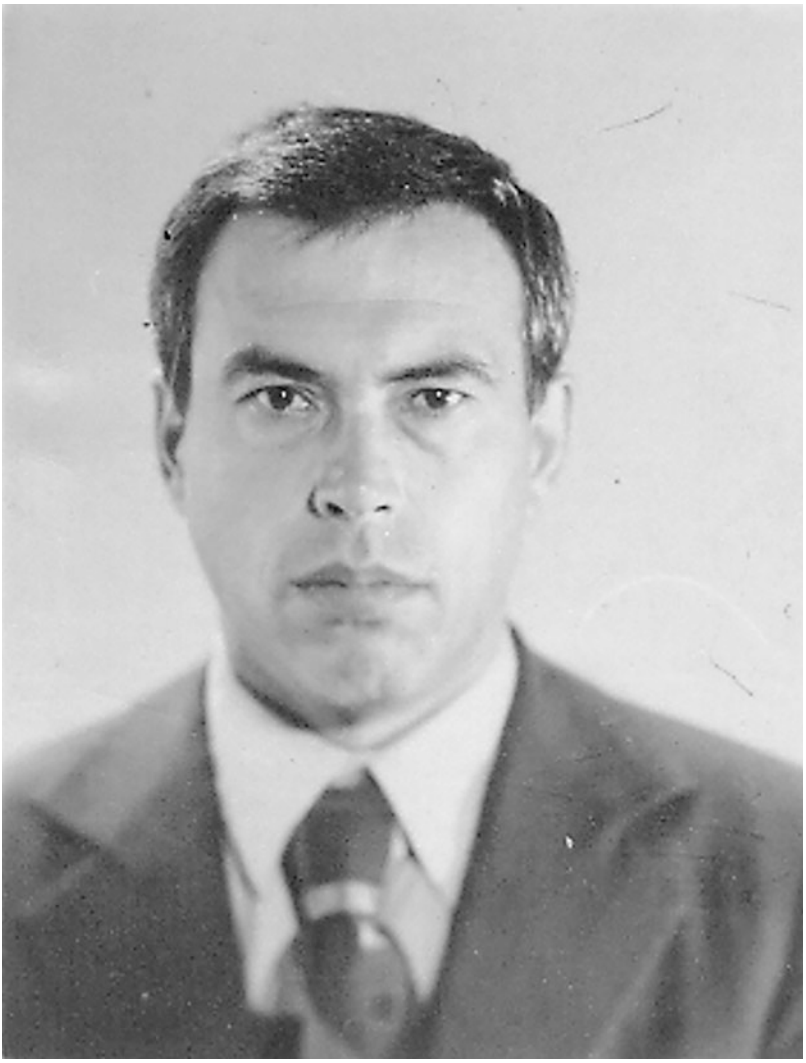
Dear Alexander!

At the present time I am in Odense on B. Toft's invitation.

I was born 25 March 1937 in Kiev. I commenced my work on Graph Theory in 1962 as a Junior Research Staff at the Institute of Mathematics in Novosibirsk, in the Department of Computing Techniques [Computer Science]. As part of my job, I had to write a program for coloring conductors in circuits. I discovered C. E. Shannon's work, dedicated to this question, published in 1949 (Russian translation was published in 1960). Having studied Shannon's work, I began to think about the precision of his bound. I knew only one type of multigraphs on which his bound was precise. This is why I assumed that for ordinary graphs (without multiple edges) Shannon's bound could be strengthened. It took a year and a half for me to prove my theorem for ordinary graphs.

In early 1964 the article was sent to "Doklady AN USSR" (AN abbreviates Academy of Sciences), but was rejected by the editorial board. In the fall of 1964, I obtained the generalization of the result to *p*-graphs and published an article about it in the anthology "Diskretnyi Analiz", issue 3 [Viz1] that was released in December 1964 in Novosibirsk (I am mailing to you a copy of this article).

In early 1964, while presenting the theorem about coloring edges of a graph at A. A. Zykov's' Seminar (present were A. A. Zykov, L. S. Melnikov, K. A. Zaretskij, V. V. Matjushkov, and others), I formulated the conjecture on the total chromatic number, which we called then *conjecture on the simultaneous coloring of vertices and edges*. Many of my colleagues in Novosibirsk attempted to prove the conjecture but without success. By the time of the publication [Viz3] of my article on unsolved problems of Graph Theory in "Uspekhi Mat. Nauk" (1968), in which I first published the conjecture, the conjecture already had a wide distribution among Soviet mathematicians. In the nearest future I will mail to you the article in Russian; the conjecture on the total chromatic number of multigraph appears on p. 131.



Vadim G. Vizing in the early 1960s, when he worked on his classic chromatic index theorem; courtesy of Vadim Vizing

Thus, Vadim G. Vizing's recollection of creating the total chromatic number problem by early 1964, verified independently by Professor Mark Goldberg, leaves no doubts about his authorship. Vizing's total chromatic number conjecture was also presented by Alexander A. Zykov at the problem session of the Manebach Colloquium in May 1967 (published in 1968 [Zyk2], p. 228). Vizing published the conjecture himself along with many other open problems in 1968 [Viz3]. In addition, any impartial expert would agree that this conjecture was a natural continuation of the train of thought emanating from Vizing's famous theorem on the chromatic index of a graph (Theorem 17.3 above).

I then looked at articles of specialists working on the total chromatic number of a graph. Hugh Hind [Hin1], [Hin2], Anthony Hilton and Hind [HH], and Amanda G. Chetwynd nearly universally credited Behzad for the conjecture. Chetwynd even "explained" what led Behzad to discover the conjecture [Che]:

> This [i.e., Brooks' Theorem and Vizing's Theorem] led Behzad to conjecture a similar result for the total chromatic number.

What is wrong with this "explanation"? Everything:

1. Reading Mehdi Behzad's 1965 thesis (Chetwynd obviously did not read it before writing about it), it is obvious that Behzad did not know Vizing's theorem: Behzad conjectures the statement of Vizing's 1964 chromatic index theorem but is able to prove it only for a trivial case of graphs of maximum vertex degree 3.
2. If Vizing's theorem led even Behzad to the total chromatic conjecture, it would have surely led (and did!) Vizing himself to formulate the conjecture. Why does then Chetwynd give no credit to Vizing?

I contacted Professor Behzad, on December 2, 2007, and asked him to present a case in support of his sole authorship. He was pleased with my commencing an investigation, and, on December 3, 2007, wrote:

> Dear Professor Soifer.
>
> I am extremely happy that after almost 40 long years eventually a genuine scholar spent time and effort to clarify an important point which unfortunately has happened many times in the past, due to different reasons.

In a few days, on December 3, 2007, he sent me a detailed statement. Behzad then asked me to replace it by the new December 14, 2007, e-mail version, which I am quoting here:

> I started to think about my Ph.D. thesis in 1963–1964, at Michigan State University, to be written under the supervision of Professor E. A. Nordhaus. In those days there was only one book in the field of graph theory in English, and no courses were offered on the subject. I was interested in vertex coloring and then line coloring. For several months, naively, I tried to solve the 4-color problem. Then I thought of combining these two types of colorings. I mentioned the notion, which was later called "total chromatic number of a graph,"[2] to Nordhaus. He liked the idea, but for several months he did not allow me to work on the notion. Later he told me this idea was so natural that he thought someone might have worked on the subject. Thanks to Professor Branko Grünbaum who resolved the problem. In my thesis I introduced this notion and presented the related conjecture. In addition, I introduced the total graph of a graph in such a way that the total chromatic number of G was equal to the vertex chromatic number of its total graph. My [doctoral] thesis was defended in the Summer of 1965. Prior to 1968, when Professor Vizing's paper entitled "Some Unsolved Problems in Graph Theory" appeared, several papers were published on topics related to total concepts; I informally talked about TCC in two of the conferences that I attended in 1965 and 1966 held at The University of Michigan, and the University of Waterloo. As I mentioned before, aside from my thesis, the Proceedings of the International Symposium in the Theory of Graphs – Rome, 1966, contains the subject and the TCC. . .

---

[2] According to Behzad, it was Nordhaus who coined the term.

> As far as I know, out of several hundred articles, theses, books, and pamphlets containing TCC, none omits my name, and very many authors provide only one reference for TCC and that is my thesis. I am not aware of a single work mentioning TCC and giving reference to Vizing alone. There are authors who have given credit to the two of us but have decided to stop doing so.

Let me first address the last argument of Professor Behzad. We, the scholars, do not determine the truth by a majority vote.

Let us further differentiate here between "total graphs" and "total chromatic number conjecture." Professor Behzad writes above "Prior to 1968, when Professor Vizing's paper entitled 'Some Unsolved Problems in Graph Theory' appeared, several papers were published on topics related to total concepts." Indeed, the following papers, authored or coauthored by Behzad, addressed "total graphs" but did not include "the total chromatic number conjecture":

M. Behzad and G. Chartrand, Total graphs and traversability, *Proc. Edinburgh Math. Soc.* (2) 15 (1966). 117–120.

M. Behzad, G. Chartrand, and J.K. Cooper Jr. The colour numbers of complete graphs, *J. London Math. Soc.* 42, (1967) 225–228.

M. Behzad, A criterion for the planarity of the total graphs of a graph, *Proc. Cambridge Philos. Soc.* 63 (1967). MR35#2771.

M. Behzad and H. Radjavi, The total group of a graph, *Proc. Amer. Math. Soc.* 19 (1968), 158–163.

I am reading Mehdi Behzad's thesis [Beh]. I was surprised to see a PhD (!) thesis without nontrivial proofs. However, the author demonstrates a good intuition: he conjectures (already published by Vizing a year earlier) Vizing's theorem on the edge chromatic number of a graph (conjecture 1, p. 18) and formulates the total chromatic number conjecture (conjecture 1, p. 44).

Mehdi Behzad and Gary Chartrand submitted their "expository article" on total graphs to the 1966 Rome Symposium, and it was published [BC1] in 1967. I read there (I am replacing their notations by contemporary ones):

> It was conjectured in [Beh1] that
> (i) $\Delta(G) \le \chi_1(G) \le \Delta(G) + 1$ and
> (ii) $\Delta(G) + 1 \le \chi_2(G) \le \Delta(G) + 2$.
>
> The conjecture (i) has been proved by Vizing [Viz1] but (ii) remains an open question.

The good news is that Behzad and Chartrand thus published the total chromatic number conjecture. As to "conjecture (i)," we are told in the above that Behzad [!] conjectured the chromatic index theorem and that Vizing proved Behzad's [sic] conjecture! In fact, Vizing's paper was already published in 1964, a year before Behzad ever conjectured this result. Of course, Vizing worked on *his* conjecture about the chromatic index much earlier, during 1962–1963, for as he says, it took him a year and a half to prove his conjecture, and, in early 1964, Vizing submitted his paper.

M. Behzad had no way of knowing about my findings that show that Vizing formulated the total chromatic number conjecture in early 1964, i.e., well before Behzad. However, Behzad knew about Vizing's 1968 paper [Viz3], where the total chromatic number conjecture was published. Surely, it took considerable time, prior to the submission of this paper, for Vizing to assemble such a large survey of unsolved problems of graph theory. Thus, the independent authorship of Vizing should not have been questioned. Yet, Behzad in the 1971 book [BC,

p. 214], jointly with Gary Chartrand, his former fellow PhD student of E.A. Nordhaus, gives the sole credit to himself for the total chromatic number conjecture. It happens yet again in the 1979 book by three authors, Behzad, Chartrand, and Linda Lesniak–Foster [BCL, p. 252].

In our phone conversation on January 2, 2008, I informed Professor Behzad of my findings and determination of the Vizing and Behzad's joint authorship of the conjecture. He expressed a pleasure that at long last somebody took the time and effort to investigate the credit for this famous conjecture and that credit was rightfully due to the two independent persons. Yet, Behzad's Wikipedia page still today (February 2, 2023) alleges that "he introduced his [sic] total coloring theory (also known as 'Behzad's [sic] conjecture' or 'the total chromatic number conjecture')."

Summing up, the total chromatic number conjecture was first formulated by Vadim G. Vizing in early 1964 and published in 1968. Mehdi Behzad independently formulated the conjecture in his unpublished thesis in the summer of 1965 and published it, jointly with Gary Chartrand, in 1967. This, in my opinion, unquestionably merits a joint credit to Vizing and Behzad.

I hope that this analysis will end editorial room bias, politicking, threats of not publishing papers that include Vadim G. Vizing's name as the conjecture's coauthor, and will restore the join credit for the conjecture. Joint credit and correct publication dates were given by Tommy R. Jensen and Bjarne Toft in their enlightened 1995 problem book [JT] and repeated from there in Reinhard Diestel's textbook [Die]. In later papers, e.g., [HMR], Hugh Hind, Michael Molloy, and Bruce Reed are now giving credit to both Vizing and Behzad for the concept of the total chromatic number and the conjecture. Yet, even the fourth edition of *Graphs & Digraphs* [CL] by Chartrand and Lesniak (Behzad was dropped from the authors), which appeared in 2005, still credits M. Behzad, and Behzad alone, for total coloring and the total chromatic number conjecture.

In the first edition of *The Mathematical Coloring Book*, I expressed my hope that, having read these lines, the authors will correct the credit in their next edition. I am pleased to report that in the 2011 fifth edition of *Graphs & Digraphs* [CLZ], Chartrand, Lesniak, and their new coauthor Ping Zhang finally gave a joint credit to Behzad and Vizing. There are still some authors, apparently, including Mehdi Behzad, who insist on Behzad's sole authorship of the conjecture – it does not surprise me, for there are those few who invent "alternative facts" and insist that the Holocaust never happened, allege that the evolution theory is a "hoax," and the coronavirus (SARS-CoV-2) pandemic was invented by the media and liberals. In my opinion, it should have been the highest honor for anyone to coauthor a conjecture with the author of the classic chromatic index theorem, Vadim G. Vizing.

Until his 1981 early retirement at the age of 45, Mehdi Behzad was a professor at Sharif University of Technology in Iran. Wikipedia informs that after 1981, Dr. Behzad "has continued to serve the Iranian scientific community in different capacities."

Following his visit to Denmark in 1995, Vadim Vizing, who lately worked on the theory of scheduling, wrote to me that he was going to renew "intensive work on graph theory." He has indeed, as his publications show. Vadim Georgievich Vizing (25 March 1937–23 August 2017) spent most of his life in the beautiful subtropical Black Sea port of Odessa, Ukraine, where my parents also spent their youth.

In spite of active work, the total chromatic number conjecture has withstood all assaults. With an ease of formulation and an apparent difficulty of proving, this conjecture now belongs to mathematics' classic open problems.

17.3 What Else Can We Color in a Graph?

We could color vertices of a graph G, which poses a question of finding the *chromatic number* $\chi_v(G)$ or simply $\chi(G)$.

We could color edges, which poses a question of finding the *edge chromatic number* $\chi_e(G)$, which is sometimes denoted by $\chi_1(G)$.

We could color vertices and edges, which poses a question of finding the *total chromatic number* $\chi_{ve}(G)$, which is sometimes denoted by $\chi_2(G)$.

As mentioned before, the total chromatic number conjecture is open, but, in the case of planar graphs with a large enough maximum vertex degree, some results are obtained. For example, Oleg Borodin [Bor2] proved the following inequalities:

Theorems on Total Coloring of Planar Graphs 17.14 (Borodin):
Every planar graph with $\Delta(G) \geq 9$ is totally $(\Delta + 2)$-colorable.
Every planar graph G with $\Delta(G) \geq 16$ satisfies $\chi_{ve}(G) = \Delta(G) + 1$.

The series of works by Borodin, Jensen, and Toft was followed by Daniel P. Sanders and Yue Zhao, who proved that every planar graph with $\Delta \geq 7$ is totally $(\Delta + 2)$-colorable. The remaining open case they formulated as a question [SZ2]:

Open Problem 17.15 (Sanders–Zhao, 1999): Is every planar graph with $\Delta = 6$ totally 8-colorable?

In the case of a plane graph, we could color vertices and faces, which poses a question of finding the *vertex-face-chromatic number* $\chi_{vf}(G)$. We could color edges and faces, which poses a question of finding the *edge-face chromatic number* $\chi_{ef}(G)$. In addition, we could color "everything": vertices, edges, and faces – the entire hardware of a graph – and pose a question of finding the *vertex-edge-face chromatic number*, also known as the *entire chromatic number* and denoted by $\chi_{vef}(G)$.

What do we know about these characteristics of a graph?

In 1965, Gerhard Ringel conjectured that the *vertex-face-chromatic number* $\chi_{vf}(G)$ of a plane graph G does not exceed 6. He also proved that seven colors suffice. In 1984, Oleg V. Borodin proved it and more [Bor3].

A graph is called a *1-plane* if it is drawn in the plane so that every edge intersects at most one other edge in an interior point. Consistently, I attribute a joint credit to the author of the conjecture and the author of its proof.

Ringel–Borodin's Six-Color Theorem 17.16 [Bor3]: Every 1-plane graph G satisfies the inequality $\chi_{vf}(G) \leq 6$.

This beautiful result is also the best possible, for the complete graph K_6 is 1-planar and thus fewer than six colors will not always suffice. Borodin's proof is noteworthy; he uses a set of 35 unavoidable reducible configurations.

A decade later, in 1995, Oleg Borodin found an even better proof [Bor4]. He reduced the set of unavoidable reducible configurations from 35 to 18. Let me include a proof of theorem 17.16 limited to only plane graphs, for it is very short and very sweet:

Theorem 17.17 (Archdeacon 1986, [Arc]): Every plane triangle-free graph G satisfies the inequality $\chi_{vf}(G) \leq 6$.

Proof [Bor2] The Grötzsch theorem 20.13 (see it later in this book) allows coloring of the vertices of G with colors 1, 2, and 3. Erase color 3 from the vertices and "blow up" the now uncolored vertices into small faces so as to preserve the adjacency of "old" faces. (Thus, every uncolored k-vertex becomes a small k-face.) Color all faces with colors 3, 4, 5, and 6, which can be done due to the dual version of the Four-Color Theorem. Finally, contract the new faces back to vertices, preserving their colors. This gives the desired coloring. ■

In 1973, Hudson V. Kronk and John Mitchem posed the following conjecture [KMi]:

Conjecture 17.18 (Kronk–Mitchem): Every plane graph G with $\Delta(G) \geq 3$ satisfies the inequality $\chi_{vef}(G) \leq \Delta(G) + 4$.

Following a series of works by a number of researchers, in 2011, Wenzhang Wang and Xuding Zhu [WZ] completed the remaining cases $\Delta = 4$ and $\Delta = 5$ of the Kronk–Mitchem conjecture, thus giving birth to the statement:

Kronk–Mitchem–Waang–Zhu's Theorem 17.21 Every plane graph G with $\Delta(G) \geq 3$ satisfies the inequality $\chi_{vef} G \leq \Delta(G) + 4$.

As always, I assign a join credit to the authors of the conjecture and its provers.

In 1974, Leonid S. Mel'nikov [Me] posed the following conjecture:

Mel'nikov's Conjecture 17.22 Every plane graph G satisfies the inequality $\chi_{ef}(G) \leq \Delta(G) + 3$.

In 1997, Mel'nikov's conjecture was proved by Daniel P. Sanders and Y. Zhao [SZ3] and, independently, by A.O. Waller [Wa].

Mel'nikov–Sanders–Zhao–Waller's Theorem 17.23 Every plane graph G satisfies the inequality $\chi_{ef} G \leq \Delta(G) + 3$.

&&&

Oleg Veniaminovich Borodin is a professor at Novosibirsk State University and head of the Graph Theory Laboratory in the Sobolev's Mathematics Institute of the Siberian Branch of Russian Academy of Sciences.

Chapter 18
Carsten Thomassen's 7-Color Theorem

One day in 1998, I was asked by *The American Mathematical Monthly* to referee a manuscript submitted by one of the world's leading graph theorists, Professor Carsten Thomassen of the Technical University of Denmark. The paper offered a fresh, purely graph-theoretic approach to finding the chromatic number of the plane. I was very impressed, asked the author to expand his too concise (for *The Monthly*) presentation, and informed him of my research that proved that Edward Nelson, and Nelson alone (without Hadwiger), was the author of the problem. Of course, I recommended *Monthly* to publish this work. In this chapter, I will present Thomassen's attempt to find the chromatic number of the plane. He has not found it – no one has – but he obtained a fine result and, in the process, showed how graph theory proper can be utilized in an assault on this problem. I will present Thomassen's proof with minor editorial revisions. The use of paper and pencil is a must while reading the proof written in Thomassen's style.

Thomassen offers a vast generalization of the popular hexagonal coloring that we used to prove the upper bound 7 (Chapter 2) to the class of colorings that he calls *nice*. He considers a graph G on a surface S that is a metric space (i.e., curve-wise connected Hausdorff space in which each point has a neighborhood homeomorphic to an open circular disc of the Euclidean plane). The graph G on the surface S creates a map $M(G, S)$, in which a region is an edge-connected component of $S \backslash G$. For his purposes, Thomassen assumes that each region that has diameter less than 1 is homeomorphic to a Euclidean disc and is bounded by a cycle in G. I choose to avoid a detour into the basics of topology and offer the unfamiliar reader to simply think that S is the plane or a sphere, i.e., the graph G is drawn on the Euclidean plane or a sphere – coloring the plane is, after all, our main goal.

The *area* of a subset A of S is the maximum number of pairwise disjoint open discs of radius ½ that can be packed in A. (If this maximum does not exist, we say that A has infinite area.) A simple, closed curve C is *contractible* if $S \backslash C$ has precisely two edgewise connected components such that one of them is homeomorphic to an open disc in the Euclidean plane. This component is called the *interior* of C and is denoted by $int(C)$. If S is a sphere, then $int(C)$ denotes any component of $S \backslash C$ of the smallest area.

Given a graph G on a surface S, Thomassen defines *nice coloring* of S as a coloring in which each color class is the union of regions (and part of their boundaries) such that the distance between any two of these regions is greater than 1.

© Alexander Soifer 2024
A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_18

Finally, I need to introduce here a *map–graph duality*, which we will use not only in this chapter but in the next chapter as well. Given a map M, we can define the graph of the map, or *map graph* $\Gamma(M)$, as the graph whose vertices are the regions of M with two vertices adjacent if and only if the corresponding regions share a part of their boundary, which is not merely a finite number of points. If the map $M(G, S)$ is induced by the graph G on the surface S, then we will simplify the notation for the map graph to $\Gamma(G, S)$.

The Thomassen 7-Color Theorem 18.1 Let G be a connected graph on a surface S. Then, every nice coloring of S requires at least seven colors, if there exists a positive integer k satisfying the following conditions (i), (ii), and (iii):

(i) Every noncontractible, simple closed curve has diameter of at least 2.
(ii) If C is a simple closed curve of diameter less than 2, then the area of $int(C)$ is at most k.
(iii) The diameter of S is at least $12\ k + 30$.

Before proving his theorem, Thomassen introduces Tool 18.2, for which he needs a few notations and definitions.

If $V(G)$ is the vertex set of a graph G and $x \in V(G)$, then $D_1(x)$ stands for the set of neighbors of x. For $n > 2$, we define $D_n(x)$ inductively as the set of vertices in $V(G) \backslash [\{x\} \cup D_1(x) \cup \ldots \cup D_{n-1}(x)]$ that have a neighbor in $D_{n-1}(x)$. A graph G is called *locally finite* if $D_1(x)$ is finite for each vertex x of G and *locally connected* if the minimal subgraph of G that contains $D_1(x)$ is connected for each vertex x of G. We call G *locally Hamiltonian* if G has a cycle with vertex set $D_1(x)$ for each vertex x of G.

Tool 18.2 Any connected, locally finite, locally Hamiltonian graph with at least 13 vertices has a vertex of degree at least 6.

Proof If no vertex of the graph G satisfying all conditions has a degree of at least 6, pick a vertex x of maximum degree. Clearly, $deg(x) \geq 3$.

Assume $deg(x) = 3$. Since G contains a cycle with the vertex set $D_1(x)$, the subgraph of G induced by $\{x\} \cup D_1(x)$ is the graph of the tetrahedron. Since the maximum degree in G is 3, $D_2(x)$ is empty. Since G is connected, G is the graph of the tetrahedron, i.e., it has just 4 vertices, in contradiction to the assumption that G has at least 13 vertices.

Assume now that $deg(x) = 4$. Since the vertices of $D_1(x)$ form a cycle, we can conclude that each vertex $y \in D_1(x)$ has at most one neighbor $z \in D_2(x)$. Since the vertices of $D_1(y)$ form a cycle, z has at least three neighbors in $D_1(x)$. Thus, there are at most four edges from $D_1(x)$ to $D_2(x)$, and, therefore, every vertex in $D_2(x)$ has at least three neighbors in $D_1(x)$. Hence, $D_2(x)$ has at most one vertex z. Since the vertices of $D_1(z)$ form a cycle, it follows that $D_3(x) = \varnothing$. Thus, G has at most six vertices, a contradiction.

Finally, assume that $deg(x) = 5$. Each vertex $y \in D_1(x)$ has at most two neighbors in $D_2(x)$ because the vertices of $D_1(x)$ form a cycle and this cycle uses up two points out of the maximum degree 5 of y. Since the vertices of $D_1(y)$ form a cycle, every neighbor z of y in $D_2(x)$ has at least two neighbors in $D_1(x)$. Observe that z cannot have two or more neighbors in $D_3(x)$ because then a cycle with vertex set $D_1(z)$ shows that z has at least two neighbors in $D_2(x)$, that is, z has a total of at least six neighbors, a contradiction. So, z has at most one neighbor in $D_3(x)$ and that neighbor has at least three neighbors in $D_2(x)$. Since there are at most 10 edges from $D_1(x)$ to $D_2(x)$, and every vertex in $D_2(x)$ has at least 2 neighbors in $D_1(x)$, it follows that $D_2(x)$ has at most 5 vertices. Hence, there are at most five edges from $D_2(x)$ to $D_3(x)$. Since each vertex in $D_3(x)$ has at least three neighbors in $D_2(x)$, it follows that $D_3(x)$ has at most one vertex, and thus $D_4(x) = \varnothing$. Hence, G has at most 12 vertices, a contradiction that completes the proof.

Now we are ready to prove the theorem. ■

Proof of Thomassen's Theorem Given a graph G on a surface S that satisfies (i), (ii), and (iii). Assume the opposite, i.e., that there is a nice coloring utilizing at most six colors. Let x be a vertex of the map graph $\Gamma = \Gamma(G, S)$, and let C_x be the cycle in G bounding the corresponding region. Let us choose an orientation of C_x, and let $x_1, x_2, \ldots, x_k, x_1$ be the vertices of $D_1(x)$ listed in the order in which they appear as we traverse C_x.

Thomassen first considers a simple case, which illustrates the idea of his proof. Assume that for each vertex x, all vertices $x_1, x_2, \ldots, x_k$ are distinct. In this case, Γ is locally Hamiltonian. Since the surface S is edgewise connected, it follows that Γ is connected. Since S has a diameter greater than 13, Γ has more than 12 vertices, and, hence, by Tool 18.2, Γ has a vertex of degree at least 6. Now, x and its neighbors must have distinct colors because x corresponds to a face of diameter < 1 on S. This contradiction completes the proof in this case.

In the general case, a vertex may appear more than once in the sequence $x_1, x_2, \ldots, x_k, x_1$ above. Omit those appearances (except possibly one) of x_i for which C_{x_i} and C_x have only one vertex in common. In other words, if x_i appears more than once in the new sequence, then we list only those appearances for which C_{x_i} and C_x share an edge. Then, any two consecutive vertices in the sequence $x_1, x_2, \ldots, x_k, x_1$ are neighbors in Γ, and thus, Γ is locally connected. It follows that $\Gamma - x$ is connected. Moreover, if y is any other vertex of Γ, then $\Gamma - x - y$ is connected unless y appears twice in the sequence $x_1, x_2, \ldots, x_k$, that is, C_x and C_y have at least two edges in common.

Now, let x and y be vertices such that C_x and C_y have at least two edges e and f, respectively, in common, i.e., $y = x_i = x_j$ for $1 \le i < j - 1 < k - 1$. Let R be a simple closed curve in the regions bounded by C_x and C_y such that R crosses each of e and f precisely once and has no other point in common with G. By (i), R is contractible. Hence, $\Gamma - x - y$ is disconnected. We say in this case that $\{x, y\}$ is a *2-separator* in Γ. For each vertex z in Γ such that C_z is in $int(R)$ and has color 1, we pick a point P_z in $int(C_z)$. By (ii), there are at most k points P_z and, hence, there are altogether at most 6 k vertices z such that $int(C_z) \subseteq int(R)$.

Let $int(\Gamma, x, y)$ stand for the subgraph of $\Gamma - x - y$ induced by all those vertices z in Γ such that C_z is in $int(R)$ for some R. Then, each connected component of $int(\Gamma, x, y)$ has at most 6 k vertices. Since S has a diameter of at least 12 k + 3, it follows that G has two vertices whose graph distance is at least 12 k + 2. Hence, $\Gamma - x - y$ has some component that is not in $int(M, x, y)$. We claim that $\Gamma - u - v$ has precisely one such component, which we call $ext(M, x, y)$. To see this, let $e_1, e_2, \ldots, e_m$ be the edges in $C_x \cap C_y$ occurring in this cyclic order on C_x. Then, $e_1, e_2, \ldots, e_m$ divide $D_1(x)\backslash\{y\}$ into m classes $A_1, A_2, \ldots, A_m$. By letting $\{e, f\} = \{e_i, e_{i+1}\}$, $1 \le i < m$ in the preceding argument, we conclude that for each $i = 1, 2, \ldots, m$, either $A_i \subseteq int(\Gamma, x, y)$ or $A_i \cap int(\Gamma, x, y) = \varnothing$. Since the former cannot hold for each $i \in \{1, 2, \ldots, m\}$, the latter must hold for some i, and, hence, the former holds for all other $i \in \{1, 2, \ldots, m\}$. Thus, we proved that for any two vertices x, y in Γ, $\Gamma - x - y$ has precisely one connected component $ext(\Gamma, x, y)$ with more than 6 k vertices.

If $\{u, v\}$ is a 2-separator in Γ such that either x or y or both are in $int(\Gamma, u, v)$, then clearly $int(\Gamma, x, y) \subset int(\Gamma, u, v)$. (To see this, we use the properties of Γ established previously and forget about S.) If no such 2-separator $\{u, v\}$ exists, then we say that $\{x, y\}$ is a *maximal 2-separator* and that xy is a *crucial edge.* Since each connected component of $int(\Gamma, x, y)$ has at most 6 k vertices, a maximal 2-separator exists (provided a 2-separator exists). Let H be the subgraph of Γ obtained by deleting $int(\Gamma, x, \mathrm{y})$ for each maximal 2-separator $\{x, y\}$. Then, $H \neq \emptyset$. Moreover, H is connected since the shortest path in Γ between two vertices in H never uses vertices in $int(\Gamma, x, y)$. Similarly, H is locally connected. We now claim that H is locally Hamiltonian. Consider again a vertex x in H and the sequence $x_1, x_2, \ldots, x_k, x_1$ in $D_1(x)$ (taken in Γ). If this sequence forms a Hamiltonian cycle in $D_1(x)$ in H, we are done. By definition of H, $k \geq 3$. So, assume that $x_i = x_j$ where $1 \leq i < j - 1 < k - 1$. Then, $\{x, x_i\}$ is a 2-separator and vertices can be reindexed so that $int(\Gamma, x, x_i)$ contains all the vertices $x_{i+1}, x_{i+2}, \ldots, x_{j-1}$. We repeat this argument for each pair i, j such that $x_i = x_j$ where $1 \leq i < j - 1 < k - 1$. Then, the vertices in $x_1, x_2, \ldots, x_k, x_1$ that remain after we delete all vertices in the interiors of the 2-separators form a cyclic sequence with no repetitions. As H is connected and locally connected and has at least three vertices (by (iii)), the preceding reduced cyclic sequence has at least two distinct vertices. It cannot have precisely two vertices u, v because then $H - u - v$ is disconnected, and, hence, $\Gamma - u - v$ is disconnected (because Γ is obtained from H by "pasting graphs on the edges of H"). Since one of the edges xu or xv is crucial (because $D_1(x)$ is smaller in H than in Γ), the maximality property of the 2-separator $\{x, u\}$ or $\{x, v\}$ implies that $ext(\Gamma, u, v)$ is the connected component of $\Gamma - u - v$ containing x. For each vertex z in that component, Γ has a path of length at most 6 k from z to x, u, or v. Hence, Γ has diameter at most 12 k + 1, a contradiction that proves that H is locally Hamiltonian.

If H has a vertex x of degree at least 6, we are done because x and its neighbors must have different colors in the nice coloring. Assume now that each vertex of H has a degree of at most 5. By Tool 18.2, H has at most 12 vertices. Hence, H has at most 30 edges. Since Γ is obtained from H by "pasting" $int(\Gamma, x, y)$ on the crucial edge xy for each crucial edge of H, we conclude that the diameter of Γ is at most 12 k + 29, a contradiction to (iii). ■

Observe: All three conditions in the theorem are essential. If *any* of these conditions (i), (ii), (iii) are dropped, then the number of colors needed may decrease:

A thin two-way infinite cylinder has a nice 6-coloring, which shows that (i) cannot be omitted.
A thin one-way infinite cylinder (with a small disc pasted on the boundary of the cylinder to form the bottom) shows that (ii) cannot be omitted.
A sphere of diameter less than 1 has a nice coloring in two colors; hence, (iii) cannot be omitted.

Later in this book, we will study a somewhat similar Townsend–Woodall's theorem, – Woodall's 5-color theorem, obtained by highly different means.

Part IV
Coloring Maps

G. D. Birkhoff once told one of the authors that every great mathematician had at some time attempted the Four Colour Conjecture, and had for a while believed himself successful.
– Hassler Whitney and W. T. Tutte[1]

The word disease is quite appropriate for a puzzle which is easy to comprehend, apparently impossible for anyone to solve, infectious, contagious, recurrent, malignant, painful, scarring, and sometimes even hereditary!
– Frank Harary[2]

If I may be so bold as to make a conjecture, I would guess that a map requiring five colors may be possible.
– H. S. M. Coxeter[3]

In this Part, we will color regions of maps. The following few definitions will help us formalize our intuitive notion of a map.

By allowing more than one edge to connect two vertices, we slightly generalize the notion of a graph: what we get is called a *multigraph*. A multigraph that *can be* drawn in the plane without intersection of its edges is called *planar*, while a multigraph that *is* drawn in the plane without intersection of its edges is called a *plane*. A multigraph is called *connected* if for any two vertices, there is a path connecting them. An edge x of a connected multigraph G is called a *bridge* if the multigraph $G - x$ is not connected.

We will call a plane drawing of a connected multigraph without bridges a *map*. A map divides the plane into *regions*. Regions are *adjacent* if they share at least one edge.

Coloring a map is an assignment of colors to each of the regions of the map such that no adjacent regions get the same color. Let n be a positive integer; a map M is called *n-colorable* if there is a coloring of M in n colors.

---

[1] [WT].

[2] From the appropriately entitled paper [Har1] "The Four Color Conjecture and Other Graphical Diseases," appropriately "supported in part by a grant from the National Institute of Mental Health."

[3] [Cox2].

A natural question then is what is the *minimum* number of colors we must use to color any map? You can easily construct an example showing that four colors are necessary. You have likely heard the puzzle and the conjecture I am going to introduce here as an overture to this map coloring part.

The following puzzle originated in discussions between the well-known mathematician August Ferdinand Möbius and his amateur mathematician friend Adolph Weiske and "was perhaps originated by Weiske" [Tie2]. Möbius shared the puzzle with the public in his 1840 lecture. It was apparently solved even by the Bishop of London, later Archbishop of Canterbury (see Chapter 20 for details). Wouldn't you like to solve it on your own and try to help the brothers?

Möbius–Weiske's Puzzle IV–1. (Circa 1840) Once upon a time in the Far East, there lived a prince with five sons. These sons were to inherit the kingdom after his death. But in his will, the prince made the stipulation that each of the five parts into which the kingdom was to be divided must border on every other . . . After the death of the father, the five sons worked hard to find a division of the land, which would conform to his wishes; but all their efforts were in vain.[4]

The following conjecture together with Fermat's Last Theorem had arguably been the two most popular open problems of mathematics.

The Four-Color Conjecture (4CC) IV–2 (Francis Guthrie, 1852, or before) Any map in the plane is 4-colorable.

My late friend Klaus Fischer of George Mason University once asked me in the early 1990s, why would I want to write about the conjecture so celebrated that everything has been written about it? Well, *everything* is never written, I replied, and every little bit helps.

[4] [Tie2].

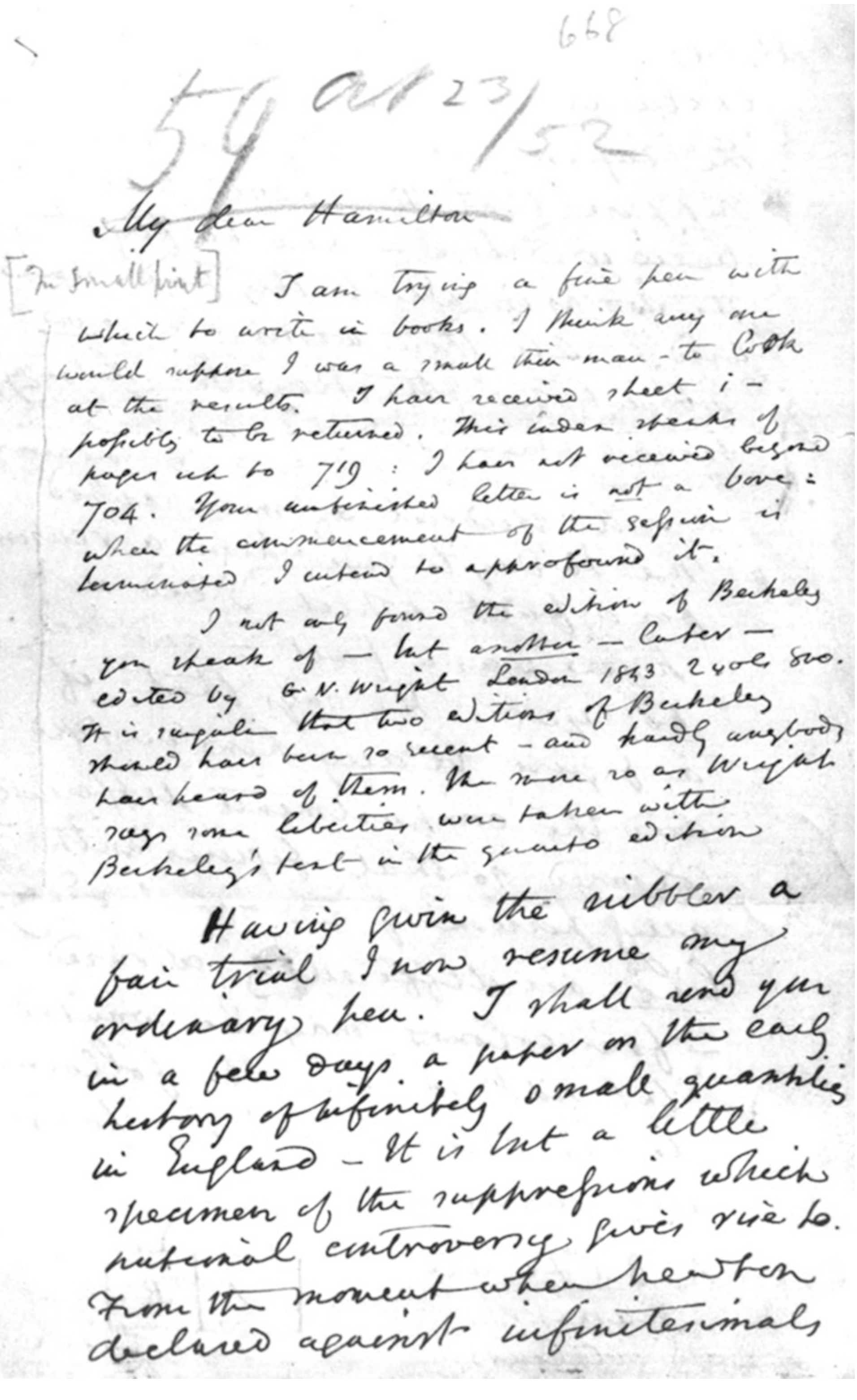
Chapter 19
How the Four-Color Conjecture Was Born

19.1 The Problem Is Born

It takes time and effort to gain access and read manuscripts. The two letters containing the first mention of the 4CC are of high importance; yet, to the best of my knowledge, their complete facsimiles have never been published before. Selected transcriptions served a purpose, but, as we will see in Section 19.2, they contained certain shortcomings. In view of this, I am reproducing here, for the first time, the facsimile of De Morgan's letter to Hamilton and the relevant fragment of Hamilton's reply. Seeing – and reading – these letters allows us to immerse in the World Victorian. Analysis of these documents and the corrected transcription of De Morgan's letter will follow. I am grateful to the Board of Trinity College Dublin, whose kind permission made reproducing these letters ([DeM1] and [Ham]) possible.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_19



668

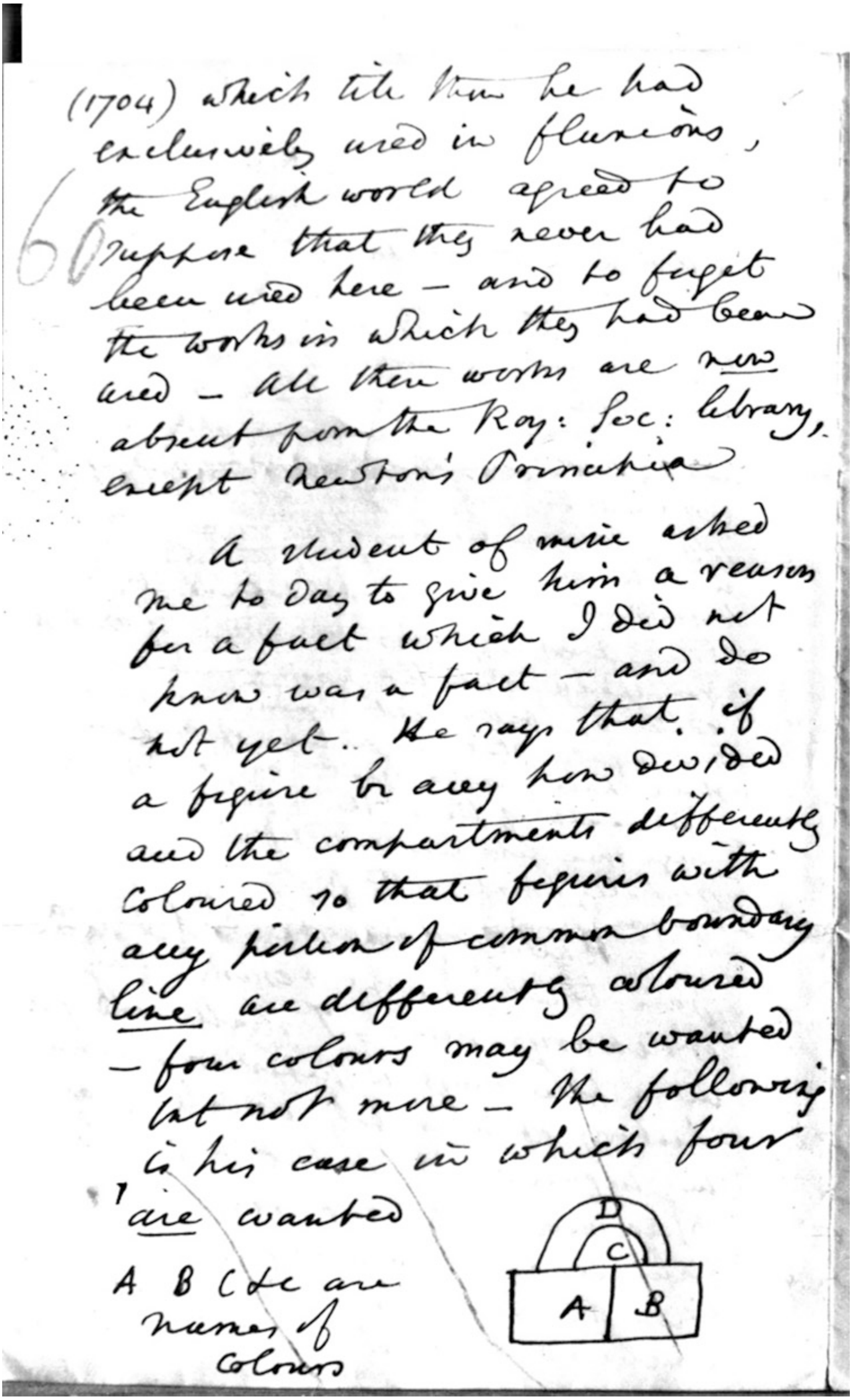
59 at 23/52

My dear Hamilton

[In small print] I am trying a fine pen with which to write in books. I think any one would suppose I was a small thin man — to look at the results. I have received sheet i — possibly to be returned. This index speaks of pages up to 719: I have not received beyond 704. Your unfinished letter is not above: when the commencement of the session is terminated I intend to approfound it.

I not only found the edition of Berkeley you speak of — but another — later — edited by G. N. Wright London 1843 2 vols 8vo. It is singular that two editions of Berkeley should have been so recent — and hardly anybody have heard of them. The more so as Wright says some liberties were taken with Berkeley's text in the quarto edition.

Having given the nibbler a fair trial I now resume my ordinary pen. I shall send you in a few days a paper on the early history of infinitely small quantities in England — It is but a little specimen of the suppressions which national controversy gives rise to. From the moment when Newton declared against infinitesimals



(1704) which title then he had exclusively used in fluxions, the English world agreed to suppose that they never had been used here — and to forget the works in which they had been used — All these works are now absent from the Roy: Soc: library, except Newton's Principia

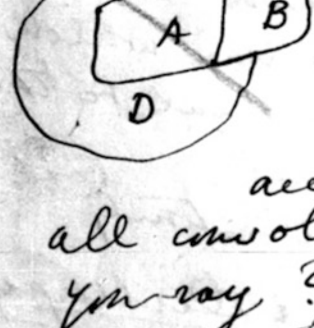
A student of mine asked me to day to give him a reason for a fact which I did not know was a fact — and do not yet. He says that if a figure be any how divided and the compartments differently coloured so that figures with any portion of common boundary line are differently coloured — four colours may be wanted but not more — the following is his case in which four are wanted

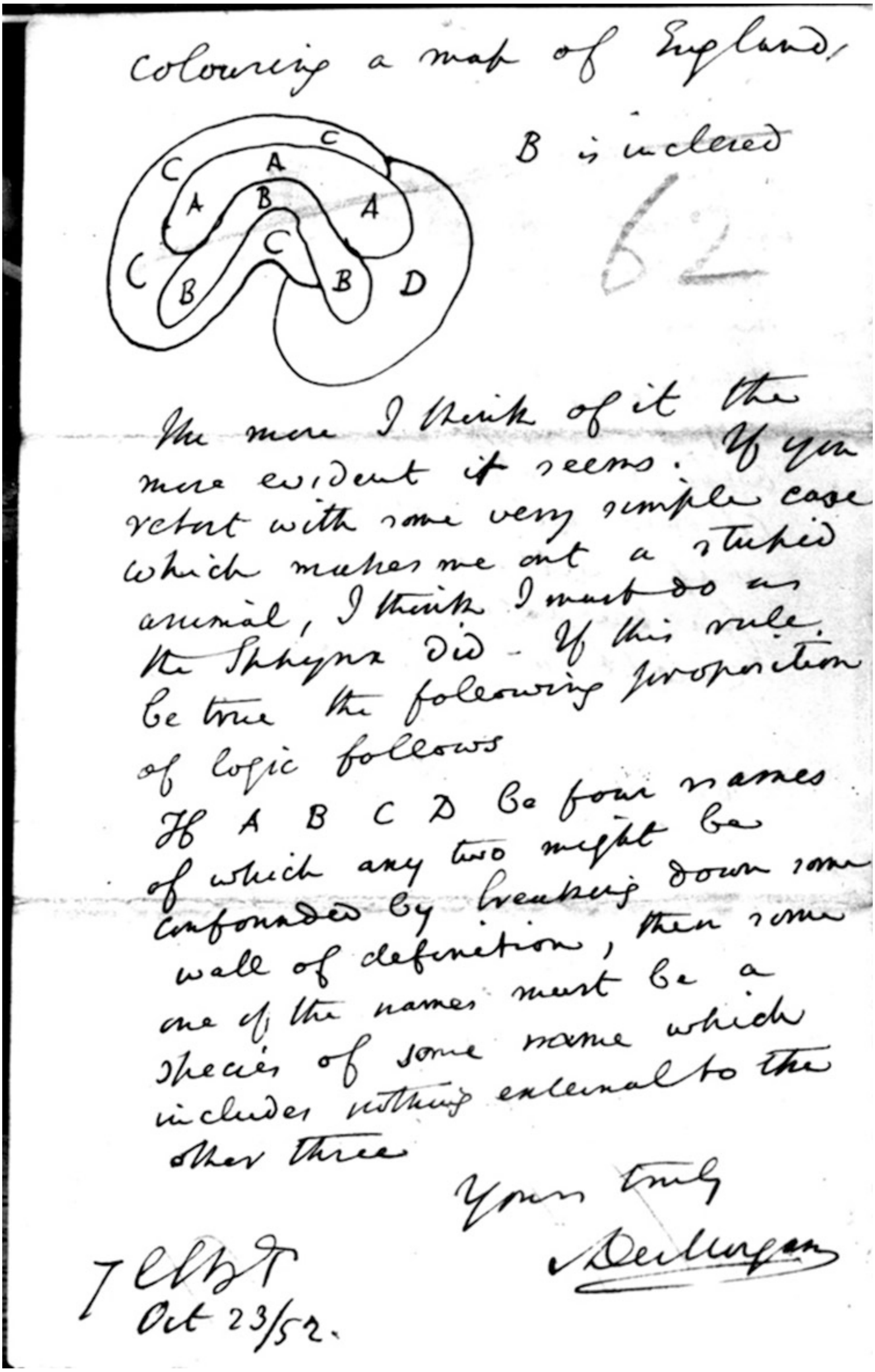
A B C &c are names of colours

Query cannot a necessity for five or more be invented?

As far as I see at this moment, if four ultimate compartments have each boundary line in common with one of the others, three of them inclose the fourth, and prevent any fifth from connexion with it. If this be true, four colours will colour any possible map without any necessity for ~~the~~ colour meeting colour except at a point.

Now it does seem that drawing three compartments with common boundary ABC two and two – you cannot make a fourth take boundary from all, except by inclosing one – But it is tricky work and I am not sure of all convolutions – What do you say? And has it, if true been noticed? My pupil says he guessed it in





colouring a map of England.

B is inclosed

62

The more I think of it the more evident it seems. If you retort with some very simple case which makes me out a stupid animal, I think I must do as the Sphynx did. If this rule be true the following proposition of logic follows

If A B C D be four names of which any two might be confounded by breaking down some wall of definition, then some one of the names must be a species of some name which includes nothing external to the other three

Yours truly

A De Morgan

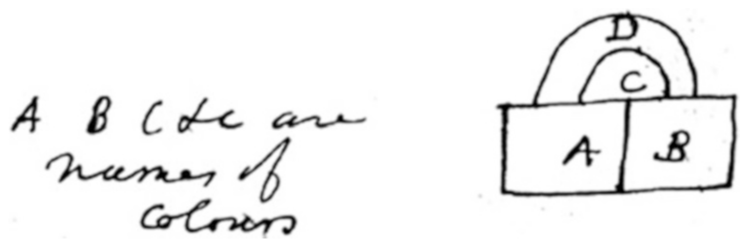
7 CBT
Oct 23/52.

Augustus De Morgan, letter to William R. Hamilton, October 23, 1852; courtesy of Trinity College Dublin

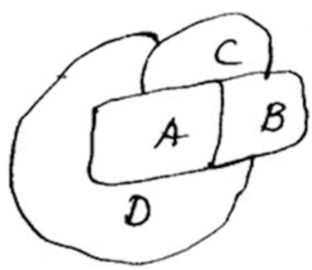
The written record of the problem begins with the October 23, 1852 letter that Augustus De Morgan, professor of mathematics at University College, London, wrote [DeM1] to Sir William Rowan Hamilton, professor of mathematics at Trinity College Dublin (the underlined words belong to the original manuscript; those in italics are mine):

My dear Hamilton[1]

. . . A student of mine asked me to day to give him a reason for a fact which I did not know was a fact – and do not yet. He says that if a figure be any how divided and the compartments differently coloured so that figures with any portion of common boundary line are differently coloured – four colours may be wanted but not more – the following is *his* case in which four are wanted [.]

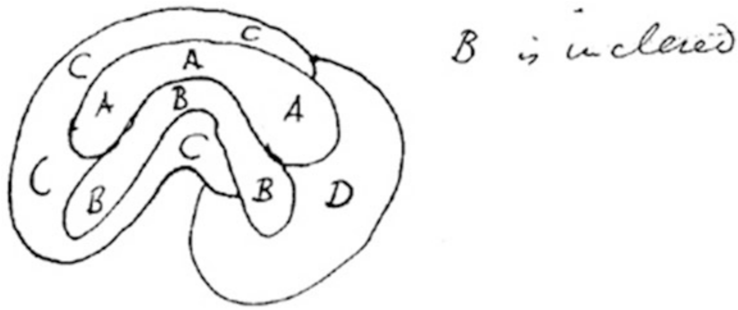


Query [:] cannot a necessity for five or more be invented [?] As far as I see at this moment, if four ultimate compartments have each boundary line in common with one of the others, three of them inclose the fourth, and prevent any fifth from connexion with it. If this be true, four colours will colour any possible map without any necessity for colour meeting colour except at a point.



[1] De Morgan, A., Letter to W.R. Hamilton, dated October 23, 1852; TCD MS 1493, 668; Trinity College Dublin Library, Manuscripts Department.

Now it does seem that drawing three compartments with common boundary A B C two and two – you cannot make a fourth take boundary from all, except by inclosing one – But it is tricky work and I am not sure of all convolutions – What do you say? And has it, if true [,] been noticed? My pupil says he guessed it in colouring a map of England [.]



The more I think of it the more evident it seems. If you retort with some very simple case which makes me out a stupid animal, I think I must do as the Sphynx did – If this rule be true the following proposition of logic follows [.]

If A B C D be four names of which any two might be confounded by breaking down some wall of definition, then some one of the names must be a species of some name which includes nothing external to the other three [.]

Yours truly

ADeMorgan [Signed]

7 CSCT[2]

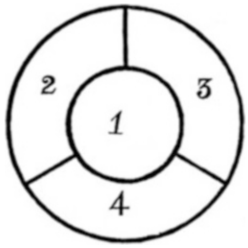
Oct 23/52

So, if Hamilton were to find a "very simple" solution to this puzzle, De Morgan "must do as the Sphynx did." What did the Sphynx do? In ancient mythology, the Sphynx offered hard riddles and viciously killed and ate those who could not solve the riddle. However, Oedipus solved the riddle, and, as a result, the Sphynx committed suicide. OK, we get it, De Morgan is the Sphynx, and if Hamilton finds an easy proof of De Morgan's four-color riddle, De Morgan would kill himself. But whom does De Morgan refer to in "student of mine asked me to day"? That student brought 4CC to De Morgan, and, thus, the student's identity may lead us to the author of 4CC. No one knew the student's name; perhaps, even De Morgan did not know or did not remember, for he had never mentioned the name, just "a student of his."

It was only 28 years later, in 1880, that Frederick Guthrie disclosed that he was the student mentioned by De Morgan in this letter. Frederick published his own account [GutFr], which

[2] These four letters must stand for De Morgan's address, which was 7 Camden Street, Camden Town.

reveals for the first time that the author of 4CC, the four-color conjecture, is Frederick's 2-year-senior brother, Francis:



Some thirty years ago, when I was attending Professor De Morgan's class, my brother, Francis Guthrie, who had recently ceased to attend them (and who is now professor of mathematics at the South African University, Cape Town), showed me the fact that the greatest necessary number of colors to be used in coloring a map so as to avoid identity of color in lineally contiguous districts is four. I should not be justified, after this lapse of time, in trying to give his proof, but the critical diagram was as in the margin.

With my brother's permission I submitted the theorem to Professor De Morgan, who expressed himself very pleased with it; accepted it as new; and, as I am informed by those who subsequently attended his classes, was in the habit of acknowledging whence he had got his information.

If I remember rightly, the proof which my brother gave did not seem altogether satisfactory to himself; but l must refer to him those interested in the subject.

Thus, we learn from the younger brother Frederick Guthrie, a professor of chemistry and physics, that 4CC was created by the 20-year-old student Francis Guthrie (of course, he may have been even younger when the conjecture first occurred to him) and that Francis Guthrie found a configuration showing that four colors are necessary and shared this simple configuration with his brother Frederick, who then passed it to De Morgan. There was likely more to Francis' proof, but it "did not seem altogether satisfactory to himself," as Frederick reports. *Will we ever learn what else Francis Guthrie, at such a tender age, deduced about this incredible mathematical conjecture?* Read on, for I am adding in Section 19.2 my new, June 7, 2020, historical conjecture about Francis' attempted proof!

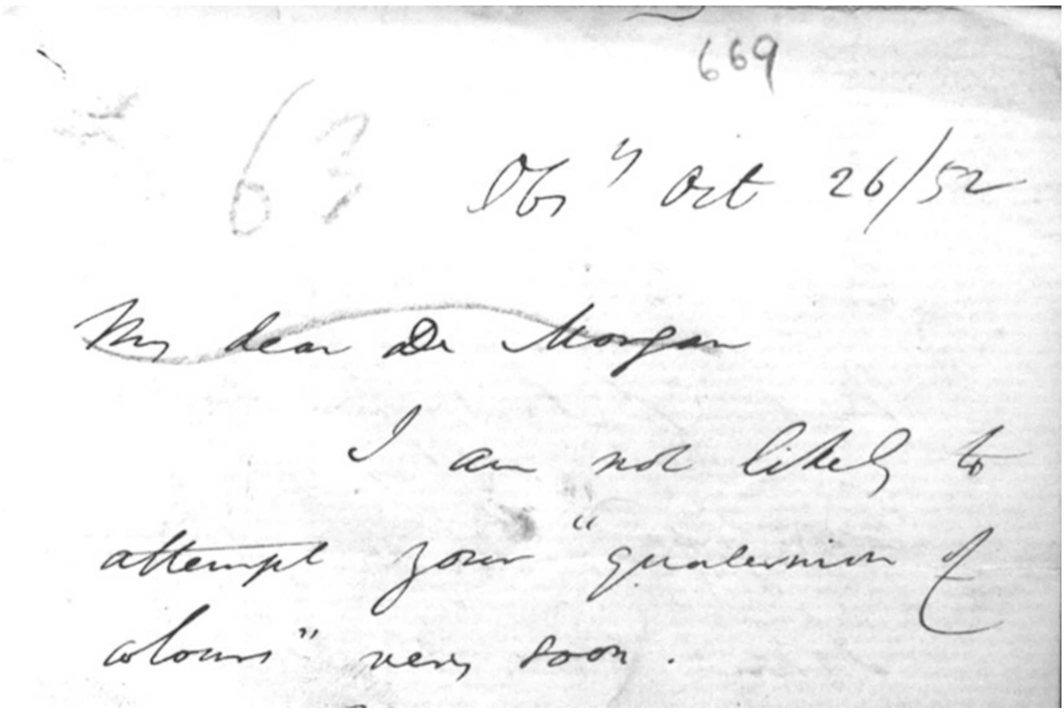
Let us roll back to De Morgan. The day he received the 4CC from Frederick Guthrie, i.e., on October 23, 1852, he *immediately* wrote about it to William Rowan Hamilton, who was not only one of the leading mathematicians at the time but also De Morgan's "intimate friend" and lifelong correspondent[3]. Hamilton's October 26, 1852, reply (the Royal Post must have worked extremely well, as there were only 3 days between the dates of De Morgan's letter and Hamilton's reply) is also preserved in the manuscript collection of Trinity College Dublin.

[3] When W.R. Hamilton died, De Morgan wrote about it in his September 13, 1865, letter to Sir J.F.W. Hershel [DeM5]: "W.R. Hamilton was an intimate friend whom I spoke to once in my life – at Babbage's, about 1830; but for thirty years we have corresponded."

Hamilton, apparently, was so obsessed with his discovery of *quaternions*[4] that he could not make himself interested in the coloring of maps [Ham]:

My dear De Morgan[5]

I am not likely to attempt your “quaternion of colors” very soon . . .


669

Oct 26/52

My dear De Morgan

I am not likely to attempt your "quaternion of colours" very soon.

William R. Hamilton, letter to A. De Morgan, October 26, 1852; reproduced by kind permission by the Board of Trinity College Dublin

That is all! Just the Victorian, most cordial way of saying “I am not interested, lay off my back.” De Morgan was left alone to keep the 4CC alive, and he succeeded. He repeatedly mentioned the problem in his lectures at University College ([GutFr]) and formulated it in his letters (we know a few such instances: [DeM1], [DeM2],[6] and [DeM3][7]). As discovered in 1976 by John Wilson, a high school teacher from Eugene, Oregon [WilJ], De Morgan also became the first to publish the problem in his April 14, 1860, unsigned long review in *The Athenaeum* [DeM4] of W. Whewell’s book *The Philosophy of Discovery*:

[4] Arguably, this obsession prevented Hamilton from inventing linear algebra.

[5] Hamilton, W.R., Letter to A. De Morgan, October 26, 1852; TCD MS 1493, 669; Trinity College Dublin Library, Manuscripts Department.

[6] Locations of both [DeM2] and [DeM3] come from N.L. Biggs [Big], who analyzed De Morgan’s contribution to the 4CC and the separation axiom.

[7] First found by Bertha Jeffreys in 1979 [JeffB].

> When a person colours a map, say the counties in a kingdom, it is clear he must have so many different colours that every pair of counties which have some common boundary *line* – not a mere meeting of two corners – must have different colours. Now, it must have been always known to map-colourers that *four* different colours are enough.

De Morgan's notion of acquaintance of cartographers with the sufficiency of 4 colors was a silly invention – there is no evidence whatsoever that cartographers (then or now for that matter) knew about it or needed to minimize the number of colors, since the juxtaposition of colors and the addition of textures create sufficient representations for scores of additional colors. While De Morgan did not advance the solution of 4CC at all, he single-handedly popularized it for decades and ensured its long life. Even the mathematician of the day, Arthur Cayley, got hooked on 4CC but was unable to prove it. In the report based on the June 13, 1878, meeting of the London Mathematical Society, we read ([Cay1] and [Cay2]):

> Questions were asked by Prof. Cayley, F.R.S. (Has a solution been given of the statement that in colouring a map of a country, divided into counties, only four distinct colours are required, so that no two adjacent counties should be painted in the same colour?)

Cayley also published a two-page article [Cay3] on the question. Does his choice of the publication, *Proceedings of the Royal Geographical Society and Monthly Record of Geography*, suggest that Cayley believed De Morgan about the usefulness of 4CC for mapmakers? Perhaps, not, but the coincidence adds a touch of humor to our Victorian story. In his paper, Cayley shows that it suffices to prove 4CC for *trivalent* maps, as they are called now (i.e., maps in which three regions meet at every vertex):

> The theorem, if it is true at all, is true under more stringent conditions: if in any case the Figure includes four or more areas meeting in a point (such as the sectors of a circle), then if (introducing a new area) we place at the point a small circular area, cut out from and attaching itself to each of the original sectorial areas, it must according to the theorem be possible with four colours only to colour the new Figure; and this implies that it must be possible to colour the original Figure so that only three colours (or it may be two) are used for the sectorial areas. And in precisely the same way (the theorem is in fact the same) it must be possible to colour the original Figure in such wise that only three colours (or it may be two) present themselves in the exterior boundary of the Figure.

Finally, Cayley tries to explain at length the difficulty of proving 4CC by a straightforward induction and that was all he was able to do. Arthur Cayley states twice – in the course of two pages – that he "failed to obtain a proof" of 4CC. These statements by one of the great mathematicians of his time must have stirred interest in 4CC. Professionals and amateurs alike jumped on the opportunity to make Cayley out to be "a stupid animal," as De Morgan put it in his letter quoted above.

The proof was very soon found and published in 1879 in a prestigious American (!) journal by Alfred Bray Kempe, a 30-year-old London barrister (lawyer) and an avid amateur mathematician and an expert on linkages[8]. We will look at his work in the next chapter.

[8] His 1877 book *How to Draw a Straight Line* was published again 100 years later in the United States by the National Council of Teachers of Mathematics, with a funny (for 1977) statement on the copyright page: "Alfred Bray Kempe, 1849–" indicating Kempe's very long life indeed.

19.2 A Touch of Historiography and a Historical Conjecture

It is extremely surprising that for more than 100 years, confusion reigned in the history of the four-color conjecture (4CC), one of the most popular problems in the history of mathematics. Truth and fiction alternated like positive and negative parts of a sin curve. Without presenting here a complete historiography of the problem, I would just mention that the Möbius–Weiske's puzzle is mixed up with 4CC countless times and, consequently, credit for 4CC is often given to Möbius. It has been happening even in relatively recent times. For example, as late as in 1958, the great geometer H.S.M. Coxeter wrote [Cox]:

> The Four-Color Theorem [sic] was first mentioned by Möbius [sic].

There are, however, authors, who present the problem's history without fantasy and "inventions." For example, Alfred Errera puts it about right in his December 1920 doctoral thesis [Err]:

> Cayley attributed the exposition of the map theorem [sic] to De Morgan, whereas Frederic Guthrie claimed, in 1880, that his brother Francis Guthrie had demonstrated [it] some thirty years earlier.

In 1965, Kenneth O. May summarized 4CC's history extremely well [May]. Apparently, he was the first to quote De Morgan's letter:

> A hitherto overlooked letter from De Morgan to Sir William Rowan Hamilton.

May then went on to quote De Morgan's October 23, 1852, letter and Hamilton's reply from the monumental three-volume edition *Life of Sir William Rowan Hamilton*, 1882–1889, written by Hamilton's close friend, the Rev. Robert Perceval Graves [Grav]. Volume 3 includes Hamilton's correspondence with De Morgan, and the letter of our interest, De Morgan to Hamilton of October 23, 1852, appears on pages 422–423. Graves was pressed for space – he wrote (vol. 3, p. v):

> The . . . larger portion of the volume [673 pp. long] consists of a selection from a very extensive correspondence between Sir W. R. Hamilton and Professor Augustus De Morgan . . . The quantity of material was so great that I have had to exclude matter that possessed inherent value, either because it was in subject unsuited to this work, or because, being mathematical, the investigations carried on were too abstruse or too extended. The general reader will perhaps complain that I have introduced more than enough of mathematical investigation; but he will, I hope, withdraw the complaint when he calls to mind that it was as scientific men that the writers corresponded, that it would be unjust to them if their correspondence as printed should not retain this character, and that the mathematical discussion did in fact most often afford suggestion to the play of thought which, passing beyond the boundaries of science, prompted the wit and the learned and pleasant gossip which the readers will enjoy.

Thus, May knew that Graves condensed letters – in fact, Graves used quotation marks to show in practically every letter that he published selections and not complete letters. Graves favored "pleasant gossip" indeed. For example, in De Morgan's letter of our interest, he keeps in De Morgan's trying "a fine pen with which to write in books," and "Having given the nibbler a fair trial, I now resume my ordinary pen." However, Graves – and consequently May – omitted all De Morgan's mathematical drawings, illustrating the first ever thoughts on the four-color conjecture, and they both omitted an important phrase. As a fine historian, May should have looked at these important letters in the manuscript at the place where Hamilton spent his life, the place that sponsored Graves' voluminous biography of W.R. Hamilton, which appeared in the *Dublin University Series*, Trinity College Dublin. He would have found there the 1900 *Catalog of the Manuscripts in the Library of Trinity College Dublin* compiled by T.K. Abbott, where I read ([Abb], p. v):

> In 1890 the Rev. Robert P. Graves presented [Trinity College, Dublin] a collection of mss. which had belonged to Sir W. R. Hamilton, including his correspondence with Sir John Herschel, Professor De Morgan, and others.

In fact, the two letters of our interest, catalogued as 668 and 669 (De Morgan's and Hamilton's, respectively) are contained in the group TCD MS 1493 of Hamilton–De Morgan correspondence manuscripts, which was donated to Trinity College Dublin in 1900, as Stuart Ó Seanór, assistant librarian of the Manuscripts Department at Trinity College, disclosed to me in a letter from March 21, 1997 [OSe]:

> TCD MS 1493 was presented by J R H O'Regan of Marlborough, Wilts in 1900 (a descendant of Hamilton's through his daughter Helen) just in time to be mentioned in T K Abbott's *Catalogue of the manuscripts in the library of Trinity College Dublin* published that year...
>
> Graves' three volume biography of Hamilton or other writings of his may reveal that Hamilton corresponded with De Morgan and even citation of them might date from before the papers were in a library.
>
> Le meas
>
> Stuart Ó Seanór [signed]

By now, you must be wondering, which important phrase is missing in Graves and May letter; I am putting it in italics:

> He says that if a figure be any how divided and the compartments differently colored so that figures with any portion of common boundary line are differently colored – four colors may be wanted but not more – *the following is his case in which four are wanted* [.]

The missing phrase was restored in 1976 by Norman L. Biggs, E. Keith Lloyd, and Robin J. Wilson in their wonderful textbook on graph theory through its history [BLW]. To do that, the authors clearly had to see the manuscript letter or its photocopy. Unfortunately, they misread a word while transcribing the missing phrase, and the wrong word appeared in various editions of their book [BLW] as follows:

> . . . the following is the [sic] case in which four are wanted [.]

In the manuscript, one can clearly see the word "his" where the authors of [BLW] put the second "the." The difference is subtle but important: "the following is the case" would have indicated that De Morgan showed to Hamilton his own example.[9] In fact, De Morgan wrote

[9] The authors of [BLW] misread another word as well: they quote De Morgan as "I am not sure of the [sic] convolutions," whereas De Morgan wrote "I am not sure of all convolutions," which makes more sense.

"the following is *his* [i.e., student's] case," i.e., De Morgan conveyed an example that four colors are wanted, which Francis Guthrie devised and passed on to De Morgan through his brother Frederick. Now, I present my new historical conjecture that appears for the first time in this expanded edition of *The Mathematical Coloring Book*:

Historical Conjecture 19.1 (Soifer, 2020) Arguments and drawings in De Morgan's October 23, 1852, letter belong to Francis Guthrie.

A Plausible Argument We have established that at least one example and one drawing in De Morgan's letter came from Francis Guthrie. *I now believe that likely all arguments and all drawings in the letter must have belonged to Francis Guthrie as well. After all, on October 23, 1852, in addition to his daily routine chores, De Morgan taught his course(s) and wrote to Hamilton his long, four-page letter with a good number of drawings and arguments. He did not have the time to ponder the map coloring problem and must have shared with Hamilton Francis Guthrie's four-color conjecture and Guthrie's attempted proof!*

We have thus established that De Morgan's contemporaneous account agrees with Frederick Guthrie's 1880 recollection: Frederick presented to De Morgan not only the four-color conjecture but also his brother Francis' "*proof*," albeit "*not altogether satisfactory to himself* [i.e., to Francis]," as Frederick put it.

The history of 4CC was enriched by the 1976 discoveries by John Wilson [WilJ] and in 1979 by Bertha S. Jeffreys of Cambridge, England [JefB], who found additional examples of De Morgan's writings about 4CC.

19.3 The Creator of the Four-Color Conjecture: Francis Guthrie

I find it fascinating to read old newspapers: yes, their life span is 1 day, and, for people of the day, they become worthless the day after their publication. However, they depict that 1 day in many ways better than any other sources. For a reader, a century or centuries later, newspapers are a treasure trove of the life's interests and people's aspirations of the day. They allow us to "touch" the distant culture and to breathe its air.

I was looking at *Cape Times* from Monday, October 23, 1899. In the column of my interest first came *The America Cup*:

> The possession of [the celebrated sailing] America Cup was decided to-day when the Columbia won her second race against the Shamrock by five minutes. The Cup therefore remains in America.
>
> All-important for South African people of the day, *Ship's Movements* came next:
>
> The Clan Macpherson left Liverpool for Algon Bay on Thursday morning.
>
> The Pombroks Castle arrived at Plymouth at two on Thursday afternoon.
>
> The Spartan left St. Vincent last night.

Following these 1.75-inch-long reports, I saw something that must have mattered to the folks of the colony of South Africa a great deal: a 22-inch-long (!) column of *The Late Professor Guthrie*. Let us read a bit of it together [Gut1]:

There has just passed away from us a man who has left a greater mark upon our Colonial life than will be readily recognized by many who did not come into contact with him; or by some who have been taught by this age of self-advertisement to suppose that no good work can be done in modesty and retirement. Professor Francis Guthrie, L.L.B., B.A., whose death on the 19th ... we briefly announced in Saturday's issue, was born in London in 1831.

Francis Guthrie; courtesy of John Webb and the Mathematics Department, University of Cape Town

We can learn much about Francis Guthrie from this eulogy [Gut1] and from [Gut2] and [Gut3].

> Born on January 22, 1831 in London, Francis Guthrie received his degree B.A. with first class honors from the University College, London. He then earned L.L.D., a law degree, and for some time was a consulting barrister in Chancery practice. In 1861 Guthrie left the old world and accepted an appointment at the newly established Graff-Reinet College in the Colony of South Africa. Following his resignation in 1875, and a brief visit to England, in 1876 Guthrie was appointed to the Chair of Mathematics in the South African College, Cape Town [presently called the University of Cape Town], from which he retired after 22 years on January 31 of 1899. Several months later, on October 19, 1899, Guthrie died in Claremont, Cape Town.

Professor Guthrie was universally liked and respected by his peers. He served on the University Council, 1873–1879, and was secretary of the Senate in 1894. He was an early member of the South African Philosophical Society (now the Royal Society of South Africa) and of its Council, a member of the Meteorological Commission, and, for many years, the examiner of the Cape University.

His several publications cover mathematics (none on 4CC), meteorology, and his true passion: botany. Guthrie and his lifelong friend Harry Bolus were pioneers in the study of *ericas* of Southern Africa. In 1873, Harry Bolus discovered a new genus on the summit (altitude 6,500 feet) of the Gnadouw–Sneeuwbergen near Graff–Reinet. Bolus named it in honor of his friend *Guthriea capensis*.

I am compelled to return to *Cape Times* [Gut1], as it conveys the life of the frontier unknown to most of us through personal experience and shows a side of Francis Guthrie that is not widely known. Guthrie was a pioneer of the frontier. He discovered not only the four-color conjecture but also routed for the railroad that determined the future of his region of South Africa:

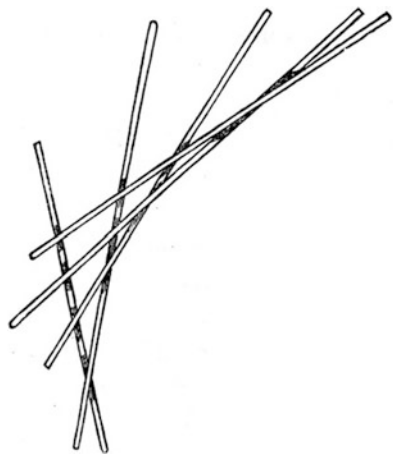
> In 1871–1873, when the agitation for railway extension was at its height and the battle of the routes was being fought, Professor Guthrie ardently espoused the Midland cause. The problem of that day was to show the Government and Parliament how, if a railway were made to Graaff-Reinet, it could get over the Sneeuwberg Mountains to the northwards. Some case had to be made out before the Government would sanction even a flying survey. Professor Guthrie, in a company with the late Charles Rubidge and some others, climbed the mountains, aneroid in hand, in search of the most available pass. Their efforts had for immediate result the construction of Forth Elizabeth and Graaff-Reinet line; and it is a tribute to the accuracy of those early amateur railway explorers that the more recent extension of that line to Middelburg follows very nearly the route over the Lootsberg which they had suggested as the most feasible. The people of Graaff–Reinet were not ungrateful, and a public banquet and laudatory addresses showed their appreciation of the efforts of Professor Guthrie and his colleagues.

This remarkable 22-inch-long eulogy ends with unattributed poetic lines, which I traced to James Shirley (1596–1666):

> Only the actions of the just
> Smell sweet, and blossom in the dust.

19.4 The Brother

While we are moving through the Victorian history of the problem, I can offer you something mathematical to do as well. Frederick Guthrie (1833–1886), who, by 1880, was a professor of chemistry and physics at the School of Science, Kensington, and the younger brother of the 4CC creator Francis Guthrie, in his letter, quoted above [GutFr], created and solved a three-dimensional analogue of 4CC that Francis allegedly neglected:



> I have at various intervals urged my brother to complete the theorem in three dimensions, but with little success.
>
> It is clear that, at all events when unrestricted by continuity of curvature, the maximum number of solids having superficial contact each with all is infinite. Thus, to take only one case n straight rods, one edge of whose projection forms the tangent to successive points of a curve of one curvature, may so overlap one another that, when pressed and flattened at their points of contact, they give $n - 1$ surfaces of contact.

Thus, Frederick Guthrie posed and solved the following problem:

Problem 19.2 (Frederick Guthrie, 1880) Is there a positive integer n such that n colors suffice for proper coloring of any three-dimensional Euclidean map?

Frederick Guthrie continues:

> How far the number is restricted when only one kind of superficial curvature is permitted must be left to be considered by those more apt than myself to think in three dimensions and knots.

Guthrie's prose is imprecise. It seems to me that he intended to pose the following problem:

Problem 19.3 (Frederick Guthrie, 1880) What is the minimum number of colors required for proper coloring of any three-dimensional Euclidean map if each monochromatic set is convex?

I am compelled to allow you the time to ponder an alternative solution to Problem 19.2 and a solution to Problem 19.3. We will return to them in Chapter 21.

Chapter 20
A Victorian Comedy of Errors and Colorful Progress

20.1 A Victorian Comedy of Errors

This period in the history of the four-color conjecture (4CC) plays itself out like a Victorian version of Shakespeare's *The Comedy of Errors.* Judge for yourself!

Alfred Bray Kempe's proof of 4CC was announced on July 17, 1879, in *Nature* [Kem1]. The proof itself was published later the same year in the *American Journal of Mathematics, Pure and Applied* [Kem2], as Kempe puts it (p. 194), "at the request of the Editor-in-Chief," i.e., James Joseph Sylvester. Being Jewish, the famous English mathematician J.J. Sylvester could not get a job at Oxford or Cambridge, as those were Christian institutions. Following many jobs at many places in England, the United States, and, again, England, Sylvester was invited to the New World as an inaugural professor of mathematics at the newly founded Johns Hopkins University, to the great benefit of the young American mathematicians. But that is another story.

Apparently, Kempe believed in mapmakers' myth of De Morgan and even expanded it:

> It has been stated somewhere by Professor De Morgan [must be a reference to Athenaeum [DeM4]] that it has long been known to map-makers as a matter of experience – an experience however probably confined to comparatively simple cases – that *four* colours will suffice in any case.

Kempe entitled his paper to fit the myth, *On the Geographical Problem of the Four Colours*.

The proof was an unqualified success. While Kempe was elected a fellow of the Royal Society based on this work on linkages, the coloring success might have been a factor in Cayley, Sylvester, and others nominating him for the honor.

Soon, a number of authors produced simplifications and variations of Kempe's proof. The first one came from William E. Story, associate editor in charge of the *American Journal of Mathematics, Pure and Applied* ([Sto]). Story's paper immediately followed Kempe's article [Kem2]. Simplifications then came from Kempe himself ([Kem3] and [Kem4]). They were followed by the new "series of proofs of the theorem that four colours suffice for a map" by Peter Guthrie Tait ([Tai1], [Tai2], and [Tai3]).

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_20

The popularity of the four-color theorem (4CT) became so overwhelming that in late 1886, the headmaster of Clifton College somehow learned about it and offered the problem as a "challenge problem" to his students:

> In colouring a plane map of counties, it is of course desired that no two counties which have a common boundary should be coloured alike; and it is found, on trial [sic] that four colours are always sufficient, *whatever the shape or number* of the counties or areas may be. Required, a good proof of this. Why *four?* Would it be true if the areas are drawn so as to cover a whole sphere?

In the funniest turn of this story, the headmaster warned the contestants that "no solution may exceed one page, 30 lines of MS., and one page of diagrams"! Published on January 1, 1887, in the *Journal of Education* [Head1], the challenge attracted a solution from such an unlikely problem-solver as the bishop of London, whose "proof" [Head2] was published in the same journal on June 1, 1889.

Let us give a compliment to the headmaster for his unexpectedly great question: "Why four?" Even today, although we have two proofs of 4CC (see Chapters 22 and 24), we still do not really know the answer to this innocent question.

Then comes the 29-year-old Percy John Heawood – and spoils the party! Almost with regret for his own discovery [Hea1], Heawood apologetically writes:

> The present article does not profess to give a proof of this original Theorem [i.e., 4CT]; in fact its aims are rather destructive than constructive, for it will be shown that there is a defect in the now apparently recognized proof.

Yes, 11 (eleven!) years after Kempe's 1879 publication [Kem2], Heawood discovered a hole in it (as well as in the later two versions of Kempe's proofs ([Kem3] and [Kem4])). Moreover, Heawood constructed an example that showed that Kempe's argument as it was, could not work. There was a positive side to Heawood's paper: he showed that *Kempe's argument* actually proved that *five* colors suffice for coloring any map.

In a gentlemanly way, Heawood informed Kempe first, and Kempe was the one who cordially accepted Heawood's findings and reported them to the London Mathematical Society at its Thursday, April 9, 1891, meeting, with "Major P. A. MacMahon, R.A., F.R. S., Vice-President, in the Chair" [Kem5]:

> Mr. Kempe spoke on the flaw in his *proof* "On the Map-Colour Theorem," which had recently been detected by Mr. P. J. Heawood and showed that a statement by the latter at the close of his paper failed. He further stated that he was unable to solve the question to his satisfaction.

The authors of [BLW] researched publications of the period at hand. They reported that they found "no complimentary references to Heawood in the popular journals, and no record of honours granted to him." Heawood's work [Hea1] and his consequent papers dedicated to map coloring are certainly major contributions and deserve better recognition. As it were, Heawood's work [Hea1] remains almost unnoticed and unquoted by his contemporaries. Long after 1890, we can find papers still giving credit to Kempe and Tait for proving 4CT (see, for example, [DR]).

While giving credit to Kempe, Tait offered his own “proofs.” It appears that the belief in Kempe’s proof is extrapolated by the contemporaries to the belief in Tait’s proofs: I was unable to find any contemporaneous refutation of Tait’s “proofs.”

Tait described his strategy as follows [Tai1]:

> The proof of the elementary theorem is given easily by induction; and then the proof that four colours suffice for a map follows almost immediately from the theorem, by an inversion of the demonstration just given.

This is true: Tait found a nice proof that his “elementary theorem” is equivalent to 4CT. The trouble is, it is not so “elementary,” and, moreover, its proof is not “given easily by induction” and, in fact, is not given at all.

The bishop of London erred too: he mistakenly believed that the Möbius–Weiske’s problem IV–1 was equivalent to 4CT. The direct refutation of his “proof” was published many years later, in 1906, by John C. Wilson [JWil]. Both De Morgan and Cayley, nearly half a century earlier, knew that 4CC was much more than a mere fact that five countries in a map cannot be mutually adjacent. Obviously, the headmaster of Clifton College and the bishop of London did not.

True to its genre, our *Comedy of Errors* had a happy end. Alfred Bray Kempe eventually became the president of the London Mathematical Society. Frederick Temple, our bishop of London, reached the highest religious title of the Archbishop of Canterbury.

Let me translate for you the ending of “The Fairytale about the Gold Cockerel” (“Сказка о золотом петушке”) by the great Russian poet Aleksandr Pushkin:

> *Tale is lie, but with a hint,*
> *Bright young lad can learn a bit.*[1]

Accordingly, our Victorian *Comedy of Errors* leaves us plenty of valuable and enjoyable mathematics. The bright ideas of Kempe, Tait, and Heawood are alive and well. Get your paper and pencil ready: in this chapter and in the next one, we will look at our British Victorian inheritance. As the Bard put it,

> *All’s well that ends well!*

20.2 2-Colorable Maps

Let us now visit some of the Victorian problems. To simplify the excursion, we will translate the Victorian problems into today’s jargon. I suggest we start with a warmup.

Problem 20.1 Prove that a map formed in the plane by finitely many circles can be two-colored (Fig. 20.1).

[1] The original Pushkin’s Russian text is as follows:
Сказка ложь, да в ней намек!
Добрым молодцам урок.

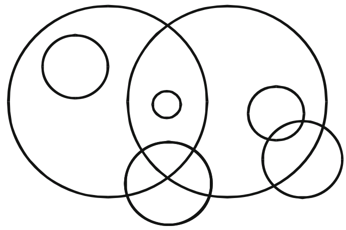


Fig. 20.1 A map formed by circles

Proof Partition regions of the map into two classes (Fig. 20.2): those contained in an even number of circles (color them gray) and those contained in an odd number of circles (leave them white). Clearly, the neighboring regions get different colors because when we travel across their common boundary line, the parity changes. ■

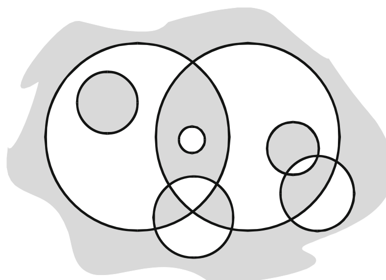


Fig. 20.2 2-coloring of a map formed by circles

I am sure you realize that the shape of a circle is of no consequence. We can replace circles in Problem 20.1 by their continuous one-to-one images, called *simple closed curves*, because the Jordan curve theorem holds for them all[2]:

Jordan Curve Theorem 20.2 A simple closed curve in the plane divides the plane into two regions (inside and outside).

Problem 20.3 Prove that a map formed in the plane by finitely many simple closed curves is 2-colorable.

We can replace simple closed curves by straight lines:

Problem 20.4 Prove that a map formed in the plane by finitely many straight lines is 2-colorable (Fig. 20.3).

[2] See its proof, for example, in [BS]

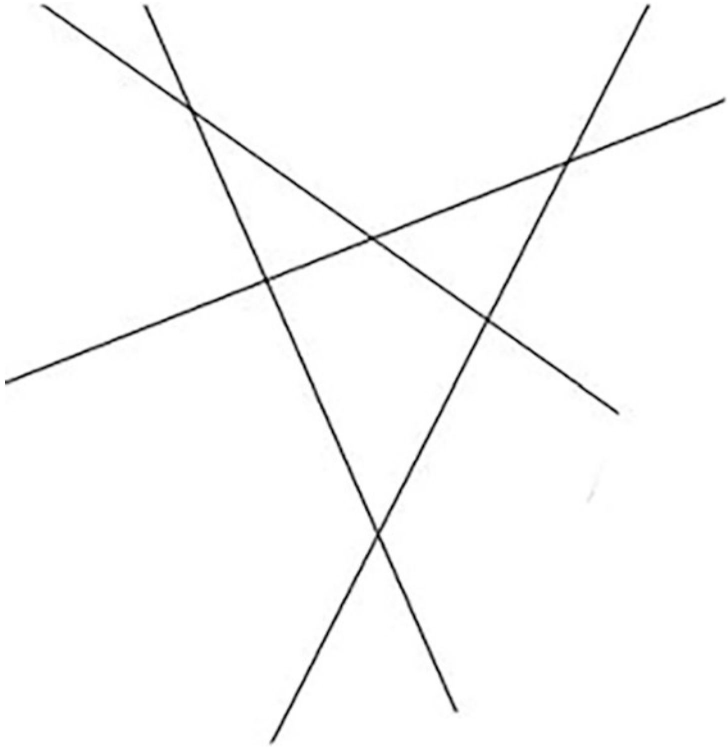


Fig. 20.3 A map formed by straight lines

An inductive proof is well-known,[3] but, as is usually the case with proofs by induction, it does not provide an insight. I found a "one-line" proof that creates similarity between simple closed curves and straight lines.

Proof Attach to each line a vector perpendicular to it (Fig. 20.4). Call the half-plane *inside* if it contains the vector and *outside* otherwise. Repeat the proof of Problem 20.1 word by word. (Fig. 20.5). ■

[3] See, for example, [DU].

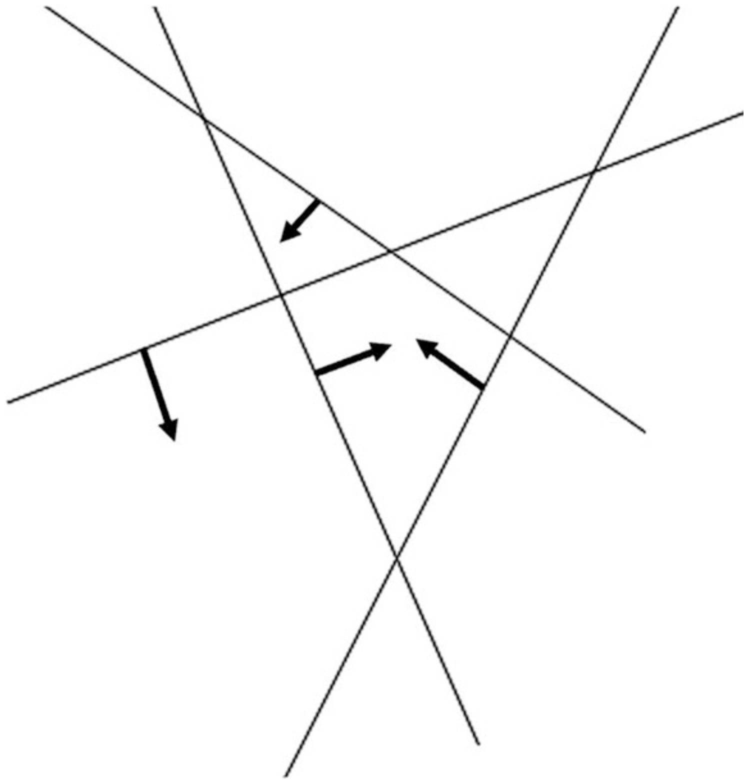


Fig. 20.4 Attaching vectors to lines

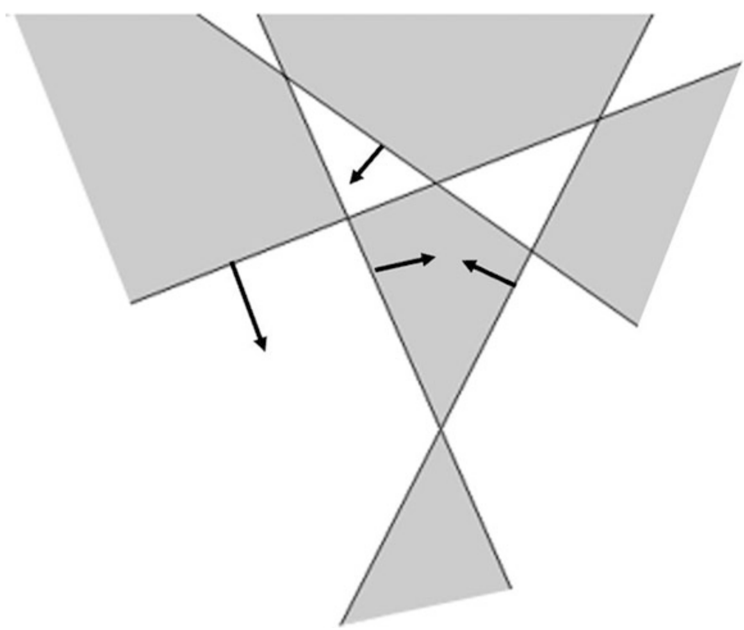


Fig. 20.5 2-coloring of a map formed by straight lines

Problem 20.5 Prove that a map formed in the plane by finitely many simple closed curves and straight lines is 2-colorable.

So, what is in common between simple closed curves and straight lines? What allows a 2-coloring to exist? Each vertex in the above maps is a result of the intersection of two or more curves or lines and, therefore, has an even degree. This fact first appears in print on the last page of the 1879 paper by Alfred Bray Kempe in which he attempted to prove 4CC [Kem2].

Kempe's Two-Color Theorem 20.6 (A.B. Kempe, 1879, [Kem2]). A map is 2-colorable if and only if all its vertices have even degrees.

Let us take another look at the map M formed by circles in Fig. 20.1. We can construct the *dual graph* $G(M)$ of the map M as follows: we represent every region by a vertex (think of the capital city) and call two vertices adjacent if and only if the corresponding two regions are adjacent, i.e., they have a common boundary line.[4] The dual graph $G(M)$ of the map M of Fig. 20.1 is presented in Fig. 20.6 (I bent and stretched the edges to make the graph look aesthetically pleasing).

Observe: The dual graph $G(M)$ of any map M is planar: We can draw its edges through common boundaries of the adjacent regions so that the edges will have no points in common, except the vertices of the graph.

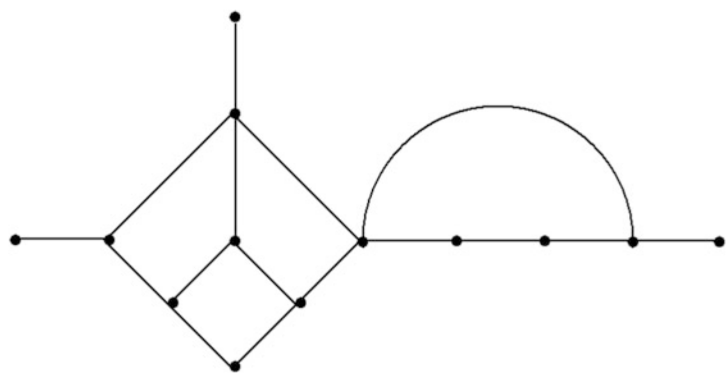


Fig. 20.6 The dual graph of the map from Fig. 20.1

Now the problem of coloring maps can be translated into the language of coloring vertices of planar graphs. But wait a second; this problem is not new to us: we have already solved it as Problem 12.2. Let us repeat it here:

Kempe's Two-Color Theorem 20.7 (*In Graph-Theoretical Language*). The chromatic number $\chi(G)$ of a graph G does not exceed 2 if and only if G contains no odd cycles.

[4] The idea of the dual graph of a map was one of the first ideas of graph theory: Leonard Euler used it in 1736 to solve the Problem of Bridges of Königsberg. The language of maps was universally used by the first researchers of 4CC. Yet, I noticed that while Kempe used the language of maps in the main body of his 1879 paper [Kem2], he did describe the construction of the dual graph on the last page of this paper.

20.3 3-Colorable Maps

It is natural to give a name to the smallest number of colors required to color a map M; let us call it the *chromatic number of map M* or the *face chromatic number* and denote it by $\chi_2(M)$.

We have an abundance of maps of chromatic number 2 around us: maps created by circles, by straight lines, and by simple closed curves (Problems 20.1–20.5). Square grids deliver us examples of large periodic maps of chromatic number 2: just recall the chessboard coloring. Can you think of a way of creating large periodic maps of chromatic number 3? You have already seen a couple of such constructions in this book, but in a totally unrelated context.

Problem 20.8 Find the chromatic number of the hexagonal map created by the old Chinese lattice in Fig. 6.7.

Solution Behold (Fig. 20.7):

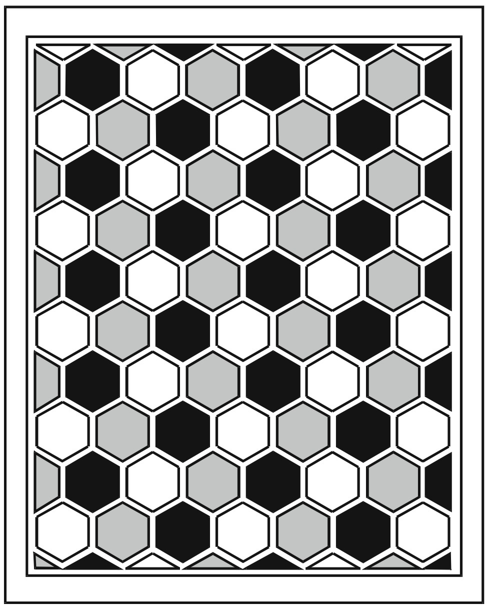


Fig. 20.7 3-coloring of a Chinese lattice

Problem 20.9 Find the chromatic number of the map in Fig. 6.6, which is formed by octagons and squares.

Solution Behold (Fig. 20.8):

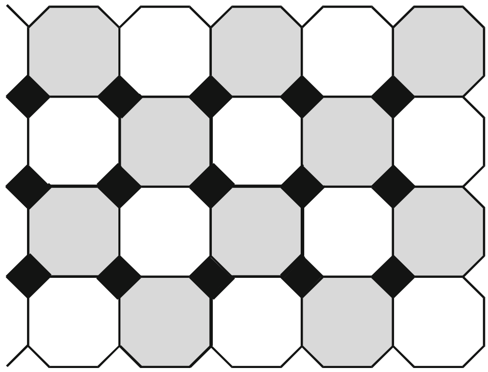


Fig. 20.8 3-coloring of the Soifer tiling of the plane from Fig. 6.6

What is special about the maps in Problems 20.8 and 20.9 that makes their chromatic number to be 3? Is it the fact that they are *cubic*, i.e., each vertex of these maps has degree 3? Or is it due to an even number of neighbors of every region?

A.B. Kempe [Kem2] repeats Cayley's argument that we can convert any map M into the trivalent map M' such that

$$\chi_2(M) \ \le \ \chi_2(M')$$

Kempe writes:

> I should show that the colours could be so arranged that only three should appear at every point of concourse. This may readily be shown thus: Stick a small circular patch, with a boundary drawn round its edge, on every point of concourse, forming new districts. Colour this map. Only three colours can surround any district, and therefore the circular patches. Take off the patches and colour the uncovered parts the same colour as the rest of their districts. Only three colours surrounded the patches, and therefore only three will meet at the points of concourse they covered.

Our maps in Problems 20.8 and 20.9 are cubic, and, for cubic maps, an even number of neighbors is the key indeed:

Kempe's Three-Color Theorem 20.10 (A.B. Kempe, 1879, [Kem2]). A cubic map M has face chromatic number 3 if and only if the boundary of each of its regions consists of an even number of edges[5].

Let us translate Kempe's three-color theorem into the language of graphs by going to the dual graph $G = G(M)$ of the map M. Of course, since M is a trivalent map, all regions of G are triangular (i.e., 3-cycles). A plane graph, whose regions are all 3-cycles, is called a *triangulation*.

[5] Kempe states only the sufficient condition, but, in my opinion, the necessary condition is easier to prove and was possibly known to him.

Kempe's Three-Color Theorem 20.11 (*Three-Color Theorem for Graphs*). Let G be a connected plane triangulation. Then the following three assertions are equivalent:

(a) The chromatic number $\chi(G)$ of G satisfies the inequality $\chi(G) \le 3$.
(b) The face chromatic number $\chi_2(M)$ of G satisfies the inequality $\chi_2(M) \le 2$.
(c) The degree of every vertex of G is even.

Proof Kempe does not prove his statement. The proof presented here is a substantially simplified version of several problems from the 1952 Russian book by Evgenii B. Dynkin and Vladimir A. Uspensky [DU].

$(a) \Rightarrow (b)$. Since $\chi(G) = 3$, we can label each vertex of G with one of the colors a, b, or c. For every face, we have one vertex of each of the colors a, b, or c. Take a face F; if the direction of going around its vertices $a \to b \to c$ is clockwise, then we color F red; otherwise, we color F blue. It is easy to see that any two adjacent faces are thus assigned different colors.

$(b) \Rightarrow (a)$. Let G be face two-colored red and blue. For every edge xy of G, we assign one direction (out of possible two: $x \to y$ or $y \to x$) such that when we travel along the assigned direction, the red triangle is on our right (and thus a blue triangle is on our left). Obviously, for any two vertices v, w of G, there is a directed path from v to w, and, while the length of such a path (i.e., the number of edges in it) is not unique, its length modulo 3 is *unique*.

Assuming that we proved this uniqueness (see next paragraph for the proof), the rest is easy. Let us call our three colors 0, 1, and 2. Pick a vertex v and color it 0; then, for any vertex w of G, we select one directed path P from v to w and the remainder upon division of the length $l(P)$ of P by 3 determines the color we assign to w. This guarantees that the adjacent vertices are assigned different colors (do you see why?), and the implication $(b) \Rightarrow (a)$ is proven.

Proof of Uniqueness Let us first prove that the length $l(P)$ of any closed directed path P is divisible by 3. Assume it is not; then, among all directed closed paths of length not divisible by 3, there is one of *minimum* length l; we call it P'. P' has no self-intersections, for, otherwise, it could be shortened (can you see how?) in contradiction to its minimum length. P' partitions the plane into two areas: the inside and the outside. We combine the outside into one region O and, as a result, get a new map M_1, all regions of which are already colored red and blue, except the region O (see Fig. 20.9).

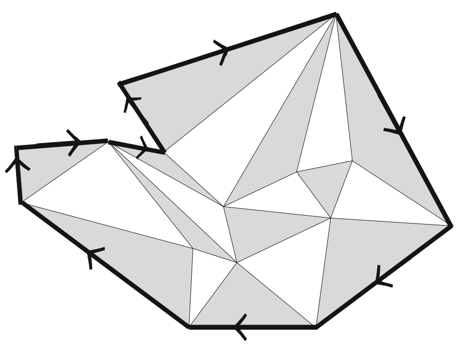


Fig. 20.9 2-colored map M_1

If the loop P' has a clockwise direction, then all triangles bordering on P' are colored red; otherwise, they are all colored blue. Since, in either case, all triangles bordering on P' are assigned the same color, say, red, we can complete the 2-coloring of the map M_1 by assigning the outside region O the opposite color – blue.

Every edge belongs to the boundary of one red and one blue region; therefore, the total numbers of edges on the boundaries of all red and all blue regions are equal. Thus, we get the following equality:

$$3r = 3b + l,$$

where r and b denote the number of red and blue triangles, respectively (and l, as you recall, is the length of P'). This equality contradicts the fact that l is not divisible by 3.

Assume now that there are two directed paths P_1 and P_2 from a vertex v to a vertex w with their lengths l_1 and l_2, respectively, such that l_1 and l_2 give different remainders upon division by 3. Let P_3 be a path of length l_3 from w to v. Then we get two different *closed* paths $P_1 + P_2$ and $P_1 + P_3$ of lengths $l_1 + l_2$ and $l_1 + l_3$, respectively. Therefore, in view of the above, both integers $l_1 + l_2$ and $l_1 + l_3$ are divisible by 3. But, then, the number

$$(l_1 + l_2) - (l_1 + l_3) = l_2 - l_3$$

is divisible by 3, and the desired uniqueness is proven.

$(b) \Leftrightarrow (c)$. This is precisely the two-color theorem we have discussed above. ■

Kempe's three-color theorem, as can be easily seen, has the following corollary:

Corollary 20.12 (P.J. Heawood, 1898, [Hea2]). Let G be a connected planar graph G. Then, the following assertions are equivalent:

(a) The chromatic number $\chi(G)$ of G satisfies the inequality $\chi(G) \leq 3$.
(b) G can be embedded (as a subgraph) into a triangulation graph G' such that the degree of every vertex of G' is even.

In his survey [Ste], Richard Steinberg describes the history of the three-color problem and its state at the time of his writing in 1993. In this otherwise wonderful historical work, Steinberg dismisses Alfred B. Kempe in a number of unjustified ways:

> The most notorious paper in the history of graph theory: the 1879 work by A. B. Kempe [Kem2] that contains the fallacious proof of the Four Color Theorem . . .
>
> Kempe's language is somewhat unclear – he was a barrister by profession.

I object to this arrogance of a "professional." Pierre de Fermat was a "barrister by profession" too. Does the mere fact that professionals are paid for services make them necessarily superior to amateurs? And when an amateur turns professional (which happens every day), does his language improve overnight?

Yes, Kempe's language is not as precise as our present standards require. But the same can be said of Tait and Heawood, yet, Steinberg approvingly quotes Gabriel Dirac's passionate but illogical argument in defense of Heawood's writing:

> Most of the assertions stated in [Hea2] are not actually proved, only made plausible, but they have since been proved rigorously by other writers, which indicates [sic] that Heawood was in possession of the necessary proofs but did not choose to include them.

As we have seen, Kempe's last page of [Kem2] contained a number of observations, including both the two-color theorem and the three-color theorem that are listed without proof, as "two special cases" of map coloring. I believe that Kempe knew the proofs but omitted them possibly because his main, if not the only, goal was to prove the much more complex four-color theorem.

Tomas L. Saati, in the title of his 1967 paper, called the attempt "The Kempe Catastrophe" [Saa1]. I cannot disagree more. As we will see in the next chapter, Alfred B. Kempe did not succeed in his goal, but what a fine try it had been, far exceeding anything his celebrated professional predecessors De Morgan and Cayley achieved in years of toying with 4CC! Moreover, today, both known successful assaults of 4CC have Kempe's approach as their foundation. Kempe came up with beautiful ideas; his *chain argument* has been used many times by fine twentieth-century professionals, such as Dénes Kőnig in his 1916 work on the chromatic index of bipartite graphs and Vadim Vizing (Chapter 18) in his famous 1964 chromatic index theorem.

For his important work on linkages, the contemporaries elected A.B. Kempe (1849–1922), a fellow of the Royal Society (1881) and president (1892–1894) of the London Mathematical Society. Kempe was knighted in 1913.

20.4 The New Life of the Three-Color Problem

In the first half of the twentieth century, it seemed that the three-color problem had been settled in the Victorian age. Since the late 1940s and the 1950s, we have witnessed the accelerating explosion of results on the relationship between the chromatic number of a graph and its small cycles (please see the discussion of it in Chapter 12). Examples of triangle-free graphs were in the mathematical air. Only one word, planar, needed to be added for revisiting the three-color problem and seeking a deeper understanding of what causes a map to be 3-colorable.

The first significant step of this new era of 3-colorable graphs was taken by the German mathematician Herbert Grötzsch in 1958–1959 [Grö].

The Grötzsch Theorem 20.13 Every triangle-free planar graph is 3-colorable; moreover, every proper 3-coloring of a 4- or 5-cycle can be extended to a 3-coloring of the whole graph.

In order to demonstrate that the restriction to planar graphs cannot be omitted, Grötzsch constructed the graph we discussed in Chapter 12 (see Fig. 12.8). His theorem, however, allowed an improvement, which was delivered by the great geometer (and *Geombinatorics'* major contributor and editor from its inception in 1991 to his passing on September 14, 2018) Branko Grünbaum [Grü] in 1963.

The Grünbaum Theorem 20.14 A planar graph with at most three 3-cycles is 3-colorable.[6]

This result is best possible, as K_4, a graph with four 3-cycles shows. Is there a life after the best possible result?

In mathematics – of course! As Valeri A. Aksionov and Leonid S. Mel'nikov observed [AMe], "Grünbaum put forth the question which determined the direction of further research." Grünbaum defined *the distance between triangles of a graph* as the length of the shortest path between the vertices of these triangles. He conjectured that if this distance is at least 1, then the planar graph is 3-colorable. Ivan Havel, who constructed a counterexample to Grünbaum's conjecture, posed and refuted his own conjecture (with a distance of at least 2) and, in the end, posed a more restrained open problem in 1969 [Hav].

Havel's Open Problem 20.15 Does there exist an integer n such that if the distance between any pair of triangles in a planar graph G is at least n, then G is 3-colorable?

Havel's problem is still open. According to Baogang Xu (e-mail of May 10, 2007), it is known that if such an n exists, it is at least 4.

Meanwhile, Richard Steinberg reasoned as follows: the restrictions on 3-cycles have been settled; but what if we were to impose no restrictions on 3-cycles but, instead, limit 4-cycles and 5-cycles? In his 1975 letter to the Russian mathematicians V.A. Aksionov and L.S. Mel'nikov, Steinberg posed his now well-known and still open problem [Ste].

Steinberg's Open Problem 20.16 Must a planar 4- and 5-cycle-free planar graph be 3-colorable?

Further research on the three-color problem was inspired by Havel's and Steinberg's open problems and often by a combination of the two of them. I am grateful to the Chinese mathematician Baogang Xu for navigating me through the labyrinth of the recent state of the problem. Let us look at the explosion of 3-coloring results.

Abbott–Zhou's Theorem 20.17 ([AZ], 1991). A planar graph without cycles of lengths 4–11 is 3-colorable.

Sanders–Zhao and Borodin's Theorem 20.18 ([SZ1], 1995; [Bor1], 1996). A planar graph without cycles of lengths 4–9 is 3-colorable.

Borodin–Glebov–Raspaud–Salavatipour's Theorem 20.19 ([BGRS], 2005). A planar graph without cycles of lengths 4–7 is 3-colorable.

Luo–Chen–Wang's Theorem 20.20 ([LCW], 2007). A planar graph without cycles of lengths 4, 6, 7, and 8 is 3-colorable.

Chen–Raspaud–Wang's Theorem 20.21 ([CRW], 2007). A planar graph without cycles of lengths 4, 6, 7, and 9 is 3-colorable.

In 2003, Oleg V. Borodin and André Raspaud [BoR] started a direction that combined Steinberg's and Havel's problems.

---

[6] A lemma used in the proof of Grünbaum's theorem was corrected and proved by Valeri A. Aksionov in 1974 [Aks].

Borodin–Raspaud's Theorem 20.22 ([BoR], 2003). A planar graph without 5-cycles and triangles of distance less than four is 3-colorable.

They also formulated two conjectures stronger than the (still open) positive answer to Steinberg's Problem 20.16. The authors called them "Bordeaux 3-Color Conjectures" – I will add the authors' names to give them credit. By *intersecting triangles*, the authors mean those with a vertex in common; by *adjacent triangles*, they call those with an edge in common.

Bordeaux 3-Color Borodin–Raspaud's Conjecture 20.23 ([BoR], 2003). A planar graph without 5-cycles and intersecting triangles is 3-colorable.

Bordeaux 3-Color Borodin–Raspaud's Strong Conjecture 20.24 ([BoR], 2003). A planar graph without 5-cycles and adjacent triangles is 3-colorable.

A proof of conjecture 20.24 in the positive would imply the validity of conjecture 20.23 and the positive answer to Steinberg's problem 20.16.

Baogang Xu strengthened Borodin–Raspaud's result 20.22:

Xu's Theorem 20.25 ([Xu2], 2007). A planar graph without 5-cycles and triangles of distance less than 3 is 3-colorable.

In a significant improvement of Borodin et al.'s Theorem 20.19, Baogang Xu proved a strong result in the direction of proving Bordeaux 3-color Borodin–Raspaud's conjecture 20.24:

Xu's Theorem 20.26 ([Xu1], 2006). A planar graph without adjacent triangles and 5- and 7-cycles is 3-colorable.

Two more results were obtained in the direction of Steinberg's and Havel's problems.

Lu–Xu's Theorem 20.27 ([LX], 2006). A planar graph without cycles of lengths 5, 6, and 9 and without adjacent triangles is 3-colorable.

Xu's Theorem 20.28 ([Xu3], 2014). A planar graph without cycles of lengths 5 and 6 and without triangles of distance less than 2 is 3-colorable.

The international group of mathematicians, Oleg V. Borodin, Aleksey N. Glebov, both from Russia, Tommy R. Jensen from Denmark, and André Raspaud from France [BGJR] put a new twist on the three-color oeuvre.

Borodin–Glebov–Jensen–Raspaud's Theorem 20.29 ([BGJR], 2006). A planar graph without triangles adjacent to cycles of lengths 4–9 is 3-colorable.

The authors have also formulated an attractive conjecture and named it after the city they posed it in:

Novosibirsk 3-Color Conjecture 20.30 ([BGJR], 2006). A planar graph without triangles adjacent to cycles of lengths 4 and 5 (or equivalently 3 and 5) is 3-colorable.

A few years later, this group produced more relevant results.

Borodin–Glebov–Montassier–Raspaud's Theorem 20.31 ([BGMR], 2009). Planar graphs without 5- and 7-cycles and without adjacent triangles are 3-colorable.

Borodin–Glebov's Theorem 20.32 ([BGl], 2011). Planar graphs without 5-cycles and with minimal distance between triangles of at least 2 are 3-colorable.

The decade that followed brought about many more results of the kind. The explosion of recent results has been so great that the field surely needed a new comprehensive survey, like the one that Richard Steinberg authored in 1993. Indeed, Oleg V. Borodin authored such a survey, *Colorings of plane graphs: A survey* in 2013 [Bor2]. I refer you to this valuable comprehensive survey for many exciting results (a number of statements are not translated correctly; I corrected the translations of those quoted above). My deep gratitude goes to Oleg Borodin for sharing his most impressive works.

It is fascinating to see how the seemingly lesser known cousin of the celebrated four-color conjecture (4CC) has flourished so beautifully and has become an exciting area of mathematical inquiry, even after 4CC was settled!

Chapter 21
Kempe–Heawood's Five-Color Theorem and Tait's Equivalence

21.1 Kempe's 1879 Attempted Proof of the Four-Color Theorem

I am compelled to present here Alfred Bray Kempe's attempted proof of the four-color conjecture (4CC). As you recall from Chapter 20, that proof contained an oversight, found a "mere" 11 years later by Percy John Heawood. Why then do I choose to present the unsuccessful attempt here? First of all because of the beautiful ideas Kempe invented. Second, because it is not easy to notice a flaw right away. Third, because P.J. Heawood did not have to do much to salvage Kempe's ideas to show that, in fact, *they* (i.e., *Kempe's ideas!*) prove the five-color theorem. Finally, because just like their contemporaries underestimated the work of Heawood, my contemporaries often underestimate the contributions of Kempe.

So it comes. Fasten your seat belts; I challenge you to find Kempe's oversight!

I will translate both the theorem and Kempe's proof into the usual nowadays language of dual graphs. The authors of *The Four-Color Problem* [SK], the first ever book on the subject, Thomas L. Saaty and my friend Paul C. Kainen, write (p. 7):

> The notion of dual graph mentioned above was introduced by Whitney (1931) and used to give an elegant characterization of when a graph is planar.

In fact, the notion of dual graphs appears on the last page of A.B. Kempe's 1879 paper [Kem2], as I mentioned after theorem 20.6 in the previous section, and Leonard Euler used it already in 1736. Kempe reinvented the notion but did not do much with it (what can one do with a promising notion that is introduced too late, on the last page of the paper?). I will use it here to make Kempe's attempted proof easier to read. I will also rearrange Kempe's proof.

The Four-Color Theorem (4CT) for Graphs 21.1 The chromatic number of any planar graph does not exceed 4.

Attempted Proof by Alfred Bray Kempe First, Kempe presented his brilliant chain argument, and, then, he rediscovered Euler's formula 21.2 and used it to find the graph theory's first ever set of *unavoidable configurations* (Tool 21.3 and the equivalent Tool 21.3'), as it is called today. We will do the latter two first.

© Alexander Soifer 2024
A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_21

Euler's Formula for Maps 21.2 For any map M in the plane, the following equality holds:

$$R + V = E + 2,$$

where R, V, and E are the number of regions, vertices, and edges of M, respectively.

Hint You can add edges to M as necessary, until you get a triangulation $T(M)$ so that Euler's formula holds for M if and only if it holds for $T(M)$, and then use induction. Let me not present here the complete proof: too many books have already done so. ■

Kempe's Tool 21.3 (A.B. Kempe, 1879, [Kem2]). Any planar map contains a vertex of degree at most 5.

Proof We can assume without loss of generality that each face is incident with at least three edges, for, otherwise, we can insert some vertices of degree 2 to remedy the situation.

We will argue by contradiction. Assume that the desired statement does not hold for a planar graph G, i.e., all V vertices of G have degrees of at least 6. Let R and E stand for the number of regions and edges of G, respectively. Since every edge is incident with two vertices, and with two regions, we get $6V \geq 2E$ and $3R \geq 2E$ or $V \geq \frac{1}{3}E$ and $R \geq \frac{2}{3}E$. Then, by Euler's formula 21.2, we get:

$$\frac{2}{3}E + \frac{1}{3}E \geq E + 2,$$

which is absurd. ■

I enjoyed translating Kempe's attempt into the contemporary terminology of unavoidable sets of reducible configurations that I found in Douglas R. Woodall's paper [Woo2]. I will present Kempe's attempted proof here in this language, for this would better prepare you for the next chapter, where we will discuss Appel and Haken's proof of 4CT.

A configuration C is called *reducible* if the minimal (in terms of the number of vertices) counterexample G to 4CT cannot contain C; i.e., should G contain C, G can be *reduced* to a smaller counterexample.

A finite set S of configurations is called *unavoidable* for a certain class Φ of maps if every map from Φ contains at least one element of S. Tool 20.3 could be reformulated in the language of unavoidable configurations as follows:

Kempe's Tool in Current Terminology 21.3' The set of four configurations in Fig. 21.1 is unavoidable, i.e., at least one of them appears in any nontrivial plane triangulation.

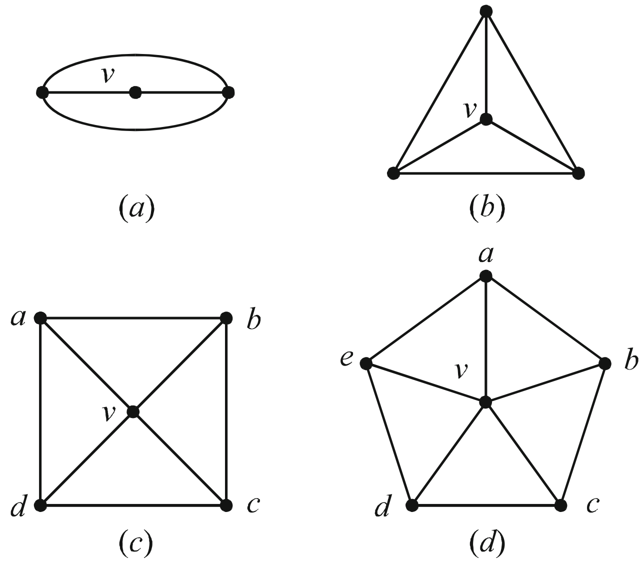


Fig. 21.1 The unavoidable set of four configurations

Kempe's Argument Kempe set out to prove that the four configurations in Fig. 21.1 form an unavoidable set of reducible configurations.

Assume that there is a planar graph that is not 4-colorable. Then, among all planar non-4-colorable graphs, there is a graph, call it G, of minimum order (i.e., minimum number of vertices). Embed G in the plane as a plane graph and add edges, if necessary, to make a triangulation T out of G. T is not 4-colorable as it may have even more restriction on coloring than G, but $T-v$ is 4-colorable for any vertex v. Fix a vertex v, and color $T-v$ in four colors. According to Tool 21.3, T contains one of the four configurations listed in Fig. 21.1, which prompts us to consider a few cases.

1. If T contains a configuration (a) or (b), the 4-coloring of $T-v$ can be easily extended to a 4-coloring of T: just assign the vertex v a color not used on the vertices adjacent to v. We got a contradiction; therefore, the assumption that T is a minimal counterexample to 4CT is false, and, thus, T can be reduced. Configurations (a) and (b) are reducible.
2. Let T contain a configuration (c). We will look at three subcases.

   (2a) If no more than three colors have been used to color the vertices a, b, c, and d, we can extend the 4-coloring of $T-v$ to a 4-coloring of T: just assign the vertex v a color not used on the vertices adjacent to v.

   (2b) Assume now that the vertices a, b, c, and d are assigned four different colors: following Kempe's taste, let these colors be red, blue, green, and yellow, respectively. Consider a subgraph T_{RG} of $T-v$ that is formed by all red and green vertices of $T-v$, with all edges connecting these vertices (we call T_{RG} a subgraph *induced* by the red and green vertices). If the vertices a and c belong to different components of T_{RG}, we interchange colors, red and green, in the component that contains the vertex c. As a result, we get a new 4-coloring of $T-v$, in which both vertices a and c are colored red. Thus, we can extend the 4-coloring of $T-v$ to a 4-coloring of T: just color the vertex v green.

(2c) Let us now assume that both vertices a and c belong to the same component of T_{RG}, i.e., there is, what we call today the *Kempe chain* C_{RG} in T_{RG} that connects a and c (see Fig. 21.2).

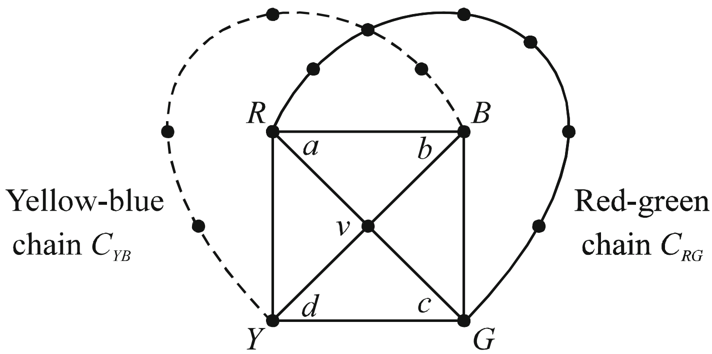


Fig. 21.2 Kempe's chains at work

Consider a subgraph T_{YB} of T–v induced by all yellow and blue vertices of T–v. Since the chain C_{RG} separates the vertices b and d, they must lie in different components of T_{YB}. We now interchange colors, yellow and blue, in the component of T_{YB} that contains the vertex b. As a result, we get a new 4-coloring of T–v, in which both vertices b and d are colored yellow. Thus, we can extend the 4-coloring of T–v to a 4-coloring of T: just color the vertex v blue.

We have thus proved in all cases that T is 4-colorable. A contradiction, therefore, the assumption that T is a minimal counterexample to 4CT is false, and, thus, T can be reduced. Configuration (c) is reducible.

3. Finally, let T contain a configuration (d). We will consider three subcases.

(3a) If no more than three colors have been used to color the vertices a, b, c, d and e, we can extend the 4-coloring of $T-v$ to a 4-coloring of T: just assign the vertex v a color not used on the vertices adjacent to v.

(3b) Assume now that the vertices a, b, c, d and e are assigned four different colors: red, blue, yellow, green, and blue, respectively. Consider subgraphs T_{RY} and T_{RG} of $T-v$ that are induced by all its red-and-yellow and red-and-green vertices, respectively. If the vertices a and c belong to different components of T_{RY}, or a and d belong to different components of T_{RG}, we interchange colors in the component that contains the vertex a. As a result, we get a new 4-coloring of $T-v$ such that the color red is not assigned to any of the vertices a, b, c, d and e. Thus, we can extend the 4-coloring of $T-v$ to a 4-coloring of T: just color the vertex v red.

(3c) Let us now assume that vertices a and c belong to the same component of T_{RY}, and a and d belong to the same component of T_{RG}, i.e., there is a Kempe chain C_{RY} in T_{RY} that connects a and c and a Kempe chain C_{RG} in T_{RG} that connects a and d (see Fig. 21.3).

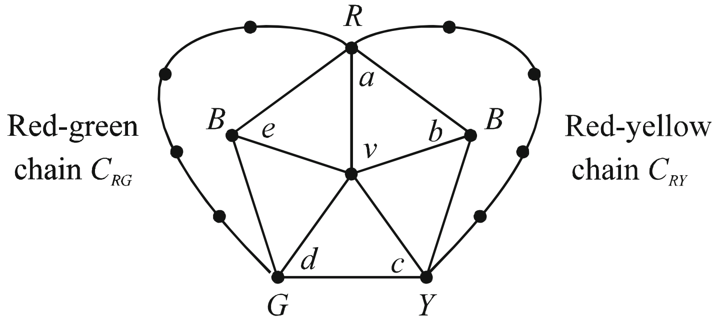


Fig. 21.3 Kempe's chains at work

Consider subgraphs T_{BG} and T_{BY} of $T-v$ induced by all its blue-and-green and blue-and-yellow vertices, respectively. The vertex b must lie in a component of T_{BG} that is different from those to which d and e belong, and e lies in a component of T_{BY} that is different from those to which b and c belong. We, therefore, interchange colors, blue and green, in the component of T_{BG} that contains b; and blue and yellow, in the component of T_{BY} that contains e. As a result, b becomes green and e yellow. Thus, we can extend the 4-coloring of $T-v$ to a 4-coloring of T: just color the vertex v blue.

We have proved in all cases that T is 4-colorable. A contradiction, therefore, the assumption that T is a minimal counterexample to the 4CT is false, and, thus, T can be reduced. Configuration (d) is reducible. ■

The four-color theorem has thus been proved or has it? In hindsight, we know that it was not. Have *you* noticed the hole? Try finding it on your own before reading the next section, in which I will reveal the hole.

21.2 The Hole

The hole occurs in subcase 3c. Everything Kempe did in the neighborhood of the vertex v was fine. He did get rid of color blue among the vertices adjacent to v and was, therefore, able to assign blue to v. *However*, while interchanging two colors in one component (as was done in subcase 2c) does create an allowable coloring of $T-v$, in subcase 3c, Kempe interchanged coloring in *two* components. Moreover, he interchanged colors in the components of T_{BG} and T_{BY} that *shared a color* (blue). Thus, there was no guarantee that what he got in the outset was an allowable coloring of $T-v$ (i.e., that everywhere in the graph adjacent vertices were assigned different colors). Thus, Kempe's attempted proof has a hole.

21.3 The Counterexample

In fact, Percy John Heawood was not only the first to find the above hole: he constructed a map such that if one follows Kempe's argument, two adjacent regions would get the same color assigned to them. Tomas L. Saati did not just translate Heawood's example into the language of graph theory but also added niceties of symmetries to his graph [Saa2, p. 9]. My assistant Phillip Emerich and I added further niceties of regular hexagons and pentagons to Saati's graph – see Fig. 21.4 for our embedding. Letters *R*, *B*, *Y*, and *G* stand for colors red, blue, yellow, and green, respectively. As a result of Kempe's recoloring, the adjacent vertices, *x* and *z*, end up with the same color assigned to them.

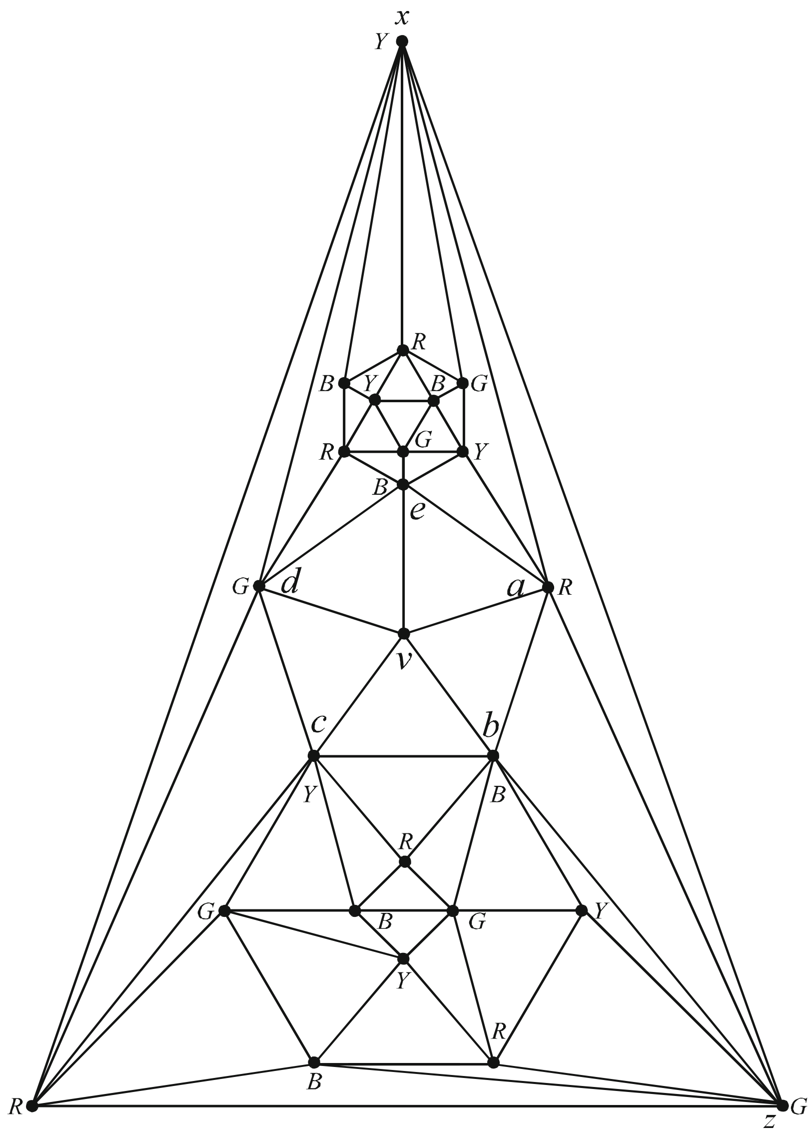


Fig. 21.4 Heawood's counter example in graph form

Heawood's counterexample is a graph of order 25. While reading Kempe's attempted proof, I found a counterexample of order just 9 that refutes Kempe's proof as written by him. In 1997, I published my counterexample in the *Mathematics Competitions* journal [Soi32].

Problem 21.4 (Soifer, 1997, [Soi32]). Construct a counterexample to Kempe's attempted proof of order not greater than 9.

Behold (Fig. 21.5):

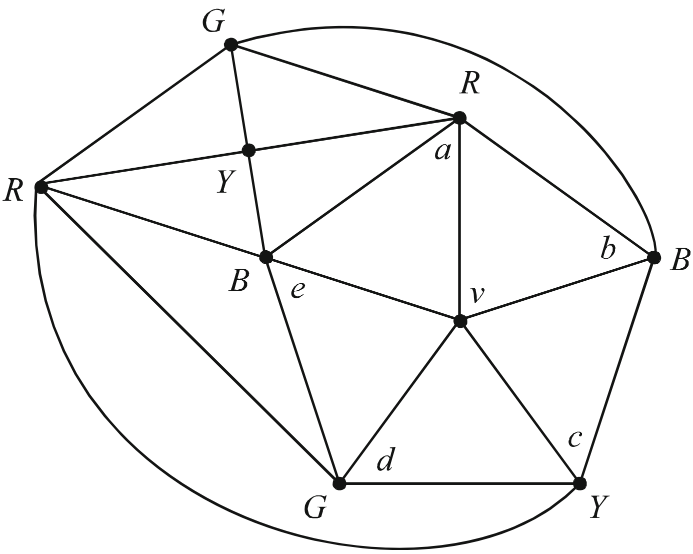


Fig. 21.5 The Soifer graph: a counterexample to Kempe of the smallest order

Moreover, in 1997, I conjectured that my counterexample is the smallest possible:

Conjecture 21.5 (Soifer, 1997, [Soi32]). For any graph of order less than 9, Kempe's argument works.

By accident, I ran into a familiar word "Soifer" in Wolfram MathWorld:

http://mathworld.wolfram.com/SoiferGraph.html

The Wolfram article included my Fig. 21.5, embedded in the plane in a slightly different way, and named it the Soifer graph (Fig. 21.6):

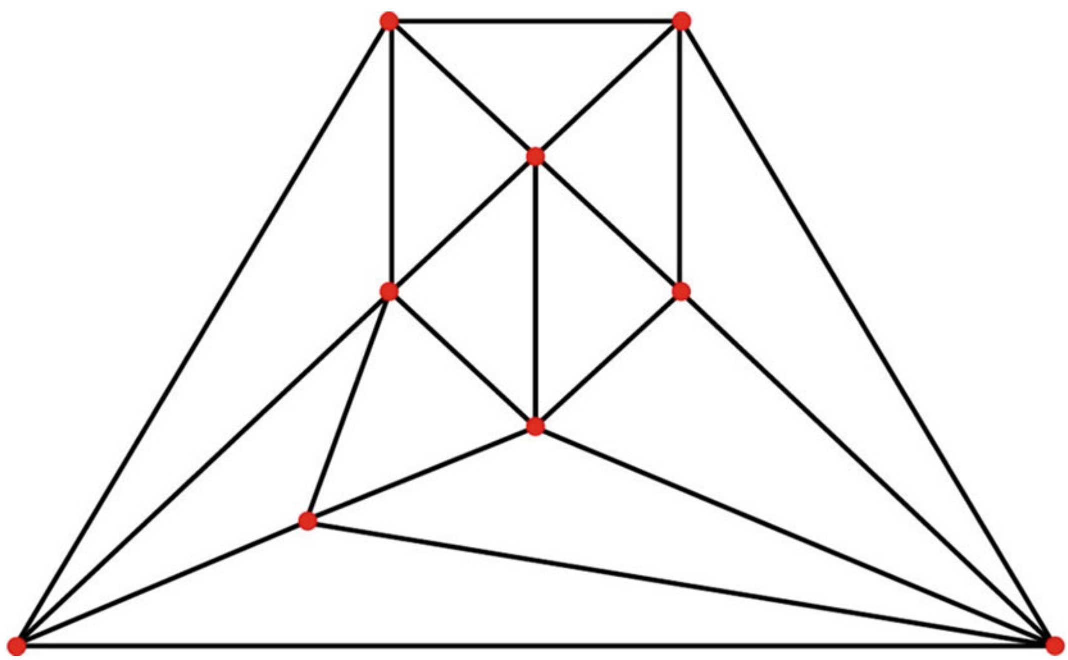


Fig. 21.6 The Soifer graph in Wolfram MathWorld

The Soifer Graph is a planar graph on 9 nodes that tangles the Kempe chains in Kempe's algorithm and thus provides an example of how Kempe's supposed proof of the Four-Color Theorem fails. As proved by Gethner and Springer, the Soifer graph is the smallest such counterexample (and is smaller than the Kittell Graph and Errera Graph).

It is implemented in the Wolfram Language as GraphData ["SoiferGraph"].

Little did I know that in 2003, Ellen Gethner and William M. Springer, II, of my University of Colorado's Denver campus named Fig. 21.5 "the Soifer graph," relabeled it, and proved my Conjecture 21.5 that the Soifer graph is the smallest possible [GS].

I also discovered that the Soifer graph was *commercially* used by Lisa Crosby Clark for making "Soifer graph bags"! Behold:

https://www.deviantart.com/lisacclark/art/Soifer-Graph-Bag-214832082

Fig. 21.7. A Soifer graph bag

I asked Lisa Crosby Clark to send me a complimentary copy of the Soifer graph bag but got no bag and no reply.

Rudolf and Gerda Fritsch created a graph [FF, p. 176], now known as the Fritsch graph, which is also of order 9, and also delivers a counterexample to Kempe's original 1879 attempted proof.

21.4 The Kempe–Heawood Five-Color Theorem

P.J. Heawood in his 1890 paper [Hea1] pointed out that Kempe's argument actually proves that five colors suffice. When we use five colors, there is no need to simultaneously interchange colors in two Kempe chains, and, thus, Kempe's chain argument works. Do verify the proof of the five-color theorem on your own.

I believe that the name often used today for this result, "the Heawood five-color theorem," is unfair. While Heawood was the first to formulate and prove the theorem, he merely adjusted an ingenious argument created by Kempe. It is therefore only fair to name the result after both inventors. I have little doubt Heawood would have agreed!

The Kempe–Heawood Five-Color Theorem 21.6 Five colors suffice to color any map in the plane.

21.5 Tait's Equivalence

Not only Augustus De Morgan but also Arthur Cayley contributed to spreading the word about the four-color conjecture. Peter Guthrie Tait is clear about it [Tai1, p. 501]:

> Some years ago, while I was working at knots, Professor Cayley told me of De Morgan's [sic] statement that four colours had been found by experience [sic] to be sufficient for the purpose of completely distinguishing from one another the various districts on a map.

When, in 1880, Alfred B. Kempe published yet another sketch of a proof [Kem5], similar to his original attempt, Tait was apparently inspired to enter the map coloring arena. In 1 year, 1880, he published a paper [Tai1] and then withdrew and replaced it with a one-page "abstract" [Tai2], which he expanded to an article [Tai3]. These papers contain some amusing statements, for example [Tai3, p. 657]:

> The difficulty in obtaining a simple proof of this theorem originates in the fact that it is not true without limitation.

One can paraphrase it to say, "It is difficult to prove, especially what is not true." Indeed, very much so! The Tait papers, however, also contain brilliant observations, such as what we call Tait's equivalence (Problem 21.8 below). Let us start our Tait review with his inductive attempt of proving 4CC.

The Four-Color Theorem 21.7 Every map on the plane is 4-colorable.

Tait's Attempted Proof ([Tai3]). Proof by induction in the number of regions. For a map with one region, 4CC holds.

Assume that any map with less than n regions is 4-colorable.

Given a map M with n regions, by Kempe's Tool 21.3, M contains a region R bounded by at most five edges. If R is bounded by two or three edges, erase one of them, say e. The resulting map can be 4-colored by the inductive assumption. Now reinstate e. At most three colors are forbidden for coloring R (one per each neighbor), and, we, therefore, use the remaining color for R.

Let R be bounded by four edges, and the adjacent regions clockwise are R_1, R_2, R_3, and R_4. [At least one of the two pairs of the opposite regions R_1 and R_3, R_2 and R_4, is nonadjacent; let R_2 and R_4 be the nonadjacent regions.] We erase a pair of opposite edges e_2 and e_4 that separate the regions R_2, R, and R_4 (Fig. 21.8)[1]. The resulting map can be 4-colored by the inductive assumption.

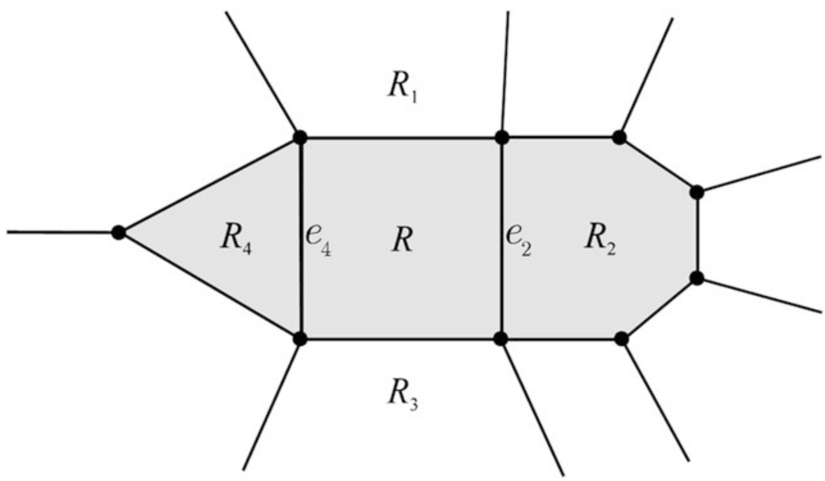


Fig. 21.8 Tait's attempted proof of the Four-Color Theorem

Now, reinstate e_2 and e_4. At most three colors are forbidden for coloring R (*because* R_2 *and* R_4 *are assigned the same color!*) and we, therefore, can use the remaining color on R. *Please observe that no Kempe chain argument was used in this case, and the proof is much shorter than in Kempe's attempt*.

Finally, let R be bounded by five edges, and erase a pair of nonadjacent edges, say e_1 and e_2. Here, Tait suddenly stops and writes:

[1]Tait in [Tai3, p. 660] wrote: "*either pair* of opposite sides of a four-sided region may be erased, and afterwards restored." This choice can cause a problem if the opposite regions are adjacent, hence I had to correct Tait's attempt by adding the previous sentence in brackets.

> But when we erase any two non-adjacent sides of a five-sided district, a condition is thereby imposed on the nomenclature of the remaining lines, with which I do not yet see how generally to deal.

Of course, Tait knew that he could continue his proof by 4-coloring the resulting map, which can be done by the inductive assumption, and then reinstate e_1 and e_2 and use the Kempe chain argument, as in [Kem2] or [Kem5]. *He did not! Why?*

The only plausible explanation, in my opinion, is that *Tait at the very least had doubts about the validity of Kempe's argument in the last case, if not realized the existence of the hole – 10 years prior to Heawood's work* [Hea1]. ■

With mathematical Olympiad-like brilliance, Tait proved the following fabulous equivalence. A fine statement meets as fine a proof. Enjoy!

The dual graph of a planar triangulation graph is a planar graph, whose vertices all have degree 3. If all vertices of a graph have the same degree 3, then we say that the graph is *regular of degree 3* or simply a *3-regular graph*.

Tait's Equivalence, Graph Version 21.8 (Tait, 1880). A planar 3-regular graph can be (vertex) 4-colored if and only if it can be edge-3-colored.

Proof [Tai2], [Tai3]: Let vertices of a planar 3-regular graph G be 4-colored in colors a, b, c, and d. We then color its edges in colors x, y, and z as follows: an edge is colored x if it connects vertices colored a and b or c and d; an edge is colored y if it connects vertices colored a and c or b and d; and an edge is colored z if it connects vertices colored a and d or b and c. We can easily verify that a proper edge coloring is thus obtained, i.e., no adjacent edges are assigned the same colors. In view of symmetry, it suffices to show it for the edges incident with a vertex colored a, which is demonstrated in Fig. 21.9.

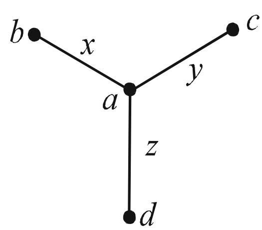


Fig. 21.9 Proof of Tait's equivalence

For the proof of the *converse* statement, Tait added points and edges to make degrees of every vertex even. Instead, I will subtract (remove) edges, which makes the argument more transparent.

Let edges of a planar 3-regular graph G be 3-colored in colors x, y, and z. Look at the subgraph G_{xy} of G induced by all its edges colored x and y.[2] Every cycle of G_{xy} must be even, as it alternates edges colored x and y. Therefore, by Kempe's two-color theorem (Problem 20.7), the vertices of G_{xy} (which precisely comprise *all* vertices of G) can be 2-colored in colors, say A and B. Similarly, we create the subgraph G_{yz} of G induced by all its edges

[2] I substantially simplified here Tait's language without changing his ideas. He talks about converting every triangular face into a four-sided one by inserting one new vertex per face inside an edge. He then throws the inserted vertices away, which is equivalent to keeping precise edges without insertions. These kept edges are then 2-colored.

colored y and z, and color all its vertices in two colors, say *1* and *2*. We thus assigned every vertex of G one of the following *four pairs* of colors: *A1*, *A2*, *B1*, or *B2*. It is easy to verify that we have ended up with the proper vertex 4-coloring of G! ■

The Tait equivalence can also be formulated in the dual language of maps:

Tait's Equivalence, Map Version 21.9 (Tait, 1880, [Tai2], [Tai3]). A map whose underlying graph is 3-regular can be (face) 4-colored if and only if it can be edge-3-colored.

21.6 Frederick Guthrie's Three-Dimensional Generalization

Have you found your own solution of Frederick Guthrie's Problems 19.1 and 19.2? As you recall, he generalized his brother's four-color problem to the three-dimensional Euclidean space and proved that no finite number of colors suffices.

Problem 21.10 For any positive integer n, there is a three-dimensional map that cannot be colored in n colors (so that regions having a common boundary – and not merely finitely many points – are assigned different colors).

In fact, unlike the Möbius–Weiske's puzzle, for any positive integer n, there are n solids such that every two have a common boundary surface.

Second Solution This solution appears in the 1905 paper of the Austrian mathematician and puzzlist Heinrich Tietze [Tie1]. As Frederick Guthrie before him, Tietze showed that in the three-dimensional space, we can easily construct $n + 1$ mutually adjacent solids. Just put $n + 1$ long enough parallelepipeds, numbered 1 through $n + 1$ on a plane, and, then, put on top of them $n + 1$ more parallelepipeds, which are perpendicular to the first ones and combine into one two solid parallelepipeds that are labeled with the same number (see Fig. 21.10). ■

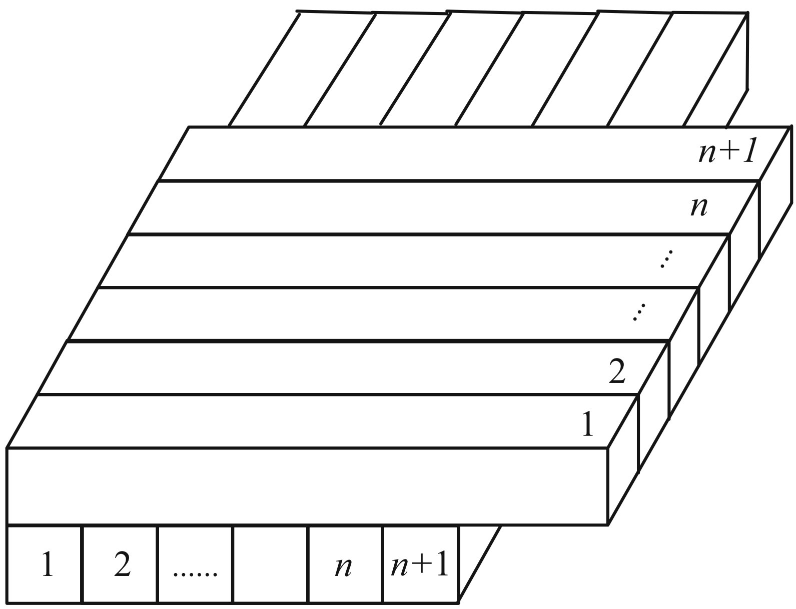


Fig. 21.10 Tietze's argument in the 3-space

Granted, this puzzle was easy to solve. However, according to Tietze, the German mathematician Paul Stäckel, who also solved the above problem, posed the same question for *convex* solids (I believe that Frederick Guthrie posed it first but did not formulate it very precisely – see the end of Chapter 19). Heinrich Tietze solved this harder problem in the same 1905 paper [Tie1].

Tietze's Theorem 21.11 (H. Tietze, 1905). For any positive integer n, there are n convex solids such that every two have a common boundary surface.

Thus, map coloring in three dimensions did not provide as lasting fun as has the two-dimensional variety of map coloring.

Chapter 22
The Four-Color Theorem

The most famous conjecture of graph theory or perhaps of the whole mathematics, the four colour conjecture, became recently the theorem of Appel and Haken.
—Paul Erdős, 1979[1]

Four–colour problem, the as yet unsolved [sic] problem of proving as a mathematical theorem that on any plane map only four colours are needed to give different colours to any regions that have a common boundary.
—Oxford English Dictionary, May 24, 2023 [!][2]

The year is 1976. I receive a notice about a meeting of the Moscow Mathematical Society in disbelief: the topic is the proof of the four-color conjecture (4CC) just obtained by two Americans, whose names do not ring my bell, but if the proof holds, are certainly destined to enter the history of mathematics, perhaps, history of culture. Every attendee is exhilarated.

So, it is happening: Kenneth Appel and Wolfgang Haken of the University of Illinois, with the aid of John Koch and some 1200 hours of fast main frame computing, convert Francis Guthrie's 4CC into 4CT, the four-color theorem.

The Four-Color Theorem 22.1 (K. Appel, W. Haken, and J. Koch [AH1–AH4]). Every planar map is 4-colorable.

In this chapter, we will look, however concisely, at the ideas of Appel–Haken's proof as presented by the authors in their monograph [AH4].

Appel and Haken's work grew from the 1879 approach discovered by Alfred B. Kempe (discussed in Chapter 20), improved in 1913 by George D. Birkhoff of Harvard University, and was brought into the realm of possibility by Heinrich Heesch of the University of Hanover through his committed work over many decades (1936–1972).

[1] [E81.16] was published in 1981 in the premier issue of *Combinatorica* and received by the editors on September 15, 1979.

[2] [OED] https://www.oed.com/view/Entry/73969?redirectedFrom=four-colour+problem#eid3604483

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_22

Birkhoff found new sets of reducible configurations [Bir], larger than those of Kempe. Heesch built on the work of his predecessors and developed a theory of reducible configurations [Hee1]:

> An investigation of the concepts of reduction has been attempted in the Heesch's "Untersuchungen zum Vierfarbenproblem" (Mannheim, 1969, Chapter I), where the concepts of A-, B-, C-, or D-reducible configurations are developed from the work of A. Errera, G. D. Birkhoff, and C. E. Winn.

Heesch was the first to utilize a computer in his pursuit [Hee2]:

> The D- or the C-reducibility of a configuration can be recognized much better by computing than by such direct calculations as have been given by the authors up to now.

Above all major technical contributions, Heinrich Heesch envisioned and conjectured the existence of a *finite set of unavoidable reducible configurations*. Appel and Haken paid their tribute to Heesch on the very first lines of their major paper that preceded their great announcement [AH0]:

> This work has been inspired by the work of Heesch [Hee1], [Hee2] on the Four-Color Problem, especially his conjecture [Hee1, p. 11, paragraph 1, and p. 216] that there exists a finite set S of Four-Color reducible configurations such that every planar map contains at least one element of S. (This conjecture implies the Four-Color Conjecture but is not implied by it.) Furthermore, in 1970 Heesch communicated an unpublished result ... which he calls a finitization of the Four-Color Problem.

In 1969, Heesch also pioneered a brilliant idea of *discharging* in search for unavoidable sets of configurations [Hee1]. His book paved the way for computer-aided pursuits of reducibility.[3]

Heesch's role is hard to overestimate. In addition to the credits we have enumerated above, Heesch personally influenced Haken and shared with him many unpublished ideas. In Appel and Haken's own words [AH2]:

> Haken, who had been a student at Kiel when Heesch gave his talk, communicated with Heesch in 1967 inquiring about the technical difficulties of the project of proving Heesch's conjecture and the possible use of more powerful electronic computers.
>
> In 1970 Heesch communicated to Haken an unpublished result which he later referred to as a finitization of the Four-Color Problem, namely that the first discharging step ..., if applied to the general case, yields about 8900 z-positive configurations (most of them not containing any reducible configurations) which he explicitly exhibited...
>
> Heesch asked Haken to cooperate on the project and, in 1971, communicated to him several unpublished results on reducible configurations.[4]

[3] Looking back, it seems highly surprising that in his 1972 42-page survey [Saa2] of various approaches to 4CC, T.L. Saati did not even mention the name of Heinrich Heesch.

[4] Haken then explains why his collaboration with Heesch ended: "The cooperation between Heesch and Haken was interrupted in October 1971 when the work of Shimamoto was thought to have settled the Four Color Problem."

To understand how the discharging works, let us look at the following simple example that I found in Douglas R. Woodall's papers [Woo2], [Woo3] where he credits K. Appel and H. Haken for it.

Problem 22.2
(K. Appel and H. Haken). The set of five configurations in Fig. 22.1 is unavoidable, i.e., at least one of them appears in any plane triangulation.

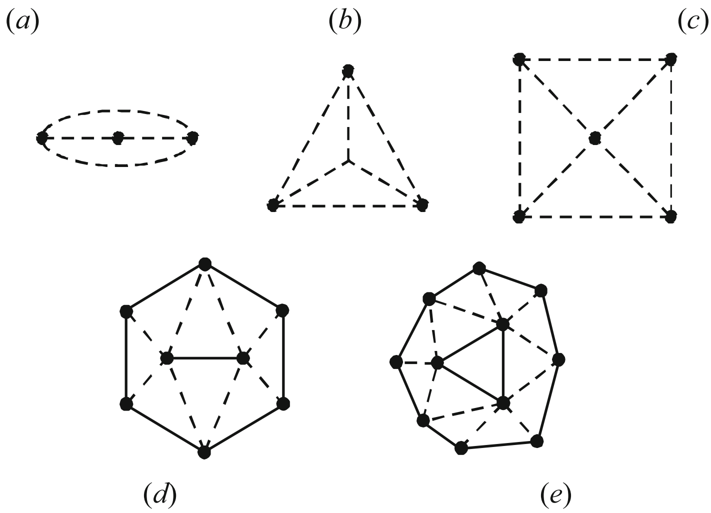


Fig. 22.1 An unavoidable set of configurations

Proof

1. **Observations.** We will argue by contradiction. Assume that there is a plane triangulation G that contains none of the configurations from Fig. 22.1. We can make the following observations:

   Observation A. G has no vertices of degree less than 5 because G contains no configurations (a), (b), and (c).

   Observation B. Every vertex v of degree 5 in G has at least three neighbors of degree 7 or greater; for, otherwise, v would have at least one neighbor of degree 5 and hence G would contain the forbidden configuration (d) or v would have at least three neighbors v_1, v_2, v_3 of degree 6. What is wrong with the latter, you may ask? In the latter case, at least two of the 6-valent neighbors of v, say, v_1 and v_2, must be neighbors of each other (in the *triangulation* G, the neighbors of v are connected to each other in a closed path; see Fig. 22.2), and, thus, the forbidden configuration (e) is contained in G.

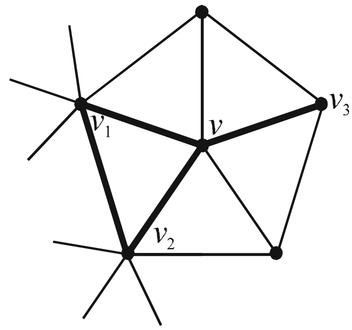


Fig. 22.2 In the triangulation G the neighbors of v are connected to each other in a closed path

Observation C. Every vertex v of degree 7 has at most three neighbors of degree 5, for, otherwise, two of its 5-valent neighbors would be neighbors of each other (it is similar to the argument in observation B above: prove it on your own), which would precisely mean that G contains a configuration (d).

Observation D. Every vertex v of degree $i \geq 8$ has at most $\left\lfloor \frac{i+1}{2} \right\rfloor$ neighbors of degree 5, where for a real number r, the symbol $\lfloor r \rfloor$ denotes the maximum integer such that $\lfloor r \rfloor \leq r$. The proof of this observation is similar to the proof of observation C above (try it on your own).

2. **Charging.** To each vertex of G of degree i, we assign an *electrical charge* equal to 6–i. This means that vertices of degree 5 receive a unit charge, vertices of degree 6 get a zero charge, vertices of degree 7 receive a charge equal to a negative one, etc. In his paper [Kem2], Kempe derived the following equality as a corollary of his rediscovering Euler's formula (Problem 20.2):

$$\sum_{i=2}^{\Delta} (6-i)\ V_i = 12, \qquad (K)$$

where V_i stands for the number of vertices of degree i. In Heesch's language of electrical charges, this equality precisely means that the sum of charges of all vertices in G, i.e., the *total charge* in G, is equal to positive 12 units.

3. **Discharging.** Let us now perform *discharging*, i.e., redistribution of charge among the vertices without changing the total charge of G. The crux of such a proof is to find the discharging that "works" for the set of configurations in question, which, in our case, is presented in Fig. 22.1, i.e., brings us the desired contradiction. Let us transfer $\frac{1}{3}$ of the charge from each vertex of degree 5 to each of its neighbors of degree 7 or greater.

As a result, every vertex of degree 5 ends up with zero or a negative charge because it has at least three neighbors of degree 7 or greater (see observation B above). Vertices of degree 6 will remain with a zero charge, as they are unaffected by discharging. A vertex v of degree 7 would not end up with a positive charge because v has at most three neighbors of degree 5, each contributing charge $\frac{1}{3}$ to v (observation C). Finally, a vertex of degree $i \geq 8$ that started with a charge $6 - i$, in view of observation D, can end up with the charge at most

$$6–i + \frac{1}{3}\left[\frac{i+1}{2}\right] < 0$$

Thus, we end up with *no* vertices of position charge, which contradicts the total charge remaining the positive 12. ■

Did you enjoy the mathematical Olympiad-like discharging argument? Then, you would enjoy proving on your own the following result first obtained without discharging in 1904 by Paul August Ludwig Wernicke from Göttingen University, who in the same year defended his doctorate under the great Hermann Minkowski.

Problem 22.3 (Wernicke, 1904, [Wer]). Prove that the set of five configurations in Fig. 22.3 is unavoidable.

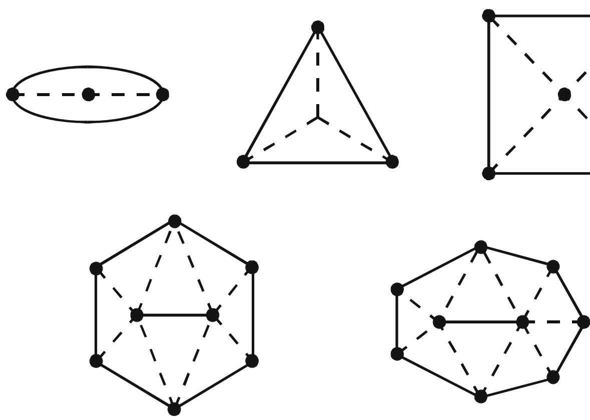


Fig. 22.3 A set of five unavoidable configurations

Let us now look at the other critical aspect of the Appel–Haken proof: *reducibility*. Appel and Haken used the so-called C- and D-reducibilities introduced by Heesch as vast extensions of the technique used by Kempe to show that a region with four neighbors could not occur in a minimal counterexample to the four-color conjecture. In fact, it suffices to restrict ourselves to configurations with vertices of degree five and greater since Kempe showed that vertices of lesser degrees cannot occur in a minimal counterexample. The authors provide the following example [AK4].

> Assume that the planar triangulation Δ is the minimal counter-example to the Four-Color Conjecture, which contains, for example, a configuration C of Fig. 22.4a. (Legend in Fig. 22.4d shows how to read the degrees of the vertices of the configuration from the diagram in Fig. 22.4.) Then the graph Δ – C obtained from Δ by removing C and edges connecting C to the rest of Δ, must be four-chromatic. A contradiction would be obtained, if we show that every four-colorings of Δ – C can be extended to a four-coloring of Δ.

Appel and Haken repeatedly used good humor while praising the use of computers. Here is one example (numbers 0, 1, 2, and 3 on the ring in Fig. 22.4c indicate the four colors we are using):

> If one were lucky, one might be able to show a fourteen-ring configuration D-reducible with only a few years of careful work. There are obviously some slackers who would not be fascinated by such a task. Such people, with an immorally low tolerance for honest hard work, tend to program computers to do this task. In fact, they find it ideally suited to computers, which are fast, meticulous, and not able to complain about the boring aspects of the work.

All humor aside, however, their solution required an enormous amount of both manual and computer work. The crux of the Appel–Haken proof was to find such a set of configurations that was *both* unavoidable and consisted of reducible configurations, the so-called *unavoidable set of reducible configurations*. In one of the early 1978 analyses of the proof [Woo3], Douglas R. Woodall assessed this critical part as follows:

> Discharging procedure and the unavoidable set of configurations were modified every time a configuration in the set turned out not to be C-reducible (or was not quickly proved to be C-reducible). It is clear that these progressive modifications relied on a large number of empirical rules, which enabled an unwanted configuration to be excluded from the unavoidable set at the expense of possibly introducing one or more further configurations. Appel and Haken carried out about 500 such modifications in all. They continued until they had excluded
>
> (i) Every configuration that contained one of three "reduction obstacles" – features that Heesch had discovered, by trial and error, to prevent configurations from being C-reducible.
> (ii) Every configuration of ring size 15 or more.
> (iii) Every configuration that was not proved to be reducible fairly quickly (in particular, within 90 minutes on an IBM 370-158 or 30 min on a 370-168).
>
> By the time they had finished (iii), they had constructed an unavoidable set all of whose configurations had been proved reducible; they had therefore proved the theorem. Probably they had excluded from the unavoidable set many configurations that are actually C-reducible but it turned out to be quicker to exclude any configuration that was *not quickly* proved reducible, and to replace it by one or more other configurations, than to carry the analysis of any one configuration to its limit.
>
> The empirical rules, upon which these progressive modifications were based, were discovered in the course of a lengthy process of trial and error with the aid of a computer, lasting over a year. By the end of this time, however, Appel and Haken had developed such a feeling for what was likely to work (even though they could not always explain why) that they were able to construct the final unavoidable set without using the computer at all. This is the crux of their achievement. Unavoidable sets had been constructed before, and configurations had been proved reducible before, but no-one before had been able to complete the monumental task of constructing an unavoidable set of reducible configurations.

It was a great achievement by Appel and Haken, for they reduced the infinity of various maps to a *finite* set of unavoidable reducible configurations, which needed to be checked. This reduction was a mathematical achievement, and it allowed the use of computers (surely, with infinitely many cases, a computer would have been useless!). The set was now finite, but very large, at first consisting of 1936 configurations. The enormous computer verification used over 1200 hours of main frame computing time (on IBM-360 and IBM-370). By 1989, when Appel and Haken produced the 741-page book [AH4] presenting their solution, they reduced the number of configurations to 1476. Such a surprising resolution of the famous problem, both in its volume of work and in use of computing, was bound to cause controversy, and it promptly did. The Appel–Haken–Koch proof of 4CT was a cultural event: it prompted debates and reassessments in many fields of human endeavors, particularly in mathematical and philosophical circles. In the next chapter, we will look at the debates and some striking views it has inspired, as well as at the new proof of 4CT, and an old but still most promising Hadwiger's conjecture.

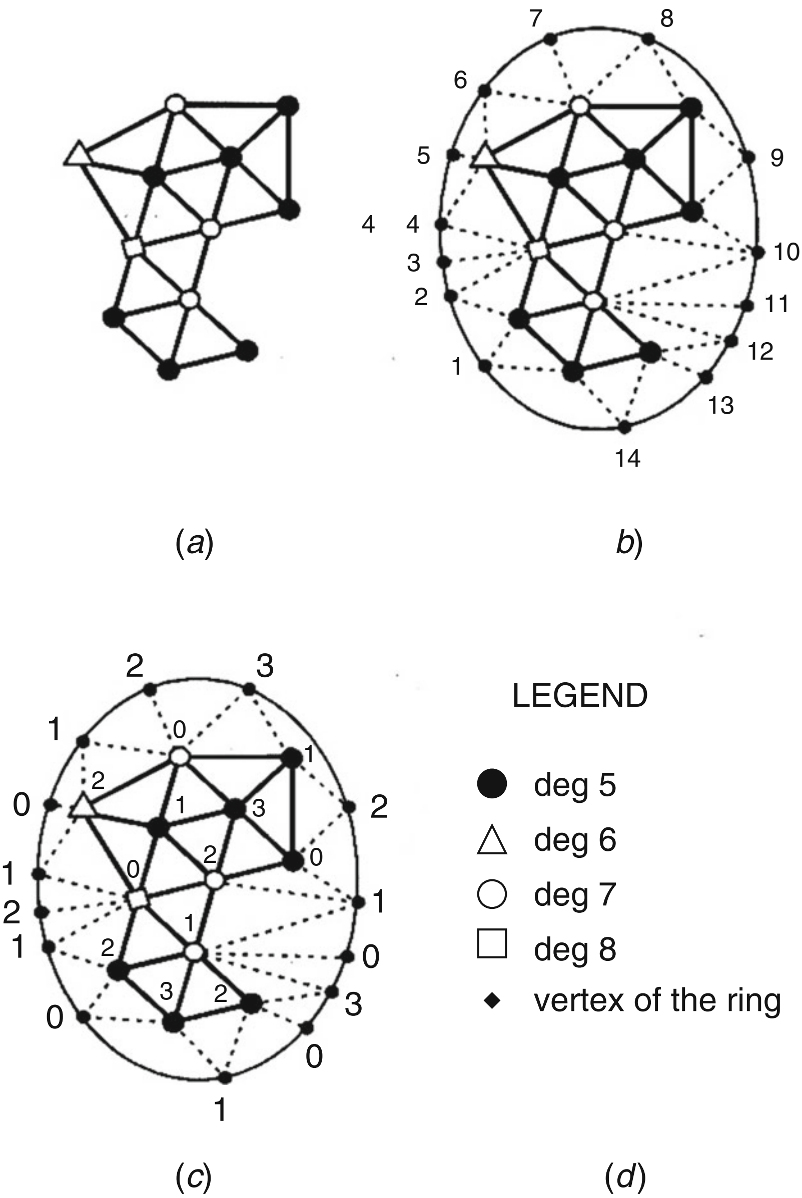


Fig. 22.4 Appel–Haken's example of a fourteen-ring configuration

In a phone interview in the fall of 1991 (before October 14, 1991), **Wolfgang Haken** shared with me brief details of his life: born on June 21, 1928, in Berlin; obtained his doctorate from the University of Kiel in 1953; came to the United States in 1962; started to work on 4CC in 1968; and came up with the first ideas of his own in October 1970. **Kenneth Ira Appel** was born on October 8, 1932, in Brooklyn, New York, and got his doctorate from the University of Michigan in 1959.

During the interview, Wolfgang Haken *accepted* my invitation to write his view of the Appel–Haken accomplishment entitled, on my recommendation, in Alexandre Dumas' style, "Fifteen Years Later." I offered to publish his complete *unedited* essay in *Geombinatorics* and include it in my *Mathematical Coloring Book*. The day after my interview with Haken, I received a phone call from Kenneth Appel, who questioned my person and my goals in asking for an essay, as if he was a guard at an armory and I was suspected of stealing explosives. Consequently, no essay came from Wolfgang. Appel and Haken probably expected – and rightly so – a great reception for their achievement. It did not quite happen and likely caused their disappointment and suspicion.

The second epigraph, from the Oxford English Dictionary [OED], shows how little attention is paid to mathematics: Oxford failed to notice even by 2023 the now 47-year-old solution of one of two most famous problems in the multi-millennial history of mathematics!

I have got to quote from the March 2005 unpublished, but web-posted, paper by Georges Gonthier, a researcher from the Programming Principles and Tools Group of Microsoft-Cambridge, UK [Gon]. With a deep insight of someone who verified a 4CT proof and came up with a "machine proof," he assessed the contributions of the players of the first successful assault of 4CC:

> Although Heesch had correctly devised the plan of the proof of the Four Colour Theorem, he was unable to actually carry it out because he missed a crucial element: computing power. The discharge rules he tried gave him a set R containing configurations with a ring of size 18, for which checking reducibility was beyond the reach of computers at the time. However, there was hope, since both the set of discharge rules and the set R could be adjusted arbitrarily in order to make every step succeed.
>
> Appel and Haken cracked the problem in 1976 by focusing their efforts on adjusting the discharge rules rather than extending R, using a heuristic due to Heesch for predicting whether a configuration would be reducible (with 90% accuracy), without performing a full check. By trial and error they arrived at a set R, containing only configurations of ring size at most 14, for which they barely had enough computing resources to do the reducibility computation. Indeed the discharging formula had to be complicated in many ways in order to work around computational limits. In particular the formula had to transfer arity between non-adjacent faces, and to accommodate this extension unavoidability had to be checked manually. It was only with the 1994 proof by Robertson *et al.* that the simple discharging formula that Heesch had sought was found.

We will discuss the Robertson–Sanders–Seymour–Thomas proof and its Gonthier's verification in one of the subsequent chapters.

Chapter 23
The Great Debate

Computers are useless.
They can only give you answers.

– Pablo Picasso

To reject the use of computers as what one may call "computational amplifiers" would be akin to an astronomer refusing to admit discoveries made by telescope.

– Paul C. Kainen, 1993 [Kai]

I would be much happier with a computer-free proof of the four color problem, but I am willing to accept the Appel–Haken proof – beggars cannot be choosers.

– Paul Erdős, 1991 Erdős' letter to A. Soifer [E91/8/14ltr]

Interest in the 4CC seems not to be high in the math literature because it is now thought to have been proven or something.

– Thomas L. Saaty, 1998 Saaty's e-mail to A. Soifer of April 13, 1998

23.1 40 Years of Debate

Forty+ years later, the controversy surrounding the Appel and Haken proof is amazingly alive and well. Even when the extraordinary in many respects Appel and Haken's proof was just announced, the President of the Mathematical Association of America Lynn Arthur Steen was very careful [Ste]: he did not write that the conjecture had been *proved* but instead used the word "verified" in describing the most important mathematical event of that year.

The proof was met with considerable confusion in the mathematical community due to the authors' extensive use of computers and a verification issue. This was the first computer-aided solution to a major, celebrated mathematical problem. As such, it naturally raised mathematical, philosophical, and psychological questions. In Table 23.1, I put together a "representative" collection of reactions – take a long look at it, then join me for a discussion.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_23

Table 23.1 Reflections on the 4CT

Steen	1976	[Ste]	The Four-Color Conjecture...was verified [sic] this summer . . .
Appel & Haken	1977	[AH1]	Our proof of the Four-Color Theorem suggests that there are limits to what can be achieved in mathematics by theoretical methods alone.
Gardner	1980	[Gar3]	The proof is an extraordinary achievement...To most mathematicians, however, the proof of the Four-Color conjecture is deeply unsatisfactory.
Halmos	1990	[Hal]	By an explosion I mean a loud noise, an unexpected and exciting announcement, but not necessarily a good thing. Some explosions open new territories and promise great future developments; others close a subject and seem to lead nowhere. The Mordell conjecture...is of the first kind; the Four-Color Theorem of the second.
Erdős	1991	[E91/8/14ltr]	I would be much happier with a computer-free proof of the four-color problem, but I am willing to accept the Appel–Haken proof – beggars cannot be choosers.
Graham	1993	[Hor]	The things you can prove may be just tiny islands, exceptions, compared to the vast sea of results that cannot be proven by human thought alone.
Kainen	1993	[Kai]	To reject the use of computers as what one may call "computational amplifiers" would be akin to an astronomer refusing to admit discoveries made by telescope.
Hartsfied & Ringel	1994	[HR]	Appel and Haken proved it by means of computer program. The program took a long time to run, and no human can read the entire proof because it is too long.
Jensen & Toft	1995	[JT]	Does there exists a short proof of the Four-Color Theorem... in which all the details can be checked by hand by a competent mathematician in, say, two weeks?
Graham	2002	[Gra4]	Computers are here to stay. There are problems for which computer helps; there are problems for which computer may help; and there are problems for which computer will never help.

The confusion of mathematicians is very clear when we read the words of Martin Garner in his celebrated *Scientific American* column [Gar3]:

> The proof is an extraordinary achievement . . . To most mathematicians, however, the proof of the Four-Color conjecture is deeply unsatisfactory.

Which is it, dear Martin, "an extraordinary achievement" or "deeply unsatisfactory"? Surely these terms are mutually exclusive. Paul Halmos, who chose to sum up the main twentieth century contributions to mathematics a decade too early (and thus missed a lot, proof of Fermat's Last Theorem, for example), writes in 1990 [Hal]:

> By an explosion I mean a loud noise, an unexpected and exciting announcement, but not necessarily a good thing. Some explosions open new territories and promise great future developments; others close a subject and seem to lead nowhere. The Mordell conjecture . . . is of the first kind; the Four-Color Theorem of the second.

A loud noise that leads nowhere, Professor Halmos? It suffices to observe that much of graph theory has been invented through the 124 years of attempts to settle 4CC. Nora Hartsfield and Gerhard Ringel [HR] paint this historic event as routine, boring, and unworthy of attention:

> Appel and Haken proved it by means of computer program. The program took a long time to run, and no human can read the entire proof, because it is too long.

Moreover, Hartsfield and Ringel [HR] cite a 2-page announcement of Appel and Haken, and not their articles and a complete 741-page proof, as if questioning the legitimacy of the proof.

There were, however, those who gave the event much thought. In 1978, the philosopher Thomas Tymozcko of Smith College illustrates the arrival of computer-aided proofs with a brilliant allegory [Tym]:

> Let us consider a hypothetical example which provides a much better analogy to the appeal to computers. It is set in the mythical community of Martian mathematicians and concerns their discovery of the new method of proof "Simon says." Martian mathematics, we suppose, developed pretty much like Earth mathematics until the arrival on Mars of the mathematical genius Simon. Simon proved many new results by more or less traditional methods, but after a while began justifying new results with such phrases as "Proof is too long to include here, but I have verified it myself." At first Simon used this appeal only for lemmas, which, although crucial, were basically combinatorial in character. In his later work, however, the appeal began to spread to more abstract lemmas and even to theorems themselves. Oftentimes other Martian mathematicians could reconstruct Simon's results, in the sense of finding satisfactory proofs; but sometimes they could not. So great was the prestige of Simon, however, that the Martian mathematicians accepted his results; and they were incorporated into the body of Martian mathematics under the rubric "Simon says."
>
> Is Martian mathematics, under Simon, a legitimate development of standard mathematics? I think not; I think it is something else masquerading under the name of mathematics. If this point is not immediately obvious, it can be made so by expanding on the Simon parable in any number of ways. For instance, imagine that Simon is a religious mystic and that among his religious teachings is the doctrine that the morally good Martian, when it frames the mathematical question justly, can always see the correct answer. In this case we cannot possibly treat the appeal "Simon says" in a purely mathematical context. What if Simon were a revered political leader like Chairman Mao? Under these circumstances we might have a hard time deciding where Martian mathematics left off and Martian political theory began. Still other variations on the Simon theme are possible. Suppose that other Martian mathematicians begin to realize that Simonized proofs are possible where the attempts at more traditional proofs fail, and they begin to use "Simon says" even when Simon didn't say! The appeal "Simon says" is an anomaly in mathematics; it is simply an appeal to authority and not a demonstration.

> The point of the Simon parable is this: that the logic of the appeals "Simon says" and "by computer" are remarkably similar. There is no great formal difference between these claims: computers are, in the context of mathematical proofs, another kind of authority. If we choose to regard one appeal as bizarre and the other as legitimate, it can only be because we have some strong evidence for the reliability of the latter and none for the former. Computers are not simply authority, but warranted authority. Since we are inclined to accept the appeal to computers in the case of the 4CT and to reject the appeal to Simon in the hypothetical example, we must admit evidence for the reliability of computers into a philosophical account of computer-assisted proofs. . .
>
> The conclusion is that the appeal to computers does introduce a new method into mathematics.

Tymoczko is correct: Appel–Haken–Koch's proof changed the meaning of the word "proof" by letting in a reliable experiment as allowable means, by taking away the absolute certainty we cherished so much for so long in the mathematical proof. Thomas L. Saaty and Paul C. Kainen, whose great timing allowed them to publish in 1977 the first ever book on The Four-Color Problem that included a discussion of its solution, were first to insightfully observe the substantial but inevitable trade-offs of the acceptance of such a proof [SK, end of part one]:

> To use the computer as an essential tool in their proofs, mathematicians will be forced to give up hope of verifying proofs by hand, just as scientific observations made with a microscope or telescope do not admit direct tactile confirmation. By the same token, however, computer-assisted mathematical proof can reach a much larger range of phenomena. There is a price for this sort of knowledge. It cannot be absolute. But the loss of innocence has always entailed a relativistic world view; there is no progress without the risk of error.

In the essay [Kai] written in 1993 on my request especially for *Geombinatorics*, Paul C. Kainen elaborates further on the above allegory:

> To reject the use of computers as what one may call "computational amplifiers" would be akin to an astronomer refusing to admit discoveries made by telescope.

This is certainly an elegant and powerful metaphor. However, one cannot argue with Tymoczko's warning about keeping the order right, that we have accepted the legitimacy of the use of computers first, and only based on this acceptance, we can claim the existence of the formal proof [Tym]:

> Some people might be tempted to accept the appeal to computers on the ground that it involves a harmless extension of human powers. On their view the computer merely traces out the steps of a complicated formal proof that is really out there. In fact, our only evidence for the existence of that formal proof presupposes the reliability of computers.

As Tymoczko rightly observes, the timing of Appel–Haken–Koch work was favorable for the acceptance of their proof [Tym]:

> I suggest that if a "similar" proof had been developed twenty-five years earlier, it would not have achieved the widespread acceptance that the 4CT has now. The hypothetical early result would probably have been ignored, possibly even attacked (one thinks of the

early reaction to the work of Frege and of Cantor). A necessary condition for the acceptance of a computer-assisted proof is wide familiarity on the part of mathematicians with sophisticated computers. Now that every mathematician has a pocket calculator and every mathematics department has a computer specialist, that familiarity obtains. The mathematical world was ready to recognize the Appel–Haken methodology as legitimate mathematics.

Douglas R. Woodall and Robin Wilson state in their 1978 essay [WW] that "there is no doubt that Appel and Haken's proof is a magnificent achievement which will cause many mathematicians to think afresh (or possibly for the first time) about the role of the computer in mathematics." Yet, they share concerns with Paul Halmos and others:

> The length of Appel and Haken's proof is unfortunate, for two reasons. The first is that it makes it difficult to verify . . . The other big disadvantage of a long proof is that it tends not to give very much understanding of why the result is true. This is particularly true of a proof that involves looking at a large number of separate cases, whether or not it uses the computer.

Paul Erdős put the state of 4CT most aptly in his August 14, 1991 letter to me [E91/8/14ltr]:

> I would be much happier with a computer-free proof of the four color problem, but I am willing to accept the Appel–Haken proof – beggars cannot be choosers.

So, what are we, mathematicians, to do? The answer, in a form of a question, comes from the Danish graph theorists Tommy R. Jensen and Bjarne Toft in their book of open coloring problems of graph theory [JT]:

> Does there exist a short proof of the Four-Color Theorem... in which all the details can be checked by hand by a competent mathematician in, say, two weeks?

Appel and Haken [AH1] apparently do not believe in the existence of a computer-free proof of the Four-Color Theorem. I beg to disagree: the mere existence of a computer-aided proof does not exclude that one day someone will find a computer-free proof of 4CT.

In 1976, many humans believed that they make less mistakes than computers. Today, 40+ years after Appel–Haken, we all are ready to stipulate that computers are much more mistake-free in pursuing tasks we assign to them than humans proving theorems "by hand."

23.2 Twenty Years Later, or Another Time – Another Proof

Twenty years later, when the familiarity with and trust in computing have dramatically improved, as did computers themselves, a new team of players came on 4CT stage: leading graph theorists Neil Robertson and Paul Seymour and their young students and colleagues Daniel Sanders and Robin Thomas. This reminded me of the Hollywood film *Seven Brides for Seven Brothers.* Only here we had *Four Mathematicians for Four Colors*. Four on four, they had to be able to handle 4CT and handle they did.

In their work on graph theory, the authors thought that in a sense, the validity of $H(6)$ (Hadwiger's Conjecture for 6 – we will formulate it later in this chapter) depended upon the validity of 4CT. Thus, they felt compelled to either verify Appel–Haken proof or find their own. They decided that the latter was an easier task.

The "Four Musketeers" undoubtedly realized that "20 years later," as Alexandre Dumas used to say (precisely the title of Alexandre Dumas' sequel to the famed *Three Musketeers*), they would have to get a much better proof than the original one by Appel and Haken, for otherwise they would be asked "why did you bother?" The remarkable thing is these authors have achieved just such a proof.

I first learned about it in February 1993 on the coast of the Atlantic Ocean during a Florida Atlantic University conference from Ron Graham, who also forwarded to me the e-mail announcement of the forthcoming March 24, 1994, DIMACS talk by Paul Seymour, who at the time worked at Bellcore. I asked Paul's coauthor Neil Robertson for the details. His May 9, 1994 reply [Rob1] due to its medium, e-mail, concisely, and instantly summarized what he thought was most important about the new proof:

> We have a new proof, along the same lines as the AHK[1] proof, relying more on the computer, and so more reliable. The unavoidable set is in the area of 600 configurations (<= 638), and we get a quadratic algorithm. Dan wrote a nice article about this for SIAM (I think). Seymour, Thomas, Sanders and I are involved. With a slightly larger unavoidable set the overall proof becomes very simple (apart from the calculations) as we avoid almost all degeneracies by using D-reduction and reducers for C-reduction from the single edge contraction minors of the given configuration. Will forward to you a copy of Dan's article.

One important point to notice here is that the proof relies more (not less!) on the computer than Appel–Haken–Koch's proof, and it makes the proof more (not less) reliable due to its clear separation of human and machine tasks. The size of the unavoidable set of reducible configurations is substantially reduced from Appel and Haken's 1476 to 636, but the greatest improvements are in the much-much simpler discharging procedures. Later the same day, Neil forwarded to me Daniel Sander's summary, entitled, in a word,

"**NEWPROOFOFTHEFOURCOLORTHEOREM**."

I have got to share with you parts of this announcement summary, as it includes the authors' assessment of the Appel–Haken proof and comparisons of the two proofs; I will add my comments as footnotes:

> Before and after Appel and Haken, many claims have been made to prove 4CT with the aid of a computer, none of which held up to the test of time. But Appel and Haken's proof has stood; for 18 years. Why? Some may say that the proof is inaccessible. It is so long and complicated; has anyone actually read every little detail? At least two attempts were made to independently verify major portions of Appel and Haken's proof [AH2, AH3], which yielded no significant problems. Appel and Haken [AH4] published a more complete (741 pages) version of their proof five years ago, but many remain hesitant.

[1] Appel, Haken and Koch.

The author of this paper, together with Neil Robertson, Paul Seymour, and Robin Thomas, announces a new proof of 4CT. The proof uses the same techniques as that of Appel and Haken: discharging and reducibility. The new proof, however, makes improvements in the complexity of the arguments. Hopefully these improvements will help people to better understand and appreciate Appel and Haken's method.

To describe the improvements in more detail requires a discussion of the discharging method. Simple reductions show that one needs [to] only consider plane triangulations of minimum degree five.

An easy manipulation of Euler's formula gives the following equality for these graphs: $\sum_{v \in V(G)}(6 - deg(v)) = 12$.[2] This value $6 - deg(v)$ has come to be known as the charge of v. The vertices of degree five are the only vertices of positive charge. The vertices of degree at least seven have negative charge and are known as major vertices.

The discharging method is to locally redistribute the positive charge from the vertices of degree five into the major vertices. The sum of the new charges will equal the sum of the old charges, and thus the[re] will be a vertex which has its new charge positive, known as an overcharged vertex.

The structure of the graph close to an overcharged vertex is determined by the rules that were used to discharge the vertices of degree five. Each possible structure that can yield an overcharged vertex must be examined [to] find within it some *configuration* that is reducible (provably cannot exist in minimal counterexample to 4CT). Thus, there are the two steps of the proof of 4CT.

Discharging: defining a set of discharging rules which in turn gives a list of configurations that a plane triangulation of minimum degree five must have.

Reducibility: showing that no minimal counterexample to 4CT can contain any of these configurations.

The two forms of reducibility that Appel and Haken use are known as C-reducibility and D-reducibility. The idea of D-reducibility is that no matter what coloring the ring (border of the configuration) has, it can be changed by Kempe chains (swapping the colors of an appropriate 2-colored subgraph) into a coloring that extends into a coloring of the configuration. C-reducibility is the same idea, except with first replacing the configuration [b]y a smaller configuration, thus restricting the possible colorings of the ring. Bernhart (see [GS]) found a new form of reducibility which can show some configurations reducible that D- and C-reducibility cannot.

Although we were able to produce six configurations which were reducible by the block count method, these configurations turned out not to be needed.

The new proof still uses only D- and C-reducibility, which were clearly defined by Heesch [Hee1] based upon ideas of Birkhoff [Bir].

The primary discharge rule that Appel and Haken use is the following:

A vertex x of degree five originally has a charge of 1. Send a charge of 1/2 from x to each major neighbor of x. Unfortunately, this simple rule is not enough to prove 4CT. It yields a list of configurations, but not all of them are reducible. So, for each

[2] This formula is due to A. B. Kempe, 1879 [Kem2].

non-reducible configuration, they define secondary discharge rules, which move the charge around a bit more.

These new rules produce the need for even more rules, and so on, but eventually the process stopped with a list of 1476 reducible configurations. The total number of secondary discharge rules that they used was 486. A better primary discharge rule permits improvement in both of these areas.[3]

Here is the primary discharge rule used in the new proof. Imagine each vertex x of degree five expelling its positive charge equally in each of the five directions around it. Thus, x will send 1/5 to each of its neighbors. The major vertices have a negative charge that attracts this positive charge that was expelled. Thus, if the neighbor y of x is major, it absorbs this 1/5. If the neighbor y is not major, the charge just keeps going, splitting half to the left and half to the right. Let p and r be the common neighbors of x and y, to the left and right the edge xy.

The left 1/10 rotates counterclockwise through the neighbors of p, while the right 1/10 rotates clockwise through the neighbors of r. If $deg(p) \geq 8$, its attraction is so great, that the 1/10 doesn't make it to the next neighbor; this charge gets absorbed by p. Otherwise, the 1/10 rotates until it reaches a major neighbor of p, unless deg $p = 7$, and it has rotated through four neighbors; in this case p absorbs it. Similarly, for r. Using this primary rule, only 20 secondary discharge rules are necessary to produce a list of 638 reducible configurations . . .

The largest size ring that Appel and Haken use is a 14-ring; their original list had 660 14-rings. The list of 638 mentioned above contains 161 14-rings. It is not known whether 14-rings can be avoided altogether, but at least 12-rings appear to be necessary . . .

Totally automating the discharge analysis allowed us to try several heuristics on how to make these choices. Having the discharge analysis automated also hinders the possibility of errors creeping in; a human error was found in Appel and Haken's discharge analysis (its correction can be found in [AH4, p. 24]).

Recently, Appel and Haken [AH4] have proven a quartic algorithm to Four-Color planar graphs using their list of 1476 reducible configurations . . . we have found a quadratic algorithm to Four-Color planar graphs . . .

The reducibility and discharging programs that were used to complete the new proof of 4CT will soon be available by anonymous ftp. The total amount of computer time required to prove 4CT on a Sun Spark 10 is less than twenty-four hours.

About two years later, on February 19, 1996, I attended Paul Seymour's plenary talk at the Southeastern International Conference on Combinatorics, Graph Theory, and Computing at Florida Atlantic University. I knew that the new proof was superior to the original one in a number of ways. Yet, I was wondering what compelled the authors to look for another computer-aided proof of 4CT. Paul Seymour addressed it right in the beginning of his talk, as I was jotting down his words:

---

[3] Better discharge rules of the new proof allowed a reduction of secondary discharge rules from 486 in Appel and Haken to just 20, which resulted is a much more accessible ideologically proof and a 43-page paper [RSST] vs. the 741-page oversized book [AH4] by Appel and Haken.

> It was difficult to believe it [Appel–Haken's proof]: you can't check it. First you need a computer. Second, non-computer part is awful: it contains hundreds of pages of notes. You can't understand. You are not quite sure that the theorem is true. Nobody checked the proof. This is a bit scary.
>
> We assumed 4CT is true in earlier work, so we had to have a sure proof. General framework is the same, but details are better.

The talk ended with questions and Paul Seymour's answers:

Erdős: Is there going to be a normal proof?
Seymour: I don't have any reason to think it is impossible. I try it from time to time.
Soifer: What are approaches to "normal proof"?
Seymour: I am not going to tell you my wrong proofs. Start with a triangle, flip it over, put a rubber band around, look for a smaller set of reducible configurations.
Soifer: Why did you use Sun Microsystems workstation in your solution?
Seymour: This is what I have in my office.
Soifer: How long does it take to verify your proof?
Seymour: Computer can verify the proof in 5 min; 6 months by hand.

Table 23.2 compares the two proofs side by side.

Table 23.2 Comparison of the Two 4CT Proofs

	Appel–Haken–Koch	Robertson–Sanders2 Seymour–Thomas
Number of secondary discharging rules	486	20
Number of unavoidable configurations	1476	638
Computer time to prove	1200 h	24 h
Computer time to verify	Not available	5 min
Speed of graph coloring algorithm	quartic	Quadratic
Number of pages in the final publication	741	43

When the paper [RSST] with the new proof was submitted on May 25, 1995, it consisted of just 43 pages – a vast improvement over the 741+XV-page oversized monograph [AH4]. With great advantages of the new proof (see Table 23.2), let us not forget, however, that Appel and Haken discovered a proof first. And let us remember that "the most notorious paper in the history of graph theory: the 1879 work by A. B. Kempe [Kem2] that contains the fallacious proof of the Four Color Theorem" [Ste], "The Kempe Catastrophe" [Saa1] – paved the way!

March 2005 brought a new development in 4CT saga, when Georges Gonthier of Microsoft-Cambridge, UK, produced "a formal proof of the famous Four Color Theorem that has been fully checked by the Coq proof assistant." "It's basically a machine verification of our proof," wrote Paul Seymour in his January 17, 2008 e-mail to me. Let me give the podium to Georges Gonthiers for the assessment of his work in the unpublished but posted on the web paper [Gon]:[4]

[4] *The Economist* (April 2–8, 2005) reported "Dr Gonthier says he is going to submit his paper to a scientific journal in the next few weeks." This, however, to the best of my knowledge has not happened.

> We took the work of Robertson et al. as our starting point, reusing their optimized catalog of 633 reducible configurations, their cleverly crafted set of 32 discharge rules, and their branch-and-bound enumeration of second neighborhoods [RSST]. However, we had to devise new algorithms and data structures for performing reducibility checks on configurations and for matching them in second neighborhoods, as the C integer programming coding tricks they used could not be efficiently replicated in the context of a theorem prover, which only allows pure, side effect free data structures (e.g., no arrays). And performance *was* an issue: the version of the Coq system we used needed three days to check our proof, whereas Robertson et al. only needed three hours . . . ten years ago! (Future releases of Coq should cut our time back to a few hours, however.)
>
> We compensated in part this performance lag by using more sophisticated algorithms, using multiway decision diagrams (MDDs) . . . for the reducibility computation, concrete construction programs for representing configurations, and tree walks over a circular zipper . . . to do configuration matching. This sophistication was in part possible because it was backed by formal verification; we didn't have to "dumb down" computations or recheck their outcome to facilitate reviewing the algorithms, as Robertson et al. did for their C programs [RSST].
>
> Even with the added sophistication, the program verification part was the easiest, most straightforward part of this project. It turned out to be much more difficult to find effective ways of stating and proving "obvious" geometrical properties of planar maps. The approach that succeeded was to turn as many mathematical problems as possible into program verification problems.

In the concluding section, "Looking ahead," Gonthier sees his success as a confirmation that the "programming" approach to theorem proving may be more effective than the traditional "mathematical" approach, at least for researchers with computer science background:

> As with most formal developments of classical mathematical results, the most interesting aspect of our work is not the result we achieved, but how we achieved it. We believe that our success was largely due to the fact that we approached the Four Colour Theorem mainly as a *programming* problem, rather than a *formalization* problem. We were not trying to replicate a precise, near-formal, mathematical text. Even though we did use as much of the work of Robertson et al. as we could, especially their combinatorial analysis, most of the proofs are largely our own.
>
> Most of these arguments follow the generate-and-test pattern... We formalized most properties as computable predicates, and consequently most of our proof scripts consisted in verifying some particular combination of outcomes by a controlled stepping of the execution of these predicates. In many respects, these proof scripts are closer to debugger or testing scripts than to mathematical texts. Of course this approach was heavily influenced by our starting point, the proof of correctness of the graph colouring function. We found that this programs-as-proof style was effective on this first problem, so we devised a modest set of tools (our tactic shell) to support it, and carried on with it, generalizing its use to the rest of the proof. Perhaps surprisingly, this worked, and allowed us to single-handedly make progress, even solving subproblems that had stumped our colleagues using a more orthodox approach.
>
> We believe it is quite significant that such a simple-minded strategy succeeded on a "higher mathematics" problem of the scale of the Four Colour Theorem. Clearly, this is

the most important conclusion one should draw from this work. The tool we used to support this strategy, namely our tactic shell, does not rely on sophisticated technology of any kind, so it should be relatively easy to port to other proof assistants (including the newer Coq). However, while the tactic shell design might be the most obvious byproduct of our work, we believe that it should have wider implications on the interface design of proof assistants. If, as this work seems to indicate, the "programming" approach to theorem proving is more effective than a traditional "mathematical" approach, and given that most of the motivated users of proof assistants have a computer science background and try to solve computer-related problems, would it not make interface of a proof assistant more similar to a program development environment, rather than strive to imitate the appearance of mathematical texts?

These Georges Gonthier's words proved to be prophetic. Near the end of this new edition, you will meet a computer scientist who used solvers to advance the chromatic number of the plane problem.

23.3 The Future that Commenced 80 Years Ago: Hugo Hadwiger's Conjecture

There are a number of conjectures that, if proved, would imply the Four-Color Theorem. In 1943, Hugo Hadwiger posed the most prominent of these conjectures [Had3].

An *edge contraction* of a graph G consists of deleting an edge and "gluing" together (i.e., identifying) its incident vertices. We say that a graph G is *contractible* to a graph H if H can be obtained from G by a sequence of edge contractions. In this case, H is called a *contraction* of G and G is said to be *contractible* to H.

We can view the Hadwiger conjecture $H(n)$ as a series of conjectures, one for every positive integer n.

The Hadwiger Conjecture *H*(*n*) 23.1 (1943, [Had3]). Every connected n-chromatic graph G is contractible to K_n.

The truth of the conjecture $H(n)$ for $n < 5$ has been proved in 1952 by G. A. Dirac [Dir]. But it is the case $H(5)$ that proved to be particularly important. Why? Because the following equivalence takes place:

Theorem 23.2 $H(5)$ is equivalent to the Four-Color Theorem ($4CT$).

$H(5) \Rightarrow 4CT$. Proof in this direction is very simple. Given $H(5)$, assume G is a planar graph that is not 4-colorable. But then by $H(5)$, G is contractible to K_5, which is absurd since any contraction of the planar G must be planar as well.

$4CT \Rightarrow H(5)$. Proof in this direction is more involved: here is its sketch. Assume $4CT$ is true. Let G be a 5-chromatic graph, not contractible to K_5, of the minimum order with respect to this property. Then, it can be shown that G is 4-connected. In 1937, before Hadwiger formulated his conjecture, K. Wagner [Wag] showed that a 4-connected graph not contractible to K_5 is planar. Thus, G is a planar graph of chromatic number 5, which contradicts $4CT$. ■

The most surprising result was published in 1993 by Neil Robertson, Paul D. Seymour, and Robin Thomas. They proved that $H(6)$ is also equivalent to $4CT$!

Theorem 23.3 ([RST], 1993). The following statements are equivalent:

(a) $4CT$;
(b) $H(5)$;
(c) $H(6)$.

The authors [RST] comment in the abstract:

> We show (without assuming the 4CC) that every minimal counterexample to Hadwiger's conjecture [for 6] is "apex", that is, it consists of a planar graph with one additional vertex. Consequently, the 4CC implies Hadwiger's conjecture [for 6], because it implies that apex graphs are 5-colourable.

Right after his plenary talk on February 19, 1996, at the Southeastern International Conference on Combinatorics, Graph Theory, and Computing in Boca Raton, Florida, I asked Paul Seymour about his and Robertson's result on the relationship between $4CT$ and Hadwiger's Conjecture. Paul replied as follows:

> I believe that all of them (Hadwiger's Conjectures for various n) are equivalent. We have a result that if 4CC is true, then for every n there is $f(n)$ such that for Hadwiger's Conjecture to be true, it suffices to check graphs of order not exceeding $f(n)$.

Paul Seymour's Conjecture 23.4 All Hadwiger's conjectures for various $n \geq 5$ are equivalent to each other and equivalent to 4CT.

It seems plausible that a computer-free proof of the Four-Color Theorem (shouldn't it exist!) will come as a consequence of a (computer-free) proof of the Hadwiger Conjecture $H(n)$ for some $n > 4$.

Another way of finding a short computer-free proof was suggested by Appel and Haken [AH1]:

> Of course, a short proof of the Four-Color Theorem may some day be found, perhaps, by one of those bright high school students.

I would love that, amen. Perhaps, by one of the winners of the Soifer (formerly Colorado) Mathematical Olympiad!

Chapter 24
How Does One Color Infinite Maps? A Bagatelle

How does one measure fun in mathematics? Certainly not by the length of exposition. This is a short chapter, a bagatelle. I hope nonetheless that you will enjoy it.

We know (4CT) that every finite map in the plane is 4-colorable. What about maps with infinitely many countries? This sounds like a natural question, which I have heard from various people at various times. In particular, Peter Winkler, then Director of Fundamental Mathematics Research at Bell Labs and now professor of mathematics at Dartmouth, asked me this question on October 11, 2003, right after my talk at the Princeton-Math Discrete Mathematics Seminar. The 4-colorability of infinite maps follows from 4CT due to De Bruijn–Erdős' Compactness Theorem 27.1, which will appear later in this book. Let us record this corollary formally.

Infinite Map Coloring Theorem 24.1 Every map with infinitely many countries is 4-colorable.

Proof Given an infinite map M. As we know, we can translate the problem of coloring M into the problem of coloring the planar graph $G(M)$. Since by 4CT every finite subgraph of $G(M)$ is 4-colorable, $G(M)$ is 4-colorable as well by De Bruijn–Erdős' Compactness Theorem 27.1. ■

In late December 2004, I was giving talks at the Mathematical Sciences Research Institute in Berkeley, California. There my old friend Professor Gregory Galperin showed me a proof of Theorem 24.1. His proof was longer and worked only for countable maps. Nevertheless, I have got to share it with you here because of its striking beauty. Plus, of course, Galperin's proof – unlike the proof above – does not require the Axiom of Choice in its full force.

Countable Map Coloring Theorem 24.2 Every map with countably many countries is 4-colorable.

Proof by G. Galperin [Gal]. Given a countable map M. Enumerate the countries of the map by positive integers: 1, 2, . . ., n, . . . Let integers 1, 2, 3, and 4 be the names of the colors to be used.

© Alexander Soifer 2024
A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_24

Let n be a positive integer. Take a map consisting of the first n countries. By 4CT, there is a 4-coloring of the submap consisting of these n countries. Let the colors assigned to these countries be $a_1, a_2, \ldots, a_n$, respectively (of course, each a_i is equal to 1, 2, 3, or 4). We represent this coloring by a number x_n in its decimal form: $x_n = 0.a_1a_2\ldots a_n$.

As we do this for each positive integer n, we end up with the sequence $S = \{x_1, x_2, \ldots, x_n, \ldots\}$ of real numbers. Since S in bounded, $S \subset [0, 1]$, by the Bolzano–Weierstrass theorem, it contains a convergent subsequence S'.

$$S' = \{x_{i_1}, x_{i_2}, \ldots, x_{i_n}, \ldots\},$$

where $i_1 < i_2 < \ldots < i_n < \ldots$ Let the limit point of S' be y, which in decimal form looks like

$$y = 0.y_1y_2\ldots y_n\ldots$$

It is easy to prove that the sequence $y_1, y_2, \ldots, y_n, \ldots$. delivers a (proper) 4-coloring of the respective regions $1, 2, \ldots, n, \ldots$ Indeed, since all the decimal digits of all x_i were 1, 2, 3, or 4, the same must be true about the decimal digits $y_1, y_2, \ldots, y_n, \ldots$ of the limit point y. Two neighboring regions could not be assigned the same color by this rule, for otherwise they would have been assigned the same color already in a coloring that we decoded by one of the x_{i_n}. ■

Chapter 25
Chromatic Number of the Plane Meets Map Coloring: Townsend–Woodall's 5-Color Theorem

In Chapter 8, I described Douglas R. Woodall's 1973 attempt to obtain a result on chromatic number of the plane under an additional condition that monochromatic sets are closed or simultaneously divisible into regions [Woo1]. Six years after his publication, Stephen P. Townsend found a logical mistake in Woodall's proof, constructed a counterexample showing that Woodall's proof cannot work and went on to discover his own proof of the following major result.

The Townsend–Woodall Theorem 25.0 [Tow2]. Every 5-colored planar map contains two points of the same color unit distance apart.

In this chapter, I will convey the story of the proof and the proof itself.

25.1 On Stephen P. Townsend's 1979 Proof

This story reminds me the famed Victorian Affair, which we discussed in Chapters 20 and 21 of this part. To concisely sum it up, in 1879 Alfred B. Kempe published a proof of the 4-Color Theorem, in which 11 years later, Percy J. Heawood found an error and constructed a counterexample to demonstrate its irreparability. Heawood salvaged Kempe's proof as the 5-Color Theorem, but the Four-Color Conjecture had to wait nearly a century more for its proof.

Our present story started with Douglas R. Woodall's 1973 publication, in which 6 years later Steven P. Townsend found an error and constructed a counterexample to demonstrate its irreparability. So far, the two stories are very similar. Unlike its Victorian counterpart, however, Townsend went on to prove Woodall's statement and so I thought the new story had a happy end – until February 11, 2007, when I asked Stephen Townsend about "the story of the proof." The surprising reply reached me by e-mail on February 20, 2007:

Story of the Proof

I first became interested in the plane-colouring problem in 1977 or 1978. At that time I was a lecturer in the Department of Mathematics at the University of Aberdeen, having just completed my doctoral thesis (in Numerical Analysis). I had read an article that

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_25

> listed some of the unsolved problems in Combinatorics at that time, and this one caught my attention.
>
> I was totally unaware of Douglas's 1973 proof, which was both my folly and my good fortune. Folly, in that I should have conducted a more exhaustive literature search before devoting time to the problem. Good fortune in that had I been aware of Douglas's paper I would not have spent any time on the problem; I certainly would not have had the temerity to check Douglas's proof for accuracy. It should be noted that I was a numerical analyst, not a combinatorialist, so my awareness of the field of combinatorics was somewhat limited, in spite of brushing shoulders at Aberdeen with some eminent contributors to the field.
>
> It was not until I had completed the proof, and was considering what references to include, that I came upon Douglas's paper. I was both devastated and puzzled. The puzzlement came from my intimate knowledge of the difficulties of certain aspects of the proof and the fact that Douglas seemed to have produced a proof that circumnavigated these difficulties. So it was with an attitude of "how did he manage this?" that I went through his proof and consequently spotted the error.
>
> A colleague at Aberdeen, John Sheehan, whom I'm sure you will have come across, encouraged me nonetheless to submit my proof for publication, but including a reference to Douglas's work. The rest I think you know.

Yes, Stephen Townsend was lucky, for not only was he first to produce a proof – he also discovered the statement of the result on his own, albeit after Woodall's publication – and this Townsend's independent discovery was a necessary condition for finding the proof.

Townsend's good luck, however, ran into a wall, when the *Journal of Combinatorial Theory*'s Managing Editor and distinguished Ramsey theorist Bruce L. Rothschild wrote thus to Townsend on April 3, 1980:

> The Journal of Combinatorial Theory – Series A is now trying very hard to reduce its large backlog, and we ask all our referees to be especially attentive to the question of the importance of the papers. In this case the referee thought that the result was not of great importance. In view of our backlog situation then, we are reluctant to publish the paper. However, since it does correct an error in a previously published paper, we would like to have a very short note about it. Perhaps, you would be willing to do the following: Write a note pointing out the error, stating the theorem (Theorem 1) (without proof) used to get around the trouble, and that the theorem must be used with care to get around the problem.

Stephen P. Townsend had satisfied the Editor (what choice did he have!) and produced a 2-page proof-free note [Tow1], which was published the following year. This is where the story was to end in 1981.

No blame should be directed at Douglas R. Woodall – we all make mistakes (except those of us who do nothing). The mistake notwithstanding, Woodall's 1973 paper has remained one of fine works on the subject. Moreover, he was the one who first alerted me to his mistake and Townsend's 2-page note. "I am a fan of your 1973 paper," I wrote to Woodall in the October 10, 1993, e-mail, in which I called [Tow1] "the Townsend's addendum." The following day Woodall replied as follows:

I will put a reprint in the post to you today, together with a photocopy of Townsend's "addendum," as you so tactfully describe it. (The fact is, I boobed, and Townsend corrected my mistake.)

However, regret is in order about the decision by the *Journal of Combinatorial Theory Series A* (*JCTA*). While they apparently (and correctly) assessed Woodall's paper as being "of great importance" (an impossible test if one interprets it literally), they labeled Townsend's paper as "not of great importance" and denied its readers – and the world – the pleasure and the profit of reading Townsend's proof of the major result.

I have corrected *JCTA*'s quarter-a-century old mistake, when I published Townsend's complete paper in the April 2005 issue of *Geombinatorics* [Tow2]. Townsend's work was preceded by my historical introduction [Soi24], a version of which you have just read. I ended that introduction with the words I would like to repeat here: It gives me a great pleasure to introduce and publish Townsend's proof. In my opinion, it *is* of great importance – judge for yourselves!

It pains me to see that most researchers in the field are still unaware of Woodall's mistake and Townsend's proof. It suffices to look at the major problem books to notice that: the 1991 book by Croft–Falconer–Guy [CFG] and the more recent 2005 book by Brass–Moser–Pach [BNP] give credit to Woodall and do not mention Townsend. I hope this chapter will inform my esteemed colleagues of the correct credit and Townsend's achievement.

Stephen Phillip Townsend was born on July 17, 1948 in Woolwich, London, England. He received both graduate degrees, Master's (1972) and doctorate (1977) from the University of Oxford. Townsend has been a faculty first in the department of mathematics (1974–1980) and then in the department of computer science (1982–present) at the University of Aberdeen, Scotland. Since 1995, he has also been Director of Studies (Admissions) in Sciences. In addition to publications in mathematics, Steven's list of publications includes "Women in the Church – Ordination or Subordination?" (1997).

25.2 Proof of Townsend–Woodall's 5-Color Theorem

In this section, I will present Stephen P. Townsend's proof. As you now know, it first appeared in 2005 in *Geombinatorics* [Tow2]. However, when I was writing this book chapter, I asked Stephen to improve the exposition, make his important proof more accessible to the reader not previously familiar with topology, and include plenty of drawings to help you visualize the proof. He did it, quite brilliantly. Thus, presented below exposition of the proof has been written by Professor Townsend especially for this book in 2007.

He starts with a few basic definitions from general (point set) topology.

Definitions

A pair of points in the Euclidean plane E^2 unit distance apart having the same color is called a *monochrome unit*.

Let S and T be subsets of E^2. S is said to *subtend T at unit distance* if T is the union of all unit circles centered on points in S.

Let A be any closed, bounded doubly connected set in E^2 containing a circle of unit radius. If the removal of any point in A renders A simply connected, then such a point is called a *cut point* of A. If A has no cut points, its interior A^0 is said to be a *unit annulus*. If A has a finite number of cut points (which must occur on a circle of unit radius), then A^0 is said to be a *finitely disconnected unit annulus* (Fig. 25.1).

A *planar map* (Fig. 25.2) is an ordered pair M(S, B) where S is a set of mutually disjoint bounded finitely connected open sets (*regions*) in E^2 and B is a set of simple closed curves (*frontiers*) in E^2 satisfying

(i) The union of the members of S and B forms a covering of E^2;
(ii) There exists a one-to-one function F:S→B such that b = F(s), s∈S, is the *exterior boundary* of s;

the *boundary* of s∈S is the union of F(s) and at most a finite number of other members of B, which are the *interior boundaries* of s.

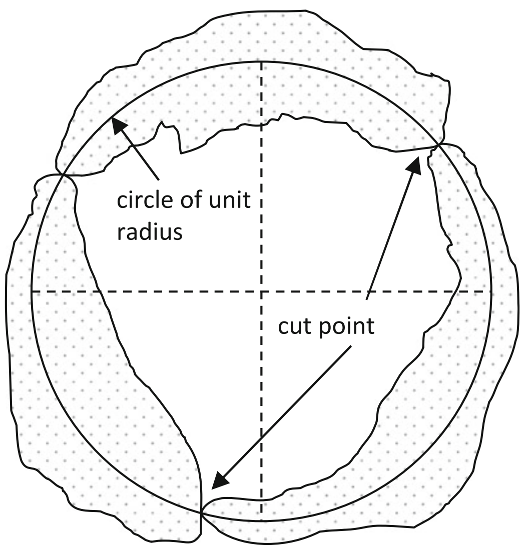


Fig. 25.1 Finitely disconnected unit annulus

A point on the boundary of s is called a *boundary point* of s. A boundary point, which lies on the boundary of *k* regions, $k \geq 3$, is called a *vertex of degree k*. A closed subset of a frontier b∈B, which is bounded by two vertices and contains no other vertices, is called an *edge* of each region for which *b* is part of the boundary. Two regions are *adjacent* if their boundaries contain a common edge or a common frontier.

The above definition is more general than the usual definition of a planar map, which requires each region s∈S to be simply connected and requires each frontier b∈B to contain at least two vertices.

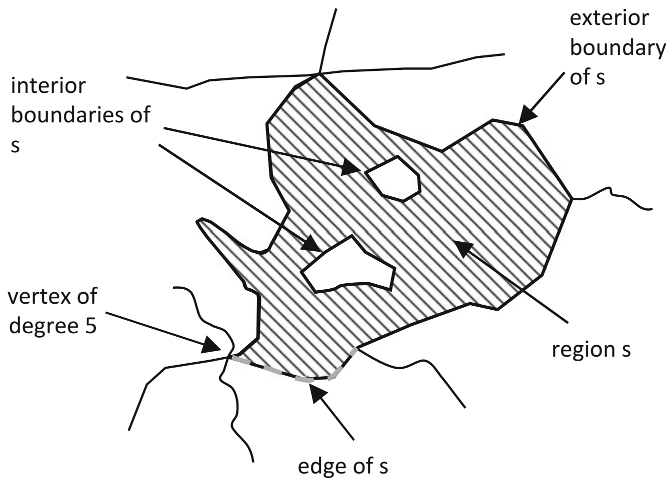


Fig. 25.2 Planar map

An *r-coloring* of a planar map is a function $C_r{:}E^2 \rightarrow \{c_1, c_2, \ldots, c_r\}$ where C_r is constant over each region in S and where a boundary point is given the color of one of the regions in the closure of which it lies.

Initial Observations

To prove that an r-colored map must contain a monochrome unit, it is sufficient to examine only those r-colored maps satisfying

(i) Each region has no interior boundaries, i.e., its closure does not contain the closure of any other region;
(ii) Different regions of the same color have no common boundary points.

This is best understood by observing that every r-colored map with no monochrome units may be simplified to an r-colored map with no monochrome units satisfying (i) and (ii) above as follows.

(a) For each region s with interior boundaries, remove these boundaries and assimilate into s all regions whose closures are contained in the closure of s.
(b) Remove any edges common to adjacent regions of the same color.
(c) For each vertex $\underline{v}$ which is a boundary point of two nonadjacent regions of the same color, choose $\varepsilon > 0$ sufficiently small and describe an ε-neighborhood whose closure contains $\underline{v}$ and whose intersection with each of the two regions is non-null, coloring this ε-neighborhood the same color as the two regions, and thus forming one new region incorporating the original two and the ε-neighborhood (see Fig. 25.3).

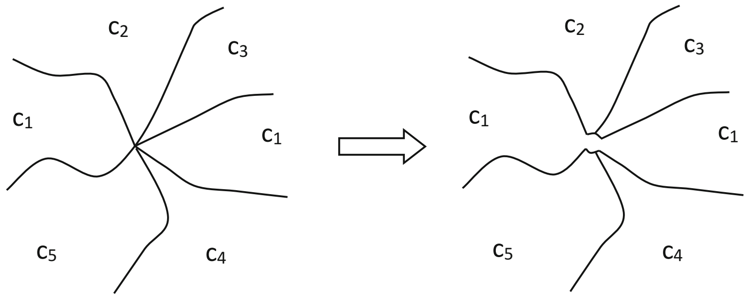


Fig. 25.3 Incorporating the original two regions and the ε-neighbourhood

Note that a consequence of (ii) is that we do not need to consider vertices of degree greater than r in an r-colored map. A sequence of theorems now follows, concluding with the main result that *every 5-colored planar map contains a monochrome unit.* Here is an outline of the proof:

(1) We show that every 4-colored planar map contains a monochrome unit;
(2) We show that every 5-colored planar map containing a vertex of degree 3 contains a monochrome unit;
(3) We show that every 5-colored planar map without a monochrome unit must contain a vertex of degree 3;
(4) For 2 and 3 both to be true, every 5-colored planar map must contain a monochrome unit.

The Proof Townsend presents the proof in stages through five theorems.

Theorem 25.2 Let A^0 be a finitely disconnected unit annulus (see Fig. 25.1), for which a circle of unit radius contained in its closure, A, has at least one arc of length greater than $\pi/3$ containing no cut points of A. Then, any 2-coloring of A^0 contains a monochrome unit.

Outline of Proof The basic argument is as follows (see Fig. 25.4).

1. We assume that A^0 is 2-colored and contains no monochrome unit.
2. Points x and y can be selected from A^0, so that they are differently colored and as close together as we want.
3. The points x and y can also be chosen so that (a) x is unit distance from at most one cut point of A, and (b) y is unit distance from no cut points of A.
4. Point x subtends an arc α of finite length in A^0, each point of which is unit distance from x, and consequently the opposite color to x. Similarly y subtends an arc β in A^0 which is the opposite color to y.
5. Arc α subtends a two-dimensional region, each point of which is unit distance from a point on α. This region intersects A^0 in a band P' of finite width, each point of which must be the same color as x. A similar region subtended at unit distance by arc β intersects A^0 in a band Q', each point of which is the same color as y.
6. Points x and y can be chosen to lie sufficiently close together to make $R = P' \cap Q'$ non-null.
7. But points in R must simultaneously have the color of x and the color of y, which is impossible. Consequently, the initial two assumptions are incompatible.

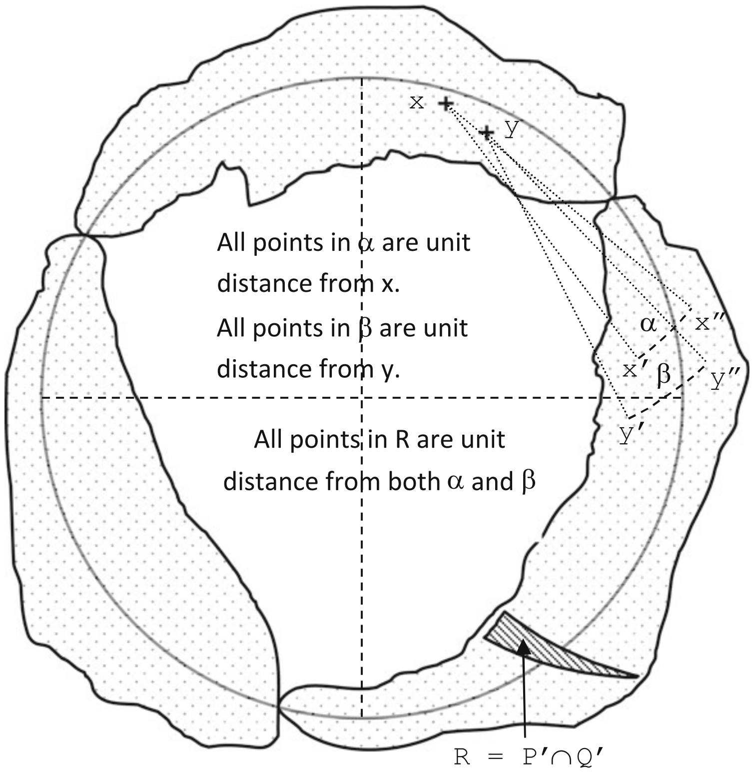


Fig. 25.4

The proof hinges on our ability to construct arcs α and β such that each does not intersect a cut point of A. This will be true if x is unit distance from at most one cut point of A and y is unit distance from no cut points of A.

Tool 25.2 Let γ be any simple arc of length L in A^0 with the following properties:

- γ contains at least two points unit distance apart.
- γ contains at most M points, each unit distance from exactly one cut point of A.
- all other points in γ are unit distance from no cut points of A.
- γ is 2-colored with no monochrome units.

Then given $\varepsilon > 0$, there exists an ε-neighborhood in γ containing a point of each color, one of which is unit distance from no cut points of A and the other of which is unit distance from at most one cut point of A.

Proof Let $d(x, y)$ be the straight line distance between two points x and y on γ, and let $\delta(x, y)$ be the distance along γ between x and y.

By assumption, there exist two points x_1 and y_1 in γ, not both the same color, with $d(x_1, y_1) = 1$. Let $\varepsilon > 0$ be given. The following algorithm uses the method of bisection to prove the lemma (see Fig. 25.5).

1. If M > 1, then from the M points in γ that are unit distance from exactly one cut point of A, select the two that are closest together measuring along γ. Let h be the distance between them along γ.
2. If $h < \varepsilon$ then set $\varepsilon = h$.
3. Set $i = 1$.
4. Let w_i be the point in γ mid-way (by arc-length) between x_i and y_i.
5. If the colors of w_i and x_i are not the same, then put $x_{i+1} = x_i$ and $y_{i+1} = w_i$ otherwise

   put $x_{i+1} = w_i$ and $y_{i+1} = y_i$.

6. If $\delta(x_{i+1}, y_{i+1}) \geq \varepsilon$ increase i by 1 and re-cycle from 4.
7. Points x_{i+1} and y_{i+1} satisfy the requirements.

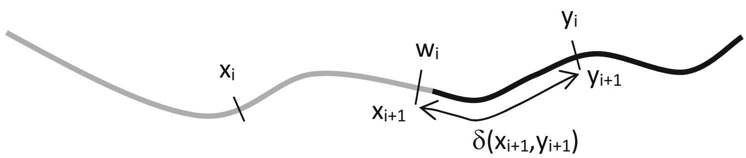


Fig. 25.5

The algorithm terminates in not more than n cycles, where n is the smallest integer such that $\varepsilon 2^n > L$. ■

Proof of Theorem 25.2

Let A^0 be 2-colored with no monochrome units. Let N be the number of cut points of A. Let C be a circle of unit radius contained in A. By assumption, C has at least one arc of length greater than π/3 containing no cut points of A; hence, C has an arc containing no cut points of A, whose end points are unit distance apart. There are at most 2N points on C in A^0 that are unit distance from a cut point of A. Some of these may be unit distance from two different cut points of A, but none can be unit distance from more than two cut points of A. By following a path sufficiently close to C, it is possible to construct a simple closed curve that, apart from the cut points of A, lies entirely within A^0 that contains at most 2N points in A^0 that are unit distance from a cut point of A, and that contains no points in A^0 that are unit distance from more than one cut point of A. (This curve can merely trace the path of C for the most part, deviating only to bypass any points on C in A^0 that are unit distance from two different cut points of A.) There exists an infinite family Γ of such simple closed curves, for each of which there is an arc of finite length containing two points unit distance apart not separated by a cut point (see Fig. 25.6). This must be so since C has two such points, and we can choose the members of Γ to be as close to C as required. For any given $\varepsilon > 0$, this arc contains an ε-neighborhood in which lies a point of each color, one of which is unit distance from at most one cut point of A, and the other of which is unit distance from no cut points of A (by Tool 25.3).

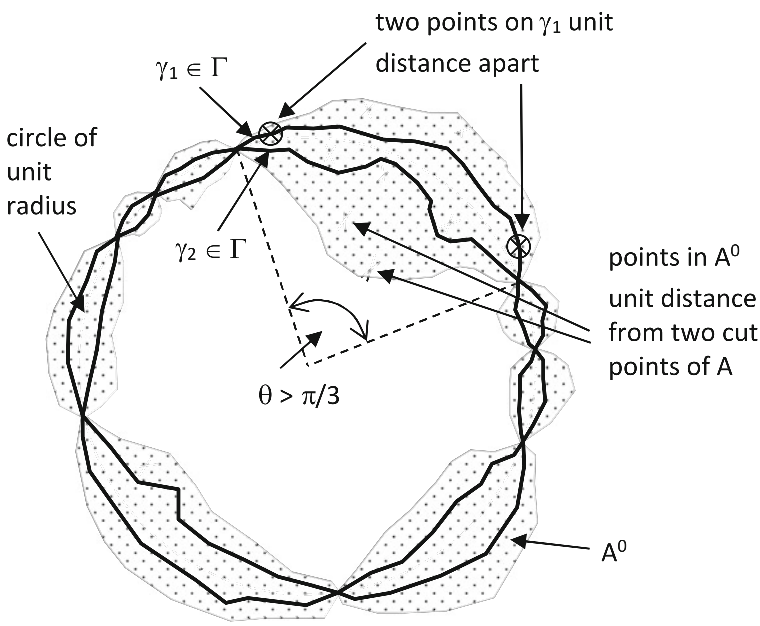


Fig. 25.6

Let γ_1 and γ_2 be members of Γ. Let x and y be two differently colored points in an ε-neighborhood on γ_1 such that x is unit distance from at most one cut point of A and y is unit distance from no cut points of A.

In A^0, there exists an arc α of unit radius and center x which intersects γ_1 at x' and γ_2 at x'' and no point of which is a cut point of A. (If x is unit distance from one cut point of A, then the arc α can be constructed on the other side of x from this cut point.) Arc α cannot be the same color as x so must be the same color as y. Similarly, there exists an arc β in A^0 of unit radius and center y which intersects γ_1 at y' and γ_2 at y'' and no point of which is a cut point of A. Arc β must be the same color as x.

Let P and Q be sets subtended at unit distance by α and β, respectively. P and Q are finitely disconnected unit annuli, each having one cut point at x and y, respectively, and each intersecting A^0 in a band of finite width between γ_1 and γ_2. Let these bands be P′ and Q′, respectively. All points in P′ must be the same color as x, and all points in Q′ the same color as y. Q′ may be considered to be the image of P′ under a homeomorphism T which depends on $|x–y|$. Defining $d(P', Q') = \sup\{|p-T(p)|:p \in P'\}$, we have $d(P', Q') \to 0$ as $|x-y| \to 0$; in this sense, we say $P' \to Q'$ as $|x-y| \to 0$. There must then exist $\varepsilon > 0$ such that for $|x-y| < \varepsilon$, $P' \cap Q' \neq \emptyset$. But all points in $P' \cap Q'$ must simultaneously be colored the same as x and y, which is impossible. Consequently, the original assumptions are incompatible, and so if A^0 is 2-colored it must contain a monochrome unit. ■

Using this result, it is possible to exclude two configurations from any 4-coloring of E^2 without monochrome units and show as a natural consequence that any 4-colored map in E^2 contains a monochrome unit.

Theorem 25.4 Let E^2 be 4-colored. If for some distinct points x and y, there exist two simple arcs with endpoints x and y, each, excepting the endpoints, being monochrome but not both the same color, then E^2 contains a monochrome unit.

Proof Let the two simple arcs be γ and δ. If $|x - y| > 1$, then both γ and δ contain a monochrome unit.

Assume $|x–y| \leq 1$. Then, the intersection of the sets subtended at unit distance by γ and δ (excluding the endpoints) is a finitely disconnected unit annulus with at most two cut points (see Fig. 25.7). This annulus is 2-colored at most, since it cannot contain the colors of γ and δ, and a circle of unit radius contained in its closure has an arc of length greater than π/3 containing no cut points, and so by theorem 25.2 the annulus contains a monochrome unit. ■

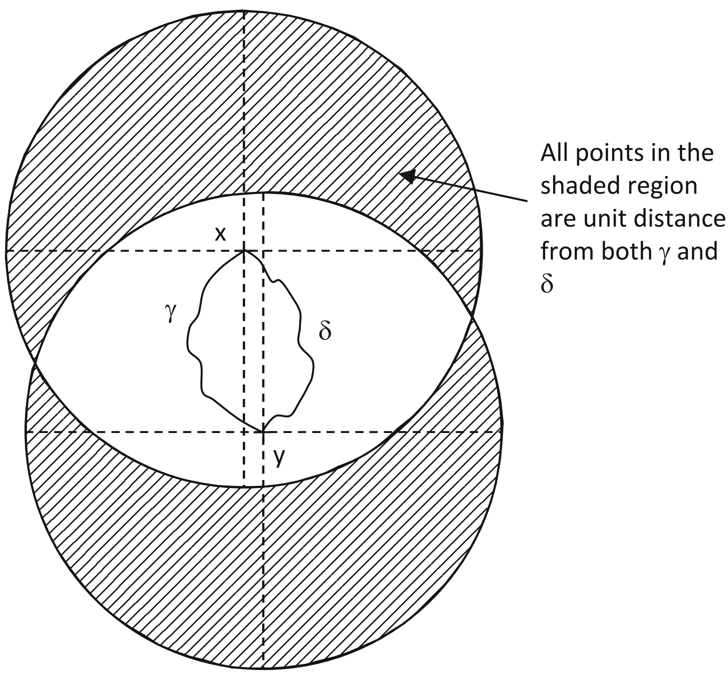


Fig. 25.7

Theorem 25.5 If a 4-coloring of E^2 contains two differently colored, bounded, open connected monochrome sets with a common boundary of finite length, then E^2 contains a monochrome unit.

Proof Let G and F be two such sets, and let x and y be two distinct points on the common boundary. Because the closure of G is a simply connected Jordan region, there is a simple arc γ with endpoints x and y which, apart from its endpoints, lies in G. There exists a similar arc δ in F. By Theorem 25.4, E^2 contains a monochrome unit. ■

Corollary Every 4-colored planar map contains a monochrome unit.

A similar result involving three sets can be proved for 5-colorings of E^2, and again the consequence is that every 5-colored planar map contains a monochrome unit, but this requires a careful proof.

Theorem 25.6 If a 5-coloring of E^2 contains three disjoint, differently colored, bounded, open, connected, monochrome sets each having a common boundary with each of the other two, and all three having one common boundary point, then E^2 contains a monochrome unit.

Proof Let v be the boundary point common to all three sets and let a_1, a_2 and a_3, respectively, be boundary points common to each pair of sets. We assume that these points are distinct and are chosen to be not more than one unit from each other. There are simple closed curves γ_1 colored c_1 containing v, a_1 and a_2; γ_2 colored c_2 containing v, a_1 and a_3; and γ_3 colored c_3 containing v, a_2, and a_3, where in each case the coloring refers to every point on the curve with the possible exception of the points v, a_1, a_2, and a_3 (see Fig. 25.8).

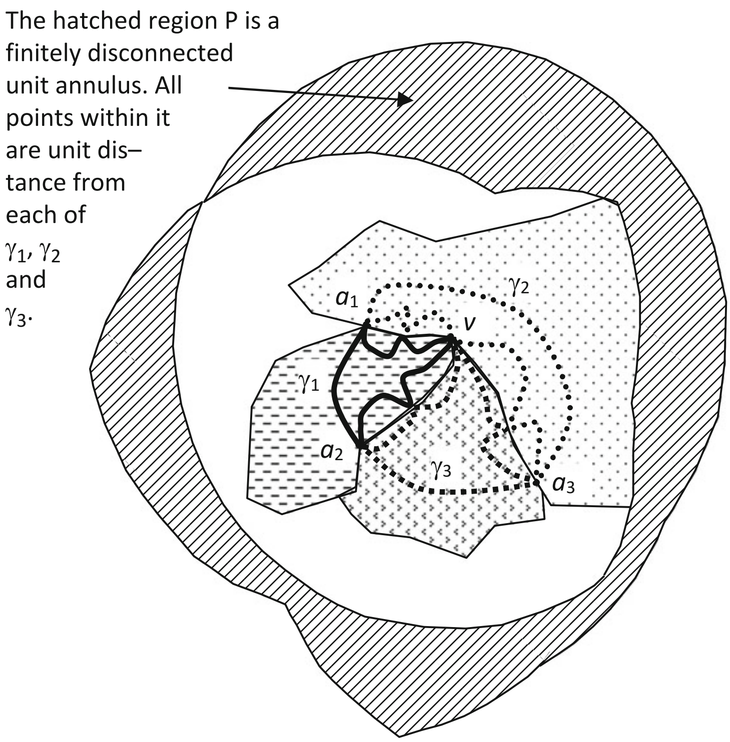


Fig. 25.8

Let P be the intersection of the sets subtended at unit distance by γ_1, γ_2, and γ_3 excepting the points v, a_1, a_2, and a_3. P is either a unit annulus or a finitely disconnected unit annulus with at most three cut points. (A necessary condition for such a cut point to exist is that a set boundary incident to v is an arc of a circle of unit radius; if the cut point exists, then it lies at the center of this circle.) P satisfies the requirements of Theorem 25.2, and since it is 2-colored (viz. not c_1, c_2, or c_3), it must contain a monochrome unit. ■

Corollary Every 5-colored planar map containing a vertex of degree 3 contains a monochrome unit.

Theorem 25.7 Every 5-colored planar map contains a monochrome unit.

Proof We show (i) that every 5-colored planar map with no monochrome units contains a vertex of degree 3 or 4 and (ii) that every such map containing a vertex of degree 4 also contains a vertex of degree 3.

(i) Let v be any vertex in a 5-colored planar map and assume that it has degree 5. Assume that the map has no monochrome units.

Let γ be the boundary of one of the regions which has v as a boundary point. Let a be a point on γ that lies on an edge connected to v. Let b be a point on γ that lies on the other edge connected to v (see Fig. 25.9). Let c be a point on the edge connected to v that is on the opposite side of va to b. Let d be a point on the edge connected to v that is on the opposite side of vb to a.

There is a simple closed curve γ_1 passing through v, a, and b all the points of which, except possibly v, a, and b, are colored c_1. There is a simple closed curve γ_2 passing through v, a, and c all the points of which, except possibly v, a, and c, are colored c_2. And there is a simple closed curve γ_3 passing through v, b, and d all the points of which, except possibly v, b, and d, are colored c_3. Let T_2 be the intersection of the sets subtended at unit distance by γ_1 and γ_2 and let T_3 be the intersection of the sets subtended at unit distance by γ_1 and γ_3 (In Fig. 25.9, T_2 is the hatched region and T_3 is the gray region).

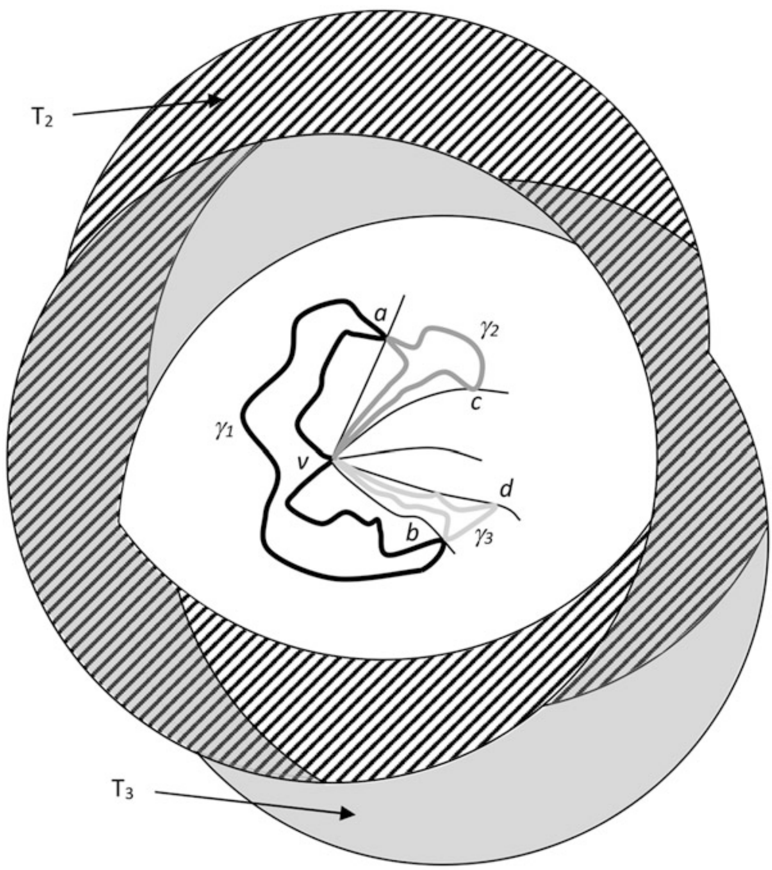


Fig. 25.9

We consider two cases.

(1) The first is when the angle θ subtended at *v* by a line from *a* to *b* (through the region enclosed by γ) is greater than π. The interiors of T_2 and T_3, T_2^0 and T_3^0, respectively, are unit annuli with no cut points, and so by Theorem 25.2 cannot be 2-colored. T_2^0 must contain regions colored c_3, c_4, and c_5, and T_3^0 must contain regions colored c_2, c_4, and c_5. The interior of $T_1 = T_2 \cup T_3$ is a 4-colored unit annulus with no cut points.

There is a vertex in T_1^0. To prove this, assume it is not so. Then, there must be edges in T_1^0 that do not intersect each other in T_1^0, each of which intersects both the interior and the exterior boundary of T_1. Any such edge, *e*, must cross both T_2^0 and T_3^0. This means that the regions on either side of *e* must be colored c_4 and c_5. Consequently, T_1^0 is a 2-colored unit annulus, containing no cut points.

(2) The second case is when the angle θ is not greater than π. It is clear, since *v* is a vertex of degree 5, that the region enclosed by γ may be chosen such that θ is not less than 2π/5. Let a_1 be a point between *v* and *a* on the edge on which *a* lies. Similarly let b_1 be a point between *v* and *b* on the edge on which *b* lies. Choose curve γ_1 so that it passes through a_1 and b_1 as well as *v*, *a*, and *b* and so that all of its points, except possibly *v*, *a*, a_1, b_1, and *b*, are colored c_1. Similarly choose γ_2 to pass through a_1 as well as *v*, *a*, and *c* and γ_3 to pass through b_1 as well as *v*, *b*, and *d*.

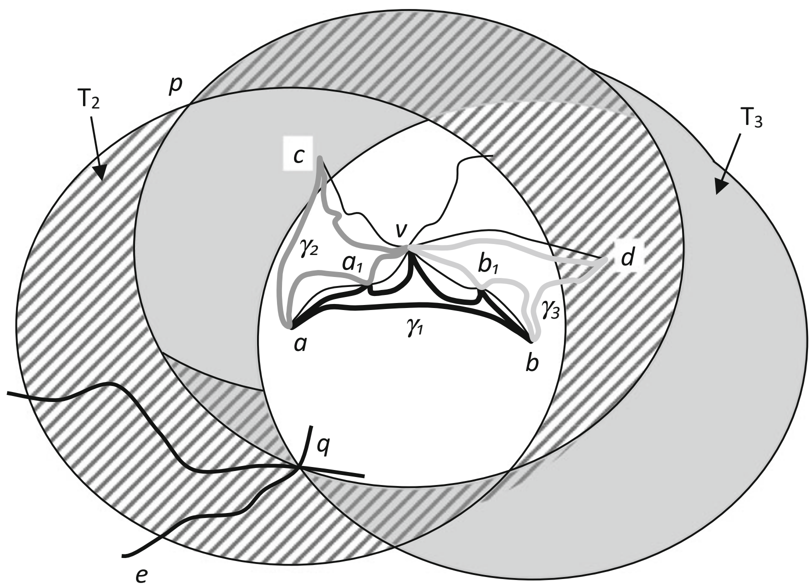


Fig. 25.10

Now each of T_2^0 and T_3^0 is a finitely disconnected unit annulus with at most one cut point (see Fig. 25.10). The single cut point in T_2^0, say *p*, only occurs in the event that *v*, *a*, and a_1 lie on the circle of unit radius centered in *p*. Similarly, the single cut point in T_3^0, say *q*, only occurs in the event that *v*, *b*, and b_1 lie on the circle of unit radius centered in *q*. The interior of $T_1 = T_2 \cup T_3$ is a 4-colored finitely disconnected unit annulus with at most one cut point. This cut point only occurs in the event that *p* and *q* are coincident, and all of

v, a, a_1, b_1, and b lie on the same circle of unit radius. If one of p and q lies on the exterior boundary of T_1 and the other lies on the interior boundary, then the length of the arc of the unit circle centered in v passing through p and q is θ radians, and this means that the distance between p and q is greater than one.

As before we assert, there is a vertex in T_1^0. To prove this, assume it is not so. Then there must be edges in T_1^0 that do not intersect each other in T_1^0, each of which intersects both the interior and the exterior boundary of T_1. Any such edge, e, must cross both T_2^0 and T_3^0 except in the case that e passes through p and remains entirely within T_3 until it reaches the opposite boundary of T_1, or e passes through q and remains entirely within T_2 until it reaches the opposite boundary of T_1. Note that such an edge e cannot pass through both p and q, since this would imply the existence of a monochrome unit in one of the regions on either side of e. Apart from these exceptional edges, every edge in T_1^0 must separate regions colored c_4 or c_5. This means that T_1^0 contains a 2-colored finitely-disconnected unit annulus, containing at most two cut points.

Clearly there is a circle of unit radius in T_1 which has an arc of length greater than $\pi/3$ containing no cut points of T_1^0. Therefore, by theorem 25.2, T_1^0 contains a monochrome unit. This is a contradiction to the initial assumption; consequently, there must be a vertex in T_1^0, and since T_1^0 is 4-colored, this vertex is at most of degree 4.

(ii) We show that every 5-colored planar map with no monochrome units containing a vertex of degree 4 also contains a vertex of degree 3.

Suppose v is a vertex of degree 4 in a 5-colored planar map. Let c_1, c_2, c_3, and c_4 be the colors of the four regions of which v is a boundary point. Let a, b, c, and d be points on the four edges incident to v. Let a_1, b_1, c_1, and d_1 be points on the edges between a and v, b and v, c and v, and d and v, respectively. Assume that the map has no monochrome units.

There exists a simple closed curve γ_1, defined in the closure of the region colored c_1, that passes through v and four of the edge points defined above, and such that every point in γ_1, except possibly v and the four edge points, is colored c_1. Similarly, there exist simple closed curves γ_2, γ_3, and γ_4, each of which contains v and four of the edge points, the points on each curve being colored, respectively, c_2, c_3, and c_4 except possibly v and the edge points. Let the order of the γ_i be chosen such that γ_2 and γ_4 have only the point v in common (see Fig. 25.11).

Let T_i, i=1, 2, 3, 4, be the intersection of sets subtended at unit distance by γ_j, j=1, 2, 3, 4, $j \neq i$. Set T_i is 2-colored with colors c_i and c_5. Define $T = \cup T_i$. The interior of T, T^0, is a unit annulus with center v, possibly finitely disconnected with at most two cut points (see Fig. 25.11).

Every point within T^0 that is on a boundary of a region of the planar map is a boundary point of at most three regions. Suppose none of these boundary points is a vertex. Then there must exist edges that pass from the interior boundary to the exterior boundary of T, which pass through either both of T_1 and T_3 or both of T_2 and T_4. It is possible for an edge to cut T and only cut one of T_1 and T_3 or one of T_2 and T_4, but such an edge must intersect the unit circle centered in v at one of at most four points, these points being cut points (if they exist) of the finitely disconnected annuli which are the interiors of $T_1 \cup T_2$, $T_3 \cup T_4$, $T_1 \cup T_4$, and $T_2 \cup T_3$. There must be edges crossing T which intersect the circle of unit radius centered in v at points other than these four cut points. (If not, then there is an arc of the circle of unit radius centered in v, of length greater than or equal to $\pi/2$, that lies in or on the boundary of a region of the map. But then this region must contain a monochrome unit.) An edge crossing both T_1 and T_3 (or both T_2 and T_4) must separate regions with different colors. But the only color common to both T_1 and T_3 (or both T_2 and T_4) is c_5. We have arrived at a contradiction. Hence, there must be vertices in T^0, and these are of degree 3.

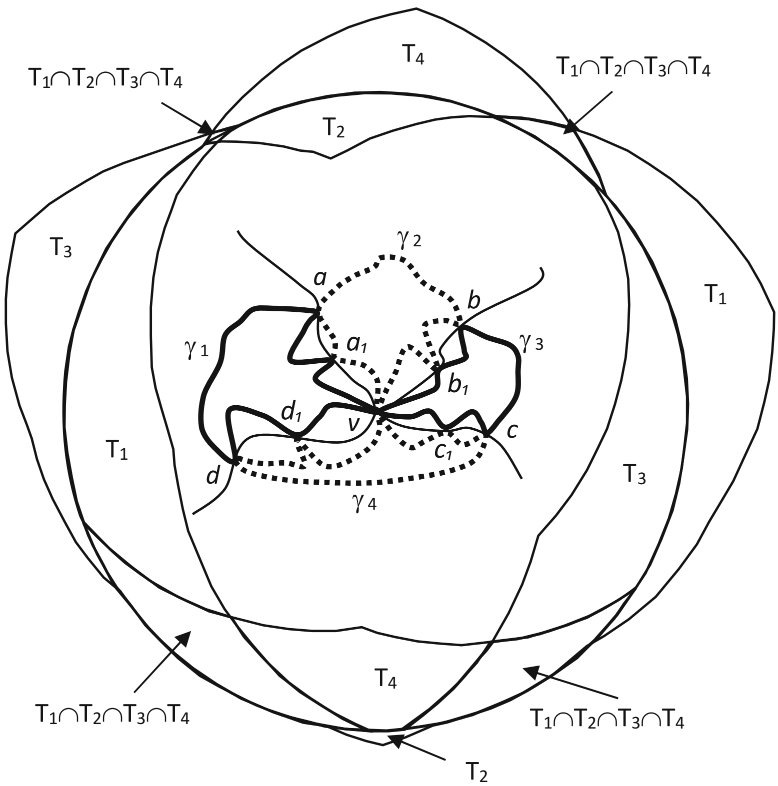


Fig. 25.11

Now, by the corollary of Theorem 25.6, our 5-colored map contains a monochrome unit! ■

Part V
Colored Graphs

Chapter 26
Paul Erdős

I hope several [of my] results will survive for centuries, but we will see.

– Paul Erdős (Talk at the Keszthely, Hungary, 1993, at the Conference dedicated to Paul Erdős' 80th birthday.)

Paul Erdős' contributions to mathematics cannot be measured through his papers alone. Over the years he has traveled extensively among the mathematical centers of the globe. Like the bumblebee, flying from flower to flower transmitting pollen, Paul Erdős has created an enormous cross-pollination effect in mathematics. An Erdős visit to a mathematical center is marked by intense work. Mathematicians gather round and discuss the current problems in their various fields. The resulting interplay of ideas is exhausting and highly productive.

– Joel H. Spencer ([Sp1])

The early involvement of Paul Erdős in problem solving at high school level had a strong influence on his own life-work, and to this day he can make the young feel close to him. This closeness to the young is determined also by another factor: the human side of Erdős, his warmth and compassion, his love of youth, his strong sense of justice, unspoilt and at times childishly naïve.

– Marta Sved ([Sve1])

26.1 The First Encounter[1]

In August 1988, I came to Budapest for an international congress. A Moscow friend suggested to contact a Hungarian mathematician Károly Mályusz, a PhD from Moscow State University, who could show me around the Hungarian capital. Károly kindly decided to be my genie.

[1]This chapter is based on my two essays published in *Geombinatorics* in 1993 [Soi28], and in 1997 [Soi30].

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_26

– Would you like me to show you something – he asked – a place to buy Hungarian souvenirs, or perhaps, a disco to meet beautiful Hungarian girls?

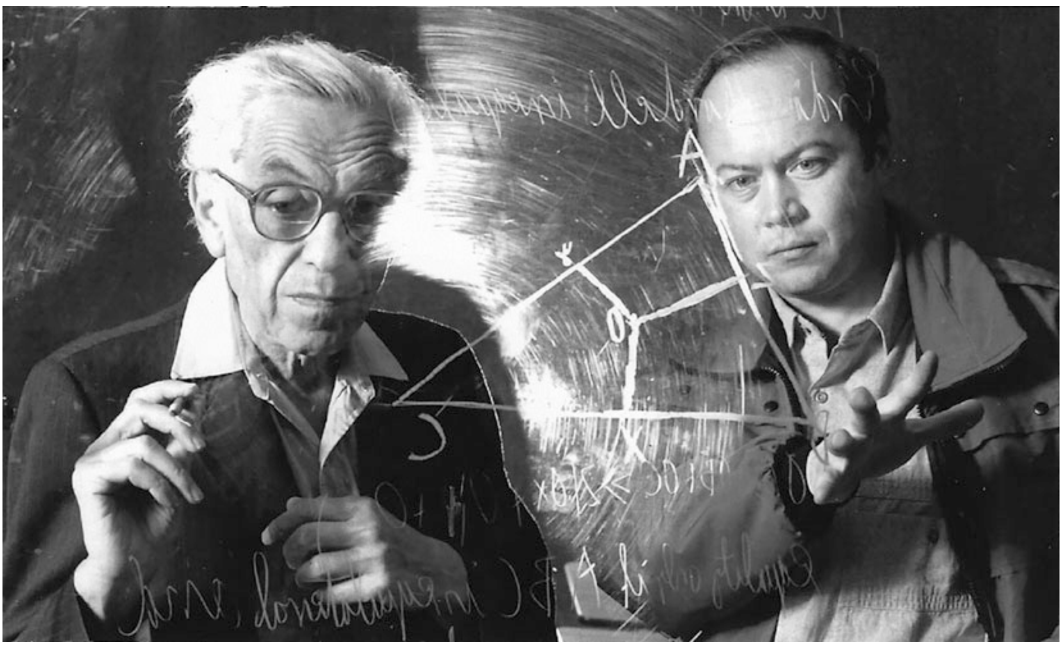
The offer sounded attractive. But, the many legends I heard about Paul Erdős came to my mind, and I replied:

– Is Erdős a real person? Is he in Budapest?
– Of course, he is a real person, but he can be anywhere in the world on any given day.
– Well – said I – you offered to grant me one wish. My choice is to meet Erdős.

The following day Károly called me early with the good news: Paul Erdős was in town and willing to see me. We found Paul in his huge office with high ceilings at the Alfréd Rényi Institute of Mathematics on Reáltanoda 13–15, speaking with two Russian mathematicians, the father Sergei Borisovich Stechkin and the son Boris Sergeevich Stechkins. I joined them. No language was known to all, but every two had a language in common: Erdős and the Stechkins spoke German, Erdős and I used English, and the Stechkins and I knew Russian.

The Russians soon left. Without looking at me, Paul opened with:

– Let x sub 1, . . ., x sub n be n points in the plane no three on a line . . .



Paul Erdős and Alexander Soifer, the first working meeting in Colorado Springs. (Photo by Tom Kimmel, March 1989, University of Colorado, Colorado Springs)

I realized that Paul was formulating a problem and started jotting it down. The problem was beautiful. But did he ask himself or me? Did he want me to solve it right there? Did he want me to offer him a problem in return? I gave him the most difficult problem of that year's Colorado Mathematical Olympiad that I created, and none of some 1,000 participating high school students solved:

– Five points lie inside a triangle of area one . . .

To my disbelief, Paul solved it on the spot! As we were parting, I received a treasure of a gift: two reprints inscribed for me.

The next day I came back. I had an idea but still no solution for his problem. My embarrassment disappeared when Paul said:

– This is an open problem, and I offer . . . dollars for its [first] solution.

The few meetings with Paul during the congress affected my life and started our very special friendship. The idea for the book *Problems of pgom Erdős* occurred to me right then, in August 1988, when for the first time I was listening to Paul presenting "some of my favorite problems" for standing only room of ca. 300 international congress delegates. Erdős' problems were legendary, and as true legends, they were passed from person to person, and sometimes changed in the process to become something else, not intended by the author. Right after the talk, I asked Paul to write such a book, but he replied, "why don't *you* write a book of my problems?" "I envision it as a book of *your* favorite open problems with *your* commentaries for each of the problems," replied I, and added, "I would be happy to publish it if you like." I had a few meetings with Paul Erdős in August 1988 in Budapest. These meetings could make for an enjoyable story elsewhere. During these visits, I convinced Paul that his "Wandering Jew" constant traveling around the world, preaching his mathematics, is important. His problem-posing papers are too. Yet, there is something most important: he ought to write a book of his favorite open problems. Soon Paul agreed to write such a book, and I agreed to publish it, and we signed a contract. However, Paul's nomadic life out of a single suitcase and a garment bag did not lend itself to writing a book, which requires less movement and more contemplation. A few years had passed, and the writing train did not leave the station.

During Paul's next stay with me in Colorado Springs (December 1991–January 1992), we were thinking math, accompanied by the gentle "noise" of Mozart (Beethoven created too much "noise" for Paul), and taking walks in the Garden of the Gods, an old American Indian sacred ground full of remarkable red vertical rocks. Paul wanted to attempt an ascent up the Pikes Peak, rising to 14,115 feet above the sea – he accomplished it once a long time before – but I talked him out of it. One night the phone rang. Ron talked with Paul, as he has done every night, and then called me to the phone. "You know, Sasha, this book will never happen unless you join Paul in writing it." "I would be very pleased if you and Ron were to join me in writing the book," Paul told me. Ron chose not to join the writing but promised to read the manuscript and give his feedback. And so, in addition to joint papers, I became Paul's coauthor on the book *Problems of pgom Erdős*. The book is not yet finished but is in the works. (Ron and Fan Chung did publish a book of Paul's problems in graph theory.)

A list of mathematicians inspired by Paul Erdős may go on for longer than the list of his ca. 1600 publications. Trajectories of his travels probably added up to a set dense at every point on the globe. Paul inscribed his reprints for me with mysterious sequence of letters after his name:

Erdős Pal, pgom, ld, ad, ld, cd.

To my inquiry, Paul explained:

pgom = poor great old man
ld = living dead (i.e., over 60 years old)
ad = archaeological discovery (> 65)
ld = legally dead (> 70)
cd = counts dead (> 75)

– “Great” I agree, “old” alas, but why “poor”? asked I.
– “All old men are poor,” replied Paul.

In July 1993 in Keszthely, Hungary, I reminded Paul that “the emergency” of adding another pair of initials had arrived. Paul thought for a moment and then declared: “nd, nearly dead.”

Paul Erdős and Alexander Soifer, the last time together, Baton Rouge, February 22, 1996

It is impossible to overestimate how much the field of this book owes to Paul Erdős. With the intuition of a genius, Paul saw the beauties to be had in what became known as *Ramsey Theory* and led our way to this Garden of Eden.

“Two Thinkers”, Paul Erdős in Colorado Springs, December 28, 1991. (Photograph by Alexander Soifer)

26.2 Old Snapshots of the Young[2]

Martha (Waksberg) Svéd (1909, Budapest – September 30, 2005, Adelaide), a member of the legendary Budapest circle of young Jewish students-mathematicians that included Paul Erdős, Paul Turán, Tibor Gallai, Esther Klein, George Szekeres, and many others, had known Erdős as few did. For Paul’s 80th birthday, she wrote on my request her warm, lyrical reminiscences especially for *Geombinatorics’* “Erdős is Eighty” special issue [Sve2]. This subsection is all hers. Martha Svéd recollects:

> Yes, E. P., this is the name: initials for Erdős Pál, Hungarian form of the name Paul Erdős, name by which we, old Hungarian friends called him and still refer to him. This is the year when he is eighty years YOUNG. At this point I recall his Cambridge lecture

[2]First published in *Geombinatorics* [Sve2].

attended by the two of us, G. [George] and M. Sved in 1959. Ahead of those formulae about the secrets of primes, Hebrew words appeared on the blackboard:

זקנה היא לא נעימה

(This lecture was held just after his extended stay in Israel.) The translation is: “Old age is not pleasant,” referring to himself. His greeting words to us (having last seen each other in Budapest in 1938) were: “SAM SURRENDERED.”

It is now necessary to give an Erdős dictionary for those who are not initiated. Since the set of those whose Erdős number is 1, has measure, there will be a large subset not needing such a dictionary, hence they may skip it. I add here for the uninitiated the definition of the “Erdős number.”[3] It is the number of links in the chain leading to the origin, E. P. himself. The aristocrats are those whose name has appeared together with E. P. in at least one publication, hence can boast of number 1. My own number is 2, but without great claims of merit. The thanks for it must go to George Szekeres for my single joint publication with him. G. Szekeres holds number 1 with high multiplicity. Since G. Sved and I have lived in Australia for more than fifty years and only for a few years in the same city as G. & E. Szekeres, my mathematical contacts have been locals, so my Erdős number would have been hard to trace. This is why I am proud and happy for being asked to add my lines to this celebratory volume. I should add here that some conjecture is floating around: if you have joint publications with at least three coauthors then your Erdős number is finite, (though it could be distressingly large!). This is my reason to leave the mathematical bits to others, restricting myself to reminiscences about our great and faithful friend whose letters still begin: “G. and M. Sved”, followed by a paragraph about his own personal jaunts across four continents, with the last paragraph beginning “Let n points...” I try to translate (not adequately) his Hungarian self-description of “not being a university professor but a world professor of mathematics,” the traveling missionary, to whom “Sam surrendered” in 1959.

Now to the Dictionary, apologizing for its incompleteness and haphazard ordering:

Epsilon: small, negligible in some respects, but the word has also another meaning, an endearing one: child generally. When talking about the offsprings of his friends, the Epsilon could be quite grown up, perhaps having Epsilons of his own, i.e., epsilon squares.

[3] Casper Goffman must have been first to define Erdős Numbers in print in 1969 [Gof]. Apparently, the concept was born in the 1960s and Paul Erdős himself did not know about it before 1968.

Paul Erdős with the Epsilon, Isabelle Soulay Soifer, Kalamazoo, Michigan, June 4, 1992. (Photograph by Alexander Soifer)

Omega: large, many.
Trivial: of a person: mean, uncaring, unjust etc., hence Triviality.
Victory: solution of a problem found. However, once during a hike he sang out: "Victory! I lost my wallet." ???
Fascism: the nastiest swearword Paul can think of, when he is clumsy and drops or mislays something.
Boss: wife or girlfriend.
Slave: husband or boyfriend.
Captured: snared into marriage or long term relationship.
Liberated slave: divorced man.
Sam: United States of America.
Joe: the late Soviet Union, abbreviated name of Joseph Stalin.
Cured: passed away, "cured" of the illness of life.[4]

I am now in the position of being able to explain that greeting in Cambridge in 1959. While E. P. is not "patriotic" in the nationalistic sense, he could never deny his

[4] This reminds me the Polish film director Krzysztof Zanussi, who entitled his 2000 film "Life as a Fatal Sexually Transmitted Disease."

Hungarian identity. During the oppressive communist regime of Hungary in the 1950s he visited Budapest and was probably the only person who was allowed to leave freely. While having lived in the States permanently after World War II, he did not acquire U.S. citizenship, moreover, refused to sign the "loyalty oath" of the McCarthy era during the "cold war." He was not expelled for this but was warned that he would not get a return visa when leaving. Nevertheless, when invited to an international mathematical conference in Holland in 1954, he took it up. His return visa was refused for years. A letter went then to President Eisenhower, signed by the greatest names of American Mathematics. This letter pointed out what the loss of Paul Erdős meant to their country. There was now a whole generation of youngsters growing up, entirely missing his inspiring influence. This was a loss which U.S. mathematics could not afford. This worked. E. P. was given the visa, for a limited period at that time, but by now he is welcome with open arms at any time when he wishes to enter and spend a very short or a long period there. In fact, he has now two main "bases." One is in Budapest at the Mathematical Research Institute of the Hungarian Academy, where all his publications are kept and which serves him as a home when in Hungary, since he has not entered his own flat since the death of his mother at an advanced age. She had been his faithful companion and secretary through all his travels. His flat is now used by visiting mathematician friends. The other base is at Bell Laboratories, where his friend Ronald Graham looks after the business matters of his life.

I go back now to early days to write reminiscences. The name Erdős Pál was well known by us, Esther Szekeres (then Klein) and myself before we met him at the university. All of us were frequent problem solvers of our beloved Hungarian magazine, the School Journal for Mathematics and Physics.

The beginnings of this journal date at the turn of the century, in a great period of prosperity, liberalism and culture. World War I and years in the era to follow washed it all away. The mathematical school journal was revived in 1925, to go into oblivion again late during World War II. Nazism, taking hold of Hungary killed not only the journal, but also its editor, Andrew Farago together with his family. There was also a number of mathematicians, most of them young and of great promise who became victims of fascism. A plaque at the Research Institute of Mathematics commemorates their names. The school journal came to vigorous life again after this second war, also new journals for secondary students were born around the world to inspire the young, but the human victims could not be brought back to life.

Esther Klein (Szekeres) was my classmate and best friend in the four final years of our secondary schooling. We had an exceptional teacher of mathematics, R. Rieger. That day in the beginning of year 1925 pictures still clearly in my mind, when our teacher appeared in class with the first issue of the revived school journal in his hand. He pointed out the two sets of problems, aimed at two levels, at the lower and higher grades of schooling. Both Esther and I became problem solvers working completely independently of each other, continuing until the time when we left school and became university students, a privilege meted to a very restricted number of Jewish youngsters.

Esther, to whom I shall refer from now on as Eps, (for, as coined by E. P. twisting the petting name used by her mother) remained really faithful to mathematics during our university years but I became somewhat wavering. My loss of dedication was only partly due to the early "capture" of each other with G. Sved whose mathematics was rich but not "pure," being an engineer. The case with G. Szekeres was different. He studied chemical engineering, to satisfy his father, in 1928 still a leather manufacturer. Yet the real love of this other George was always mathematics, with some theoretical physics thrown in, as his later contributions to relativity theory show. Nevertheless, he completed his course successfully, and worked for some years in the leather industry, in Hungary first, then during the early war years in Shanghai. He found time during engineering studies to join our little circle. His name together with those of T. P. (Paul Turán) and G. T. (Tibor Gallai) and of course, E. P. known to us through that journal which published not only the names of successful problem solvers, specially printing (with some editing) the best solutions, but supplied also at the end of each school year the photos of the most frequent contributors of solutions.

T. P. (Paul Turán) was in the same year as ourselves, but E. P. and G. T., younger than the three of us, appeared on the scene two years later, E. P. as mover and shaker, G. T. as sharp critic. E. P. seldom graced the lecture room with his presence. I am not sure now whether he was even enrolled, like the rest of us for secondary teachers training, running in conjunction with our courses in mathematics and physics. He certainly missed the fifth and final year required, consisting of teaching practice in one of the schools officially prescribed and ending with the examination in philosophy and theory of education. In that fifth year of academic education he was already in Manchester with a scholarship, working as a post-doctoral fellow with Mordell, having gained his Ph. D. with L. Fejer being his supervisor. His doctoral thesis was based on results obtained on the distribution of prime numbers. Still as an undergraduate he obtained new results, (elementary proof of Csebisev's theorem). To this day primes are one of his prime concerns.

However, I must go back now to those earlier years, when E. P. was holding court, in the students' common room, or being one of our crowd in our City Park, where we were tackling problems set in the then new and by now classic collection of problems in analysis and number theory by G. Polya and G. Szego. In those years with E. P. at the university, Esther and I had to take charge of him to ensure that required enrolment and semester end formalities were satisfied. Our rewards were rich. Our group held together strongly then and in some later years, with some of us already graduates (though not holding teachers' appointments, with our teaching work being confined to private tutoring). We shared hikes at our charming hills near our city and continued our mathematical meetings at a site around the Statue of Anonymous in our city park.

Paul's parents adopted all his young friends. The Erdős home became our second home. The parents, Louis and Anne were both mathematics teachers, but in our time only Louis was active as a schoolteacher. His mother was sacked in 1919, in the days following the upheavals after the war. Louis, who had been on active war service and had returned after long years in Russian camps for war prisoners, could not be dismissed. Fascism (the word had not yet come into existence) in those days was "mild" in comparison with what came at the end of World War II. Louis, outstanding as a teacher, a man of wisdom, vision combined with a sense of humour, was a delightful company for us. Taking breath in that warm and stimulating atmosphere created by Anne and Louis nurturing Paul was a gift. During the years after 1939 E. P. was in the U.S. and ourselves in Australia, being able to keep some contact. I was shaken when I read the news in one of Paul's letters that his father died a natural death during those war years. Then I was comforted by the thought that he was spared the dangers, degradations and humiliations to be meted out at the end of that war.

Since E. P. and his mathematics, (the two are being inseparable), form the pivot of such large collection of mathematicians (with their bosses or slaves), the number of stories surrounding him together with those histories of mathematical problems solved or still in states of conjecture is also of an impressive multitude, I want to add here my own story about a problem I witnessed at birth.

Paul calls it "the happy [end] problem." My friend Eps, not long after her return from Göttingen, in those days the world centre of mathematics, posed the following question: given 5 points in the plane, no 3 collinear, conjecture: it is always possible to select 4 to form the vertices of a convex quadrilateral. It was a problem of unusual flavour, but my own waverings did not point in that direction. All the more were E. P. and Gy. Szekeres aroused. As Gy. S. confessed later, his attraction to the problem was sparked by the person proposing it. Actually, Eps found the proof, and efforts to generalize began. They resulted in the first Erdős–Szekeres joint publication to appear about two years later. The authors were not aware at the time that they solved and extended an old theorem by Ramsey. The significance of this publication was that it yielded life time results for the Trio involved: Eps and Gyu (pronounce Dew, and I am not giving here a linguistic lecture to explain this) became the couple of mathematicians bearing the name G. and E. Szekeres; P. Erdős and G. Szekeres became life-long coworkers, though in fields different from that first joint paper; E. P. became the originator of a new field in mathematics: combinatorial geometry, one of the new chapters created by him.

The youth of E. P. is of a lifetime duration. His approach to problems is "elementary," his best working pals are the young, his games, hobbies and relaxations do not belong to the world of the old, and ignoring social conventions are those of a child. He has remained the Peter Pan of mathematics.

&&&

Márta Svéd passed away on September 30, 2005. She outlived her dear friends Esther Klein and George Szekeres by just two days.

Chapter 27
De Bruijn–Erdős' Theorem and Its History

27.1 The De Bruijn–Erdős Compactness Theorem

They were both young. On August 4, 1947, the 34-year-old Paul Erdős, in a letter to the 29-year-old Nicolaas Govert de Bruijn of Delft, The Netherlands, offers the following conjecture [E47/8/4ltr]:

> Let G be an infinite graph. Any finite subset of it is the sum of k independent sets (two vertices are independent if they are not connected). Then G is the sum of k independent sets.

Paul adds in parentheses "I can only prove it if $k = 2$." In his 5-page August 18, 1947, reply [Bru1], de Bruijn reformulates the Erdős conjecture in a way that is very familiar to us today:

> Theorem. Let G be an infinite graph, any finite subgraph of which can be k-colored (that means that the nodes are coloured with k different colours, such that the two connected nodes have different colours). Then G can be k-coloured.

Following a nearly 3-page long transfinite induction proof of the "Theorem," de Bruijn observes [Bru1]:

> I am sorry that this proof takes so much paper; its idea, however, is simple. Perhaps, you do not call it a proof at all, because it contains "Wellordering," but we can hardly expect to get along without that.

This is a very insightful observation, as de Bruijn and Erdős rely on the Axiom of Choice or equivalent (like Well-Ordering Principle or Zorn's Lemma) very heavily. When in early 2004 Professor de Bruijn received from me a reprint of Shelah–Soifer 2003 paper (to be discussed closer to the end of this book) which analyzed what happens with the de Bruijn–Erdős Theorem in the absence of the Axiom of Choice, de Bruijn replied to me on January 27, 2004, as follows [Bru6]:

> About the axiom of choice, I remember a conversation with Erdős, during a walk around 1954. I told him that I hated the axiom of choice, and that I wanted to do analysis without it, maybe except for the countable case. He was surprised and said: but you were always so good at it. Indeed, I had loved transfinite induction, just because it worked exactly the same way as ordinary induction.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_27

This invaluable de Bruijn's e-mail also contained the conclusion of the story of the de Bruijn–Erdős Theorem [Bru5]:

> Erdős and I did not take any steps to publish the k-coloring theorem. In 1951 I met Erdős in London, and from there we went together by train to Aberdeen, which took a full day. It was during that train ride that he told me about the topological proof of the k-coloring theorem. Not long after that, he wrote it up and submitted it for publication. I do not think I had substantial influence on that version.

Let us look at a proof of this celebrated theorem, which we have formulated without proof and used in Chapter 5.

The De Bruijn–Erdős Compactness Theorem 27.1 (August 18, 1947; pub. [BE2], 1951). An infinite graph G is k-colorable if and only if every finite subgraph of G is k-colorable.[1]

In what follows, we will need a few definitions from set theory.

Given a set A, any subset R of the so-called Cartesian product $A \times A = \{(a_1, a_2) : a_1, a_2 \in A\}$ is called a *binary relation* on A. We write $a_1 R a_2$ to indicate that the *ordered pair* (a_1, a_2) is an element of R.

Poset, or *partially ordered set*, is a set A together with a particularly "nice" binary relation on it, i.e., a *relation* that *satisfies* the following three *properties*:

1. *Reflexivity*: $a \le a$ for all $a \in A$;
2. *Antisymmetry*: If $a \le b$ and $b \le a$ for $a, b \in A$, then $a = b$;
3. *Transitivity*: If $a \le b$ and $b \le c$ for $a, b, c \in A$, then $a \le c$.

A *chain*, or *totally ordered set*, is a poset that satisfies the fourth property:

4. *Comparability*: For any $a, b \in A$, either $a \le b$ or $b \le a$.

Let A be a set with a *partial ordering* $\le$ defined on it, and B a *subset* of A. An *upper bound* of B is an element $a \in A$ such that $b \le a$ for every $b \in B$.

Let $\le$ be a partial *ordering* on a set A, and $B \subseteq A$. Then, we say that $b \in B$ is a *maximal element* of B if there exists no $x \in B$ such that $b \le x$ and $x \ne b$.

In 1935, Max Zorn (1906, Germany–1993, USA) introduced the following important tool, which he called *maximum principle*. It was shown by Paul J. Campbell that, in fact, a number of famous mathematicians – Hausdorff, Kuratowski, and Brouwer – preceded Zorn, but Zorn's name got as attached to this tool as, say, Amerigo Vespucci's name to America.

Zorn's Lemma 27.2 If S is a nonempty partially ordered set in which every chain has an upper bound, then S has a maximal element.

During the summer of 2005, I supervised at my University of Colorado a research month of Dmytro "Mitya" Karabash, who had just completed his freshman year at Columbia University. One of my assignments for him was to prove the de Bruijn–Erdős theorem 27.1 and then to write a proof as well. After going through several revisions, Mitya produced a fine proof, which follows here, slightly edited by me.[2]

---

[1] This theorem requires the Axiom of Choice or equivalent.

[2] You can also read the original proof in [BE2]; a nice proof by L. Pósa in the fine book [Lov2] by László Lovász; and a clear insightful proof of the countable case in the best introductory book to Ramsey Theory [Gra2] by Ronald L. Graham.

Proof of theorem 27.1 by D. Karabash We say that a graph G has the property P and write $P(G)$ if every finite subgraph of G is *k-colorable*. For a graph G, we write $G = (V, E)$, where V is the vertex set and E is the edge set of G. Now let S be the set of all graphs with the property P which are obtained from G by the addition of edges, i.e.,

$$S = \{(V, F) | E \subseteq F \text{ and } P(V, F)\}.$$

Let S be partially ordered by the inclusion of edge sets. Observe that for every chain A_i in S, its union $A = (V, \bigcup_i E(A_i))$ is also in S [here $E(A_i)$ stands for the edge set of the graph A_i]. Indeed, every finite subgraph F of A must be contained in some A_i (because F is finite), and therefore, F is k-colorable. Since A has property P, A is in S, as desired.

We have proved that in S, every chain has an upper bound. Therefore, by Zorn's Lemma, S contains a maximal element, call it M. Since M is in S, M has property P; since M is maximal, no edges can be added to M without violating property P.

We will now prove that *non-adjacency* (here to be denoted by the symbol $\neg adj$) is an equivalence relation on M, i.e., for every $a, b, c \in V(M)$, if $a \neg adj\ b$ and $b \neg adj\ c$, then $a \neg adj\ c$. Let us consider all finite subgraphs of M that contain a and b, and all k-colorings on them. Since $a \neg adj\ b$, there must be a subgraph M_{ab} for which the colors of a and b are the same for all k-colorings of this subgraph, for otherwise we could add the edge ab to M while preserving property P and attain a contradiction to M being a maximal element of S. Construct a subgraph M_{bc} similarly. The subgraph $M_{ab} \cup M_{bc}$ is finite and, thus, k-colorable; it contains subgraphs M_{ab} and M_{bc}; therefore, by construction of M_{ab} and M_{bc}, any coloring of $M_{ab} \cup M_{bc}$ must have pairs (a,b) and (b,c) colored in the same color. Thus, a and c have the same color for all k-colorings of the subgraph $M_{ab} \cup M_{bc}$, and therefore, a is not adjacent to c.

From the fact that the non-adjacency is an equivalence relation on M, we conclude that the edge-complement M' of M is made of some number of disjoin complete graphs K_i because in M' adjacency is an equivalence relation. Therefore, $a \in K_i$, $b \in K_j$, $i \neq j$ implies $a \neg adj\ b$ in M' or equivalently $a\ adj\ b$ in M.

Suppose there is more than k disjoint complete subgraphs K_i in M'. Then pick $k + 1$ vertices, all from distinct $V(K_i)$. Since all of the vertices are located in distinct $V(K_i)$, they must be all pairwise non-adjacent in M' and thus form a complete graph M_{k+1} on $k + 1$ vertices in M. We obtained a finite subgraph M_{k+1} of M which is not k-colorable, in contradiction to M having property P. Therefore, M' consists of at most k complete subgraphs $V(K_i)$, $i = 1, \ldots, k$. Now we can color each subgraph $V(K_i)$ in a different color. Since no two vertices of $V(K_i)$ are adjacent in M, this is a proper k-coloring. Since G is a subgraph of M, G is k-colorable, as desired. ■

Corollary 27.3 Compactness Theorem 5.1 is true.

The proof of theorem 27.1 is much more powerful than you may think. It works not only for graphs, but even for their important generalization – hypergraphs. Permit me to burden you with a few definitions.

As you recall from Chapter 12, a *graph* $G = G(V, E)$ is a non-empty set V (of vertices) together with a family E of 2-element subsets (edges) of V. If we relax the latter condition, we will end up with a hypergraph.

A *hypergraph* $H = H(V, E)$ is a non-empty set V (of *vertices*) together with a family E of subsets (*edges*) of V, each containing *at least two* elements. Thus, an edge e of H is a subset of V; its elements are naturally called *vertices of the edge e* (or *vertices incident with e*).

Let n be a positive integer. We would say that a hypergraph H is *n-colored* if each vertex of H is assigned one of the given n colors. If *all* vertices of an edge e are assigned the same color, we call e a *monochromatic edge*.

The *chromatic number* $\chi(H)$ of a hypergraph H is the smallest number of colors n for which there is an n-coloring of H without monochromatic edges.

A hypergraph $H_1 = H_1(V_1, E_1)$ is called a *subhypergraph* of a hypergraph $H = H(V, E)$, if $V_1 \subseteq V$ and $E_1 \subseteq E$.

Compactness Theorem for Hypergraphs 27.4 The chromatic number $\chi(H)$ of a hypergraph H is equal to the maximum chromatic number of its finite subhypergraphs.

Proof Repeat word-by-word the proof of Theorem 27.2, just replace "graph" by "hypergraph"). ■

Chapter 28
Nicolaas Govert de Bruijn

Ever since 1995, I have exchanged numerous e-mail messages – and sometimes letters – with the Dutch mathematician N. G. de Bruijn. His elegant humor, openness in expressing views even on controversial issues, his eyewitness accounts of post-World War II events in Holland made this correspondence fascinating and enjoyable for me. We also shared interest in finding out who created the conjecture on monochromatic arithmetic progressions, which was proved by B. L. van der Waerden (see the rigorously argued answer later in this book). Yet, for years, I have been asking Professor de Bruijn to share with me his autobiography to no avail. For a long while, I did not even know what "N. G." stood for. On October 29, 2005, I tried to be a bit more specific in my e-mail. I wrote:

> May I ask you to describe your life – and any participation in political affairs – during the occupation, May 1940–1945, and during the first post-war years, 1945 up to your Sep–1952 appointment to replace Van der Waerden at Amsterdam?

De Bruijn understood my maneuver but provided the desired reply on November 1, 2005 [Bru11]:

> You are asking for an autobiography in a nutshell.
>
> I was born in 1918 [on July 9th, in Den Haag], so I just left elementary school in 1930 when the great depression broke out. I managed to finish secondary school education in 4 years (the standard was 5 or 6). After that, I could not get any job, and could not get any financial support for university education. I used my next two years (1934–1936) to study mathematics from books, without any teacher. I passed the examinations that qualified me as a mathematics teacher in all secondary schools in the Netherlands. But there weren't any jobs. Yet I had some success: I could get a small loan that enabled me to study mathematics and physics at Leiden University. In the academic year 1936–1937 I attended courses in physics and astronomy, and in 1937–1938 courses in mathematics on the master's degree level. That was all the university education I had. The most inspiring mathematician in those days at Leiden was H. D. Kloosterman.
>
> In 1939 I was so lucky to get an assistantship at Delft Technical University. It didn't pay very much, but it left me plenty of time to get involved in various kinds of mathematical research. It was quite an inspiring environment, and actually it was the

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_28

only place in the Netherlands that employed mathematical assistants (Delft had about 8 or 9 of them). In 1940 the country was occupied, and from then on the main problem was to avoid being drawn into forced labour in Germany. In that respect my assistantship was a good shelter for quite some time.

All the time I lived with my parents in The Hague, not so safe as it seemed. We were hiding a Jewish refugee (a German boy, a few years younger than me), who assisted my brother in producing and distributing forbidden radio material, like antennas that made it possible to eliminate the heavy bleep-bleep-bleep that the Germans used in the wavelengths of the British Radio. And later, when radios were forbidden altogether, my brother built miniature radios, hidden in old encyclopedia volumes. All this activity ended somewhere in the beginning of 1944 when our house was raided by the *Sicherheitspolizei*. My brother and his Jewish assistant were taken into custody, but by some strange coincidence they came back the next day. Nevertheless, they had to leave to a safer place, where both of them survived the war. A few months later, I got my first real job. It was at the famous Philips Physical Labs at Eindhoven. The factory worked more or less for German war production, just like most factories in the country, but the laboratories could just do what they always did.

Four months later, Eindhoven was occupied by the allied armies, in their move towards the battle of Arnhem. From then on we were cut off from the rest of the country, where people had a very bad time.

So this was about my life during the war. Compared to others, I had been quite lucky. I had even managed to get my doctorate at the [Calvinist] Free University, Amsterdam [March 1943], just a few weeks before all universities in the country were definitely closed (Leiden University had already been closed in 1940, because of demonstrations against the dismissal of Jewish professors).

In 1946 I got a professorship at Delft Technical University. I had to do quite elementary teaching, leaving me free to do quite some research, mainly in analytical number theory. It got me into correspondence with Erdős, and around 1948 he visited us at Delft.

In 1951 I made a mathematical trip abroad for the first time in my life. There I had contact with Erdős too. We had a long train ride together from London to Edinburgh.

In 1952 I got that [Van der Waerden's] professorship at Amsterdam, at that time the mathematical Mecca of the Netherlands. I stayed there until 1960, when I got my professorship at Eindhoven Technological University, where I retired in 1984. After that, I always kept a place to work there.

I think this is all you wanted to know.

In fact, on November 1, 2005, I asked for a few additional details:

I know you are one of the most modest men. Yet, I would think you were not just an observer when your family hid a Jewish boy and your brother did activities not appreciated by the occupiers. Would you be so kind to share with me your role is these activities during 1940–1945?

What were the names of your brother and his Jewish-German assistant? What was the difference in age between you and your brother?

Two days later, my questions were answered [Bru12]:

I hardly ever participated in my brother's activities. At most three times I delivered an antenna or a radio to some stranger. My brother was a year and a half older than I. His name was Johan.

The Jewish boy's name was Ernest (Ernst) Goldstern. He was born 24 December 1923 (in Muenchen, I believe). His family came to Holland in the late 1930s, where Ernst just completed his secondary school education in Amsterdam. He lived with us in The Hague from 1940 to 1944. I helped him to study advanced mathematics, which he could use after the war. He went into Electrical Engineering and got his degree in Delft. He died 19 January 1993. Johan died in 1996.

My treasured and admired correspondent Nicolaas Govert de Bruijn passed away on February 17, 2012. He was 93.

Chapter 29
Edge-Colored Graphs: Ramsey and Folkman Numbers

29.1 Ramsey Numbers

In this chapter, we will see that no matter how edges of a complete graph K_n are colored in two or, more generally, finitely many colors (each edge in one color), we can guarantee the existence of the desired monochromatic subgraph as long as we choose n to be large enough.

Naturally, when we talk about edge-colored graphs, we call a subgraph *monochromatic* if all its edges are assigned the same color.

Frank Harary told me that he was once asked to suggest problems for the W. L. Putman Mathematical Competition, and he suggested to use a problem that had already existed in the mathematical folklore:[1]

Problem 29.1 (W. L. Putnam Mathematical Competition, March 1953). Prove that no matter how the edges of the complete graph K_6 are colored in two colors, there is always a monochromatic triangle K_3.

Proof Let v_0 be a vertex of K_6, whose five incident edges are colored red and blue. Then, v_0 is incident with at least three edges of the same color, say, red (Fig. 29.1).

[1] Stanisław Radziszowski advises me that already in 1947 this problem was offered in a competition in Hungary (e-mail from 07/04/2020).

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_29

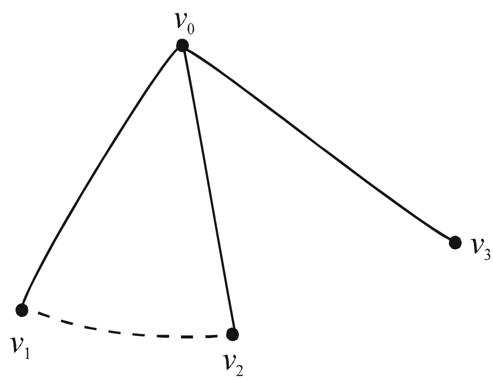


Fig. 29.1 Vertex v_0 is incident with at least 3 red edges

If any two of the vertices v_1, v_2, v_3, say, v_1 and v_2 are connected by a red edge, then we are done: v_0, v_1, and v_2 is a red monochromatic triangle. Otherwise, all three edges v_1,v_2, v_2, v_3, and v_3, v_1 are blue, and we are done as well. ■

Problem 29.2 Show that in the statement of Problem 29.1, 6 is the best possible number, i.e., there is a way to color the edges of K_5 in two colors without creating any monochromatic triangles.

Solution Behold (Fig. 29.2):

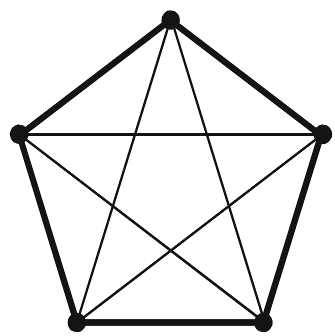


Fig. 29.2 2-colored K_5 without monochromatic triangles

For positive integers m and n, the *Ramsey number* $R = R(m, n)$ is the smallest positive integer such that *any* red and blue edge coloring of the complete graph K_R contains a red monochromatic K_m or a blue monochromatic K_n.

Problems 29.1 and 29.2 together prove, for example, that

$$R(3,3) = 6.$$

You do not need more than definitions to prove the following two equalities.

Problem 29.3 For any two positive integers m and n

$$R(m,n) = R(n,m).$$

Problem 29.4 For any positive integer n

$$R(2,n) = n.$$

When and who coined the term “Ramsey number”? The publication search readily proves that it did not exist in print before 1966. “Ramsey number” makes its first appearance in January 1966 in the remarkable Ph.D. thesis *Chromatic Graphs and Ramsey’s Theorem* by James (Jim) G. Kalbfleisch [Ka2] at the University of Waterloo, Ontario, Canada. He proves a good number of new upper and lower bounds, uniqueness of certain colorings, and the exact value $R(3, 6) = 18$ (which was also proved independently by G. Kéry). Kalbfleisch may have been the first to use computer programs in aid of his Ramsey numbers research. And, Kalbfleisch was the first to use “Ramsey number” term in print (his thesis, as was typical in mathematics in North America, was not published), in his 1966 paper [Ka3], submitted for publication on February 26, 1966. Nearly half a year later, on July 13, 1966, Jack E. Graver and James Yackel’s paper [GY] was communicated by Victor Klee. Both papers, [Ka3] and [GY], proudly displayed the new term “Ramsey number” in their titles. The term took hold and was used in an enormous number of publications since.

As to Jim Kalbfleisch, following a number of fine Ramsey number-related publications, he “defected” to statistics. Kalbfleisch served the University of Waterloo for 37 years, 1963–2000, as a student, professor, dean of mathematics, academic vice-president, and provost. On 31 December 2000, at 60, he retired to enjoy his artistic hobbies, *Daily Bulletin* reported on Thursday, 5 October 2000:

> “There’s never a good time to go,” he [Kalbfleisch] said, but after 14 years, “I feel the need for a break.” He said he is looking forward to a chance to get back to stained glass work (his long-time hobby) and enjoy music, bridge, and some travel that isn’t just for business.

The following year Kalbfleisch was awarded the title “Provost Emeritus,” a rare distinction indeed.

I would like to compute a few Ramsey numbers with you. For this, we will need the following simple but useful tool.

Basic Tool 29.5 For any graph G with p vertices $v_1, v_2, \ldots, v_p$ and q edges, $deg\ v_1 + deg\ v_2 + \cdots + deg\ v_p = 2q$.

The following Ramsey numbers were first found in 1955 by Robert E. Greenwood of the University of Texas and Andrew M. Gleason of Harvard.

Problem 29.6 (R.E. Greenwood and A. M. Gleason, [GG]). Prove that $R(3, 4) = 9$.

Proof Let the edges of a complete graph K_9 be colored red and blue. We will consider two cases.

Case 1 Assume there is a vertex, say v_0, of K_9 that is incident with at least four red edges (Fig. 29.3). Then should any two of the vertices v_1, v_2, v_3, and v_4 be connected by a red edge, we get a red triangle. Otherwise, we get a blue monochromatic K_4 on the vertices v_1, v_2, v_3, and v_4.

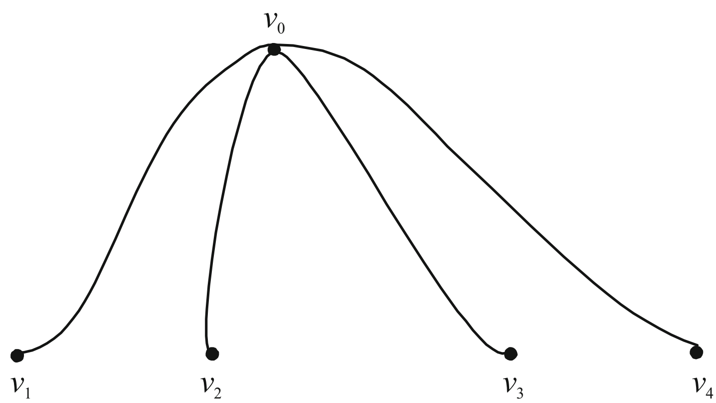


Fig. 29.3 Vertex v_0 is incident with at least 4 red edges

Case 2 Every vertex of K_9 is incident with at least five blue edges. The nine vertices of K_9 with all blue edges form a graph G. The degree of each vertex of G may not be equal to five because we would get an odd $5.9 = 45$ in the left side of the equality of tool 29.5 with an even $2q$ in the right side. Therefore, at least one vertex, say v_0, of K_9 is incident with at least six blue edges (Figure 29.4).

Due to Problem 29.1 applied to the complete graph K_6 on the vertices $v_1, v_2, \ldots, v_6$, K_6 contains a monochromatic triangle K_3. If K_3 is red, we are done. If K_3 is blue, then the three vertices of K_3 plus v_0 form a blue monochromatic graph K_4, and we are done again.

Thus, we proved the inequality $R(3, 4) \leq 9$.

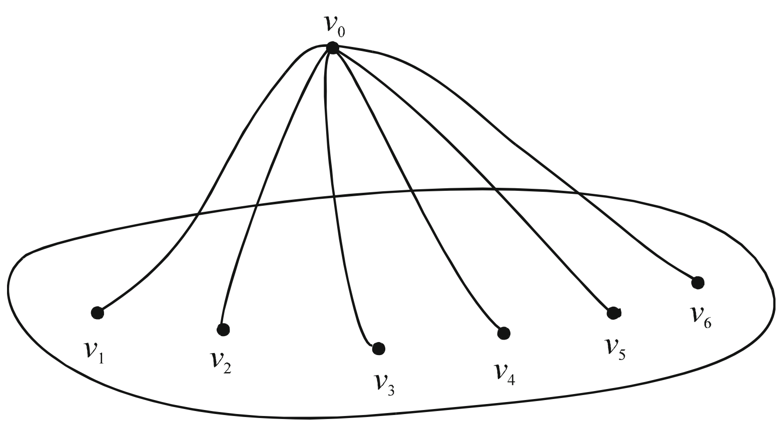


Fig. 29.4 Vertex v_0 is incident with at least 6 blue edges

Figure 29.5 shows all red edges of K_8. The edges that are not drawn, we color blue. It is easy to verify that this 2-coloring of the edges of K_8 creates neither a red monochromatic K_3 nor a blue monochromatic K_4. ■

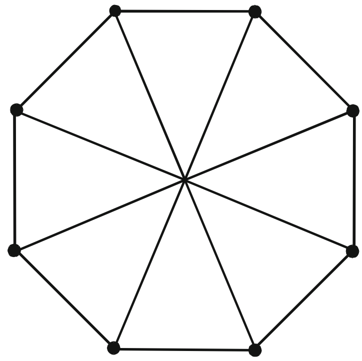


Fig. 29.5 All red edges of K_8

Problem 29.7 (R.E. Greenwood and A. M. Gleason, [GG]). Prove that $R(4, 4) = 18$.

Proof First, we will prove the inequality $R(4, 4) \leq 18$.

Let the edges of a complete graph K_{18} be colored red and blue, and v_0 be a vertex of K_{18}. Since v_0 is incident with 17 edges, by the Pigeonhole Principle v_0 must be incident with at least 9 edges of the same color.

If these 9 edges are red, we apply the equality $R(3, 4) = 9$ of Problem 29.6 to the 9-element set $S = \{v_1, v_2, \ldots, v_9\}$. If S contains a blue monochromatic K_4, we are done. If S contains a red monochromatic triangle T, then T together with v_0 and three red edges between them composes a red monochromatic K_4.

If the 9 edges are blue, we apply the equality $R(4, 3) = 9$ to the 9-set $S = \{v_1, v_2, \cdots, v_9\}$ and reason similarly to the above "red" case. Thus, the inequality $R(4, 4) \leq 18$ is proven.

Now, we have to prove that $R(4, 4) > 17$. Figure 29.6 shows all red edges of the red–blue edge coloring of K_{17} (all missing edges are blue). It is easy to verify that our K_{17} contains no monochromatic K_4.■

You can now solve on your own the following couple of problems, of which the first one gives a rare exact value of a Ramsey number, while the second problem is just an exercise.

Problem 29.8 (R.E. Greenwood and A.W. Gleason, [GG]). Prove that $R(3, 5) = 14$.

Problem 29.9 Prove that $R(4, 5) \leq 32$ and $R(5, 5) \leq 64$.

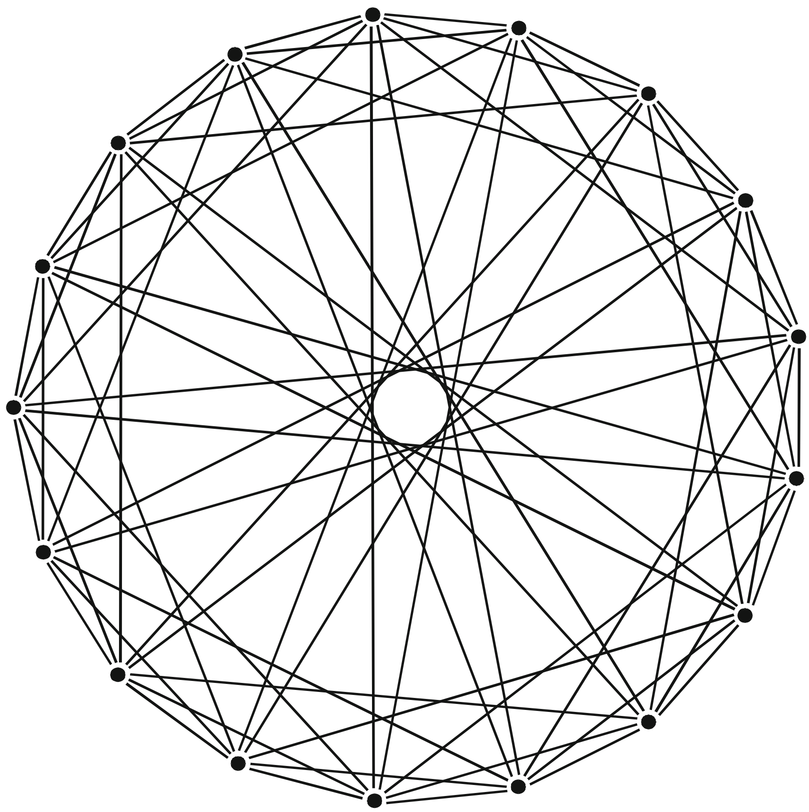


Fig. 29.6 All red edges of the red-blue edge coloring of K_{17}

In fact, the problem of calculating $R(4, 5)$ was settled completely by Brendan D. McKay of the Australian National University and Stanisław P. Radziszowski of the Rochester Institute of Technology, originally of Poland.

Result 29.10 (B. D. McKay and S. P. Radziszowski, [MR4]). $R(4, 5) = 25$.

This remarkable result took years of computing to achieve, with the happy end taking place right in front of my eyes. I attended Stanisław Radziszowski's talk in early March of 1993 at the Florida Atlantic University conference. During the talk, he mysteriously remarked that the value of $R(4, 5)$ may be established very soon. Imagine, in a matter of days, I received his e-mail, announcing the birth of the result up to a hundredth of a second (this is what computer-aided communication delivers):

From: MX%"spr@cs.rit.edu"
To: ASOIFER
Subj: R(4, 5) = 25
From: spr@cs.rit.edu (Stanisław P Radziszowski)
Message–ID: <9303191824.AA22893@rit.cs.rit.edu>

Subject: R(4, 5) = 25
To: jackkasz@utxvm.cc.utexas.edu, asoifer@happy.uccs.edu, goldberg@turing.cs.rpi.edu
Date: Fri, 19 Mar 1993 19–MAR–1993 11:24:29.37 (EST)

R(4, 5) = 25

Brendan D. McKay, Australian National University
Stanisław P. Radziszowski, Rochester Institute of Technology

The Ramsey number $R(4, 5)$ is defined to be the smallest n such that every graph on n vertices has either a clique of order 4 or an independent set of order 5. We have proved that $R(4, 5) = 25$. Previously it was only known that $R(4, 5)$ is one of the four numbers 25–28. Our proof is computational.

For integers s,t define an (s, t, n)-graph to be an n-vertex graph with no clique of order s or independent set of order t. Suppose that G is a (4, 5, 25)-graph with 25 vertices. If a vertex is removed from G, a (4, 5, 24)-graph H results; moreover, the structure of H can be somewhat restricted by choosing which vertex of G to remove. Our proof consists of constructing all such structure–restricted (4, 5, 24)-graphs and showing that none of them extends to a (4, 5, 25)-graph. In order to reduce the chance of computational error, the entire computation was done in duplicate using independent programs written by each author. The fastest of the two computations required about 3.2 years of cpu time on Sun workstations.

A side result of this computation is a catalogue of 350866 (4, 5, 24)-graphs, which is likely to be most but not all of them.

We wish to thank our institutions for their support. Of particular importance to this work was a grant from the ANU Mathematical Sciences Research Visitors Program.

— bdm@cs.anu.edu.au and spr@cs.rit.edu; March 19, 1993.

Imagine, how quickly the amount of computation increases in these "small" Ramsey numbers: "The fastest of the two computations required about 3.2 years of cpu time on Sun Workstations," and "A side result of this computation is a catalogue of 350,866[2] (4, 5, 24)-graphs"!

What about the value of the next Ramsey number, $R(5, 5)$? In the historical summary included in [MR5], we see that the lower bound of $R(5, 5)$ increased slowly from 38 (Harvey Leslie Abbott in his impressive 1965 Ph.D. thesis [Abb]) to 42 (Robert W. Irving, 1974 [Irv2]), to finally 43 (Geoffrey Exoo, 1989 [Ex4]): Exoo produced a K_5-free 2-coloring of the edges of K_{42}.

I included an easy upper bound in problem 29.10 just as an exercise – already in 1965 J. G. Kalbfleisch [Ka1] knew better when he came up with the upper bound 59. The first half of the 1990s saw a rapid improvement due to the works by Brendan D. MacKay and Stanisław P. Radziszowski: 53 (1992), 52 (1994), 50 (1995, an implication of the $R(4, 5)$ result above), and 49 (1995, [MR5]). Finally, Vigleik Angeltveit and McKay reduced it by 1 to 48 in 2016 [AnM1].

---

[2]Later in 1993 this number grew to 350,904.

Thus, today's world records in lower and upper bound competitions for the value of $R(5, 5)$ are due to Geoffrey Exoo, and Vigleik Angeltveit and Brendan McKay, respectively:

Best Bounds 29.11 ([Ex4], [AnM1]). $43 \leq R(5, 5) \leq 48$.

And when the great expert of lower bounds Geoffrey Exoo and the great experts of upper bounds Brendan McKay and Stanisław Radziszowski agree that there is evidence for a "strong conjecture," we'd better listen – and record:

McKay–Radziszowski–Exoo's Conjecture 29.12 [MR5]

$$R(5,5) = 43.$$

It may take decades or even a century to settle this number – when done, we will see whether the three coauthors of the conjecture are right. In fact, Paul Erdős liked to popularly explain the difficulties of this problem [E94.21]: "It must seem incredible to the uninitiated that in the age of supercomputers $R(5, 5)$ is unknown. This, of course, is caused by the so-called combinatorial explosion: there are just too many cases to be checked." Paul even made up a joke about it, which I have heard during his talks in a few different variants:

> Suppose aliens invade the earth and threaten to destroy it in a year if human beings do not find $R(5, 5)$. It is, probably, possible to save the earth by putting together the world's best minds and computers. If, however, the invaders were to demand $R(6, 6)$, the human beings might as well attempt a preemptive strike without even trying to ponder the problem.[3]

Ever since 1994 Stanisław Radziszowski has maintained and revised 16 times a major compendium of "world records" in the sport of small Ramsey numbers [Radz1]. This is an invaluable service to the profession. I will present here Table 29.1 of all known non-trivial classic 2-color small Ramsey numbers and their best lower and upper bounds. Where lower and upper bounds do not coincide, they both are listed in the appropriate cell.

The cells below the main diagonal are left empty because filling them in would be redundant due to the symmetry of the Ramsey function $R(m, n) = R(n, m)$, (problem 29.3). In Table 29.2, you will find references for the results listed in Table 29.1 – see the rest in Stanisław Radziszowski's 116-page compendium, revision #16, January 15, 2021 [Radz1], readily available on the Internet. You will find there a wealth of other fascinating small Ramsey-related world records, Ramsey numbers (understood broader than here), Ramsey number inequalities, and a bibliography of over 500 referenced items (Table 29.3).

In the standard 1990 text on *Ramsey Theory* [GRS2, pp. 89–90], a tiny "Table 4.1" of known values and bounds is presented, accompanied by quite a pessimistic prediction:

[3] Alternative versions appear in [E93.20] and [E94.21].

Table 29.1 World records in classical 2-color small Ramsey numbers: known nontrivial values, lower bounds (2020) and upper bounds (2017) for two color Ramsey numbers $R(k, l) = R(k, l; 2)$, for $k \leq 10$, $l \leq 15$. For the best known upper bounds (2020) with $k \geq 4$ see Table 29.3

l *k*	3	4	5	6	7	8	9	10	11	12	13	14	15
3	6	9	14	18	23	28	36	40 42	47 50	53 59	60 68	67 77	74 87
4		18	25	36 41	49 61	59 84	73 115	92 149	102 191	128 238	138 291	147 349	158 417
5			43 48	58 87	80 143	101 216	133 316	149 442	183 633	203 848	233 1138	267 1461	269 1878
6				102 165	115 298	134 495	183 780	204 1171	262 1804	294 2566	347 3703	5033	401 6911
7					205 540	219 1031	252 1713	292 2826	405 4553	417 6954	511 10578	15263	22112
8						282 1870	329 3583	343 6090	10630	16944	817 27485	41525	873 63609
9							565 6588	581 12677	22325	38832	64864		
10								798 23556	45881	81123			1313

> Table 4.1 gives all known exact bounds [values] and some upper and lower bounds on the function R. It is unlikely that substantial improvement will be made on this table.

Just compare their Table 4.1 to Table 29.1 above, and you would agree with me that the researchers in small Ramsey numbers have dramatically exceeded expectations of the authors of [GRS2] in short 40 years. We have a race here: combinatorial explosion vs. improvements in computers and computational methods. It seems that computers and mathematicians in this field have held their own and gained some!

What would happen if we were to color edges of a complete graph K_n in more than two colors? Can we then guarantee the existence of, say, a monochromatic triangle K_3? Yes, we can.

Problem 29.13 (R.E. Greenwood and A. M. Gleason, 1955, [GG]). Prove that for any positive integer *r,* there is a positive integer $n(r)$ such that any *r*-coloring of edges of a complete graph $K_{n(r)}$ contains a monochromatic triangle K_3.

Proof We will prove this statement by induction. For $r = 1$ (i.e., 1-color edge coloring), we can certainly choose $n = 3$: 1-colored edges of K_3 form a monochromatic triangle. The statement for $r = 2$ has been proven as Problem 29.1.

Assume that for a positive integer *r,* there is $n(r)$ such that any *r*-coloring of edges of the complete graph $K_{n(r)}$ contains a monochromatic triangle K_3. We need to find the value of the

Table 29.2 References for a part of Table 29.1

l / k	4	5	6	7	8	9	10	11	12	13	14	15
3	GG	GG	Kéry	Ka2 GrY	GR McZ	Ka2 GR	Ex5 GoR1	Ex20 GoR1	Kol1 Les	Kol1 GoR1	Kol2 GoR1	Kol2 GoR1
4	GG	Ka1 MR4	Ex19 MR5	Ex3 Mac	ExT Mac	Ex16 Mac	HaKr1 Mac	ExT Spe4	SuLL Spe4	ExT Spe4	ExT Spe4	Tat Spe4
5		Ex4 AnM1	Ex9 HZ1	CaET HZ1	HaKr1 Spe4	Kuz Mac	ExT Mac	Kuz HW+	Kuz HW+	Kuz HW+	Kuz HW+	ExT HW+
6			Ka2 Mac	ExT HZ1	ExT Mac	Kuz Mac	Kuz Mac	Tat HW+	Kuz HW+	Kuz HW+	HW+	XXR HW+
7				She2 Mac	Tat HZ1	Kuz HZ2	Kuz Mac	XXER HW+	XSR2 HW+	XuXR HW+	HW+	HW+
8					BurR Mac	Kuz Ea1	Kuz HZ2	XXER HW+	HW+	XXER HW+	HW+	XXER HW+
9						She2 ShZ1	XSR2 Ea1	HW+	HW+	HW+		
10							She2 Shi2	HW+	HW+			XXER

All upper bounds for $k \geq 4$, $l \geq 6$ were improved in 2019 [AnM2]; HW+ abbreviates HWSYZH, as enhanced by Boza [Boza5]

Table 29.3 Upper bounds for $R(k, l)$, $k \geq 4$, $l \geq 5$

l / k	5	6	7	8	9	10	11	12	13	14	15
4	25	40	58	79	106	136	171	211	257	307	364
5	48	85	133	194	282	381	511	673	861	1082	1342
6		161	273	427	656	949	1352	1865	2510	3308	4305
7			497	840	1379	2134	3216				
8				1532	2683	4432	7647				

All of them were obtained by Angeltveit and McKay [AnM2] in 2019, except $R(4, 5)$ [MR4], and they improve over previously best-known bounds presented in Table 29.1

function $n(r + 1)$ such that any $(r + 1)$-coloring of edges of a complete graph $K_{n(r+1)}$ contains a monochromatic triangle K_3.

Let us define the value of the function $n(r + 1)$ as

$$n(r+1) = (r+1)(n(r)-1) + 2.$$

Assume that the edges of $K_{n(r+1)}$ are $(r+1)$-colored, and v_0 is a vertex of $K_{n(r+1)}$. Since v_0 is incident with $(r+1)(n(r)-1) + 1$ edges, by the Pigeonhole Principle there is a color, say color A, such that v_0 is incident with $n(r)$ edges of color A (Fig. 29.7).

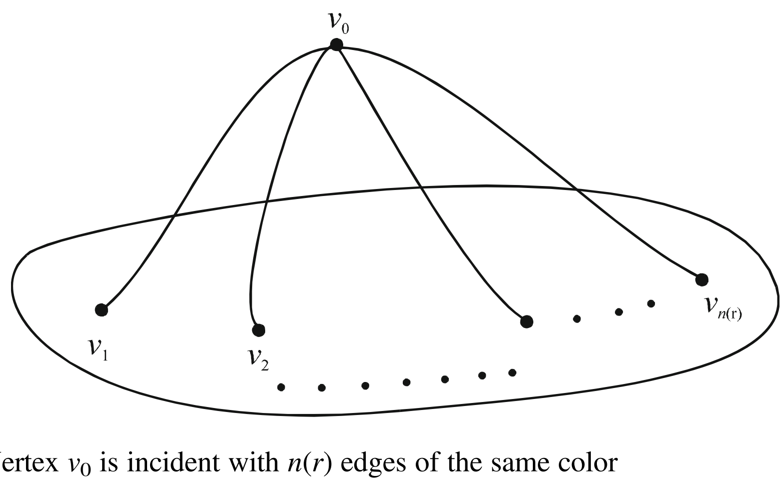


Fig. 29.7 Vertex v_0 is incident with $n(r)$ edges of the same color

If any two of the vertices $v_1, v_2, \ldots, v_{n(r)}$ are connected by an A-colored edge, these two vertices plus v_0 form an A-colored monochromatic triangle. Otherwise, we have a complete graph $K_{n(r)}$ on vertices $v_1, v_2, \ldots, v_{n(r)}$, whose edges are *r-colored*. By the inductive assumption, $K_{n(r)}$ contains a monochromatic triangle.

Note that in fact it is easy to prove by induction that $n(r) \leq \lfloor r!\,e \rfloor + 1$, where e is the base of the natural logarithms, $e = 2.718281828459045\ldots$ ■

We computed some particular Ramsey numbers and looked at ideas of proofs. Surprisingly, they are fairly recent. Even more surprising to me is that general existence results came first. The foundation for this beautiful direction in mathematics, now called *Ramsey Theory*, was laid by the young British mathematician Frank P. Ramsey. We will discuss his impressive work and short life in Chapters 30 and 31, respectively. Here, I will only formulate particular cases, graph-theoretic diagonal versions of Ramsey's Theorems.

Ramsey's Theorem, Infinite Diagonal Graph Version 29.14 Every complete infinite graph with 2-colored edges contains a complete infinite monochromatic subgraph.

Ramsey's Theorem, Finite Diagonal Graph Version 29.15 For any positive integers n and k, there is an integer $R(n; k)$ such that if $m > R(n; k)$ and the edges of a complete graph K_m are k-colored, then K_m contains a complete monochromatic subgraph K_n.

These theorems should sound familiar to you. We have solved some particular cases of Problem 29.15 earlier in this chapter and even found the values of $R(n; k)$, which for $k = 2$ we will simply denote as $R(n, n)$.[4] For example, Problems 29.1 and 29.2 show that $R(3, 3) = 6$; Problem 29.6 gives us $R(3, 4) = 9$; Problem 29.13 demonstrates that $R(3; k)$ exists for any positive integer k.

Instead of demonstrating 29.15, we will prove a stronger pair of results, 29.16 and 29.17, obtained in early 1933 and published two years later by two young unknown Hungarian university students, Pál (Paul) Erdős and Gjörgy (George) Szekeres.

Problem 29.16 (P. Erdős and G. Szekeres, [ES]) Assume that the Ramsey number $R(m, n)$ exists for every pair of positive integers m and n. Then for any integers, $m \geq 2$ and $n \geq 2$

$$R(m, n) \leq R(m-1, n) + R(m, n-1).$$

Proof Let $L = R(m-1, n) + R(m, n-1)$. We have to prove precisely that if the edges of a complete graph K_L with L vertices are colored red and blue, K_L contains a K_m with all red edges or a K_n with all blue edges. Indeed, let v_0 be a vertex of K_L whose edges are colored red and blue. We consider two cases and use an approach that proved successful in Problems 29.7 and 29.8.

Case 1 Let v_0 be incident with at least $R(m-1, n)$ red edges. Then by the definition of $R(m-1, n)$, the vertex set $S=\{v_1, v_2, \ldots, v_{R(m-1,n)}\}$ contains a blue monochromatic K_n (and we are done) or a red monochromatic K_{m-1}. In the latter case, K_{m-1} together with v_0 and $m-1$ red edges connecting them forms a red monochromatic K_m.

Case 2 Let v_0 be incident with less than $R(m-1, n)$ red edges. Since v_0 is incident with $L-1 = R(m-1, n) + R(m, n-1) - 1$ edges, each colored red or blue, we see that in this case v_0 is incident with at least $R(m, n-1)$ blue edges.

By the definition of $R(m, n-1)$, the vertex set $S = \{v_1, v_2, \ldots, v_{R(m,n-1)}\}$ contains a red monochromatic K_m (and we are done), or a blue monochromatic K_{n-1}. In the latter case, K_{n-1} together with v_0 and $n - 1$ blue edges connecting them forms a blue monochromatic K_n. ■

Problem 29.17 (P. Erdős and G. Szekeres, [ES]). For every two positive integers m and n, the Ramsey number $R(m, n)$ exists, and moreover,

[4] Please, do not overlook the significant difference between $R(n,k)$ and $R(n;k)$.

$$R(m,n) \leq \binom{m+n-2}{m-1}.$$

Proof We will use induction on $k = m + n$. We have equality when one of the numbers m, n equals 1 or 2 and the other is arbitrary (see Problems 29.4 and 29.3, and observe that $R(1, n) = 1$). Therefore, the inequality is true for $k \leq 5$, and we can assume that $m \geq 3$ and $n \geq 3$.

Assume further that $R(m–1, n)$ and $R(m, n–1)$ exist and that

$$R(m–1,n) \leq \binom{m+n-3}{m-2}$$

and

$$R(m,n–1) \leq \binom{m+n-3}{m-1}.$$

Then by Problem 29.17 and Pascal binomial equality, we get

$$R(m,n) \leq R(m–1,n) + \mathrm{R}\,(m,n–1) \leq \binom{m+n-3}{m-2} + \binom{m+n-3}{m-1} = \binom{m+n-2}{m-1}$$

as desired. We are done: $R(m, n)$ exists and satisfies the required inequality. ■

In the same paper, Paul Erdős and Gjörgy Szekeres also proved in similar spirit the Monotone Subsequence Theorem, which we will discuss in Chapter 31.

What can we learn about large Ramsey numbers if we could compute only some small Ramsey numbers? Nothing at all as far as the exact values are concerned. We can, however, aspire to estimate their growth, strive for asymptotics. This is precisely what interested Paul Erdős the most. Paul traces the developments in this direction at the 1980 Graph Theory conference at Kalamazoo [E81.20]:

It is well known that

$$c_1 n 2^{\frac{n}{2}} < R(n,n) < c_2 \binom{2n-2}{n-1}, (c_2 < 1).$$

He reports an improvement in the upper bound a few years later [E88.28]:

$$c_1 n 2^{\frac{n}{2}} < R(n,n) < \frac{c_2 \binom{2n}{n}}{(\log n)^{\varepsilon}} \qquad (*)$$

Every time Erdős speaks on this subject, he offers the same important conjecture, which still remains open today:

Erdős' $100 Conjecture 29.18 $\lim_{n \to \infty} R(n,n)^{\frac{1}{n}} = c.$

Paul adds [E88.28]: "I offer 100 dollars for a proof [of this conjecture] and 10,000 dollars for a disproof. I am sure that [the conjecture] holds." He continues with the problem of determining the limit in conjecture 29.18.

Erdős' $250 Problem 29.19 Determine c in conjecture 29.18.

Paul also gives a hint [E88.28]: "$\sqrt{2} \le c \le 4$ follows from (*), perhaps $c = 2$?" Let us record it formally.

Erdős' Open Problem 29.20 Prove or disprove that $\lim_{n \to \infty} R(n,n)^{\frac{1}{n}} = 2.$

These problems matter a great deal to Paul Erdős, for he repeats these problems in his many problem talks and papers, for example [E81.20], [E88.28], [E90.23], and [E93.20]. He even offers rare for Erdős *unspecified* compensation [E90.23]:

> Any improvement of these bounds [$\sqrt{2} \le c \le 4$] would be of great interest and will receive an "appropriate" financial reward. ("Appropriate" I am afraid is not the right word, I do not have enough money to give a really appropriate award.)

He is pessimistic about finding the asymptotic formula any time soon [E93.20]:

> An asymptotic formula for $R(n)$ would of course be very desirable, but at the moment this looks hopeless.

Yet, Erdős poses a number of other problems related to the Ramsey numbers' asymptotic behavior. Let me mention here just two examples.

Erdős' Open Problem 29.21 ([E91.31]). Is it true that for every $\varepsilon > 0$ and $n > n_0(\varepsilon)$

$$R(4,n) > n^{3-\varepsilon}?$$

In fact, probably

$$R(4,n) > \frac{cn^3}{(\log n)^{\alpha}}.$$

Erdős' Open Problem 29.22 ([E91.31]). Is it true that

$$R(n+1, n+1) > (1+c)R(n,n)?$$

In fact, it is not even known that

$$R(n+1, n+1) - R(n,n) > cn^2.$$

In the mid-1990s, when I was looking for the author of the term *Ramsey Theory* (you will find out the answer in Chapter 32), on February 19, 1996 in Baton Rouge, Louisiana, the famous graph theorist Frank Harary half-wrote, half-dictated to me a letter [Har3], which is most relevant to this chapter on Ramsey numbers, and so I will transcribe it here in its entirety (see facsimile of the opening lines in Figure 29.8):

To A Soifer 19 Feb 96

In 1965, I looked into the ramsey nos. of C_4 and $2K_2$ [two copies of K_2] and found (proved) their values are 6 and 5 resp.

Before then only ramsey nos. of complete graphs had been studied; e.g.

$r(K_3) = 6$ (folklore)

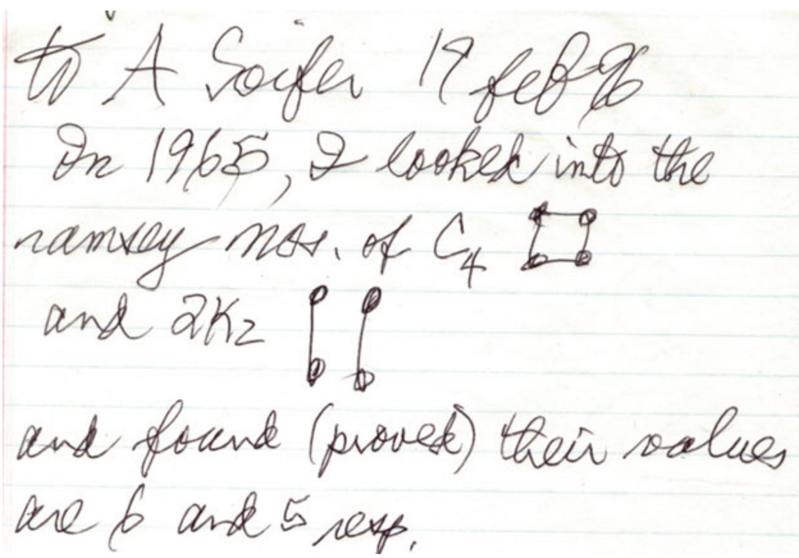
$r(K_4) = 18$ (A. Gleason + R. Greenwood)

$r(K_5) = ?$ ($100 from FH for 1st *exact* solution)

I called $r(C_4) = 6$ and $r(2K_2) = 5$ generalized ramsey nos. for graphs.

In November 1970, V. Chvatal defended his Ph.D. thesis at U Waterloo on ramsey nos. of hypergraphs. Erdős was visiting professor at the same time at Waterloo. He saw me drinking tea and grabbed my elbow saying "You must hear this doctorate defense, as Chvatal is brilliant." The next night Chvatal invited me to dinner at his house, and I proposed a series of papers to him. He accepted gladly and we had a good time writing them. I told Erdős that this was part of my big research project on Ramsey Theory.

I saw that he [Chvatal] and I would be able to carry out my research project of calculating the Ramsey Numbers of all the small G[graphs]. We wrote a series of papers "Generalized Ramsey theory for graphs" I, II, III, and maybe IV. I then continued the series to XVII. I referred early to this as the study of ramsey theory for graphs.



to A Soifer 19 feb 96
In 1965, I looked into the
ramsey nos. of C_4
and $2K_2$
and found (proved) their values
are 6 and 5 resp.

Fig. 29.8 First lines of Frank Harary's letter

Thus, Frank Harary and Václav Chvátal introduced the term *Generalized Ramsey Theory for Graphs* and started an impressive series of papers under this title. They generalized the notion of Ramsey number by including in the study the existence of monochromatic subgraphs other than complete graphs. This is a flourishing field today, and I refer you to Radziszowski's compendium [Rad1] for a summary of many achievements of this direction of research. The authors Graham–Rothschild–Spencer of [GRS2, p. 138] write:

> A major impetus behind the early development of Graph Ramsey theory was a hope that it will eventually lead to methods for determining larger values of the classical Ramsey numbers $R(m, n)$. However, as so often happens in mathematics, this expectation has not been realized; rather, the field has evolved into a discipline of its own. Asymptotic results obtained in Graph Ramsey theory may prove to be more valuable than knowing the exact value of $R(5, 5)$ [or even $R(m, n)$].

Do they prefer asymptotic results? All right. As we all know, the first exponential lower bound for diagonal Ramsey numbers was obtained by Paul Erdős in 1947 [E47.09]:

First Asymptotic Lower Bound 29.23 (Erdős, 1947 [Erd30]). $2^{n/2} \le R(n, n)$.

Only in 1975, Joel Spencer improved this bound [Sp5]:

Spencer's Asymptotic Lower Bound 29.24 (Spencer, 1975 [Sp5].

$$\sqrt{2}e^{-1}n2^{n/2}(1+o(1)) \; < \; R(n,n).$$

A group of scholars, Marcelo Campos, Simon Griffiths, Robert Morris, and Julian Sahasrabudhe, has just on March 16, 2023, at 17:38:08 uploaded [CGMS] in arXiv best exponential upper bounds for diagonal and nondiagonal Ramsey numbers. "This is the first exponential improvement over the upper bound of Erdős and Szekeres, proved in 1935," the authors write.

New Asymptotic Upper Bound 29.25 ([CGMS], 2023). There exists $\varepsilon > 0$ such that

$$R(n, n) \le (4-\varepsilon)^n$$

for all sufficiently large $n \in N$.

All right, "there exists ε," but wouldn't you like to know the value of ε? The authors obligate:

> Let us mention here for the interested reader that we will give two different proofs of Theorem1.1, the first (which is a little simpler) with $\varepsilon = 2^{-10}$, and the second with $\varepsilon = 2^{-7}$. It will be clear from the proofs that these constants could be improved further with some additional (straightforward, but somewhat technical) optimization.

While revisiting Paul Erdős' relevant essay [E90.23], I read (slightly changing notations):

> It would be very desirable to get an asymptotic formula for $R(3, k)$ (an exact formula might "not exist" in the same sense as there is no exact (and useful) formula for the nth prime). Also
>
> $$R(4, k) \; > \; \frac{c_1 k^3}{(\log k)^{c_2}}$$

should be proved. I offer for both of these problems \$250. The current best result $R(4, k) > ck^{5/2}$ is due to Joel Spencer.

Let us record it formally:

Paul Erdős' $250 Conjecture 29.26 [90.23]. Prove that $R(4, k) > \frac{c_1 k^3}{(\log k)^{c_2}}$ for sufficiently large *k*.

Imagine my shock when early in the morning (today is June 27, 2023), I received a link from Staszek Radziszowski taking me to an incredible result [MV] uploaded by Sam Mattheus and Jacques Verstraete to arXiv first on June 6, second version June 16, and the final version (so far) on June 21, 2023. Sam Mattheus informs me that they submitted their paper to *Annals of Mathematics*, an important bimonthly of Princeton-Math and the Institute for Advanced Study.

The Mattheus–Verstraete Theorem 29.27 There exist constants $c_1, c_2 > 0$ such that for all $t \geq 3$,

$$c_1 \frac{t^3}{\log^4 t} \leq r(4, \ t) \leq c_2 \frac{t^3}{\log^2 t}.$$

This theorem, of course, proves Paul Erdős' $250 conjecture 29.26. The authors "believe the upper bound in theorem 29.27 is closer to the truth."

Under [20] in their Bibliography, Mattheus–Verstraete [MV] cite "R.L. Graham, B. Rothschild, J. Solymosi, J. Spencer, *Ramsey Theory* (3rd ed.), New York: John Wiley and Sons (2015)." This is puzzling, for I knew directly from Ron Graham that the idea of producing the 3$^{\text{rd}}$ edition with the additional coauthor Jószef Solymosi had been in the works for some time, but has never materialized.

The distinguished authors of [GRS2] write about "value" (see their quote above), but how does one measure and compare values of discovering small Ramsey numbers and figuring out their asymptotic behavior? It seems that the relationship between theoretical and numerical directions of inquiry has been, as so often happens in mathematics, a marriage made in heaven. Numerical results provided a foundation for theoretical generalizations and asymptotics, while theoretical results allowed to dramatically reduce sprawling computational explosion and thus make numerical results possible. Moreover, numerical results can contain beauties both in mathematical arguments and in extreme graphs they uncover – just look at the graph in Fig. 29.9 below! I hope the rest of this chapter will illustrate my point of view. We will look at one related train of thought and the direction it has inspired, Folkman numbers. We will see that Paul Erdős was very interested in small Folkman numbers, and the authors of the above quotation, among others, have contributed their talents, energies, and results to this cause.

29.2 Folkman Numbers

In 1967, Paul Erdős and András Hajnal take the first step in a typically subtle Erdős style: they pose a particular problem.

Problem 29.28 (Erdős–Hajnal [EH2]). Construct a graph G which does not contain K_6 such that every 2-coloring of its edges contains a monochromatic K_3.

They follow up with a more general conjecture:

Erdős–Hajnal's Conjecture 29.29 [Erdős–Hajnal EH2]. For every positive integer r, there is a graph G which contains no K_4 such that every r-coloring of its edges contains a monochromatic K_3.

In the same year, 1967, Jon H. Folkman [Fol] generalizes Erdős–Hajnal conjecture. Folkman, a winner of the 1960 William Lowell Putnam Mathematical Competition and University of California Berkeley graduate before joining Rand Corporation, tragically leaves this world in 1969. He was only 31 years of age. Before I formulate Folkman's theorem in contemporary terminology, I need to introduce a few terms that have recently become standard.

Given positive integers m, n, l, an *edge Folkman graph G* is a graph without a K_l subgraph, and such that if its edges are 2-colored, there will be a subgraph K_m with all edges of color 1 or a subgraph K_n with all edges of color 2.

The *edge Folkman number* $F_e(m, n; l)$ is defined as the smallest positive integer k such that there exists an k-vertex Folkman graph G.

More generally, given positive integers $n, m_1, m_2, \ldots, m_n, l$, an *edge Folkman graph G* is a graph without a K_l subgraph, and such that if its edges are n-colored, there will be a subgraph K_{m_i} with all edges of color i for at least one value of i, $1 \le i \le n$.

The *edge Folkman number* $F_e^n(m_1, m_2, \ldots, m_n; l)$ is defined as the smallest positive integer k such that there exists a k-vertex Folkman graph G.

In this terminology, Folkman's result can be formulated as follows:

The Folkman Theorem 29.30 [Fol]. For all positive integers m, n, l; $l > \max(m, n)$, edge Folkman numbers $F_e(m, n; l)$ exist.[5]

Folkman ends [Fol] with a far-reaching generalization of Erdős–Hajnal's conjecture 29.24:

Folkman's Conjecture 29.31 [Fol]. For all positive integers $m_1, m_2, \ldots, m_n, l$, $l > \max(m_1, m_2, \ldots, m_n)$, edge Folkman numbers $F_e^n(m_1, m_2, \ldots, m_n; l)$ exist.

Vertex Folkman graphs and *vertex Folkman numbers* $F_v^n(m_1, m_2, \ldots, m_n; l)$ are defined similarly for coloring of vertices instead of edges. When $n = 2$, we omit the superscript n.

In his lyrical (a rare quality for mathematical prose) paper [Sp2], Joel Spencer recalls that in 1973, during the Erdős 60th birthday conference in Keszthely[6], Hungary, the Erdős–Hajnal conjecture 29.24 was given to the Czech mathematician Jaroslav Nešetřil and his student Vojtěch Rödl, who proved it during the conference. Moreover, they came up with pioneering results, so general, that they could be considered as principles, not unlike the Ramsey Theorems!

A *clique number* ω (G) of a graph G is the order n of its largest complete subgraph K_n.

[5] Of course, Jon Folkman did not use the term "Folkman number," which seems to have appeared first in 1993, and has since become standard.

[6] Spencer misidentifies the town as Balatonfüred; both towns are on lake Balaton.

The Edge Nešetřil–Rödl Theorem 29.32 [NR]. Given a positive integer n and a graph G, there exists a graph H of the same clique number as G, such that if edges of H are n-colored, H has an edge-monochromatic subgraph isomorphic to G.

The Vertex Nešetřil–Rödl Theorem 29.33 [NR]. Given a positive integer n and a graph G, there exists a graph H of the same clique number as G, such that if vertices of H are n-colored, H has a vertex-monochromatic subgraph isomorphic to G.

These remarkable theorems, in our terminology, imply the following:

Corollary 29.34 For any positive integers $m, n,$ the edge Folkman numbers $F_e^n(m, m, \ldots, m; m+1)$ and vertex Folkman numbers $F_v^n(m, m, \ldots, m; m+1)$ exist, where m inside the parentheses repeats n times.

Erdős–Hajnal problems, Folkman's paper, and Nešetřil–Rödl theorems inspired a new direction in Ramsey theory as well as new exciting problems on Ramsey-like numbers.

How should these new numbers be called? Several names were used at first: *Restricted Ramsey, Erdős–Hajnal*, *Graham–Spencer* [HN1], [Irv1]. Nešetřil and Rödl [NR] use such names as "*Galvin–Ramsey*" and "*EFGH*" (which, I guess, stood for Erdős–Folkman–Graham–Hajnal). In 1993, some researchers in these new numbers seem to have started, suddenly and simultaneously, to use the name *Folkman Numbers*: see, for example, Jason I. Brown and Vojtěch Rödl [BrR]; and Martin Erickson [Eri]. Slowly, through the decade that followed, this name has won out and became standard.

Obviously, if $l > R(m, n)$, where $R(m, n)$ is a Ramsey number, then $F_e(m, n; l) = R(m, n)$. The real challenge in calculating Folkman numbers occurs when $l \leq R(m, n)$, even in the simplest case $F_e(3, 3; l)$. As could be expected, the lower l is the harder is the problem (except for trivially small values of l).

In 1968 Ronald L. Graham [Gra0] published a solution of the first problem, 29.23. In fact, Graham found the smallest order graph that does the job for $F_e(3, 3, 6)$.

Graham's Result 29.35 [Gra0]. The graph $G = K_8 - C_5 = K_3 + C_5$ satisfies the conditions of problem 29.23, i.e.,

$$F_e(3, 3; 6) = 8.$$

Graham was not alone working on this new kind of a problem, as he wrote in the conclusion of [Gra0]:

> To the best of the author's knowledge, the first example of a graph satisfying the conditions [29.28] of Erdős and Hajnal was given by J. H. van Lint; subsequently L. Pósa showed the existence of such a graph containing no complete *pentagon* [$l = 5$] and Jon Folkman constructed such a graph containing no complete *quadrilateral* [$l = 4$] (all unpublished).

Paul Erdős' problem, reported in 1971 [GS1, p. 138], in the current terminology, reads simply as follows:

Paul Erdős' Open Problem 29.36 Compute edge Folkman numbers.

The simplest unknown edge Folkman number was at that time $F_e(3,3;5)$. Its first upper bound of 42 was established by M. Schäuble in 1969, which two years later was reduced to

23 by Graham and Spencer [GS1], who conjectured that 23 was the exact value. In 1973, Robert W. Irving [Irv1] reduced the upper bound to 18 and thus disproved the Graham–Spencer conjecture. A year earlier, the best lower bound of 10 was established by Shen Lin [Lin]. Joel Spencer, in his review of Irving's paper, wrote (MR0321778):[7]

> It is now known that $10 \leq F_e(3, 3; 5) \leq 18$. The determination of $F_e(3, 3; 5)$ appears to be extremely difficult.

Then, the Bulgarian mathematicians took over the problem. In 1979, N. G. Hadziivanov and N. D. Nenov [HN1] reduced the upper bound to 16. A year later, Nenov [Nen1] increased the lower bound to 11, and in 1981 he reduced the upper bound to 15 [Nen2]. In 1985, Hadziivanov and Nenov [HN2] increased the lower bound to 12.

In 1999, Stanisław P. Radziszowski, Konrad Piwakowski, and Sebastian Urbánski [PRU] increased the lower bound to match Nenov's upper bound at 15 and thus closed the problem: $F_e(3, 3; 5) = 15$.

Exact value of an edge Folkman number 29.37 ([Nen2], [PRU]). $F_e(3, 3; 5) = 15$.

In [PRU], the authors also proved that $F_v(3,3;4) = 14$ and found a unique bi-critical 14-vertex Folkman graph without a K_4 subgraph, such that any vertex 2-coloring contains a monochromatic triangle K_3 (Figure 29.9). They, as well as some of their predecessors, observed that by adding a new vertex adjacent to all 14 vertices of this graph, they get a 15-vertex Folkman graph without a K_5 subgraph and such that any edge 2-coloring contains a monochromatic triangle K_3.

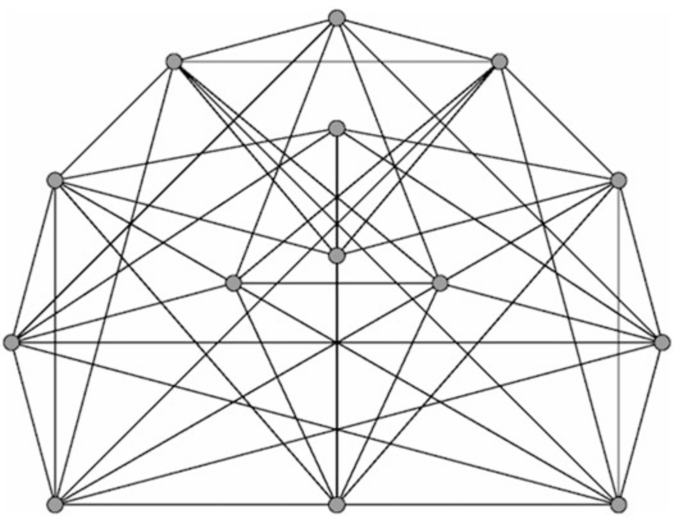


Fig. 29.9 The Piwakowski–Radziszowski–Urbánski Graph is a unique 14-vertex bi-critical $F_v(3,3;4)$ graph. (This graph is so striking, that I chose it to decorate the cover of the April–2007 issue of *Geombinatorics* XVI(4) that contained Stanisław P. Radziszowski's paper)

[7] He used the Greek α in place of not yet established Folkman symbol.

In 2017, *Geombinatorics* carried new bounds for a vertex Folkman number, obtained by Aleksandar Bikov and Nedyalko Nenov:

New bounds for a vertex Folkman number 29.38 ([BN1]). $20 \le F_v(2, 3, 3; 4) \le 24$.

Later the same year, these authors improved their algorithms and found more lower bounds:

Lower bounds for vertex Folkman numbers 29.39 ([BN3]).

$$19 \le F_v(2, 2, 2, 4; 5)$$
$$29 \le F_v(7, 7; 8)$$
$$F_v(a_1, \ldots, a_s; m) \ge m + 12, \text{if } \max\{ a_1, \ldots, a_s\} = 7.$$

This fine paper [BN3] appears to have been a Ph.D. thesis for Aleksandar Bikov, I venture to conjecture, supervised by Nedyalko Nenov in Nedyalko Nenov's Laboratory of *Sofia University “St. Kliment Ohridski.”*

In 1974 in Prague, Paul Erdős gave a talk of a special kind [E75.33]:

> I discuss some of the problems which occupied my collaborators and myself for a very long time. I tried to select those problems which are striking and which are not too well known.

One of the striking problems posed in Paul’s talk dealt with the next Folkman’s number, $F_e(3, 3; 4)$. In 1975, we knew very little about it, and Paul Erdős summarized the state of the problem as follows:

> Folkman’s upper bound for $F_e(3, 3; 4)$ is enormous (it is much bigger than $10^{10^{10^{10^{10^{10^{10}}}}}}$, the same holds for the bound of Nešetřil and Rödl.

Erdős then offered an unusual, mathematically defined price, max(100 dollars, 300 Swiss francs) for the specific bound.

Paul Erdős’ max($100, 300 SF) Problem 29.40 [E75.33]. Prove or disprove the inequality

$$F_e(3, 3; 4) < 10^{10}.$$

Dozen years later, in 1986, Frankl and Rödl [FR2] came close, within a factor of 100 from Paul Erdős’ conjectured upper bound: they used a probabilistic proof to show that $F_e(3,3; 4) \le 10^{12}$. Soon after, in 1988, Joel H. Spencer in a paper proudly called *Three Hundred Million Points Suffice* [Sp3] squeezed out of the probabilistic approach a better bound: $F_e(3, 3; 4) < 3 \times 10^8$. A mistake found in Spencer’s proof by Mark Hovey of MIT prompted Spencer in 1989 [Sp4] to increase his bound to $F_e(3, 3; 4) < 3 \times 10^9$ and change the title to *Three Billion Points Suffice*, which miraculously was still within Paul Erdős’ limit for the cash prize. The probabilistic techniques proved the existence of a Folkman graph of order 3×10^9 without actually constructing it.

Spencer's Upper Bound 29.41 ([Sp3], [Sp4], 1988–1989). $F_e(3,3;4) < 3{,}000{,}000{,}000$.

The first lower bound is a consequence of Lin's 1972 results [Lin], $10 \le F_e(3, 3; 5) \le F_e(3, 3; 4)$. In the paper published in *Geombinatorics* [RX], Stanisław P. Radziszowski and the Chinese mathematician Xiaodong Xu remark that the analysis in the cited above 1999 result $F_e(3, 3; 5) = 15$ [PRU] allows to devise a better lower bound: all 659 15-vertex graphs that have no K_5 subgraph and in every 2-coloring of edges contain a monochromatic K_3, have a subgraph K_4, hence $16 \le F_e(3, 3; 4)$. This 2007 paper [RX] contains a computer-free proof that $18 \le F_e(3, 3; 4)$, and a further computer-aided improvement to $19 \le F_e(3, 3; 4)$, which remained the best known lower bound for 10 years. In 2017, *Geombinatorics* carried the new lower bound obtained in a computer-aided proof by Aleksandar Bikov and Nedyalko Nenov [BN1]. In 2020, the same authors increased their lower bound to the best known today 21.

Bikov–Nenov's Lower Bound 29.42 ([BN4], 2020). $21 \le F_e(3, 3; 4)$.

Let us now trace the history of the upper bound after Spencer with the aid of [XR]. In 2008, L. Lu [Lu] dramatically lowered Spencer's upper bound: he showed that $F_e(3, 3; 4) \le 9697$ by constructing a certain family of K_4-free circulant[8] graphs.

During the same year, Andrzej Dudek and Vojtěch Rödl [DuR] developed a strategy to construct new Folkman graphs by approximating the maximum cut of a related graph. They substantially reduced the upper bound to 941.

Alexander Lange, Stanisław Radziszowski, and Xiaodong Xu first reduced the upper bound to 860 and then further to the best known bound of 786 with the MAX-CUT semidefinite programming relaxation as in the Goemans–Williamson algorithm. While the results were obtained by 2012 and presented in May 2012 at the 25th Cumberland Conference, Johnson City, TN, they were published only in 2014 [LRX].

Lange–Radziszowski–Xu's Upper Bound 29.43 [LRX]. $F_e(3, 3; 4) \le 786$.

Thus, the state of the problem today is this:

$$21 \le F_e(3,3;4) \le 786.$$

Enormous efforts to find $F_e(3, 3; 4)$ made me ask the expert, Stanisław Radziszowski, what would happen if we replace K_4 by $K_4 - e$ usually denoted by J_4? Would it be harder? No, $F_e(3, 3; J_4)$ does not exist. Here, any two triangles can share at most a vertex (not edge), thus we can always 2-color the edges of all triangles with 2-1 split between the colors. What about $K_5 - e = J_5$? The bounds we have today are quite wide:

$$15 = F_e(3,3;K_5) \le F_e(3,3;J_5) \le F_e(3,3;K_4) \le 786.$$

Not much is known about "neighboring" Folkman numbers. Radziszowski and Xu mention one such number obtained by L. Lu in 2008 [Lu]:

[8] The *circulant* graph can be drawn (possibly with crossings) so that its vertices lie on the corners of a regular polygon, and every rotational symmetry of the polygon is also a symmetry of the drawing.

$$F_e(K_4–e; K_4–e; K_4) \leq 30193.$$

During the 2012 SIAM Conference on Discrete Mathematics in Halifax, Nova Scotia, Ronald Graham announced a $100 award for determining whether $F_e(3, 3; 4) < 100$.

Graham's $100 Problem 29.44 Is it true that $F_e(3, 3; 4) < 94$?

As you can see, the gap between the best known lower and upper bounds of $F_e(3, 3; 4)$ is much smaller, but still significant. The authors of [RX] report:

> Geoffrey Exoo suggested to look at the well-known Ramsey coloring of K_{127} defined by Hill and Irving [HI] in 1982 in order to establish the bound $128 \leq R(4, 4; 4)$.
>
> Exoo suggested that even a 94-vertex induced subgraph, obtained by removing 3 disjoint independent sets of order 11, may still work. If true, this would imply $F_e(3, 3; 4) \leq 94$.

During Staszek Radziszowski's 8 March 2007, talk at the Florida Atlantic University conference, I hinted that a prize for a dramatic improvement in the upper bound would be in order and the speaker obligated by offering $500. Better yet, in his 22 March 2007 e-mail to me, Staszek offered two $500 prizes, for proof or disproof of the lower bound 50 and for the upper bound 127 of $F_e(3, 3; 4)$. "I believe that both of these bounds are true," he added in the e-mail.

Radziszowski's Double $500 Conjecture 29.45 $50 \leq F_e(3, 3; 4) \leq 127$.

In his 2007 talk, Radziszowski mysteriously hinted that an upper bound he conjectured to be 127 may be proved in the year 2013. On May 18, 2021, Staszek informed me that he did not conquer the upper bound, for "it is too hard." The same day I checked with another wizard of the computer-aided mathematics, Geoffrey Exoo. Here is Geoff's reply:

> Dear Alexander,
>
> Yes I did work on the Folkman problem, and I conjectured a long time ago that the cubic residue graph on 127 vertices was a Folkman graph. In fact, I believe that that graph has subgraphs of order approximately 80 that also cannot be two edge colored without a monochromatic triangle. Proving that is approximately as hard as $R(5, 5)$.

As we all know, determining the exact value of the Ramsey number $R(5, 5)$ is really hard.

Recently, Stanisław Radziszowski edited his double $500 conjecture, which now requires a harder effort on the upper bound side and reads as follows:

Radziszowski's New Double $500 Conjecture 29.46 $50 \leq F_e(3, 3; 4) \leq 94$.

In looking at recent work on Folkman numbers, I encountered an explosion. A group Zohair Raza Hassan, Yu Jiang, David E. Narváez, Stanisław Radziszowski, Xiaodong Xu produced many new results [HJNRX] (2021–2022 in arXiv, and the 2023 acceptance by the journal *Graphs and Combinatorics*.

Part VI
The Ramsey Principles

Chapter 30
From Pigeonhole Principle to Ramsey Principle

30.1 Infinite Pigeonhole and Infinite Ramsey Principles

The Infinite Pigeonhole Principle states:

Infinite Pigeonhole Principle 30.1 Let k be a positive integer. If elements of an infinite set S are colored in k colors, then S contains an infinite monochromatic subset S_1.

I am sure you will have no difficulties in proving it.

Do you see anything in common between this simple principle and Infinite Diagonal Graph Version of the Ramsey Theorem 29.14? Both say that if we have enough objects, then we can guarantee the existence of something: in the Pigeonhole Principle, it is an infinite subset; in Ramsey Theorem 29.14, we get an infinite subset of edges, i.e., the subset of two-element sets of vertices of the graph (since an edge is a pair of vertices). This connection is very close; both results are particular cases of the so-called Ramsey Theorem, one result for $r = 1$ and the other for $r = 2$. Let me formulate it here under the new, more appropriate in my opinion name:

The Infinite Ramsey Principle 30.2 For any positive integers k and r, if all r-element subsets of an infinite set S are colored in k colors, then S contains an infinite subset S_1 such that all r-element subsets of S_1 are assigned the same color.

I have always felt that something was wrong with the title "Ramsey Theorem." To see that, it suffices to read the leader of the field Ronald L. Graham, who in 1983 wrote [Gra2]:

> The generic [sic] result in Ramsey Theory is due (not surprisingly) to F. P. Ramsey.

Exactly: a "generic result," compared to much more specific typical examples, such as the Schur Theorem (Chapter 34) and the Baudet–Schur–Van der Waerden Theorem (Chapter 35). The Ramsey Theorem occupies a unique place in Ramsey Theory. It is a powerful tool. It is also a philosophical principle stating, as Theodore S. Motzkin put it, that "complete disorder is an impossibility. Any structure will necessarily contain an orderly substructure"[1]. It is, therefore, imperative to call the Ramsey Theorem by a much better fitting name: *The Ramsey Principle*.

[1] Quoted from [GRS2].

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_30

The original double induction proof of the Infinite Ramsey Principle 30.2 by F. P. Ramsey is crystal clear – read it in the original [Ram2]. I choose to present here the proof by Ronald L. Graham from [Gra1]. This is not only a beautiful proof: It demonstrates a method that has worked very productively in Ramsey Theory. Keep it in your mathematical toolbox.

Proof of 30.2 by Ronald L. Graham [Gra1]. For $r = 1$, we get the Infinite Pigeonhole Principle, which is true.

Without loss of generality, we can assume that our infinite set S coincides with the set of positive integers N. (Every infinite set S contains a countable subset equivalent to N, and N is sufficient for us to select the required in the problem subset S_1.)

We first treat the case $r = 2$ since it is easy to visualize. We can identify the 2-element subsets of N with edges of the infinite complete graph K_N with the vertex set $N = \{1, 2, \ldots, n, \ldots\}$. Let the edges of K_N be colored in k colors. It is convenient to denote the color of an edge $\{x, y\}$ by $\chi\{x, y\}$.

(1) Consider the edges of the form $\{1, x\}$, i.e., the edges incident with the vertex 1. There are infinitely many of them and only k colors; therefore, by the Infinite Pigeonhole Principle, infinitely many of these edges $\{1, x_1\}, \{1, x_2\}, \ldots, \{1, x_n\}, \ldots$ are assigned the same color, say c_1. Denote $X = \{x_1, x_2, \ldots, x_n, \ldots\}$; and let x_1 be the smallest number in X. Note that $\chi\{1, x\} = c_1$ for any x in X.
(2) Consider the edges of the form $\{x_1, x\}$ where $x \in X$; i.e., the edges incident with the vertex x_1 with the other endpoint x being an element of the set X. Once again, by the Infinite Pigeonhole Principle, infinitely many of these edges $\{x_1, y_1\}, \{x_1, y_2\}, \ldots, \{x_1, y_n\}, \ldots$ are assigned the same color, say c_2. Denote $Y = \{y_1, y_2, \ldots, y_n, \ldots\}$; and let y_1 be the smallest number in Y. Note that $\chi\{x_1, y\} = c_2$ for any y in Y.
(3) Consider the edges $\{y_1, y\}$, where $y \in Y$, i.e., edges incident with the vertex y_1 with the other endpoint y being an element of Y. By the Infinite Pigeonhole Principle, infinitely many of these edges $\{y_1, z_1\}, \{y_1, z_2\}, \ldots, \{y_1, z_n\}, \ldots$ are assigned the same color, say c_3. Denote $Z = \{z_1, z_2, \ldots, z_n, \ldots\}$; and let z_1 be the smallest number in Z. We have $\chi\{y_1, z\} = c_3$ for any z in Z, etc.

We can continue this construction indefinitely. As a result, we get the infinite set $T = \{1, x_1, y_1, z_1, \ldots\}$. It has one key property: *for any two elements t, t' from T the color of the edge $\{t, t'\}$ depends only on the value of $\min\{t, t'\}$*. Consequently, our edge coloring χ on T uniquely determines vertex coloring χ^* on T as follows:

$$\chi^*(t) = \chi\{t, t'\} \text{ for } t' > t.$$

Thus, we get the set T colored in k colors. By the Infinite Pigeonhole Principle, some infinite subset S_1 of T must be monochromatic under χ^*, i.e., all colors $\chi^*(s)$ for s from S_1 are the same. But by the definition of χ^*, this means precisely that all edges $\{s, s'\}$ of S_1 have the same color under χ. This proves the Infinite Ramsey Principle for $r = 2$.

As an example of the method, let me sketch the proof for $r = 3$. The given k-coloring χ of 3-element subsets of N uniquely determines k-coloring χ_1 of the 2-element subsets of $X = N \setminus \{1\}$ by $\chi_1\{x, x'\} = \chi\{1, x, x'\}$. By the Infinite Ramsey Principle for $k = 2$, X contains an infinite subset X' monochromatic under χ_1, (i.e., all values $\chi_1\{x, x'\}$ are the same for $x, x' \in Y$), say having color c_1, and the smallest element x_1. Next, the original k-coloring χ uniquely defines k-coloring of 2-element subsets of $Y = X' \setminus \{x_1\}$ by $\chi_2\{y, y'\} = \chi\{x_1, y, y'\}$. Once again, by the Infinite Ramsey Principle for $r = 2$, Y contains an infinite subset Y' monochromatic under χ_2, having color c_2 and the smallest element y_1. We next observe that the original k-coloring χ uniquely defines k-coloring χ_3 of 2-element subsets of $Z = Y' \setminus \{y_1\}$ by $\chi_3\{z, z'\} = \chi\{y_1, z, z'\}$, etc.

Similarly to the case $r = 2$ above, we end up with the infinite set $T = \{1, x_1, y_1, z_1, \ldots\}$ which by construction has the property that the color of any triple $\{t, t', t''\}$ depends only on $min\{t, t', t''\}$. Thus, the original k-coloring χ of 3-element subsets of T uniquely defines the k-coloring χ^* of the vertices of T as follows:

$$\chi^*(t) = \chi(\{t, t', t''\} \text{ for } t'' > t' > t.$$

By the Infinite Pigeonhole Principle, some infinite subset S_1 of T is monochromatic under χ^*. By the definition of χ^*, this means that all 3-element subsets of S_1 have the same color under χ. We are done for $r = 3$.

The inductive step for the general case follows *exactly* the same lines. ■

Have you heard of the famous Helly Theorem? I noticed in 1990 that the Helly Theorem and its variations are ready for the marriage to the Infinite Ramsey Principle. This could be a new observation: not just I, but the world-leading expert Branko Grünbaum, a coauthor of the monograph *Helly's Theorem and its Relatives* [DGK], written jointly with Ludwig Danzer and Victor Klee, tells me that he has not heard of such a marriage. Here is a plane version of the Helly Theorem for the case of infinitely many figures.

Helly's Theorem for Infinite Family of Convex Figures in the Plane 30.3 Given an infinite family of closed convex figures in the plane, one of which is bounded. If every 3 of them have a point in common, then the intersection of all figures in the family is non-empty.

We can obtain the following result, for example, by combining the Helly Theorem and the Infinite Ramsey Principle. I am naming it the *Helly × Ramsey*.

The Helly × Ramsey Theorem 30.4 (Soifer, 2007). Let $F_1, F_2, \ldots, F_n, \ldots$ be a family of closed convex figures in the plane, and F_1 be bounded. If among any 4 figures, there are 3 figures with a point in common, then infinitely many figures of the family have a point in common.

Proof Consider the set $S = \{F_1, F_2, \ldots, F_n, \ldots\}$. We color a 3-element subset $\{F_i, F_j, F_k\}$ of S red if $F_i \cap F_j \cap F_k \neq \emptyset$ and blue otherwise. By the Infinite Ramsey Principle, S contains an infinite subset S_1 such that all 3-element subsets of S_1 are assigned the same color. This color cannot be blue because every 4-element subset of S_1 contains a 3-element subset $T = \{F_i, F_j, F_k\}$, such that $F_i \cap F_j \cap F_k \neq \emptyset$, i.e., T is colored red. Thus, all 3-element subsets of S_1 are red. By the Helly Theorem 30.3, all figures of the infinite subset S_1 have a point in common. ■

The statement of theorem 30.4 remains true if we replace 4 by any larger integer n.

In 1990, Paul Erdős informed me in a letter that a stronger statement was conjectured (he was not sure by whom).

Conjecture 30.5 Given an infinite family of closed convex figures in the plane, one of which is bounded. If among any 4 figures, there are 3 figures with a point in common, then there is a finite set S (consisting of n points), such that every given figure contains at least one point from S.

Moreover, n is an absolute constant (i.e., it is one and the same for all families that satisfy the above conditions).

Vladimir Boltyanski and I first published this conjecture in 1991 [BS]. Eighteen years later, on September 26, 2008, while reading the manuscript of the forthcoming new expanded 2009 Springer edition of [BS], Branko Grünbaum resolved this conjecture in the negative: I mailed him the $25 prize. Grünbaum showed that Conjecture 30.5 does not hold even for line E.

Grünbaum's Counterexample 30.6 (e-mail to A. Soifer, September 26, 2008). Define the sets as follows:

$F_0 = \{0\}$;
$F_n = \{x \in E : x \geq n\}$, for every positive integer n.

Of course, all conditions of Conjecture 30.5 are satisfied, while for any finite set S of reals, there is an integer n that is greater than any number from S. By definition, F_n does not contain any element from S. ■

On September 29, 2008, I asked Branko Grünbaum whether he can "save" conjecture 30.5 and the following day he sent me his saving recipe:

> Yes, I conjecture that Erdős problem may be resuscitated by requiring two (instead of just one) of the sets to be compact. But I do not see any easy proof.

Grünbaum's Conjecture 30.7 (e-mail to A. Soifer, September 30, 2008). Given an infinite family of closed convex figures in the plane, *two* of which are compact. If among any four figures, there are three figures with a point in common, then there is set S consisting of $N = N(n)$ points, such that every given figure contains at least one point from S.

On 3 July 2014, I received an email from Pablo Soberón from the University of Michigan:

> Dear Alexander,
> Recently I've been working with some colleagues on Conjecture 30.7 from your (very nice) mathematical coloring book, regarding Helly's theorem. In case it is of interest to you, I've copied the arXiv link to the manuscript below,
> Kind regards,
> Pablo

This paper [MMRS] by Amanda Montejano, Luis Montejano, Edgardo Roldán-Pensado, and Pablo Soberón, dated inside the paper "September 15, 2018," cites the Boltyanski–Soifer conjecture that first appeared in 1990 in our joint book [BS], then the story of my exchange with Branko Grünbaum that I've just shared with you. Then they cite Tobias Müller [Mül], who in 2013 disproves conjecture 30.7, and observe that "a natural extension of Müller's work refutes the possibility of 'saving' the conjecture by replacing the condition of two bounded sets by any number."

30.2 Pigeonhole Principle and Finite Ramsey Principle

Let us take another look at Frank P. Ramsey's pioneering 1930 paper [Ram2]. Having disposed of the infinite case, Ramsey proves the finite one [Ram2, Theorem B]. As a methodology of the new theory, it ought to be elevated to the status of a principle.

The Finite Ramsey Principle 30.8 For any positive integers r, n, and k, there is an integer $m_0 = R(r, n, k)$ such that if $m > m_0$ and all r-element subsets of an m-element set S_m are colored in k colors, then S_m contains an n-element subset S_n such that all r-element subsets of S_n are assigned the same color.

Proof The Finite Ramsey Principle follows from the Infinite Ramsey Principle 30.2 by the de Bruijn–Erdős Compactness Theorem 27.1.

A clearly written direct proof, without the use of compactness argument, can be found in the original 1930 paper by F. P. Ramsey [Ram2]; it is also reproduced in full in [GRS2]. ■

As you surely noticed, the (finite) Pigeonhole Principle is a particular case of the Finite Ramsey Principle for $r = 1$.

The Pigeonhole Principle 30.9 Let n and k be positive integers. If elements of a set S with at least $m_0 = (n - 1)k + 1$ elements are colored in k colors, then S contains a monochromatic n-element subset.

Since edges can be viewed as 2-element subsets of the vertex set of a graph, by plugging in $r = 2$ in the Finite Ramsey Principle, we get the result we encountered in the previous chapter: Finite Diagonal Graph Version of Ramsey Theorem 29.15.

It is amazing to me how swiftly the news of the Ramsey Principle travel in the times that can hardly be called the Information Age. Ramsey's paper appears in 1930. Already in 1933, the great Norwegian logician Thoralf Albert Skolem (1887–1963) publishes his own proof [Sko] of the Ramsey Theorem (with a reference to Ramsey's 1930 publication!). In 1935, yet another proof (in the graph-theoretic setting) appears in the paper [ES1] by two young Hungarians, Pál (Paul) Erdős and Gjörgy (George) Szekeres. We will look at this remarkable paper in the next chapter. The authors then consider generalizations of Erdős–Grünbaum conjectures, which are valuable, but lie outside of the scope of this book.

Chapter 31
The Happy End Problem

31.1 The Problem

During the winter of 1932–1933, two young friends, mathematics student Paul (Pál) Erdős, aged 19, and chemistry student George (György) Szekeres, 21, solved the problem posed by their young lady friend Esther Klein, 22, but did not send it to a journal for a year and a half. When Erdős finally sent this joint paper for publication, he chose J. E. L. Brouwer's journal *Compositio Mathematica*, where it appeared in 1935 [ES1].

Erdős and Szekeres were first to demonstrate the power and striking beauty of the Ramsey Principle when they solved the problem. Do not miss G. Szekeres' story of this momentous solution later in this chapter. In the process of working with Erdős on the problem, Szekeres actually rediscovered the Finite Ramsey Principle before the coauthors ran into the 1930 Ramsey publication [Ram2].

The Erdős–Szekeres Theorem 31.1 [ES1]. For any positive integer $n \geq 3$, there is an integer m_0 such that any set of at least m_0 points in the plane in general position[1] contains n points that form a convex polygon.

To prove the Erdős–Szekeres Theorem, we need the following two tools.

Tool 31.2 (Esther Klein, Winter 1932–33). Any 5 points in the plane in general position contain 4 points that form a convex quadrilateral.

In fact, in anticipation of the proof of the Erdős–Szekeres Theorem, it makes sense to introduce an appropriate notation $ES(n)$ for the Erdős–Szekeres function. For a positive integer n, $ES(n)$ will stand for the minimal number such that any $ES(n)$ points in the plane in general position contain n points that form a convex n-gon. Esther Klein's result can then be written as follows:

Result 31.3 (Esther Klein). $ES(4) = 5$.

[1] I.e., no three points lie on a line.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_31

Proof Surely, $ES(4) > 4$. Given 5 points in the plane in general position, consider their convex hull H.[2] If H is a quadrilateral or a pentagon, we are done. If H is a triangle, two of the five given points, call them a, b lie inside H. The line ab does not intersect one of the sides of triangle H, let it be the side de. Then, we get a convex quadrilateral formed by the four points $a, b, d,$ and e. ■

Tool 31.4 (P. Erdős and G. Szekeres, [ES1]). Let $n \geq 3$ be a positive integer. Then n points in the plane form a convex polygon if and only if every 4 of them form a convex quadrilateral.

Paul Erdős told me that two members of his Budapest circle E. Makai and Paul Turán established (but never published) one more exact value of $ES(n)$:

Result 31.5 (E. Makai and P. Turán). $ES(5) = 9$.

Erdős mentioned the authorship of this result numerous times in his problem papers and talks. However, I know only one instance when he elaborated on it. During Paul Erdős' stay with me in March 1989, he gave two lectures at the University of Colorado at Colorado Springs. In his first lecture, Paul mentioned that Makai and Turán found proofs of 31.5 *independently*. Paul said that Makai proof was lengthy and shared with us Turán's short Olympiad-like proof. Turán starts along Esther Klein's lines, by looking at the convex hull of the given 9 points. Let me stop right here to allow you the pleasure of finding a proof on your own.

We are now ready to prove the Erdős–Szekeres Theorem asserting the existence of the function $ES(n)$.

Proof of Theorem 31.1 by P. Erdős and G. Szekeres Let $n \geq 3$ be a positive integer. By the Ramsey Principle 30.8 (we set $r = 4$ and $k = 2$), there is an integer $m_0 = R(4, n, 2)$ such that if $m > m_0$ and all 4-element subsets of an m-element set S_m are colored in 2 colors, then S_m contains a n-element subset S_n with all 4-element subsets of S_n assigned the same color.

Now let S_m be a set of m points in the plane in general position. We color a 4-element subset of S_m red if it forms a convex quadrilateral and blue if it forms a concave (i.e., non-convex) quadrilateral. Thus, all 4-element subsets of S_m are colored red and blue. Hence, S_m contains an n-element subset S_n such that all 4-element subsets of S_n are assigned the same color. This color cannot be blue because in view of tool 31.2 any 5- or more element set contains a red 4-element subset! Therefore, all 4-element subsets of S_n are colored red, i.e., they form convex quadrilaterals. By tool 31.4, S_n forms a convex n-gon. ■

I must show you a beautiful alternative proof of the Erdős–Szekeres Theorem 31.1, especially since it was found by an undergraduate student, Michael Tarsi of Israel. He missed the class when the Erdős–Szekeres solution was presented and had to come up with his own proof under the gun of the exam! Tarsi recalls (e-mail to me of December 12, 2006):

> Back in 1972, I took the written final exam of an undergraduate Combinatorics course at the Technion – Israel Institute of Technology, Haifa, Israel. Due to personal circumstances, I had barely attended school during that year and missed most lectures of that particular course. The so-called Erdős–Szekeres Theorem was presented and proved in

---

[2] Convex hull of a set S is the minimal convex polygon that contains S. If you pound a nail in every vertex of S, then a tight rubber band around all nails would produce the convex hull.

class, and we have been asked to repeat the proof as part of the exam. Having seen the statement for the first time, I was forced to develop my own little proof.

Our teacher in that course, the late Professor Mordechai Levin, had published the story as an article, I cannot recall the journal's name, the word 'Gazette' was there, and it dealt with Mathematical Education.

I was born in Prague (Czechoslovakia at that time) in 1948 but was raised and grew up in Israel since 1949. Currently I am a professor of Computer Science at Tel Aviv University, Israel.

Proof of Theorem 31.1 by Michael Tarsi Let $n \geq 3$ be a positive integer. By the Ramsey Principle 30.8 ($r = 3$ and $k = 2$), there is an integer $m_0 = R(3, n, 2)$ such that, if $m > m_0$ and all 3-element subsets of an m-element subset S_m are colored in 2 colors, then S_m contains an *n-element* subset S_n such that all 3-element subsets of S_n are assigned the same color.

Let now S_m be a set of m points in the plane in general position labeled with integers 1, 2, . . ., m. We color a 3-element subset $\{i, j, k\}$, where $i < j < k$, red if we travel from i to j to k in a clockwise direction and blue if counterclockwise. By the above, S_m contains an n-element subset S_n such that all 3-element subsets of S_n are assigned the same color, i.e., have the same orientation. But this means precisely that S_n forms a convex n-gon! ■

In their celebrated paper [ES1], P. Erdős and G. Szekeres also discover the Monotone Subsequence Theorem.

A sequence $a_1, a_2, \ldots, a_k$ of real numbers is called *monotone* if it is increasing, i.e., $a_1 \leq a_2 \leq \ldots \leq a_k$, or decreasing, i.e., $a_1 \geq a_2 \geq \ldots \geq a_k$ (we use weak versions of these definitions that allow equalities of consecutive terms).

The Erdős–Szekeres Monotone Subsequence Theorem 31.6 [ES1]. Any sequence of $n^2 + 1$ real numbers contains a monotone subsequence of $n + 1$ numbers.

I would like to show here how the Ramsey Principle proves such a statement with, of course, much worse upper bound than $n^2 + 1$. I haven't seen this argument in literature before.

Problem 31.7 Any long enough sequence of real numbers contains a monotone subsequence of $n + 1$ numbers.

Solution Take a sequence S of $m = R(2, n + 1, 2)$ numbers $a_1, a_2, \ldots, a_m$. Color a 2-element subsequence $\{a_i, a_j\}$, $i < j$ red if $a_i \leq a_j$ and blue if $a_i > a_j$. By the Ramsey Principle, there is an $(n + 1)$-element subsequence S_1 with every 2-element subsequence of the same color. This subsequence is monotone! ■

In [ES1] P. Erdős and G. Szekeres generalize Theorem 31.6 as follows:

The Erdős–Szekeres Monotone Subsequence Theorem 31.8 Any sequence S: $a_1, a_2, \ldots, a_r$ of $r > mn$ real numbers contains a decreasing subsequence of more than m terms or an increasing subsequence of more than n terms.

A quarter of a century later, in 1959, Abraham Seidenberg of the University of California, Berkeley, found a brilliant "one-line" proof of theorem 31.8, thus giving it a true Olympiad-like appeal.

Proof of Theorem 31.8 by A. Seidenberg [Sei]. Assume that the sequence S: $a_1, a_2, \ldots, a_r$ of $r > mn$ real numbers has no decreasing subsequence of more than m terms. To each a_i assign a pair of numbers (m_i, n_i), where m_i is the largest number of terms of a decreasing subsequence beginning with a_i and n_i the largest number of terms of an increasing subsequence beginning with a_i. This correspondence is an injection, i.e., distinct pairs correspond to distinct terms a_i, a_j, $i < j$. Indeed, if $a_i \leq a_j$, then $n_i \geq n_j + 1$, and if $a_i > a_j$, then $m_i \geq m_j + 1$.

We get $r > mn$ distinct pairs (m_i, n_i), they are our pigeons, and m possible values for the first coordinates m_i, since $1 \leq m_i \leq m$, they are our m pigeonholes. By the Pigeonhole Principle, there are at least $n + 1$ pairs (m_0, n_i) with the same first coordinate m_0. Terms a_i corresponding to these pairs (m_0, n_i) form an increasing subsequence! ■

Erdős and Szekeres note that the result of their theorem 31.8 is the best possible:

Problem 31.9. ([ES1]) Construct a sequence of mn real numbers such that it has no decreasing subsequence of more than m terms *and* no increasing subsequence of more than n terms.

Proof Here is a sequence of mn terms that does the job:

$$m, m-1, \ldots 1; 2m, 2m-1, \ldots, m+1; \ldots; nm, nm-1, \ldots, (n\text{–}1)\, m+1. ■$$

H. Burkill and Leon Mirsky in their 1973 paper [BM] observe that the Monotone Subsequence Theorem holds for countable sequences as well.

Countable Monotone Subsequence Theorem 31.10 [BM]. Any countable sequence $S{:}\ a_1, a_2, \ldots, a_r, \ldots$ of real numbers contains an infinite increasing subsequence or an infinite strictly decreasing subsequence.

Hint Color the 2-element subsets of S in two colors. ■

The authors "note in passing [without proof] that the same type of argument enables us to show" the following cute result:

Curvature Preserving Subsequence Theorem 31.11 [BM]. Any countable sequence $S{:}\ a_1, a_2, \ldots, a_r, \ldots$ of real numbers possesses an infinite subsequence which is convex or concave.

Hint Recall Michael Tarsi's proof of the Erdős–Szeklers Theorem, and color the 3-element subsets of S in two colors! ■

The results of this chapter reminded me about the celebrated Helly Theorem, this time its main, finite version.

The Helly Theorem 31.12 Let $F_1 \ldots, F_m$ be convex figures in *n-dimensional* space E^n. If every $n + 1$ of these figures have a common point, then the intersection of them all $F_1 \cap \ldots \cap F_m$ is non-empty.

In particular, for $n = 2$, we get the Helly Theorem for the plane.

The Helly Theorem for the Plane 31.13 A finite family $F_1, \ldots, F_m$ of convex figures is given in the plane. If every 3 of them have a non-empty intersection, then the intersection $F_1 \cap \ldots \cap F_m$ of all these figures is non-empty as well.

The structure of the Helly Theorem appears to me similar to one of the theorem 31.1. This is why I believe that the Helly Theorem and its numerous beautiful variations are a fertile ground for applications of the powerful tool, the Finite Ramsey Principle 30.8. To the best of my – and Branko Grünbaum's – knowledge, this marriage of Helly and Ramsey has not been noticed before. To illustrate it, I have created a sample problem. Its result is not important, but the method may lead you to discovering new theorems.

Problem 31.14 (A. Soifer, 2007). Let m be a large enough positive integer, $m \geq R(3, 111, 2)$ to be precise, and $F_1, \ldots, F_m$ convex figures in the plane. If among every 37 figures there are 3 figures with a point in common, then there are 111 figures with a point in common.

Hint. The fact that $37 \times 3 = 111$ is a diversion, it has absolutely nothing to do with the solution: the statement of problem 31.14 remains true if we replace 37 and 111 by arbitrary positive integers l and n, respectively, as long as $l \leq n$.

Solution Let $m \geq R(3, 111, 2)$, and $F_1, F_2, \ldots, F_m$ be convex figures in the plane. Consider the set $S = \{F_1, F_2, \ldots, F_m\}$. We color a 3-element subset $\{F_i, F_j, F_k\}$ of S red if $F_i \cap F_j \cap F_k \neq \emptyset$, and blue otherwise. By the Finite Ramsey Principle 30.8, there is a 111-element subset S_1 of S such that all its 3-element subsets are assigned the same color. Which color can it be? Surely not blue, for among every 37 Figures, there are 3 figures with a point in common, thus forming a red 3-element subset. Thus, all 3-element subsets of S_1 are red. Therefore, by the Helly Theorem 31.13, the intersection of all 111 figures of the set S_1 is non-empty. ■

31.2 The Story Behind the Happy End Problem

On Paul Erdős' 60th birthday, his lifelong friend George (György) Szekeres gave Paul and us all a present of magnificent reminiscences, allowing us a glimpse into Erdős and Szekeres' first joint paper [ES1] and the emergence of a unique group of young unknown Jewish Hungarian mathematicians in Budapest, many of whom were destined to a great mathematical future. My request to reproduce these remarkable reminiscences, George Szekeres answered in the March 5, 1992 letter:

> Dear Alexander, … Of course, as far as I am concerned, you may quote anything you like (or see fit) from my old reminiscences in "The Art of Counting"… But of course it may be different with MIT Press, that you have to sort out with them.

I am grateful to George Szekeres and the MIT Press for their kind permissions to reproduce George's memoirs here. His *Reminiscences* are sad and humorous at the same time and warm above all. George kindly shared with me two photographs, allowing us an extremely rare glimpse at young handsome George and Esther, and also young Paul Turán with his first wife Edit (Klein) Kóbor and their son Robert. György (George) Szekeres recollects [Szek]:

> It is not altogether easy to give a faithful account of events which took place forty years ago, and I am quite aware of the pitfalls of such an undertaking. I shall attempt to describe the genesis of this paper, and the part each of us played in it, as I saw it then and as it lived on in my memory.
>
> For me there is a bit more to it than merely reviving the nostalgic past. Paul Erdős, when referring to the proof of Ramsey's theorem and the bounds for Ramsey numbers given in the paper, often attributed it to me personally (e.g., in [E42.06]), and he obviously attached some importance to this unusual step of pinpointing authorship in a joint paper. At the same time the authorship of the "second proof" was never clearly identified.

György Szekeres and Esther Klein, Bükk Mountains, Northern Hungary, 1938 (shortly after their 1937 marriage). (Courtesy of György Szekeres)

I used to have a feeling of mild discomfort about this until an amusing incident some years ago reassured me that perhaps I should not worry about it too much. A distinguished British mathematician gave a lunch hour talk to students at Imperial College on Dirichlet's box principle, and as I happened to be with Imperial, I went along. One of his illustrations of the principle was a beautiful proof by Besicovitch of Paul's theorem (2nd proof in [ES1]), and he attributed the theorem itself to "Erdős and someone whose name I cannot remember." After the talk I revealed to him the identity of Paul's co-author (incidentally also a former co-author of the speaker) but assured him that no historical injustice had been committed as my part in the theorem was less than ε.

The origins of the paper go back to the early thirties. We had a very close circle of young mathematicians, foremost among them Erdős, Turán and Gallai; friendships were forged which became the most lasting that I have ever known and which outlived the upheavals of the thirties, a vicious world war and our scattering to the four corners of the world. I myself was an "outsider," studying chemical engineering at the Technical University, but often joined the mathematicians at weekend excursions in the charming hill country around Budapest and (in summer) at open air meetings on the benches of the city park.

Paul Erdős, early 1930s, Budapest. (Courtesy of Paul Erdős)

Paul, then still a young student but already with a few victories in his bag, was always full of problems and his sayings were already a legend. He used to address us in the same fashion as we would sign our names under an article and this habit became universal among us; even today I often call old members of the circle by a distortion of their initials.

"Szekeres Gy., open up your wise mind." This was Paul's customary invitation – or was it an order? – to listen to a proof or a problem of his. Our discussions centered around mathematics, personal gossip, and politics. It was the beginning of a desperate era in Europe. Most of us in the circle belonged to that singular ethnic group of European society which drew its cultural heritage from Heinrich Heine and Gustav Mahler, Karl Marx and Cantor, Einstein and Freud, later to become the principal target of Hitler's fury. Budapest had an exceptionally large Jewish population, well over 200,000, almost a quarter of the total. They were an easily identifiable group, speaking an inimitable jargon of their own and driven by a strong urge to congregate under the pressures of society. Many of us had leftist tendencies, following the simple reasoning that our problems can only be solved on a global, international scale and socialism was the only political philosophy that offered such a solution. Being a leftist had its dangers and Paul was quick to spread the news when one of our number got into trouble: "A. L. is studying the theorem of Jordan." It meant that following a political police action A. L. has just verified that the interior of a prison cell is not in the same component as the exterior. I have a dim recollection that this is how I first heard about the Jordan curve theorem.

Apart from political oppression, the Budapest Jews experienced cultural persecution long before anyone had heard the name of Hitler. The notorious "numerus clausus" was operating at the Hungarian Universities from 1920 onwards, allowing only 5% of the total student intake to be Jewish. As a consequence, many of the brightest and most purposeful students left the country to study elsewhere, mostly in Germany, Czechoslovakia, Switzerland, and France. They formed the nucleus of that remarkable influx of Hungarian mathematicians and physicists into the United States which later played such an important role in the fateful happenings towards the conclusion of the second world war.

For those of us who succeeded in getting into one of the home universities, life was troublesome and the outlook bleak. Jewish students were often beaten up and humiliated by organized student gangs and it was inconceivable that any of us, be he as gifted as Paul, would find employment in academic life. I myself was in a slightly better position as I studied chemical engineering and therefore resigned to go into industrial employment, but for the others even a high school teaching position seemed to be out of reach.

Paul moved to Manchester soon after his Ph.D. at Professor Mordell's invitation and began his wanderings which eventually took him to almost every mathematical corner of the world. But in the winter of 1932/33 he was still a student; I had just received my chemical degree and, with no job in sight, I was able to attend the mathematical meetings with greater regularity than during my student years. It was at one of these meetings that a talented girl member of our circle, Esther Klein (later to become Esther Szekeres), fresh from a one-semester stay in Göttingen, came up with a curious problem: given five points in the plane, prove that there are four which form a convex quadrilateral. In later years this problem frequently appeared in student's competitions, also in the *American Mathematical Monthly* (53(1946)462, problem E740). Paul took up the problem eagerly and a generalization soon emerged: is it true that out of $2^{n-2} + 1$ points in the plane one can always select n points so that they form a convex n–sided polygon? I have no clear recollection how the generalization actually came about; in the paper we attributed it to Esther, but she assures me that Paul had much more to do with it. We soon realized that a simple-minded argument would not do and there was a feeling of excitement that a new type of geometrical problem emerged from our circle which we were only too eager to solve. For me the fact that it came from Epszi (Paul's nickname for Esther, short for ε) added a strong incentive to be the first with a solution and after a few weeks I was able to confront Paul with a triumphant "E. P., open up your wise mind." What I really found was Ramsey's theorem, from which it easily followed that there exists a number $N < \infty$ such that out of N points in the plane it is possible to select n points which form a convex n-gon. Of course, at that time none of us knew about Ramsey. It was a genuinely combinatorial argument and it gave for N an absurdly large value, nowhere near the suspected 2^{n-2}. Soon afterwards Paul produced his well known "second proof" which was independent of Ramsey and gave a much more realistic value for N; this is how a joint paper came into being.

I do not remember now why it took us so long (a year and a half) to submit the paper to the *Compositio.* These were troubled times and we had a great many worries. I took up employment in a small industrial town, some 120 km from Budapest, and in the following year Paul moved to Manchester; it was from there that he submitted the paper.

I am sure that this paper had a strong influence on both of us. Paul with his deep insight recognized the possibilities of a vast unexplored territory and opened up a new world of combinatorial set theory and combinatorial geometry. For me it was the final proof (if I needed any) that my destiny lay with mathematics, but l had to wait for another fifteen years before I got my first mathematical appointment to Adelaide. I never returned to Ramsey again.

Paul's method contained implicitly that $N > 2^{n-2}$, and this result appeared some thirty-five years later [ES2] in a joint paper, after Paul's first visit to Australia. The problem is still not completely settled, and no one yet has improved on Paul's value of

$$N = \binom{2n-4}{n-2} + 1.$$

Of course, we firmly believe that $N = 2^{n-2} + 1$ is the correct value. ■

These moving memories prompted me to ask for more. George Szekeres replied on November 30, 1992:

Dear Sasha, . . . Marta Svéd rang me some time ago from Adelaide, reminding me of an article that I was supposed to write about the old Budapest times . . . From a distance of 60 years, as I approach 82, these events have long lost their "romantic" freshness . . . My memories of those times are altogether fading away into the remote past, even if they are occasionally refreshed on my visits to Budapest. (I will certainly be there to celebrate Paul's 80-th birthday.)

The following year George Szekeres and Esther Klein did come to Keszthely, Hungary, located on the shores of beautiful Lake Balaton. We met for an outdoor dinner during the unforgettable conference dedicated to Paul Erdős' 80th birthday. George and Esther shared with me unique memories of Tibor Gallai, a key member of their Budapest group. See them in a later chapter, dedicated to the Gallai Theorem. During the outdoor dinner, George and Esther passed to me all complimentary shot glasses with something Hungarian resembling Russian vodka.

31.3 An Early Photograph of Turán Pál and His First Family

As you know from the previous chapter, György (George) Szekeres shared with me a rare early photograph of an essential member of the group of young Jewish mathematicians of Budapest, Turán Pál, or Paul Turán. This photograph George Szekeres and Esther Klein carried with them from Hungary to their exile in China and then to their new home in Sydney, Australia. George wanted me to have this photograph, and so I owed it to him to publish it.

I did not include it in the first edition of this book because I could not find the names of Paul's first wife and son. Almost nobody knew, and those few who did, would not answer me, as if the first marriage of Paul Turán is a classified national security secret, and I – and you, my reader – do not have security clearance to learn about Paul Turán's first family.

Paul Turán, his first wife Edit (Klein) Kóbor, and their son Róbert Turán, ca. 1940s. (Courtesy of György Szekeres)

Finally, this photograph for the first time is seeing the light of day, and you get to meet young Paul Turán with his beautiful first wife Edit (Klein) Kóbor, and their son Róbert Turán.

We know that Hungary, collaborating with Nazi Germany, was not a bed of roses for the Jews. For a while, Paul Turán could not get an academic job. In the years 1940–1944, he was involuntarily sent to labor camps. In the process of searching for Paul Turán's first family, I obtained from Hungary information that I did not see in print, at least not in mathematical texts. My inquiry to Magyar Zsidó Múzeum (Hungarian Jewish Museum) was answered on July 22, 2019, by Dr. Zsuzsanna Toronyi:

> Dear Prof. Soifer,
>
> Thank you for your inquiry about our former director's father who was a world-famous mathematician. Róbert Turán was the director of the Hungarian Jewish Museum and Archives between 1994 and 2008. He himself is a writer, with amazing stories about the family as well. His homepage is available at: http://turanrobert.hu/index.html. His e-mail address is <snip>.
>
> Let me draw your attention to Turán Pál's other son as well: he is Tamas Turán, now lives in Israel and works for the Hebrew University as well as for our Jewish Studies Department at the ELTE University, Budapest. He also started as a mathematician but studies philosophy as well. His publications at the academia.edu are available here: https://mta.academia.edu/TamasTuran.
>
> His mother is Vera Sós, also a professor of mathematics.[3] Tamas's e-mail address is <snip>.
>
> Turán Pál was active in the Jewish Community and as a member of the General Assembly, his portrait is available in our collection http://collections.milev.hu/items/show/33044 as part of huge tableaux displaying the members of the General Assembly in 1950: http://collections.milev.hu/items/show/32853.

I contacted Róbert Turán several times without receiving a reply. Tamas Turán replied that the family did not wish to share information.

31.4 Progress on the Happy End Problem

In May 1960, when Paul Erdős visited George Szekeres in Adelaide, Australia, they improved the lower bound in the Happy End Problem [ES2].

Lower Bound 31.15 (P. Erdős and G. Szekeres [ES2]). $2^{n-2} \le ES(n)$, where $ES(n)$ is the Erdős–Szekeres function, i.e., the smallest integer such that any $ES(n)$ points in general position contain a convex n-gon.[4]

It is fascinating how sure Erdős and Szekeres were of their conjecture. In one of his last, posthumously published problem papers [E97.18], Paul Erdős attached the prize and modestly attributed the conjecture to Szekeres: "I would certainly pay $500 for a proof of Szekeres' conjecture."

[3] Vera T. Sós, Paul Turán's second wife, is a famous mathematician, member of the Hungarian Academy of Sciences, a frequent collaborator with Paul Turán and Paul Erdős.

[4] Erdős and Szekeres actually proved a strict inequality.

The Erdős–Szekeres Happy End $500 Conjecture 31.16

$$ES(n) = 2^{n-2} + 1.$$

Their confidence is a bit surprising[5] because the foundation for the conjecture was quite thin, just results 31.2 and 31.5:

$$ES(4) = 5,$$
$$ES(5) = 9.$$

Of all people, I should not be surprised, for my general conjecture for the chromatic number of the Euclidean n-dimensional space (see it near the end of the book) also has a slim foundation, but my firm belief. Btw, both conjectures utilize an exponential function with the base 2.

Computing exact values of the Erdős–Szekeres function $ES(n)$ proved to be a very difficult matter. It took over 70 years to make the next step. In 2006, George Szekeres (posthumously) and Lindsay Peters, with the assistance of Brendan McKay and heavy computing, established one more exact value in the paper [SP] written "In memory of Paul Erdős":

Result 31.17 (G. Szekeres and L. Peters [SP]). $ES(6) = 17$.

In his surveys [Gra7], [Gra8],[6] Ronald L. Graham offered $1000 for the first proof – or disproof – of the Erdős–Szekeres Happy End Conjecture 31.16.

George Szekeres was, of course, correct when he wrote in his 1973 reminiscences that their 1935 upper bound

$$ES(n) \le \binom{2n-4}{n-2} + 1$$

had not been improved. In fact, it withstood all attempts of improvement until 1997 when Fan Chung and Ron Graham [CG] willed it down by 1 point to

$$ES(n) \le \binom{2n-4}{n-2}.$$

In my late 2006 New Year's greetings, I asked Ron Graham to tell me how this progress came about, and on 29 December 2006, received his reply:

Hi Sasha,

Happy New Year to you as well! Regarding your questions, here is the story.

For the Happy End Theorem, I thought it would be nice to improve the bounds on this problem before I saw Szekeres in 1996 at the Erdős ceremony in Budapest. So, Fan and I

[5] In fact, Paul Erdős repeated $500 offer for the proof of the conjecture in [E97.21] but offered there "only 100 dollars for a disproof," thus amplifying his belief in the conjecture.

[6] I thank Ron Graham for kindly providing the preprints.

worked a bit on it and did manage to lower the known upper bound by 1! (but even this took some new ideas). We circulated the preprint and very soon Kleitman and Pachter lowered the upper bound by a linear term, and shortly thereafter (still in 1997), Toth and Valtr lowered it basically by a factor of 2. More recently, this bound has been lowered by another 1! I offer $1000 to prove that

2^(n – 2) + 1 is the correct bound, and $100 to prove an upper bound of O((4–c)^n) for some c > 0 . . .

Best regards,

Ron

Let me add the bounds that Ron mentions in his e-mail. Fan Chung and Ron Graham were first to improve the bound and offer a fresh approach which started an explosion of improvements. Then the upper bound was improved by Daniel J. Kleitman and Lior Pachter [KP] to

$$ES(n) \leq \binom{2n-4}{n-2} + 7 - 2n.$$

Géza Tòth and Pavel Valtr [TV1] came next with

$$ES(n) \leq \binom{2n-5}{n-2} + 2.$$

These developments happened so swiftly, that all three above papers appeared in the same 1998 issue of *Discrete Computational Geometry*! In 2005, Tòth and Valtr came again [TV2] with the best known upper bound

$$ES(n) \leq \binom{2n-5}{n-2} + 1.$$

which is about half of the original Erdős–Szekeres upper bound.

Paul Erdős' trains of thought are infinite—they never end, and each problem solved gives birth to a new problem, or problems. The Happy End Problem is not an exception. Paul writes about the Aftermath of the Happy End Problem with his vintage humor and warmth [E83.03]:

Now there is the following variant which I noticed when I was once visiting the Szekereses in 1976 in Sydney, the following variant which is of some interest I think. It goes as follows. $n(k)$ is derived as follows, if it exists. It is the smallest integer with the following property. If you have $n(k)$ points in the plane, no three on a line, then you can always find a convex k-gon with the additional restriction that it doesn't contain a point in the interior. You know this goes beyond the theorem of Esther, I not only require that the k points should form a convex k-gon, I also require that this convex k-gon should contain none of the [given] points in its interior. And surprisingly enough this gives a lot of new difficulties. For example it is trivial that $n(4)$ is again 5, that is no problem. Because if you have a convex quadrilateral, if no point is inside we are happy; if from the five points one of them is inside you draw the diagonal AC (Fig. 31.1):

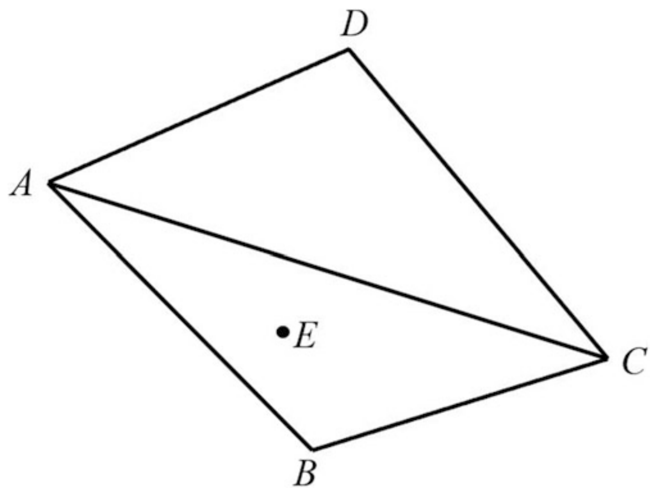


Fig. 31.1

> And you join these (*AE, EC*) and now this convex quadrilateral (*AECD*) contains none of the points. And if you have four points and the fifth point is inside then you take this quadrilateral. This is convex again and has no point in the inside. And Harborth proved that $n(5) = 10$. $f(5)$ was 9 in Esther Klein's problem but here $n(5)$ is 10. He dedicated his paper to my memory when I became an archeological discovery. When you are 65 you become an archeological discovery. Now, nobody has proved that $n(6)$ exists. That you can give, for every t, t points in the plane, no three on a line and such that every convex hexagon contains at least one of the points in its interior. It's perfectly possible that can [be done]. Now Harborth suggested that maybe $n(6)$ exists but $n(7)$ doesn't. Now I don't know the answer here.

Indeed, in 1978, Heiko Harborth [Harb] of Braunschweig Technical University, Germany, and a *Geombinatorics* editor, proved that $n(5) = 10$. In 1983, J. D. Horton [Hort] of the University of New Brunswick, Canada, proved Harborth's conjecture that $n(t)$ does not exist for $t \geq 7$. This left a mystifying gap:

Open Problem 31.18 Does $n(6)$ exist? If yes, find its value.

This new rich train of thought now includes many cars. I would like to share with you my favorite, the beautiful 2005 result by Adrian Dumitrescu of the University of Wisconsin-Milwaukee.

Dumitrescu's Theorem 31.19 [Dum].[7] For each finite sequence $h_0, h_1, \ldots, h_k$, with $h_i \geq 3$ ($i = 0, \ldots, k$), there is an integer $N = N(h_0, h_1, \ldots, h_k)$ such that any set S of at least N points in general position in the plane contains either an empty convex h_0-gon (i.e., a convex h_0-gon that contains no points of S in its interior) or k convex polygons $P_1, P_2, \ldots, P_k$, where P_i is an h_i-gon such that P_i strictly contains P_{i+1} in its interior for $i = 1, \ldots, k - 1$.

---

[7] Adrian mistakenly credits 1975 Erdős's paper with the birth of the problem about empty convex polygons. In the cited story Erdős clearly dates it to his 1976 visit of the Szekereses.

31.5 The Happy End Players Leave the Stage as Shakespearian Heroes

Paul Erdős named it *The Happy End Problem*. He explained the name often in his talks. On June 4, 1992, in Kalamazoo I took notes of his talk:

> I call it The Happy End Problem. Esther captured George, and they lived happily ever after in Australia. The poor things are even older than me.

This paper also convinced George Szekeres to become a mathematician. For Paul Erdős, the paper had a happy end too: it became one of his early mathematical gems and Paul's first of the numerous contributions to and the leadership of Ramsey Theory and, as Szekeres put it, of "a new world of combinatorial set theory and combinatorial geometry."

I always wanted to know the membership of the amazing 1930 Budapest group of young Jewish mathematicians. On May 28, 2000, during a dinner in the restaurant of the Rydges North Sydney Hotel,[8] I asked George Szekeres and Esther Klein to name the members of their group, so to speak the Choir of the Happy End Play. Esther produced – and signed "Esther Szekeres 28-5-00 SYDNEY" – the following list of the young participants, of which according to her "half a dozen usually met":

> Paul Erdős, Tibor Grünwald (Gallai), Géza Grünwald (Gergőr), Esther Klein (Szekeres), Lily Székely (Sag), George (György) Szekeres, Paul Turán, Martha Wachsberger (Svéd), and Endre Vázsonyi. Miklós Ság, and László Molnár occasionally joined the group too.

Esther Klein, Alexander Soifer, and George Szekeres, Sydney, Australia, May 28, 2000

[8] Esther wrote the list on the letterhead of the hotel, thus preserving the location of our meeting.

It was a warm, unforgettable conversation, full of reminiscences. George told me that his father was a cantor. He too was a musician: George played viola in orchestra and also played a violin.

George Szekeres told me that night “my student and I proved Esther’s Conjecture for 17 with the use of computer.” “Which computer did you use?” asked I. “I don’t care how pencil is made,” answered George.

The personages of The Happy End Problem appear to me like heroes of Shakespeare’s plays. Paul, very much like *Tempest’s* Prospero, gave up all his material possessions, including books, to be free. George and Esther were so close that they ended their lives together, like Romeo and Juliette. In the late summer 2005 e-mail, Tony Guttmann conveyed to the world the sad news from Adelaide:

> George and Esther Szekeres both died on Sunday morning [August 28, 2005]. George, 94, had been quite ill for the last 2–3 days, barely conscious, and died first. Esther, 95, died an hour later. George was one of the heroes of Australian mathematics, and, in her own way, Esther was one of the heroines.

Esther was not ill. She must have seen no sense in living without her lifelong love . . .

Chapter 32
The Man Behind the Theory: Frank Plumpton Ramsey

I harmony by algebra confirmed,
And only then, in science sophisticated,
Surrendered to the bliss of dream creative.
– Alexander Pushkin, *Mozart and Salieri*
(Translated from the Russian by Alexander Soifer for this book.)

Knowledge is a correspondence between idea and fact.
– Frank Plumpton Ramsey

32.1 Frank Plumpton Ramsey and the Origin of the Term "Ramsey Theory"

Who was "Ramsey," the man behind the theory named for him by others?

Let us start with the introduction to Ramsey's collected works [Ram3], assembled and edited right after his passing in 1930 by Ramsey's friend and disciple Richard Bevan Braithwaite, then Fellow of King's College and later the Knightbridge Professor of Philosophy at the University of Cambridge, who opens as follows:

> Frank Plumpton Ramsey was born on 22nd February 1903, and died on 19th January 1930 [a jaundice attack prompted by an unsuccessful surgery]. The son of the President of Magdalene [College], he spent nearly all his life in Cambridge, where he was successively Scholar at Trinity, Fellow at King's [at 21], and Lecturer in Mathematics in the University [at 23]. His death at the height of his powers deprives Cambridge of one of its intellectual glories and contemporary philosophy of one of its profoundest thinkers. Though mathematical teaching was Ramsey's profession, philosophy was his vocation.

The celebrated British philosopher, Cambridge "Professor of Mental Philosophy and Logic" and Fellow of Trinity College, George Edward Moore writes in the preface for Ramsey's book [Ram3]:

> He [Ramsey] was an extraordinarily clear thinker: no-one could avoid more easily than he the sort of confusions of thought to which even the best philosophers are liable, and

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_32

> he was capable of apprehending clearly and observing consistently, the subtlest distinctions. He had, moreover, an exceptional power of drawing conclusions from a complicated set of facts: he could see what followed from them all taken together, or at least what might follow, in cases where others could draw no conclusions whatsoever. And, with all this, he produced the impression of also possessing the soundest common sense: his subtlety and ingenuity did not lead him, as it seems to have led some philosophers, to deny obvious facts. He had, moreover, so it seemed to me, an excellent sense of proportion: he could see which problems were the most fundamental, and it was these in which he was most interested and which he was most anxious to solve. For all these reasons, and perhaps for others as well, I almost always felt, with regard to any subject that we discussed, that he understood it much better than I did, and where (as was often the case) he failed to convince me, I generally thought the probability was that he was right and I was wrong, and that my failure to agree with him was due to lack of mental power on my part.

Indeed, Ramsey's philosophical essays impress me immensely by their depth, clarity, and common sense – a combination that reminds me the great Michel de Montaigne. Here is my favorite quotation from Ramsey [Ram5, p. 53]:

> *Knowledge is a correspondence between idea and fact.*

Frank P. Ramsey's parents were Arthur Stanley Ramsey and Agnes Mary Wilson. In addition to Magdalene College's presidency, Arthur S. Ramsey was a tutor in mathematics. Frank was the oldest of four children. He had two sisters and a brother, Arthur Michael Ramsey, who much later became The Most Reverend Michael Ramsey, Archbishop of Canterbury (1961–1974). In 1925, Frank P. Ramsey married Lettice C. Baker, and their marriage produced two daughters. It is surprising to find in one family two brothers, Michael, the head of the Church of England and Frank, "a militant atheist," as Lettice described her husband.

The great economist John Maynard Keynes (1883–1946), then a Fellow at King's College and a close friend of Frank Ramsey, writes in March 1930 about Ramsey's contribution to economics [Key]:

> He [Ramsey] has left behind him in print (apart from his philosophical papers) only two witnesses of his powers – his papers published in the *Economic Journal* on "A Contribution to the Theory of Taxation" in March 1927, and on "A Mathematical Theory of Saving" in December 1928. The latter of these is, I think, one of the most remarkable contributions to mathematical economics ever made, both in respect of the intrinsic importance and difficulty of its subject, the power and elegance of the technical methods employed, and the clear purity of illumination with which the writer's mind is felt by the reader to play about its subject.

Keynes also draws for us a portrait of Ramsey the man (ibid):

> His bulky Johnsonian frame, his spontaneous gurgling laugh, the simplicity of feelings and reactions, half-alarming sometimes and occasionally almost cruel in their directness and literalness, his honesty of mind and heart, his modesty, and the amazing, easy efficiency of the intellectual machine which ground away behind his wide temples and

> broad, smiling face, have been taken from us at the height of their excellence and before their harvest of work and life could be gathered in.

This portrait reminds me Frank Plumpton's joking about his size while favoring human emotion over all issues of the universe (February 28, 1925):

> Where I seem to differ from some of my friends is in attaching little importance to physical size. I do not feel the least humble before the vastness of the heavens. The stars may be large, but they cannot think of love; and those are qualities which impress me far more than the size does. I take no credit for weighing nearly seventeen stones.

By kind permission of the Provost and Scholars of King's College, Cambridge, I am sharing with you here two photographs of the gentle giant, Frank Plumpton Ramsey. As Jacqueline Cox, Modern Archivist of King's College Library advises in her November 21, 1991, letter to me [Cox], "Both photographs come from the J. M. Keynes Papers (ref. JMK B/4). The first is a portrait of him aged 18 in 1921. The second shows him sitting on the ground in the open air reading a book aged 25 in 1928. The photographers are not indicated, but in the case of the second photograph a note records that it was taken in the Austrian Tyrol in August 1928."

Frank Plumpton Ramsey, aged 18. (Reproduced by kind permission of the Provost and Scholars of King's College, Cambridge)

Considering his short life, Ramsey produced an enormous amount of work in logic, foundations of mathematics, mathematics, probability, economics, decision theory, cognitive psychology, semantics, and of course philosophy. Ramsey manuscripts, held in the Hillman Library of the University of Pittsburg, fill 7 boxes and number about 1500 pages[1] [Ram5]. Probability fare is worthy of our attention. In his February 27, 1978, BBC radio broadcast (reprinted as an article [Mel] in 1995), Emeritus Professor of Philosophy at Cambridge D. H. Mellor explains:

> The economist John Maynard Keynes, to whom Braithwaite introduced Ramsey in 1921, published his *Treatise on* Probability in August of that year . . . It did not satisfy Ramsey, whose objections to it – some of them published before he was nineteen – were so cogent and comprehensible that Keynes himself abandoned it.

Frank Plumpton Ramsey, aged 25, Austrian Tyrol, August 1928. (Reproduced by kind permission of the Provost and Scholars of King's College, Cambridge)

[1]In *A Tribute to Frank P. Ramsey* [Har2], Frank Harary writes: "At her home, she [Mrs. Lettice Ramsey, the widow] showed me box upon box of notes and papers of Frank Ramsey and invited me to pore through them. As they dealt mostly with philosophy, I had to decline." As "a tribute," could Professor Harary have shown more interest and curiosity?

In fact, my friend, the Princeton Double Professor Emeritus of both Mathematics and of Economics Harold W. Kuhn tells me that Keynes decided against continuing with mathematics because Ramsey was so much superior in it. Mellor continues:

> In this paper [Ram4], after criticizing Keynes, Ramsey went on to produce his own theory. This starts from the fact that people's actions are largely determined by what they believe and what they desire – and by strength of those beliefs and desires. The strength of people's beliefs is measured by the so-called 'subjective probability' they attach to events. . . Subjective utility measures the strength of people's desires just as subjective probability measures the strength of their beliefs.
>
> The problem is how to separate these two components of people's actions ... One of the things Ramsey's paper did was to show how to extract people's subjective utilities and probabilities from the choices they make between different gambles; and by doing so it laid the foundations for the serious use of these concepts in economics and statistics as well as in philosophy.
>
> It took a long time, however, from this 1926 paper of Ramsey's to bear fruit. Only after the publication in 1944 of a now classic book by John von Neumann and Oskar Morgenstern, *The Theory of Games and Economic Behavior* [NM], did utility theory begin to catch on and be applied in modern decision theory and games theory. And for many years no one realized how much of it had been anticipated in Ramsey's 1926 paper.

I am looking at the first 1944 edition of the classic [NM] that Mellor mentions above, written by two celebrated Institute for Advanced Study and Princeton University people, John von Neumann (1903–1957) and Oskar Morgenstern (1902–1977), respectively, and its later editions (Fine Library of Princeton-Math is very fine). The authors cite in their book many colleagues, classics of the past and contemporaries: Daniel Bernoulli, Dedekind, Kronecker, D. Hilbert, F. Hausdorff, E. Zermelo, G. Birkhoff, E. Borel, W. Burnside, C. Carathéodory, W. Heisenberg, A. Speiser – even Euclid. One name is missing that merits credit the most, that of Frank P. Ramsey. Harold W. Kuhn tells me that in a 1953 letter, he asked von Neumann why the latter gave no credit to Ramsey for inventing subjective probability. Indeed, this question and von Neumann's answer are reflected in H. W. Kuhn and A. W. Tucker's 1958 memorial article about John von Neumann [KT, pp. 107–108]:

> Interest in this problem as posed [measuring "moral worth" of money] was first shown by F. P. Ramsey [Ram4] who went beyond Bernoulli in that he defined utility operationally in terms of individual behavior. (Once von Neumann was asked [by H. W. Kuhn] why he did not refer to the work of Ramsey, which might have been known to someone conversant with the field of logic. He replied that after Gödel published his papers on undecidability and the incompleteness of logic, he did not read another paper in symbolic logic.)[2]

[2] Indeed, von Neumann and Morgenstern probably did not expect Ramsey to publish on a topic far away from the foundations, such as economics, and thus might not have known about Ramsey's pioneering work by the time of the first 1944 edition of their celebrated book. However, in 1953 if not earlier they learned about Ramsey's contribution from Harold W. Kuhn, and yet new editions, which came out in 1947, 1953, 1961, etc., did not give Ramsey credit either.

Ramsey's priority was discovered and acknowledged in print by others. In already mentioned D. H. Mellor's broadcast, the philosopher of probability Richard Carl Jeffrey (1926–2002; Ph.D. Princeton 1957; Professor of Philosophy at Princeton 1974–1999) says:

> It was when Leonard Savage, statistician, was working on his book on subjective probability theory, and he wished to find out what if anything the philosophers had to say on the subject, he went to Ramsey article [Ram4] and read it, and he found that what he [Ramsey] had done was to a great extend fairly describable as rediscovering another aspect of Ramsey's work in that article – the foundations of the theory of subjective probability. It was Savage's book, *The Foundations of Statistics*, that was published in 1954, that made subjectivism a respectable sort of doctrine for serious statistician to maintain; and the remarkable thing is that Ramsey in this little paper to the Moral Sciences Club in 1926 has done all of that already.

Indeed, Leonard Jimmie Savage (1917–1971) writes in 1954 [Sav, pp. 96–97]:

> Ramsey improves on Bernoulli in that he defines utility operationally in terms of the behavior of a person constrained by certain postulates...
>
> Why should not the range, the variance, and the skewness, not to mention countless other features, of the distribution of some function join with the expected value in determining preference? The question was answered by the construction of Ramsey and again by that of von Neumann and Morgenstern.

Richard Jeffrey writes [Jef, p. 35]:

> This method of measurement [of desirability] was discovered by F. P. Ramsey and rediscovered by von Neumann and Morgenstern, through whose work it came to play its current role in economics and statistics.

More importantly, much of Jeffrey's 1965 book *The Logic of Decision* [Jef] is based on Ramsey's ideas, while Chapter 3 is simply called *Ramsey's Theory.*

Ramsey's first mathematical paper, *Mathematical Logic* [Ram1] appeared in 1926, in the midst of the *Grundlagenstreit* (Crisis in the Foundations), the confrontation between the two giants, David Hilbert and L. E. J. Brouwer, over the foundations of mathematics. Ramsey, who always addresses the most important issues of his day, does not shy away from this one either. He does not, however, take either side. Ramsey does not agree with the intuitionist approach:

> Weyl has changed his view and become a follower of Brouwer, the leader of what is called the intuitionist school, whose chief doctrine is the denial of the Law of Excluded Middle, that every proposition is either true or false. This is denied apparently because it is thought impossible to know such a thing *a priori*, and equally impossible to know it by experience... *Brouwer would refuse to agree that either it was raining or it was not raining, unless he had looked to see.*

Neither does Ramsey support Hilbert:

> I must say something of the system of Hilbert and his followers, which is designed to put an end to such skepticism once and for all. This is to be done by regarding higher mathematics as the manipulation of meaningless symbols according to fixed rules. We

start with certain symbols called axioms: from these we can derive others by substituting certain symbols called constants for others called variables, and by proceeding from the pair of formulae p, if p then q to the formula q.

Mathematics proper is thus regarded as a sort of game, played with meaningless marks on paper rather like noughts and crosses; but besides this there will be another subject called metamathematics, which is not meaningless, but consists of real assertions about mathematics, telling us what this or that formula can or cannot be obtained from the axioms according to the rules of deduction . . .

Now, whatever else a mathematician is doing, he is certainly making marks on paper, and so this point of view consists of nothing but the truth; but it is hard to suppose it is the whole truth. There must be some reason for the choice of axioms, and some reason why the particular mark $0 \neq 0$ is regarded with such abhorrence. This last point can, however, be explained by the fact that the axioms would allow anything whatever to be deduced from $0 \neq 0$, so that if $0 \neq 0$ could be proved, anything whatever could be proved, which would end the game for ever, which would be very boring for posterity. Again, it may be asked whether it is really possible to prove that the axioms do not lead to contradiction, since nothing can be proved unless some principles are taken for granted and assumed not to lead to contradiction.

Summing up both Hilbert and Brouwer–Weyl approaches, Ramsey concludes:

We see then that these authorities, great as they are the differences between them, are agreed that mathematical analysis as originally taught cannot be regarded as a body of truth, but is either false or at best a meaningless game with marks on paper.

What then was a mathematician to do? Ramsey was in favor of using the Axiom of Infinity. "As to how to carry the matter further, I have no suggestion to make; all I hope is to have made it clear that the subject is very difficult," wrote Ramsey in the end. Four years later Ramsey would take a finitist view of rejecting the existence of any actual infinity.

Ramsey comes back with a specific approach in his second mathematical paper *On a Problem of Formal Logic* [Ram2], submitted on November 28, 1928 and published posthumously in 1930. This paper gives a clear and unambiguous start to what was later named *Ramsey Theory.* What is the aim of this work? Fortunately, Ramsey answers this question right at the start of the paper:

This paper is primarily concerned with a special case of one of the leading problems of mathematical logic, the problem of finding a regular procedure to determine the truth or falsity of any given logical formula. But in the course of this investigation it is necessary to use certain theorems on combinations which have an independent interest and are most conveniently set out by themselves beforehand.

Indeed, Ramsey solves the problem in the special case, as he promises. However, little does he – or for that matter anyone else – expect that the next year, in 1931, another young genius, the 25-year-old Kurt Gödel will shake the mathematical world by publishing the (Second) Incompleteness Theorem [Göd1] that shows that Hilbert–Ackermann's *Entscheidungsproblem*, "the leading problem of mathematical logic" as Ramsey calls it, cannot have a solution in general case. Ramsey continues:

> The theorems which we actually require concern finite classes only, but we shall begin with a similar theorem about infinite classes which is easier to prove and gives a simple example of the method of argument.

Yes, the infinite case here – as happens often – is easier than the finite one but is very well worth of the presentation (in fact, the finite case follows from the infinite by the de Bruijn–Erdős Compactness Theorem, as we have seen in Chapter 28). Later in the paper, Ramsey also observes that his infinite case requires the use of the Axiom of Choice:

> Whenever universe is infinite, we shall have to assume the axiom of selection.

In fact, some 40 years later, in 1969, Eugene M. Kleinberg [Kle] will prove that Ramsey's Theorem is independent from *ZF*, the Zermelo–Fraenkel Set Theory. (More precisely, if *ZF* is consistent, then Ramsey's theorem is not provable in *ZF*.)

Ramsey realizes – and clearly states – that his new pioneering method and his "theorems on combinations have an independent interest." Indeed, Ramsey's theorems deliver the principles and the foundation to the new field of mathematics, *Ramsey Theory*. Now, this requires a clarification.

Three Ramsey Theory results appeared before Frank P. Ramsey erected its foundation, and this is why I combine these three early results under the name *Ramsey Theory before Ramsey*. They are the Hilbert Theorem of 1892, the Schur Theorem of 1916, and the Baudet–Schur–Van der Waerden Theorem of 1927. These classic results, which we will discuss in great detail in the next Part, discovered particular properties of colored integers or colored spaces in particular circumstances. These theorems constituted a real "meat" of Ramsey Theory, real applications of the Ramsey Principle in particular contexts before Ramsey formulated it.

Ramsey's amazing logical-philosophical gift allowed him to abstract the idea from any particular context, to formulate his theorems as a *method*, a *principle* of the new theory – a great achievement indeed. Surely, Ramsey fully deserves his name to be placed on the new theory, whose principle he so clearly formulated and proved, but could anyone point out who and when coined the term *Ramsey Theory*?

32.2 What's in a Name? That Which We Call a Rose by Any Other Name Would Smell as Sweet

Yes, Shakespeare believes that the name does not matter. With greatest admiration, I beg to disagree with the Bard and undertake research into the authorship of *Ramsey Theory* name. We have already seen *Ramsey's Theory of Decision* in Richard Jeffrey's 1965 book [Jef]. But we are after *Ramsey Theory*, a new and flourishing branch of combinatorial mathematics. On 21 July 1995, I asked a leader of Ramsey Theory, Ronald L. Graham. Here is our brief exchange of the day:

> Dear Ron:
> Who and when coined the name "Ramsey Theory"?
> Yours, Sasha

> Sasha,
> Beats me! Who first used the term Galois theory?

Ron

On January 22, 1996, I asked Ron again and received another concise reply the same day:

Dear Sasha,
I would imagine that Motzkin may have used the term Ramsey Theory in the 60's. You might check with Bruce Rothschild at UCLA who should know.

Still the same day I received a reply from Bruce Rothschild:

Dear Alexander,
This is a good question, to which I have no real answer. I do not recall Motzkin using the phrase,[3] though he might have. I also don't recall hearing Rota use it when I was at MIT in the late 60's. My best recollection is that I began using the term informally along with Ron sometime in the very early '70's . . . But I could be way off here.

Frank Harary was less concise. On February 19, 1996, during the Southeastern International Conference on Combinatorics, Graph Theory, and Computing in Baton Rouge, Louisiana, he gave me a multipage statement (you saw it in Chapter 29), suggesting that Frank Harary and Václav Chvátal were first to introduce the term *Generalized Ramsey Theory for Graphs* in their series of papers that started in 1972.

I am looking at the first paper [CF] of the series: Chvátal, Václav, and Harary, Frank, *Generalized Ramsey theory for graphs.* The authors generalize the notion of *Ramsey Number* by including it in the study graphs other than complete graphs. By doing so, Harary and Chvátal open a new, now flourishing chapter, *Graph Ramsey Theory*. However, *Ramsey Theory* as we understand it today stands for so much broader a body of knowledge, including the Schur, the Baudet–Schur–Van der Waerden, and the Hales–Jewett Theorems, that it does not fit inside Graph Theory. Thus, my search for the true birth of the name continued.

One 1971 survey [GR2], by Ronald L. Graham and Bruce L. Rothschild, shows a clear realization that a new theory has been born and needs an appropriate new name. Following a recitation of the Ramsey Theorem and the Schur Theorem, the authors write:

> These two theorems are typical of what we shall call a Ramsey theorem and a Schur theorem, respectively. In this paper we will survey a number of more general Ramsey and Schur theorems which have appeared in the past 40 years. It will be seen that quite a few of these results are rather closely related, e.g., van der Waerden's theorem on arithmetic progressions [Wae2], [Khi4], Rado's work on regularity and systems of linear equations [Rad1], [Rad2], the results of Hales and Jewett [HJ] and others [Garsia, personal communication] on arrays of points and Rota's conjectured analogue of Ramsey's Theorem for finite vector spaces, as well as the original theorems of Ramsey and Schur.

Yes, I agree, the new theory has been born by 1971, and the choice of its name was between two well-deserving candidates: *Schur Theory,* in honor of the main early contributor Issai Schur and his School (Schur's work was continued by his students Alfred Brauer and Richard Rado); and *Ramsey Theory,* in honor of Frank P. Ramsey who formulated the

[3] Motzkin did not use "Ramsey Theory" in his 1960s articles, as I have verified shortly after.

principles of the new theory. Soon Graham and Rothschild make the decision, and in their 1974 survey publish the first announcement of their choice [GR3]:

> Recently a number of striking new results have been proved in an area becoming known as RAMSEY THEORY. It is our purpose here to describe some of these. Ramsey Theory is a part of combinatorial mathematics dealing with assertions of a certain type, which we will indicate below. Among the earliest theorems of this type are RAMSEY's theorem, of course, VAN DER WAERDEN's theorem on arithmetic progressions and SCHUR's theorem on solutions of $x + y = z$.

It seems that *Ramsey Theory* has been suddenly and rapidly shaping through the 1970s, and the central engine of this process was new results and the above-mentioned surveys. In 1980, the long life of the name was assured when it appeared as the title of the book *Ramsey Theory* [GRS1] by three of the leading researchers of the field, Ronald L. Graham, Bruce L. Rothschild, and Joel H. Spencer. A decade later, the authors produced the second, updated edition [GRS2]. This book not only assured the acceptance of the name but it has also become the standard text on the new field of mathematics.

And now is the time to share a bit of information about the co-creators of the term *Ramsey Theory*, who of course contributed much-much more to the field than just its name.

32.3 Bruce Lee Rothschild

Bruce L. Rothschild was born on August 26, 1941, in Los Angeles. Following BS degree from the California Institute of Technology in 1963, he earned PhD degree from Yale in 1967 with the thesis *A Generalization of Ramsey's Theorem and a Conjecture of Rota*, supervised by the legendary Norwegian graph theorist Øystein Ore (1899–1968). After two years 1967–1969 at MIT, Rothchild became a professor at the University of California, Los Angeles, where he worked for many decades. In 1972, Graham, Rothschild, and Leeb shared the Polya Prize of SIAM with Hales and Jewett. Professor Rothschild's papers, many joint with Graham, made a major contribution to Ramsey Theory.

32.4 Remembering Ronald Lewis Graham

(October 31,1935–July 6, 2020)[4]

> *When Paul Erdős left this world in 1996, Ron became our universally admired Captain. Now I do not see anyone who could fill these shoes. I feel like and orphan in Ramsey Theory.*
>
> *Ron touched uncountably many people, and deserves countless remembrances. This is one of them, my personal brief tribute.*

[4] First published in [Soi48].

The Royal Couple of Mathematics, Fan Chung and Ronald L. Graham, June 4, 1992, Kalamazoo, MI. (Photo by Alexander Soifer)

My phone rang, “This is Ron Graham, can I talk to Paul?” This was my first encounter with Ron. He called Paul Erdős every night, so that Ron could tell numerous enquirers where Paul was on any given day. Love and admiration for Paul united Ron and me from the start. Once, after chatting with Paul, Ron asked for me: “Why don’t you come to the conference at Florida Atlantic University in Boca Raton, Florida? We’ll finally meet in person.” As so we did. The time stamp: February 1992.

Ron has been the most influential mathematician of the United States for many decades. He served as President of the American Mathematical Society, President of the Mathematical Association of America, member of the American Academy of Arts and Sciences, member of the Hungarian Academy of Sciences, Chief Scientist of Bell Labs, Endowed Chair at the University of California San Diego, etc. Ron was not only a member of the National Academy of Sciences – he served two terms as its Treasurer. Ron showed me once his impressive Treasurer’s office at the Academy, across a narrow street from the U.S. Department of State. All these accolades did not spoil Ron’s personality one bit. He has always been open, friendly, curious, generous to a fault, sprinkling conversations with lovely humor.

I do not know whether young Ron competed in mathematical Olympiads. I do remember him serving as the Chair of the Jury of the 2001 International Mathematical Olympiad in

Washington, D.C. (I was a coordinator). For representatives of ca. 90 countries – for this virtually herd of cats – Ron was an easy going, humorous, yet principled shepherd.

Paul Erdős "preached" (his term) to professionals, young and old. His open problems inspired generations. But he was not a natural lecturer, whereas Ron certainly was. I attended many of Ron's addresses, always with paper and pencil. His lectures were well composed – as pieces of music – elegant, deep, inspiring, yet lightened by humor. Ron excited every audience, from professional mathematicians to high school students, who in 2002 came to MIT to compete in the USA Mathematical Olympiad. In 2014, Ron gave a brilliant keynote address at the Congress of World Federation of National Mathematics Competitions (WFNMC) in Barranquilla, Colombia. All his audiences got their share of intriguing open problems and conjectures, some with cash prizes for first solutions.

WFNMC Keynote speakers Ronald Graham and Alexander Soifer, Barranquilla, July 22, 2014

Center for Discrete Mathematics and Theoretical Computer Science, DIMACS, is a joint venture of universities, such as Princeton, Rutgers, Columbia, and industrial research giants AT&T Labs – Research, Avaya Labs, IBM Research, Microsoft Research, NEC Laboratories America, Nokia Bell Labs, and Perspecta Labs. For three years between 2002 and 2007, I was a long-term visiting scholar at DIMACS and a visiting fellow at Princeton. I left Princeton and Rutgers in 2007, I thought forever. However, a year later, Executive Director of DIMACS Fred Roberts invited me back to create an international 3-day workshop on Ramsey Theory.

He offered plenty of money to my discretion (which I used to pay plenary speakers' travel and registration for graduate students). I accepted, on the condition that Ron Graham would be one of my plenary speakers. Ron agreed, and we had an incredible workshop that I titled *Ramsey Theory: Yesterday, Today, and Tomorrow*, where everyone attended every talk. Springer commemorated the workshop by publishing a book of talks [Soi49] under the same title in its prestigious series *Progress in Mathematics*.

Ron's support of colleagues was most inspirational. In 2009, Ron traveled coast to coast, from California to Florida with a copy of the original edition of *The Mathematical Coloring Book*, a heavy book I should add. He asked me to inscribe his copy. *To our Captain Ron Graham*, I wrote on the title page. Ron signed for me a poster created by his daughter and started our conversation with a question I myself thought about for a while:

– Can you write another book like this one?
– No, Ron.
– I did not think you could: so much blood, sweat, and tears. I bought three copies of it.
– Why, Ron?
– I have three offices, and I want to have your book handy wherever I am.

Ron was incredibly modest. As you know, in my research for *The Mathematical Coloring Book*, I wanted to establish who gave the new 100-year-old theory its name, *Ramsey Theory*. The famous graph theorist Frank Harary readily stated to me that he was the author. Ron Graham and Bruce L. Rothschild each wrote to me that they had no idea who coined the term. Imagine, I determined that Ron and Bruce in their 1970s joint papers in fact coined the name of this exciting new area of mathematics, *Ramsey Theory*!

I asked Ron many times to write his autobiography for the new edition of *The Mathematical Coloring Book.* I was sure his reminiscences about his first interest in mathematics, his numerous achievements, and leadership in a number of fields would inspire mathematicians young and old. In addition, I admired his elegant humorous writing and lecturing style. Ron refused every time. Here is Ron's 3 July 2019 reply to my request:

Hi Sasha,
I am too modest for such a profile!
Ron

Ron's interests were not limited to mathematics and computer science. He was a fine gymnast, ping pong player, and professional-level juggler. His passion for sports was clear when he went to Canada to watch Winter Olympics in person. In 2014, when we both were keynote speakers in Colombia, I recommended Ron to watch *Leviathan*, a cinematic masterpiece directed by my Moscow friend Andrey Zvyagintsev. Ron enjoyed the film and followed its fate through various competitions (the film won the Golden Globe and narrowly lost Oscar). Knowing that *Leviathan* painted a very grim picture of the present Russia, Ron asked me on August 27, 2016, "By the way, how is *Leviathan* now received in Russia?"

As you know, Paul Erdős offered cash prizes for first solutions of his open problems. These ranged from \$10 up to several thousand dollars for problems that were both difficult and mathematically significant. When in 1996 Paul "left," Ron offered to pay on behalf of Paul for many solutions, as he explained on February 13, 2007:

Hi Sasha,

I am willing to pay all the prizes offered by Paul that are listed in the book that Fan and I wrote: *Erdős on Graphs: His Legacy of Unsolved Problems*. These we have checked. The others (e.g., in number theory or set theory) are not (automatically) part of the offer . . .

Best regards,
Ron Graham

Ron presented his personal checks for a good number of great achievements; $500 to Saharon Shelah and $1000 to Timothy Gowers come to mind.

Ron presented a $1000 check to Aubrey de Grey for the construction of the first 5-chromatic unit-distance graph – what should Ron "buy" next, I was thinking. I felt that I knew Ron's style of problem posing and had a sense of his intuition. I also knew that we were friends, that Ron with his great sense of humor would not get angry at me by announcing *Ron's New $1000 Open Problem* without clearing it with him first.

And so I did in my March 14, 2019 talk at Florida Atlantic University, with a disclaimer "subject to Ron's approval." My audience filled the room with laughter. Immediately after the conference, on March 16, 2019, I sent Ron an e-mail:

Dear Ron,

Unfortunately, I did not see you at my talk in Boca, where I premiered *your new $1000 problem*, of course, subject to your approval. :) The audience loved it.

The New Graham's $1,000 Problem. Prove or disprove the existence of a unit-distance 6-chromatic graph.

So . . . please, reply with a yes, or a no, or your different related open problem(s) – for the inclusion in the second expanded edition of *The Coloring Book*.

Yours always,
Sasha

Ron replied the same day:

Hi Sasha,

I had to check out by 11 so unfortunately I couldn't make your talk! :(

I approve of the new $1000 problem!

Ron

In the time of pandemic, I was concerned about Ron's and Fan's health. My last inquiry on June 14, 2020, remained unanswered . . . Ron "left" on July 6, 2020.

Life Is a Fatal Sexually Transmitted Disease, so named his film the Polish director Krzysztof Zanussi. Yes, fatal. . . The Giants of *Geombinatorics'* Editorial Board are leaving the stage, Erdős, Grünbaum, Graham. I will keep them on our Editorial Board as Emeriti. Forever.

* * *

This certainly does not cover Ron's excellence in juggling ("juggling is a metaphor," he used to say), fluency in Mandarin, close friendship with Paul Erdős, etc. See all those on "Ronald Graham's special page" created by Ron's wife, coauthor, and well-known mathematician in her own rights Fan Chung at http://math.ucsd.edu/~fan/ron/.

32.5 Reflections on Ramsey and Economics, by Harold W. Kuhn

In the fall 2006, upon my return to Princeton University, I asked myself, who could best evaluate Frank P. Ramsey's works on economics? It would take an expert in mathematics and economics. I chose two great Princeton scholars to write their views for this book: John Forbes Nash, Jr. and Harold William Kuhn. They both agreed to write. A week later, during Princeton-Math coffee hour conversation with me, John Nash deferred the task to Harold:

– I am sorry, but I do not think I could write about economics of Ramsey.
– Why not, John?
– I am not an economist.
– Let me try to understand: You are a Nobel Laureate for Economics but not an economist?
– I am a mathematician. I can study economics of Ramsey for you, but I am not sure I would be able to write something meaningful about it. On the other hand, Harold will do a fine job.
– Of course, John.

The closest friend of my Princeton years, Harold Kuhn, already knew from John Nash that the "jingle" was his. "Give me two months, and I will write for your book an essay on economics of Frank Ramsey," said Harold to me.

John F. Nash, Jr. and Alexander Soifer at Princeton-Math, July 2007

Why do sometimes people work together for decades and remain strangers, and other times friendship arrives at the instant of first sight? This is a question for psychologists to ponder. Harold W. Kuhn and I became instant friends in early 2003 when I arrived in Princeton-Math., just as in 1988 an instant friendship linked Paul Erdős and me. It has always been intellectually stimulating to discuss with Harold any subject, and we shared passion for many, from mathematics to the cinema of Michelangelo Antonioni, from the scholars in the Third Reich to the Princeton-Math historic personalities, from the Fang Art of Equatorial Africa to Pierre Bonnard's drawings. Wheelchair-bound, Harold wrote to me on July 20, 2013, email, "I do not want to write anything like a foreword now." To cheer him up, I replied immediately, "That's alright. Maybe after supper? :-)" In a week, Harold sent his foreword for my book "The Scholar and the State" [Soi47] to Springer and to me. It was his last essay. My dear friend Harold passed away peacefully on July 2, 2014, in his Manhattan condo with his family by his side.

32.5.1 Harold William Kuhn

Harold W. Kuhn was born in Santa Monica, California on July 29, 1925. Following BS degree in 1947 from the California Institute of Technology, he earned PhD degree from Princeton University in 1950, while also serving as Henry B. Fine Instructor in the Mathematics Department of Princeton, 1949–1950. Following a professorship at Bryn Mawr, 1952–1958, Harold has been a Professor of Mathematical Economics at Princeton's two departments, Mathematics and Economics, becoming Emeritus in 1995. His honors include presidency of the Society for Industrial and Applied Mathematics (1954–1955), service as Executive Secretary of the Division of Mathematics of the National Research Council (1957–1960), John von Neumann Theory Prize of the Operation Research Society of America (1982; jointly with David Gale and A. W. Tucker), and Guggenheim Fellowship (1982). It was Harold Kuhn who nominated John F. Nash Jr. for the Nobel Prize (awarded in 1994) and presided over the Nash Nobel Prize Seminar in Stockholm.

From the beginning, Harold decided to not just contribute an essay about the economics of *Frank P. Ramsey* but also touch on *John von Neumann* and *John F. Nash Jr.*, both of whom he knew personally very well. In all that follows in this chapter, the podium – shall I say, the pages – belongs to Harold W. Kuhn. I join my dearest late friend Harold W. Kuhn (July 1925–July 2, 2014) in thanking Sylvia Nasar, the author of *A Beautiful Mind*, for meeting with Harold over a Manhattan lunch and copy-editing this treasure of an essay. "Writing this essay is the most important thing I did this year," wrote Harold Kuhn in his Christmas 2007 report to friends and family. I hope you will enjoy this remarkable, deep, and supremely informed Triptych as much as I have! ■

Harold W. Kuhn is passing to Alexander Soifer his Triptych, Princeton, July 2007. (Photo by John Morgan, Professor of Mathematics at Columbia University)

Although mathematics became the lingua franca of twentieth century economics, only a handful of mathematicians have exerted a direct and lasting influence on the subject. They surely include Frank Plumpton Ramsey, John von Neumann, and John Forbes Nash Jr. The similarities and differences in their life trajectories are striking. Ramsey died at 26 years of age after an exploratory liver operation following a bout of jaundice, while Nash's most productive period ended when he fell prey to schizophrenia at the age of 30. Von Neumann's original work on game theory and growth models was done before he was 30 years old. For all three, the work in economics appears as a sideline. Ramsey's friend and biographer, Richard Braithwaite wrote: "Though mathematical teaching was Ramsey's profession, philosophy was his vocation," without mentioning his contributions to economics at all or including the three papers on economics in the posthumous "complete" works that Braithwaite edited. The contributions of von Neumann to mathematical economics are but one chapter in the seven chapters comprising the memorial issue devoted to von Neumann's research and published as a special issue of the Bulletin of the American Mathematical Society. Regarding Nash, John Milnor considered "... Nash's [Nobel Economics] prize work [to be] an ingenious but not surprising application of well-known methods, while his subsequent mathematical work was much more rich and important."

Ramsey, von Neumann, and Nash came from very different backgrounds and had very different relationships to the economics and the economists of their day. Ramsey, an intimate friend of Bertrand Russell and Wittgenstein, was a Cambridge man by birth. He appears to have been interested in economics from the age of 16 and wrote his first published piece on economics at 18. He had close personal and professional contacts with such well-known economists as John Maynard Keynes, Arthur Pigou, Piero Sraffa, and Roy Harrod. He served as an advisor to the Economic Journal, where Keynes took his counsel most seriously. He was well acquainted with the trends in the economic theory of his day.

Von Neumann, the scion of a Jewish banking family in Budapest, had a wide circle of intellectual friends from Budapest, Berlin, and Vienna that included economists such as William Fellner (who was a friend from gymnasium days) and Lord Nicholas Kaldor, who gave von Neumann a reading list in contemporary economics in the 20s, and who arranged for an English translation of von Neumann's growth model to be published in the Review of Economic Studies in 1945. Thus, there is ample evidence that von Neumann was well informed of the state of economics throughout his life.

The case of John Nash, who grew up in the coal mining and railroad town of Bluefield, West Virginia, is very different. When he came to Princeton to do graduate work in mathematics at the age of 20, he had taken one undergraduate course in economics (on International trade) at Carnegie Tech, taught by an Austrian émigré, Bert Hoselitz. His major contribution on bargaining, which appears to have had its origin in this course, has two boys (Bill and Jack) trading objects such as a whip, a bat, a ball, and a knife. This was the work of a teenager. There is no evidence that Nash had read any contemporary economist outside the required readings of his one undergraduate course. Of course, later in his life, in the period when he was on the faculty at the Massachusetts Institute of Technology, he had contact with Paul Samuelson and Robert Solow, Nobel Prize winners in Economics, who knew of his work in game theory. Nash's only later excursion into economics is a theory of "ideal money," an idea that appears to have been anticipated in part by Friederich Hayek.

Now that game theory has become part of the economist's tool kit, anyone who takes an introductory economics course learns about the contributions of von Neumann and Nash. Ramsey's work, however, is less well known and the principal reason for this note is to give the reader an appreciation for the contributions of Ramsey to economics. Between the ages of 18 and 27, Ramsey wrote four papers, which we shall discuss in detail below.

(A) "The Douglas Proposals," *The Cambridge Magazine*, Vol. XI, No. 1, January 1922, pp. 74–76.

Ramsey's first work related to economics (A) was published when he was 18. He was no common 18-year-old; here is how Keynes described him: "From a very early age, about sixteen I think, his precocious mind was intensely interested in economic problems." The Cambridge Magazine was edited by C. K. Ogden, a Fellow of Magdalene College where Ramsey's father was President, from 1912 to 1922. Ramsey and Ogden met while Ramsey was still a student in his public school, Winchester, and Ogden persuaded him to study the then much-discussed social credit proposals of a certain Major Douglas. I. A. Richards recalled the upshot: "Soon after he'd done the Douglas credit thing, you know, A. S. Ramsey, his father, called up Ogden and said 'What have you been doing to Frank?', and Ogden said 'What's he been doing?'. 'Oh he's written a paper on Douglas Credit which would have won him a Fellowship in any University anywhere in the world instantly. It's a new branch of mathematics.'"

Who was this Major Douglas? Briefly, he was one of those crackpots who exist on the fringe of academic economics and whose theories promise a redistribution of wealth that appealed to a large part of the public (including, in Douglas's case, Ezra Pound and T. S. Eliot). Like many of those offering a panacea for the Great Depression, he was also an anti-Semite who invoked the theses expounded in the Protocols of the Elders of Zion in defense of his economic theories.

What was Major Douglas's heresy that Ramsey demolished? It is centered on the so-called A + B "theorem" (called by Keynes "mere mystification"). In producing a good, price is made up of two parts of the cost paid out by the producer: A equals the amount paid out for raw materials and overhead and B equals the sums paid out in wages, salaries, and dividends. According to Douglas, the amount B, paid to the consumers, is never sufficient to buy all of the good, whose cost (and price) is A + B. Therefore, the state should make up the difference through "social credit."

Ramsey first provides a verbal argument that shows that, in a stationary state, the total rate of distribution of purchasing power (taking into account payments originating in intermediate goods) equals the rate of flow of costs of consumable goods. He then writes:

> ". . . it is possible, using some complicated mathematics to show that the ratio is unity under much wider conditions which allow for changes in the quantity of production, in the rate of wages, in the productivity of labour, and in the national wealth." The "complicated mathematics", other than Ramsey's curiously rigid set of modeling assumptions, consists of the use of "integration by parts," a technique taught to every beginning student of the calculus.

(B) "A Contribution to the Theory of Taxation," *The Economic Journal*, Vol. XXXVII, March 1927, pp. 47–61.

The young Ramsey assisted A. C. Pigou, who was the successor to Alfred Marshall in the chair of Political Economy at Cambridge, on a number of occasions beginning before 1926. After providing Pigou with a mathematical proposition and examples for two articles, one on credit and one on unemployment, Ramsey assisted Pigou with changes in the third edition of *The Economics of Welfare*, published in 1929. However, it appears that Ramsey's work on taxation (B) was inspired by questions raised in Pigou's *A Study in Public Finance*.

The problem posed by Ramsey in (B) was to find an optimal system of taxation of commodities so as to raise a given quantity of revenue. For Ramsey in (B), "optimal" means minimizing aggregate sacrifice. Using this objective function, he shows that the production of each commodity should be reduced in the same proportion, thus a system of differential taxation. The mathematics employed is rather standard, namely, optimization under equality constraints using Lagrange multipliers which was taught to mathematicians of this period by treatises such as de la Vallee Poussin's *Cours d'Analyse*. The treatment is careful for the period, and Ramsey includes a number of examples of potential applications of his results. Of particular interest is a discussion of the application of income tax to savings, a subject that I believe was part of a larger research agenda that Ramsey had formulated.

(C) "A Mathematical Theory of Saving," *The Economic Journal,* Vol. XXXVIII, December 1928, pp. 543–549.

Papers (B) and (C) were published in the Economic Journal which Keynes controlled with an iron hand. Keynes wrote of (C) that it "is, I think, one of the most remarkable contributions to mathematical economics ever made, both in respect of the intrinsic importance and difficulty of its subject, the power and elegance of the technical methods employed, and the clear purity of illumination with which the writer's mind is felt by the reader to play about its subject." The article (C) is concerned with the derivation of optimal saving programs under a variety of conditions. Samuelson captures the spirit of the paper in the society in which it was created when he wrote: "Frank Ramsey, living in a happier age and being a Cambridge philosopher assumed society would last forever and seek to maximize the utility of its consumption over all infinite time." A major stumbling block immediately presents itself in that the "utility of its consumption over all infinite time" is an improper integral which, in general, will not have a finite maximum value. Ramsey proposed an elegant device to get around this problem. He assumed that there was a maximum amount of attainable utility (called "bliss"), and instead of maximizing the improper integral, he minimized the deviation from bliss over the infinite horizon.

Ramsey then derives a result that is easy to express in common English, namely: "The optimal rate of saving multiplied by the marginal utility of consumption should always equal the difference between bliss and the actual rate of utility enjoyed." The paper contains a derivation of this result by simple verbal reasoning provided by Keynes (which does not apply to the most general cases considered by Ramsey but which does give the non-mathematically adept the feeling of "understanding the result"). Contemporary mathematical economists will instantly recognize the problem as one to which the calculus of variations applies and, indeed, over 30 years after Ramsey wrote (C) such techniques took over the theoretical models of growth. We can say with real justice that Ramsey was "ahead of his time."

Recently, three economic historians (D. A. Collard, M. Gaspard, and P. C. Duarte) have put forth a very persuasive theory (based largely on unpublished notes of Ramsey that are archived at the University of Pittsburg) that Ramsey's two papers on taxation and savings were not isolated works of a mathematician answering questions put to him by economists but were rather part of an over-arching research program that Ramsey had clearly in mind. If this plausible theory is true, it makes his early death even more tragic.

(D) "Truth and Probability," in R. B. Braithwaite (ed.), *The Foundations of Mathematics and Other Logical Essays,* London: Routledge and Kegan Paul, 1931, pp. 156–198. Reprinted in H. E. Kyburg and H. E. Smokler (eds.) *Studies in Subjective Probability*, New York: Wiley 1964, pp. 61–92.

In modeling the decisions of an individual who chooses an alternative from a set of uncertain outcomes, it has long been the tradition to introduce a numerical function to measure the objective of the individual involved. When von Neumann first formulated "the most favorable result" for a player in a strategic game, he identified "the most favorable result" with "the greatest expected monetary value," remarking that this or some similar assumption was necessary in order to apply the methods of probability theory. While doing so, he was well aware of the objections to the principle of maximizing expected winnings as a prescription for behavior but wished to concentrate on other problems. The St. Petersburg paradox illustrates in clear terms the fact that the principle of maximizing expected winnings does not reflect the actual preferences of many people.

To resolve this paradox, Daniel Bernoulli suggested that people do not follow monetary value as an index for preferences but rather the "moral worth" of the money. He then proposed a quite serviceable function to measure the moral worth of an amount of money, namely, its logarithm. Whatever the defects of this function as a universal measure of preferences, and they are many, it raises the question of the existence of a numerical index which will reflect accurately the choices of an individual in situations of risk. Interest in this problem was first shown by Ramsey in (D) in which he defined utility operationally in terms of individual behavior. As Mellor writes: "In this paper (D), after criticizing Keynes, Ramsey went on to produce his own theory. This starts from the fact that people's actions are largely determined by what they believe and what they desire – and by strength of those beliefs and desires. The strength of people's beliefs is measured by the so-called subjective probability they attach to events… Subjective utility measures the strength of people's desires just as subjective probability measures the strength of their beliefs. The problem is how to separate these two components of people's actions... One of the things Ramsey's paper did was to show how to extract people's subjective utilities and probabilities from the choices they make between different gambles; and by doing so it laid the foundations for the serious use of these concepts in economics and statistics as well as in philosophy."

The bible of game theory, *The Theory of Games and Economic Behavior* by von Neumann and Morgenstern, which confronts similar problems contains no reference to the work of Ramsey. When von Neumann was queried about this omission, he explained it by saying that, after Goedel published his papers on undecidability and the incompleteness of logic, he did not read another paper in symbolic logic. Although his excuse is strengthened by the fact that (D) first appeared in the volume that Braithwaite edited after Ramsey's death, no such excuse exists for Morgenstern, when he wrote *"Some Reflections on Utility"* in 1979 and cited two articles by J. Pfanzagl while overlooking Ramsey's paper (D) and Savage's *The Foundations of Statistics*.

Aside from Ramsey's paper on Major Douglas, which was an exemplary mathematical model refuting errant nonsense, he has clear precedence in four major themes of twentieth century economics. The paper on taxation (B) was a source for both public finance theorists and for monetary economists who have characterized inflation as a tax on money holdings and have formulated optimal inflation policies as optimal taxation schemes. The paper on savings (C) has become the touchstone for economists working on growth. The fourth area is the theory of expected utility and decisions under risk which is used in an essential way in Ramsey's insights on subjective probability in (D).

I have been a friend of John Nash since he arrived in Princeton in 1948. I knew John von Neumann from 1948 until his death in 1957. I very much regret not having known Frank Ramsey. Given the modernity of his work, it is hard to grasp the fact that he died over 77 years ago.

* * *

While in Cambridge, I tried – and failed – to find the grave of Frank Plumpton Ramsey. My son Mark Samuel Soifer succeeded.

Grave of Frank Plumpton Ramsey and his parents, Cambridge, photo by Mark S. Soifer. The tombstone reads: “In loving memory of Mary Agnes Ramsey 8 Jan. 1876–15 Aug. 1927. Also of Frank Plumpton Ramsey 22 Feb. 1903–19 Jan. 1930. Also of Arthur Stanley Ramsey 12 Sept. 1869–31 Dec. 1954”

Part VII
Colored Integers: Ramsey Theory Before Ramsey and Its AfterMath

History will be written many different ways. Look out, the Chinese are coming, the Chinese are coming and they will write history from their perspective and many things we believe are important facts will not matter to them.

– Thomas L. Saaty[1]

[1] Thomas L. Saaty e-mail to A. Soifer, Aril 13, 1998.

Chapter 33
Ramsey Theory Before Ramsey: Hilbert's Theorem

A new theory often appears in an attempt to answer new questions or to shed a new light on old problems. It is not usually born overnight. Before its birth, a new mathematical theory usually grows unnoticed within old and well-established branches of mathematics. Ramsey Theory is no exception. Its roots go back decades before the 1930 pioneering paper of Frank Plumpton Ramsey saw the light of day after his untimely passing at the age of 26. As far as we know today, the first Ramseyan-type result appeared in 1892 as a little-noticed assertion in [Hil]. Its author was the great David Hilbert. In this work, Hilbert proved the theorem of our interest merely as a tool for his study of irreducibility of rational functions with integral coefficients. The tool is known as "Hilbert's Cube Lemma": We will call it here "The Hilbert Theorem."

A set $Q_n(a, x_1, x_2, \ldots, x_n)$ of integers is called an *n-dimensional affine cube* if there exist $n + 1$ positive integers a, x_1, …, x_n such that

$$Q_n(a, x_1, x_2, \ldots, x_n) = \left\{a + \sum\nolimits_{i \in F} x_i : \varnothing \neq F \subseteq \{1, 2, \ldots, n\}\right\}.$$

In this chapter and the rest of the book, it is convenient to use the symbol $[n]$ for the starting segment of positive integers:

$$[n] = \{1, 2, \ldots, n\}.$$

This theorem, which precedes Schur's Theorem (Chapter 34) and the Baudet–Schur–van der Waerden Theorem (Chapter 35), reads as follows.

The Hilbert Theorem 33.1 For every pair of positive integers r, n, there exists a least positive integer $m = H(r, n)$ such that in every r-coloring of $[m]$, there exists a monochromatic n-dimensional affine cube.

Proof easily follows from the Baudet–Schur–van der Waerden Theorem (see it in Chapter 35): the arithmetic progression $\{a, a+x, a+2x, \ldots, a+nx\}$ is precisely the cube $Q_n(a, x_1, x_2, \ldots, x_n)$ with $x_1 = x_2 = \ldots = x_n = x$. Of course, this is not Hilbert's proof, for his proof precedes van der Waerden's paper by 35 years. ■

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_33

This is a Ramseyan theorem, as it asserts a property invariant under all r-colorings of a certain set, in this case, the initial segment $[m]$ of the set of positive integers. You can see Hilbert's proof of Hilbert's Cube Lemma, its use in the proof of Hilbert's Irreducibility Theorem, and well researched and presented through history train of thought spanning from Gauss to Hilbert and beyond in the paper [VGR] by Mark B. Villarino, William Gasarch, and Kenneth W. Regan.

Nearly 100 years later, in 1989, Paul Erdős, András Sárkösy, and Vera T. Sós published [ESS] a generalization of the Hilbert Theorem. They called it aptly "a density version" of the Hilbert Theorem.

Density Version of Hilbert's Theorem 33.2 [ESS]. For every positive integer n, there is a number $m_0 = H(n)$ such that for any $m > m_0$, $B \subseteq [m]$ with $|B| > 3m^{1-2^{-n}}$, there exist distinct positive integers $a, x_1, x_2, \ldots, x_n$ such that all 2^n sums forming the n-dimensional affine cube $Q_n(a, x_1, x_2, \ldots, x_n)$ belong to B.

Hilbert's place as one of the world's leading mathematicians around the turn of the XX century had certainly not been won by this result. He did not come back to Ramseyan-style mathematics (unlike Issai Schur, as we will see in the following few chapters). Nevertheless, the style of this book calls for a brief essay on Hilbert's life. I refer you to the celebrated Hilbert's biography by Constance Reid [Reid] and Herman Weil's paper *David Hilbert and his mathematical work* for a much worthier narrative.

David Hilbert was born near Königsberg (currently Kaliningrad, Russia) in Wehlau (currently Znamensk). In 1885, he earned his PhD degree at the University of Königsberg under Ferdinand von Lindemann. Following 10 years at Königsberg, he moved to the University of Göttingen where he remained for the rest of his life.

Hilbert made major contributions to numerous areas of mathematics and physics. In 1900, at the International Congress of Mathematicians in Paris, he presented a set of problems, known as *The 23 Hilbert's Problems* (during the talk he was able to articulate 10 of them) that profoundly influenced the development of mathematics in the twentieth century. The problems included questions related to Cantor's Continuum Hypothesis and Zermelo's Axiom of Choice (problem 1), the provability of the consistency of axioms for logic (problem 2), the possibility of the axiomatization of physics (problem 6), and The Riemann Hypothesis (problem 8).

Following Felix Klein, Hilbert made Göttingen the world's premier center of mathematics. He lived to see Göttingen's superiority collapse, when following Hitler's 1933 ascent to power, many leaders of mathematics and physics were forced to leave the University and the country.

Constance Reid [Reid] conveys how Hilbert must have felt:

> Sitting next to the Nazi's newly appointed minister of education [Bernard Rust] at a banquet, he [Hilbert] was asked, "And how is mathematics at Göttingen now that it has been freed of the Jewish influence?"
>
> "Mathematics in Göttingen?" Hilbert replied. "There is really none any more."

Hilbert passed away in Göttingen on February 14, 1943.

Chapter 34
Ramsey Theory Before Ramsey: Schur's Coloring Solution of a Colored Problem and Its Generalizations

34.1 Schur's Masterpiece

Nobody remembered – if anyone even noticed – Hilbert's 1892 lemma by the time the second Ramseyan type result appears in 1916 in number theory as another little noticed lemma. Its author is Issai Schur. Our interest here lies in the result he obtained during 1913–1916 when he worked at the University of Bonn as the successor to Felix Hausdorff.[1] There he wrote his pioneering paper [Sch]: *Über die Kongruenz* $x^m + y^m \equiv z^m$ (mod. p). In it, Schur offers another proof of a theorem by the American number theorist Leonard Eugene Dickson from [Dic1], who was trying to prove Fermat's Last Theorem. For use in his proof, Schur creates, as he put it, "a very simple lemma, which belongs more to combinatorics than to number theory."

Nobody then asked questions of the kind Issai Schur posed and solved in his 1916 paper [Sch]. Consequently, nobody appreciated this result much when it was published. Now it shines as one of the most beautiful, classic theorems in the history of mathematics. Its setting is positive integers, colored in finitely many colors. The beautiful solution I am going to present utilizes coloring as well. I have got to tell you how I received this solution (see [Soi9] for more details).

In August 1989, I taught at the International Summer Institute in Long Island, New York. A fine international contingent of gifted high school students for the first time included a group from the Soviet Union. Some members of this group turned out to be mathematics Olympiads' "professionals," winners of the Soviet Union National Olympiads in Mathematics and in Physics. There was nothing in the Olympiad genre that they did not know or could not solve. I offered them – and everyone else in my class – an introduction to certain areas of

[1] Both Alfred Brauer [Bra2] and Walter Ledermann [Led1] reported the year of 1911 as the time when Schur became an *Extraordinarius* in Bonn, while Schur's daughter Mrs. Hilde Abelin–Schur [Abe1] gave me 1913 as the time her family moved to Bonn. The Humboldt University's Archive contains personnel forms (Archive of Humboldt University at Berlin, document UK Sch 342, Bd. I, Bl.25) filled up by hand by Issai Schur himself, from which we learn that he worked at the University of Bonn from April 21, 1913, until April 1, 1916, when he returned to Berlin.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_34

combinatorial geometry. We quickly reached the forefront of mathematics, full of open problems. Students shared with me their favorite problems and solutions as well. Boris Dubrov from Minsk, Belarus, told me about a visit to Moscow by the American mathematician Ronald L. Graham. During his interview with the Russian physics and mathematics magazine *Kvant*, Graham mentioned a beautiful problem that dealt with 2-colored positive integers. Boris generalized the problem to n-coloring, strengthened the result, and proved it all! He gave me this generalized problem to use in the Colorado Mathematical Olympiad.

This problem was in fact the Schur Theorem of 1916, rediscovered by Boris, with his own proof that was more beautiful than Schur's original proof, but which had already been known for ca. 18 years. Paul Erdős received this proof from Vera T. Sós and considered it important enough to include it in his talk at the 1970 International Congress of Mathematicians in Nice, France [E71.13]. Chances of receiving a solution of such a problem during the Olympiad were slim. Yet, the symbolism of a Soviet kid offering an astonishingly beautiful problem to his American peers was so great, that I decided to include it as an additional problem 6 (Colorado Mathematical Olympiad usually offers 5 problems of increasing difficulty).

The Schur Theorem 34.1 ([Sch])**.** For any positive integer n, there is an integer $S(n)$ such that any n-coloring of the initial positive integers array $[S(n)]$ contains a monochromatic triple a, b, c, such that $a + b = c$.

Proof of the Schur Theorem Let all positive integers be colored in n colors $c_1, c_2, \dots, c_n$. Due to Problem 29.13, there is $S(n)$ such that any n-coloring of edges of the complete graph $K_{S(n)}$ contains a monochromatic triangle K_3.

Construct a complete graph $K_{S(n)}$ with its vertices labeled with integers from the initial integers array $[S(n)] = \{1, 2,\dots, S(n)\}$. Now color the edges of $K_{S(n)}$ in n colors as follows: let i and j, $(i > j)$, be two vertices of $K_{S(n)}$, color the edge ij in precisely the color of the integer $i - j$ (remember, all positive integers were colored in n colors!). We get a complete graph $K_{S(n)}$ whose edges are colored in n colors. By Problem 29.13, $K_{S(n)}$ contains a triangle ijk, $i > j > k$, whose all three edges ij, jk, and ik are colored in the same color (Fig. 34.1).

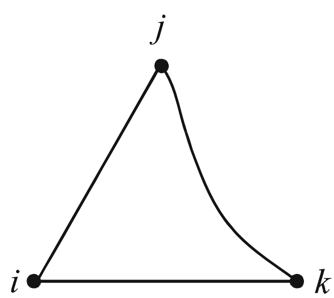


Fig. 34.1 Edge-monochromatic triangle

Denote $a = i - j$; $b = j - k$; $c = i - k$. Since all three edges of the triangle ijk are colored in the same color, the integers a, b, and c are colored in the same color in the original coloring of the integers (this is how we colored the edges of $K_{s(n)}$). In addition, we have the following equality:

$$a + b = (i–j) + (j–k) = i–k = c$$

We are done! ■

The result of the Schur Theorem can be strengthened by an additional clever trick in the proof.

Stronger Version of the Schur Theorem 34.2 For any positive integer *n,* there is an integer $S^*(n)$ such that any *n*-coloring of the initial positive integers array $[S^*(n)]$ contains a triple of *distinct* integers *a, b, c*, such that $a + b = c$.

Proof Let all positive integers be colored in *n* colors $c_1, c_2, \ldots, c_n$. We add *n* more colors c_1', c_2', . . . , c_n' different from the original *n* colors and construct a complete graph $K_{S(2n)}$ with the set of positive integers $\{1, 2, \ldots, S(2n)\}$ labeling its vertices (see the definition of $S(2n)$ in the proof of Theorem 34.1). Now we are going to color the edges of $K_{S(2n)}$ in $2n$ colors.

Let i and j, $(i > j)$, be two vertices of $K_{S(2n)}$, and c_p be the color in which the integer $i - j$ is colored, $1 \le p \le n$ (remember, all positive integers are colored in *n* colors $c_1, c_2, \ldots, c_n$). Then, we color the edge ij in color c_p if the number $\left\lfloor \frac{i}{i-j} \right\rfloor$ is even, and in color c_p' if the number $\left\lfloor \frac{i}{i-j} \right\rfloor$ is odd (for a real number r, the symbol $\lfloor r \rfloor$, as usual, denotes the largest integer not exceeding r).

We get a complete graph $K_{S(2n)}$ whose edges are colored in $2n$ colors. By Theorem 34.1, $K_{S(2n)}$ contains a triangle ijk, $i > j > k$, whose all three edges ij, jk, and ik are colored in the same color (see Fig. 34.1).

Denote $a = i - j$; $b = j - k$; $c = i - k$. Since all three edges of the triangle ijk are colored in the same color, from the definition of coloring of edges of $K_{S(2n)}$, it follows that in the original coloring of positive integers, the integers a, b, and c were colored in the same color. In addition, we have

$$a + b = (i–j) + (j–k) = i–k = c.$$

We are almost done. We only need to show (our additional pledge!) that the numbers a, b, c are all distinct. In fact, it suffices to show that $a \ne b$. Assume the opposite: $a = b$ and c_p is the color in which the number $a = b = i - j = j - k$ is colored. But then

$$\left\lfloor \frac{i}{i-j} \right\rfloor = \left\lfloor 1 + \frac{j}{i-j} \right\rfloor = 1 + \left\lfloor \frac{j}{i-j} \right\rfloor = 1 + \left\lfloor \frac{j}{j-k} \right\rfloor,$$

i.e., the numbers $\left\lfloor\frac{i}{i-j}\right\rfloor$ and $\left\lfloor\frac{j}{j-k}\right\rfloor$ have different parity; thus, the edges ij and jk of the triangle ijk must have been colored in different colors. This contradiction to the fact that all three edges of the triangle ijk have the same color proves that $a \neq b$. Theorem 34.2 is proved. ■

34.2 Schur's Numbers

Solve the following problem:

Problem 34.3 Can integers 1; 2; . . .; 581,130,733 be colored in 19 colors without creating a monochromatic triple x, y, z, such that $x + y = z$?

Are you scared?

You should! :)

Let me help you. *Hint*: for $n = 19$, $(3^n - 1)/2 = 581, 130, 733$.

This is a good illustration that the general case may be easier to prove than a particular one. Issai Schur proved this general case in his 1916 paper [Sch]. The Olympiad spirit of the problem prompted me to offer it to mid- and high school students at the 36th Soifer Mathematical Olympiad as problem 5:

Schur's Lower Bound 34.4 For any positive integer n, find an n-coloring of integers 1, 2, . . ., $(3^n - 1)/2$ such that there is no monochromatic triple x, y, z, such that $x + y = z$.

Solution. Coloring will be constructed by induction. The case $n = 1$ is trivial. Assume that there is a n-coloring of the set T: 1, 2, . . ., $(3^n - 1)/2$ not creating a monochromatic triple x, y, z, such that $x + y = z$. Partition the set R: 1, 2, . . ., $(3^{n+1} - 1)/2$ into 3 subsets:

$$1, 2, \ldots, (3^n - 1)/2;$$

$$(3^n - 1)/2 + 1, \ldots, (3^n - 1)/2 + 2, \ldots, 3^n;$$

$$3^n + 1, 3^n + 2, \ldots, \left(3^{n+1} - 1\right)/2.$$

The first subset can be properly colored due to the inductive assumption. We assign color $n + 1$ to the entire second subset. Since $(3^{n+1} - 1)/2 - (3^n + 1) + 1 = (3^n - 1)/2$, the third subset has exactly the same number of elements as the first one, and we color it by the translation of the coloring of the first subset by 3^n: if a of the first subset is assigned color m, we color $a + 3^n$ of the third subset in color m. Let us now prove that we created no monochromatic triple x, y, z with $x + y = z$.

If x, y both belong to the first subset, and their sum $x + y$ is in the first subset, then by the inductive assumption, the triple is not monochromatic. If x, y both belong to the first subset and $x + y$ is in the second subset, then the sum is in color $n + 1$ and the triple is not monochromatic. If x, y both belong to the first subset, $x + y$ cannot belong to the third subset – in all cases, we get no monochromatic triple.

The sums of any two numbers from the second subset belong to the third subset, thus again preventing a monochromatic triple.

If x belongs to the first subset and y and $x + y$ belong to the third subset, we do not get a monochromatic triple. Indeed, in this case, $y - 3^n$ has the same color as y (by our definition of colors in the third subset). And if the triple x, y, $x + y$ is monochromatic, then the triple x, $y - 3^n$, $x + (y - 3^n)$ is monochromatic and entirely in the first subset, which contradicts our inductive assumption.

Finally, if x, y both belong to the third subset, their sum $x + y$ lies outside of it. ■

There are two known definitions of the Schur Number, differing by 1. Let us choose to define the *Schur Number* as the largest integer $S(n)$, such that there is a n-coloring of $[S(n)]$ that forbids a monochromatic triple x, y, z with $x + y = z$.

As we have observed, Schur established the lower bound $S(n) \geq (3^n - 1)/2$. This lower bound is sharp for $n = 1, 2, 3$, which is easy to prove: $S(1) = 1$, $S(2) = 4$, and $S(3) = 13$.

For $n = 4$, Schur's formula gives 40; however, in 1965, using computer, Leonard D. Baumert and Solomon W. Golomb showed [BG] that in fact $S(4) = 44$.

Finding the exact value of $S(5)$ appeared to be very hard. In the 1970s, best-known bounds for $S(5)$ were $157 \leq S(5) \leq 321$, the lower bound obtained in 1979 by Harold Fredricksen [F] and the upper bound in 1973 by Earl Glen Whitehead [W].

Only ca. two decades later, in 1994, Geoffrey Exoo proved [Ex18] that $S(5) \geq 160$. Moreover, Geoffrey shares with us valuable comments (ibid):

> We have found approximately 10,000 different partitions [colorings] of [1, 160]; of these, four are symmetric [palindromal]. These 10,000 partitions are all 'close' to each other. In other words, one can begin with one of the partitions, move an integer from one set to another, and obtain a new partition. This can be contrasted with the situation for partitions of [1, 159] where we found over 100,000 partitions, most of which were not close in this sense. It is tempting to conclude that there are far fewer sum-free partitions of [1, 160] than of [1, 159].

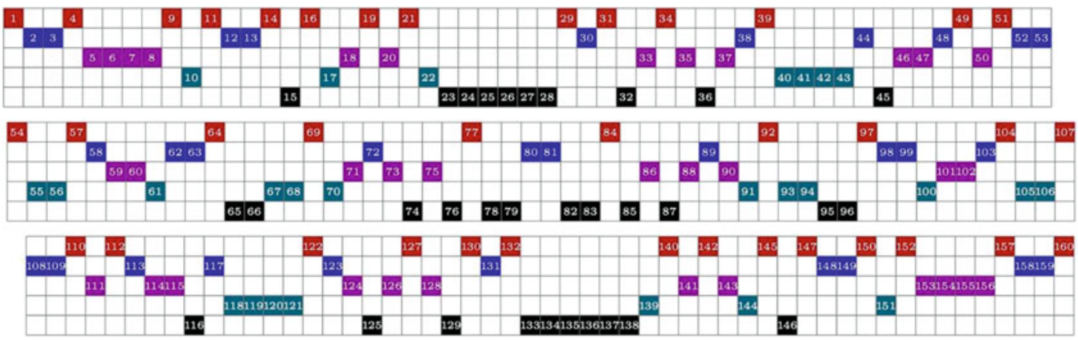
In 2017, the computer scientist Marijn J.H. Heule became interested in this problem. His result [Heu1] also appeared [Heu2] in AAAI (Submitted Tue, 21 Nov 2017 22:54:59):

$$S(5) = 160.$$

It was a very significant result, for until this publication, the upper bound of $S(5)$ stood at 315. Marijn writes:

We obtained the solution, $n = 160$, by encoding the problem into propositional logic and applying massively parallel satisfiability solving techniques on the resulting formula. We constructed and validated a proof of the solution to increase trust in the correctness of the multi-CPU-year computations. The proof is two petabytes in size and was certified using a formally verified proof checker, demonstrating that any result by satisfiability solvers—no matter how large—can now be validated using highly trustworthy systems.

As you already know, the coloring of integers from 1 to 160 in 5 colors without a monochromatic pair and its sum was first demonstrated by Geoffrey Exoo. He even produced a number of palindromal colorings, i.e., colorings where numbers i and $160 - i$ are assigned the same color. I am showing here a palindromal coloring found by Marijn Heule:



Summing up, we record:

Schur Number 5 by Heule 34.5 (Heule [Heu1], [Heu2]). $S(5) = 160$.

The asymptotic lower bound was slightly improved from Schur's exponential base 3. Following Abbott and Moser 1966 [AM], Abbott and Hanson 1972 [AH], Exoo's result allowed for the lower bound of $S(n) \geq c(315)^{\frac{n}{5}} \approx c(3.15981831)^n$ for $n > 5$ and a constant c [Ex18]. Heule's result [Heu1] raised it higher: $S(n) \geq c(321)^{\frac{n}{5}} \approx c(3.17176503)^n$.

In 2000, Harold Fredricksen and Melvin M. Sweet [FS] constructed colorings that proved new lower bounds $S(6) \geq 536$ and $S(7) \geq 1680$.

34.3 Generalized Schur

It is fitting that the Schur Theorem was generalized by one of Schur's best students – Richard Rado. Rado calls a linear equation

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b \qquad (*)$$

regular, if for any positive integer r, any r-coloring of positive integers N contains a monochromatic solution of the equation (*). As before, we say that a solution $x_1, x_2, \ldots, x_n$ is *monochromatic*, if all numbers $x_1, x_2, \ldots, x_n$ are assigned the same color.

For example, the Schur Theorem 34.1 proves precisely that the equation $x + y - z = 0$ is regular. In 1933, Richard Rado, among other results, found the following criterion:

The Rado Theorem 34.6 (A particular case of [Rad1]). Let E be a linear equation $a_1x_1 + a_2x_2 + \ldots + a_nx_n = 0$, where all $a_1, a_2, \ldots, a_n$ are integers. Then E is regular if and only if some non-empty subset of the coefficients a_i sums up to zero.

For example, the equation $x_1 + 3x_2 - 2x_3 + x_4 + 10x_5 = 0$ is regular because $1 + 3 - 2 = 0$.

Corollary 34.7 The Schur Theorem 34.1 follows from The Rado Theorem.

Richard Rado found regularity criteria for systems of homogeneous equations as well. His fundamental contributions to and influence on Ramsey Theory are hard to overestimate. I have given you just a taste of his theorems here. For more of Rado's results read his papers [Rad1], [Rad2], and others, and the monograph [GRS2]. Instead of a formal biographical data, I prefer to include here a few passages about Richard Rado (1906, Berlin – 1989, Henley-on-Thames, Oxfordshire) written by someone who knew Rado very well – Paul Erdős – from the latter's paper *My joint work with Richard Rado* [E87.12]:

> I first became aware of Richard Rado's existence in 1933 when his important paper *Studien zur Kombinatorik* [Rad1, Rado's Ph.D. thesis under Issai Schur][2] appeared. I thought a great deal about the many fascinating and deep unsolved problems stated in this paper, but I never succeeded to obtain any significant results here and since I have to report here about our joint work I will mostly ignore these questions. Our joint work extends to more than 50 years; we wrote 18 joint papers, several of them jointly with A. Hajnal, three with E. Milner, one with F. Galvin, one with Chao Ko, and we have a book on partition calculus with A. Hajnal and A. Mate. Our most important work is undoubtedly in set theory and, in particular, the creation of the partition calculus. The term partition calculus is, of course, due to Rado. Without him, I often would have been content in stating only special cases. We started this work in earnest in 1950 when I was at University College and Richard at King's College. We completed a fairly systematic study of this subject in 1956, but soon after this we started to collaborate with A. Hajnal, and by 1965 we published our GTP (Giant Triple Paper – this terminology was invented by Hajnal) which, I hope, will outlive the authors by a long time. I would like to write by centuries if the reader does not consider this as too immodest . . .
>
> I started to correspond with Richard in late 1933 or early 1934 when he was a [Jewish] German refugee in Cambridge. We first met on October 1, 1934 when I first arrived in Cambridge from Budapest. Davenport and Richard met me at the railroad station in Cambridge and we immediately went to Trinity College and had our first long mathematical discussion . . .
>
> Actually, our first joint paper was done with Chao Ko and was essentially finished in 1938. Curiously enough it was published only in 1961. One of the reasons for the delay was that at that time there was relatively little interest in combinatorics. Also, in 1938,

[2]Two years later, Rado obtained his second Ph.D. degree at Cambridge under G. H. Hardy.

> Ko returned to China, I went to Princeton, and Rado stayed in England. I think we should have published the paper in 1938. This paper "Intersection theorems for systems of finite sets" became perhaps our most quoted result.

It is noteworthy to notice how differently people can see the same fact. For Richard Rado, the Schur Theorem was about monochromatic solutions of a homogeneous linear equation $x + y - z = 0$ and so Rado generalized the Schur Theorem to a vast class of homogeneous linear equations (Rado's Theorem 34.6 above) and systems of homogeneous linear equations [Rad1]. Three other mathematicians, totally unaware of the existence of each other, saw the Schur Theorem quite differently. This group consisted of Jon Folkman, a young Rand Corporation scientist; Jon Henry Sanders, the last Ph.D. student of the legendary Norwegian graph theorist Øystein Ore at Yale (B.A. 1964 Princeton University; Ph.D. 1968, Yale University); and Vladimir I. Arnautov, a 30-year old Moldavian topological ring theorist. Arnautov's contribution to combinatorics was totally unknown in the West until I carefully looked at his ring theory paper. I shared my giving Arnautov credit with our leader Ronald L. Graham, who asked me in an e-mail "Who is Arnautov?"

For the three, the Schur Theorem spoke about monochromatic sets of symmetric sums

$$\{a_1, a_2, a_1 + a_2\} = \left\{ \sum_{i=1,2} \varepsilon_i a_i : \varepsilon_i = 0, 1; \varepsilon_1 \varepsilon_2 \neq 0 \right\}.$$

Consequently, the three proved a generalization of the Schur Theorem, quite different from Rado's kind, and paved the way for further important developments. I see therefore no choice at all but to name the following fine theorem by its three independent inventors. This may surprise readers accustomed to different attributions. I will address their concerns later in this chapter.

The Arnautov–Folkman–Sanders Theorem 34.8 ([San1], [Arn]). For any positive integers m and n, there exists an integer $AFS(m,n)$ such that any m-coloring of the initial integers array $[AFS(m,n)]$ contains an n-element subset $S \subset [AFS(m,n)]$ such that the set $\left\{ \sum_{x \in F} x : \emptyset \neq F \subseteq S \right\}$ is monochromatic.

Corollary 34.9 Both the Hilbert Theorem 33.1 and the Schur Theorem 34.1 follow from the Arnautov–Folkman–Sander Theorem 34.8.

On 25 April 2009, i.e., half a year after the first edition of *The Mathematical Coloring Book* was published, I received the following e-mail from Dr. Jon Henry Sanders:

> Dear Prof. Soifer: What has been referred to throughout the literature as the Graham–Rothschild conjecture (resolved by Hindman) was first posed by me (in the more general form for an arbitrary finite number of colors) in my disertation [spelling corrected in the next e-mail], *A Generalization of Schur's Theorem*, Yale '68. Attached is a photocopy of pgs 9 and 10 of my dissertation – Theorem 2' is the conjecture. Since Rothschild was one of two readers of my dissertation (Plummer the other) it is strange that this misattribution has existed for so long.

The same day, Dr. Sanders sent me one more e-mail:

Dear Prof. Soifer: I am sending this again to

A. correct the spelling of the word 'disertation' to 'dissertation'.

B. explain that I was pleased to see your correct efforts at attribution of the 'Arnautov-Folkman-Sanders' Theorem in the 'Mathematical Coloring Book' (I only have read some excerpts of the latter but I just purchased it on-line and look forward to reading it) and that this prompted me to try to help clarify the origins of the countable version.

C. Share with you the anecdote that when I first proposed the theorem to Prof. Ore early in '67 to see if it was known and if he thought it would be a reasonable dissertation topic, he told me that it was not known to him and to go ahead and try to prove it (he did not mention the existence of van der Waerden's theorem) – When I first learned of van der Waerden's theorem, after my dissertation had been accepted and right before graduating, I was worried that it might easily imply my theorem and somehow negate or trivialize it. Folkman's original proof using van der Waerden's theorem when I learned of it turned out to be short but not trivial and I was happy that the conferring of my degree was not jeopardized.

Best Regards,
Jon Henry Sanders

While working on the original edition of *The Mathematical Coloring Book* [Soi44], I verified J. H. Sanders' proof of Theorem 34.8 in his dissertation (where it is called "Theorem 2"), but I failed to notice the conjecture. In 2011 [Soi33], I verified Dr. Sanders' priority and changed the credit for the conjecture. Looking at the 1968 dissertation [San1] again, I see the conjecture listed as "Theorem 2'" and preceded by the words "It is natural [!] to ask whether either Lemma 1 or Theorem 2 generalize in the following way":

The Sanders Conjecture 34.10 [San1, p. 9]. Let the positive integers be divided into t classes $A_1, A_2, \ldots, A_t$, [t a positive integer]. Then there exists an infinite [countable] sequence $a_1, a_2, \ldots$ of positive integers and a number l, $1 \le l \le t$, such that $\sum_{i \in I} a_i \in A_l$ for all (non-empty) finite sets I of positive integers.

A general problem with the American system of awarding doctorates is in play here. One does not have to publish the results of a thesis. In the Soviet Union, on the other hand, main results had to be published in (refereed) journals before a candidate for doctorate earned the right to defend the doctoral thesis. Jon H. Sanders published a relevant part of his 1968 dissertation, but nearly 50 years later, on 11 December 2017, in arXiv [San2]. A Russian proverb observes, "Spoon is good for dinner," i.e., not after it. Dr. Sanders has been justly unhappy for half a century that his authorship of the conjecture was never acknowledged and his coauthorship of the theorem was acknowledged not always.

In their important 1971 paper [GR], Ron Graham and Bruce Rothschild vastly generalized a number of Ramsey type theorems and formulated this Conjecture 34.10 three years later than Jon H. Sanders and only for a division into 2 classes. Thus, the credit for the conjecture does belong to Jon Henry Sanders. Paul Erdős gave a high praise to the conjecture during his 1971 talk in Fort Collins, Colorado, published in 1973 [E73.21]:

> Graham and Rothschild ask the following beautiful question: split the integers into two classes. Is there always an infinite sequence so that all the finite sums $\sum \varepsilon_i a_i$, $\varepsilon_i = 0$ or 1 (not all $\varepsilon_i = 0$) all belong to the same class? ... This problem seems very difficult.

Graham–Rothchild's and Erdős' papers made the conjecture well known and its proof very desirable. In his impressive paper submitted in 1972 and published in 1974 [Hin], Neil Hindman proved Conjecture 34.10. As I have tried to do uniformly throughout this monograph and my other writings, I am giving credit for this result to both the author of the conjecture and the author of its first proof.

The Sanders–Hindman Theorem 34.11 ([San1], [Hin]). For any positive integer n and any n-coloring of the set of positive integers N, there is an infinite subset $S \subseteq N$ such that the set $\left\{ \sum_{x \in F} x : \varnothing \neq F \subset S; |F| < \aleph_0 \right\}$ is monochromatic.

Only on 19 July 2020, did I notice that in their fine but nearly impossible to obtain problem book [EG] published in 1980 by the Université de Genève, Paul Erdős and Ron Graham report (p. 13) that Hindman proved the theorem "*answering a conjecture of Graham and Rothschild and Sanders*." Thus, Sanders' authorship of the conjecture was known to the leaders of the field and should have been acknowledged all along.

Let us now go back and establish the most appropriate credit for Theorem 34.8. It is called "Folkman–Rado–Sanders' Theorem" in [GRS1], [Gra2], and [EG] and "Folkman's Theorem" in [Gra1] and [GRS2]. Most of the other authors have simply copied attribution from these works. Which credit is justified? In one publication only [Gra2], Ronald L. Graham gives the date of Jon Folkman's personal communication to Graham that contained the relevant proof: 1965. In one late, 1981 publication [Gra1], Graham includes Folkman's proof that uses the Baudet–Schur–Van der Waerden Theorem (see Chapters 35 and 37). Thus, Folkman merits a credit. In the standard text on Ramsey Theory [GRS2] by Graham–Rothschild–Spencer, I find an argument for credit to Folkman alone that disagrees with the first edition [GRS1] of the same book by the same authors:

> Although the result was proved independently by several mathematicians, we choose to honor the memory of our friend Jon Folkman by associating his name with the result.

Jon H. Folkman left this world tragically in 1969. He was 31. Jon was full of great promise. Sympathy and grief of his friends are understandable and noble. Yet, do we, mathematicians, have the liberty to award credits based on something other than mathematics? In this case, how can we deny Jon Henry Sanders a credit, when Sanders' independent authorship is absolutely clear and undisputed (he could not have been privy to the above-mentioned personal communication)? Sanders formulates and proves Theorem 34.8 in his 1968 Ph.D. dissertation [San1]. Moreover, Sanders proves it in a different way than Folkman: he does not use the Baudet–Schur–Van der Waerden Theorem but instead generalizes the Ramsey Theorem to what he calls in his dissertation "Iterated Ramsey Theorem" [San1, pp. 3–4].

Vladimir Ivanovich Arnautov's discovery is even more striking. Living in the Soviet Union, he was certainly not privy to the Folkman's private communication nor to the unpublished Sanders' 1968 thesis at Yale. Arnautov's paper is much closer in the presentation style to Schur's classic 1916 paper, where the Schur Theorem appears as a useful tool, "a very simple lemma," and is immediately used for obtaining a number-theoretic result, related to Fermat's Last Theorem. Arnautov formulates and proves Theorem 34.8 but treats it as a useful tool and calls it simply "Lemma 2" (in the proof of Lemma 2, he uses the Baudet–

Schur–Van der Warden Theorem). He then uses Lemma 2 and other Ramseyan tools to prove that every (not necessarily associative) countable ring allows a nondiscrete topology. This brilliant paper was submitted to *Doklady Akademii Nauk USSR* on August 22, 1969 and on September 2, 1969 was recommended for publication by the celebrated topologist Pavel S. Aleksandrov.[3] We have no choice but to savor the pleasure of associating Vladimir I. Arnautov's name with Theorem 34.8.

What about Rado, I hear you asking? As Graham–Rothschild–Spencer [GRS2] observe, Theorem 34.8 "may be derived as a corollary of Rado's theorem [Rad1] . . . by elementary, albeit nontrivial, methods."[4] In my opinion, this is an insufficient reason to attach Rado's name to Theorem 34.8. Arnautov, Folkman, and Sanders envisioned a generalization of Schur in the direction different from that of Rado and paved the way for Sanders' conjecture proved by Hindman. In fact, Erdős came to the same conclusion as I in 1973 [E73.21] when he put Rado's name in parentheses:

> Sanders and Folkman proved the following result (which also follows from the earlier results of Rado [Rad1]).

I knew Paul Erdős well enough to be certain that had he have seen Arnautov's paper, he would have definitely added Vladimir I. Arnautov's name to the authors of Theorem 34.8.

Vladimir Ivanovich Arnautov, who is to turn 84 on 30 July 2023, has been a professor at the Institute of Mathematics and Computer Science, and a member of the Academy of Sciences of Moldova. He served as the director of the Institute of Mathematics of the Moldavian Academy of Sciences, and a member of the presidium of the Academy.

34.4 Nonlinear Equations or Pythagoras Meets Ramsey

A number of mathematicians studied regularity of nonlinear equations. A special attention was paid to the Pythagorean quadratic equation. Ron Graham writes to me on January 26, 2007:

> Hi Sasha,
>
> Here is the info on the meeting in Georgia (2005). I don't [remember] when Erdős and I first published the x^2 + y^2 = z^2 partition regularity question. We both certainly mentioned it in talks for quite a while (especially in connection with the positive result of Rödl for the regularity of 1/x + 1/y = 1/z (which was never published)
> . . .
> Best regards,
> Ron

[3] *Doklady* published only papers by full and corresponding members of the Academy of Sciences of USSR. To be published, a non-member's paper had to be recommended for publication by a full member of the Academy. I had this honor once, when A.N. Kolmogorov wrote diagonally on my manuscript "Recommend, A.N. Kolmogorov." In print, right below my article's title I see "Presented by the Academy member A.N. Kolmogorov 10/11/1973."

[4] Theorem 34.8 also follows from Graham and Rothschild's results published in 1971 [GR1].

In fact, in his 2005 talk, published as [Gra7], Graham estimates that this Erdős–Graham problem "has been opened for over 30 years" and remains open as Graham reports in [Gra7], [Gra8], where he offers $250 for the first solution:

$250 Pythagoras-Meets-Ramsey Problem 34.12 (R. L. Graham and P. Erdős, before 1975). Determine whether the Pythagorean equation $x^2 + y^2 = z^2$ is partition regular, i.e., whether for any positive integer *k,* any *k*-coloring of the set of positive integers contains a non-trivial monochromatic solution x, y, z of the equation.

"There is actually very little data (in either direction) to know which way to guess," Graham remarks [Gra7], [Gra8]. I recall the following story.

In May 1993 in a Budapest hotel, right after Paul Erdős' 80th Birthday Conference in Keszthely, Hungary, Hanno Lefmann from Bielefeld University, Germany, told me that he and Arie Bialostocki from the University of Idaho, Moscow, generated by computer, with an assistance of a student, a coloring of positive integers from 1 to over 60,000 in two colors that forbade monochromatic solutions x, y, z of the equation $x^2 + y^2 = z^2$. They must have made a mistake: see Theorem 34.13 below.

In the first edition of this book, I wrote:

> This could be a basis for conjecturing a negative answer to Problem 34.12, but of course the problem remained open, awaiting new approaches.
>
> While these results were a step forward, they remain in a little studied vast area of Ramsey Theory. It deserves its own Richard Rado.

Indeed, it happened when the problem was over 40 years old. In 2016, three computer scientists, using innovative computer science machinery, solved the Pythagoras–Ramsey problem of Graham and Erdős *for 2 colors*. Marijn J. H. Heule, the leader of the group that also included Oliver Kullmann and Victor W. Marek, answered the call on 3 May 2016 [HKM]. Everyone was in awe of the size of the solution; "200 terabytes is unbelievable," Ron Graham said. Many popular scientific magazines published essays about it. On my request Marijn Heule, sums-up his view of the problem and its solution in his 10 September 2020 e-mail to me:

> The Pythagorean Triples Problem asks whether any coloring of the positive numbers with two colors, let's call them red and blue, results in a monochromatic solution of the Pythagorean equation $a^2 + b^2 = c^2$. Our approach focused on answering the question positively by looking for a range of numbers $\{1, \ldots, n\}$ for which the answer is positive. We translated the problem into propositional logic by introducing a Boolean variable for each number in the range. Assigning a Boolean variable to true means that the corresponding number is colored red, while assigning it to false means that the number is colored blue. For each solution of $a^2 + b^2 = c^2$ with a, b, and c within the range, one clause enforces that at least one of a, b, and c must be colored red and another clause enforces that at least one of a, b, and c must be colored blue. The formula with $n = 7824$ is satisfiable (a monochromatic solution can be avoided), while the formula with $n = 7825$ is unsatisfiable. State-of-the-art local search solvers can compute the former in about a minute, while proving the latter requires years of computation on a single computer. We did not only solve the problem, but also generated a proof, which can be validated using high-trustworthy systems. Below I will briefly recapitulate some highlights.

Let's start with the elephant in the room: the size of the proof. After all, this is what got the most attention in the media. In outlets around the world, from *Nature News* to *Der Spiegel*, everybody was talking about "the largest proof ever." And yes, it is big. The proof of the Pythagorean Triples Problem is 200 terabytes in size, or, as one blog commentator wrote, "400 PlayStations 4." Yet in my opinion, the proof size is really not that interesting. It would have been trivial to generate a substantially larger proof by simply using heuristics that are less effective. The size of 200 terabytes is the result of heavily optimizing the heuristics used in the automated reasoning. After about three weeks of optimization, the computation became within reach using the computational resources at my disposal at the time. If I would have had more resources, I might have stopped earlier, resulting in a larger proof. With fewer resources, I would [have] tried even harder to trim the computational costs and the size of the proof. A reduction of about 50% is probably within reach.

I consider the ability to generate and validate this enormous proof using highly trustworthy proof checkers as the most interesting result of the project. On its face, it is only natural to doubt the correctness of a computation that takes years to complete on a single computer. Various other mathematical problems that have been solved using automated reasoning even failed to produce a proof altogether. Authors of such results often argued that storing and checking a very large proof would be infeasible, that it could not be done. I refuse to accept such defeatism. To counter their argumentation, I looked for a problem that would require enormous computation resources to solve. Once solved, I also insisted on producing and validating a proof of that computation. The Pythagorean Triples Problem turned out to be a suitable problem for this purpose.

Of course, storing 200 terabytes of proof can be challenging if not impossible for many researchers. However, the size of the proof is measured in the proof format that is supported by most automated reasoning tools. The storage on disk is two orders of magnitude smaller due to heavy compression. The checker reads small chunks of the proof at a time to reduce memory consumption. Proof validation can also be performed in parallel. Most importantly, everybody can easily check if I did my work correctly: A variety of tools is available to check the proof's validity. In addition, a similar tool chain can be used to validate many other automated reasoning results.

A technique called cube-and-conquer was crucial to obtain the result. The use of cube-and-conquer realized linear time speedups even when using hundreds (or even thousands) of cores. The key ingredient of this technique is that it partitions the original problems into billions of subproblems that are solved independently. In many cases it can be challenging to partition a problem is such a way that the total runtime of solving the subproblems is similar or smaller than the runtime to solve the original problem. For the Pythagorean Triples Problem this was made possible by the above-mentioned heuristics.

It is still up in the air whether there exists a compact proof for the *infinite* Pythagorean Triples Problem: Will any coloring of the positive integers with two colors result in a monochromatic Pythagorean triple? The current proof shows the exact point of transition: one can avoid a monochromatic Pythagorean triple when bi-coloring all positive numbers up till 7824. However, this is not possible for all positive numbers up to 7825 (and higher). Notice that by searching for the smallest counterexample (the numbers up to 7825), the reasoning – and thus an infinite number of Pythagorean triples – is reduced to only 9472 Pythagorean triples.

> I also experimented with formulas that express whether the numbers up to 100,000 can be bi-colored while avoiding a monochromatic Pythagorean triple. The impossibility of that statement is significantly smaller, about a terabyte (in the same proof format). Thus, there may be a reasonably short proof if the range of numbers is extremely large. Therefore, a humanly understandable argument for the infinite Pythagorean Triples Problem might be out there. However, I consider it unlikely that there exists a short proof for the inability to bi-color the positive integers up to 7825. The shortest proof of that statement is likely multiple terabytes in size.
>
> While the Pythagorean Triples Problem is now solved for two colors, the problem is still wide open for more colors. Although I expect that the problem [the problem's answer] holds for any finite number of colors, it is highly unlikely that the method that was successful for two colors can be used for three colors or more. Some experiments showed that it is easy to color positive numbers up to 10 million using three colors without creating a monochromatic Pythagorean triple. Hence the transition point for three colors will be high. The transition point for four colors is probably astronomically large, and this might even be the case for three colors. I don't expect that we will ever establish them.
>
> For more details, please, see Marijn's exposition "Everything's Bigger in Texas":
>
> https://www.cs.utexas.edu/~marijn/ptn/

The Heule–Kullmann–Marek Theorem 34.13 Any 2-coloring of the set of all positive integers from 1 through 7825 contains a monochromatic solution x, y, z of the Pythagorean Equation $x^2 + y^2 = z^2$. Moreover, 7825 is the smallest number for which the statement is true.

I find it amazing that Heule did not only solve the old and hard problem (albeit for 2 colors) but also found the exact minimal value, 7825, that guarantees a monochromatic triple. His conjecture is of a fundamental importance:

Heule's Conjecture 34.14 The Pythagorean equation $x^2 + y^2 = z^2$ is partition regular, i.e., for any positive integer k, any *k-coloring* of the set of positive integers contains a monochromatic solution x, y, z of the equation.

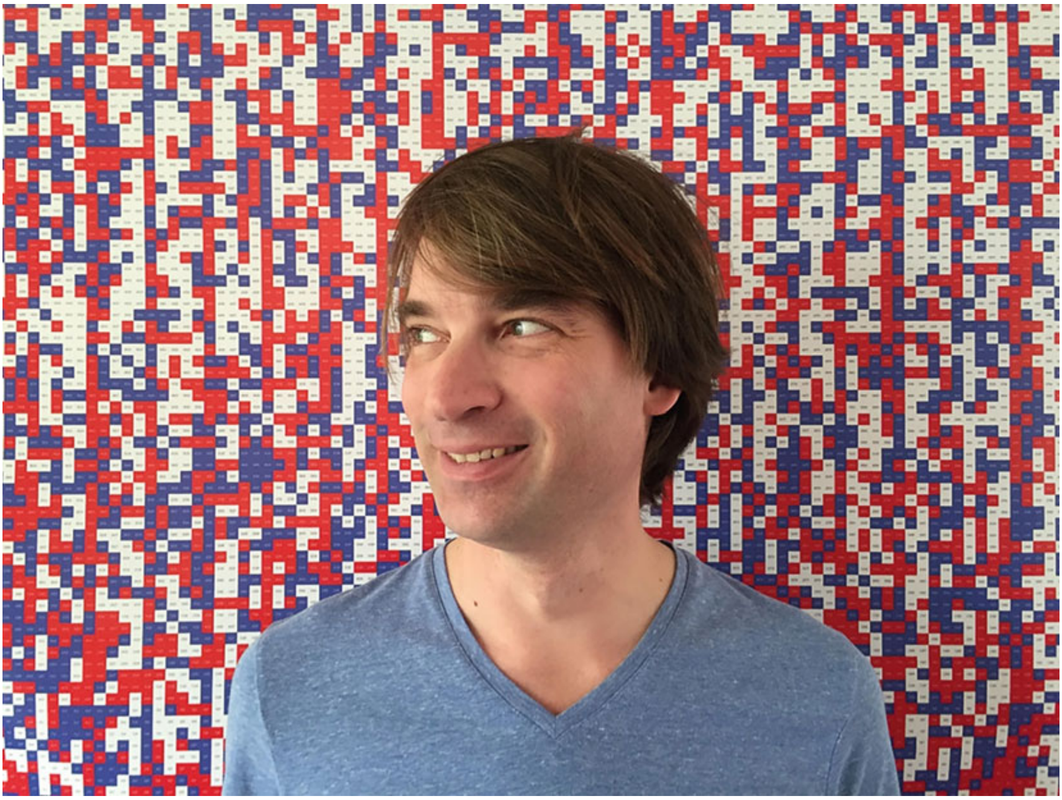
For $k \geq 3$ colors, the solution seems to be out of reach of today's computing. And yet, let us create a definition and pose a problem.

For a positive integer k, **the Pythagorean Number $P(k)$** is the smallest integer, such that any k-coloring of the set of all positive integers from 1 to $P(k)$ contains a monochromatic solution x, y, z of the Pythagorean equation $x^2 + y^2 = z^2$. In this language, Theorem 34.13 looks particularly concise:

Theorem 34.15 $P(2) = 7825$.

And now an open problem that begs to be posed and not likely to be solved in "real time" of our lives:

Open Problem 34.16 For each positive integer k, find $P(k)$ if it exists. At least find the upper and lower bounds of $P(k)$ for small values of $k \geq 3$.



Marijn J.H. Heule

34.5 Marienus Johannes Hendrikus “Marijn” Heule

Marijn Heule was born on 12 March 1979, in Rijnsburg, the Netherlands. He solved his first 100-piece puzzle before he could walk. Still, a knack for mathematics does not run in the family, even though his dad stubbornly claims all the credits because his final high school grade in math was 5 (out of 10), while his mom’s was a meager 4. More than anything, Marijn became passionate about mathematical problems thanks to inspiring teachers in high school – the Rijnlands Lyceum in the neighboring town of Oegstgeest.

Marijn went to Delft University of Technology, the Netherlands, to study Applied Computer Science. There, the undergraduate course Computational Logic and Satisfiability, taught by Hans van Maaren, piqued his interest. In short, Satisfiability (SAT) asks whether a propositional formula can be made true. Many problems in application areas such as verification, planning, and mathematics can be expressed as a SAT problem. There exist very efficient tools, called SAT solvers, to solve these problems automatically. The course got Marijn hooked on improving SAT solving. Hans van Maaren, who later became his Ph.D. advisor, brought Marijn with him to the SAT 2002 conference as an undergraduate, where that initial spark quickly ignited into a burning passion that is still the main focus of his research.

The topic of Marijn's Ph.D. research was look-ahead solvers. This type of solvers was practically abandoned by the community in 2004, when he started his Ph.D.Marijn was able to make look-ahead solving efficient for many applications and his solver won many awards at various international SAT competitions. Look-ahead techniques became the key enabler for his later research on parallelizing SAT solving.

During his Ph.D., Marijn worked as a co-editor with Armin Biere, Hans van Maaren, and Toby Walsh on the Handbook of Satisfiability. The first edition (900+ pages) was published in 2009 and has become the go-to reference for SAT research. The second edition (1500+ pages) was published in 2022.

In April 2011, Marijn started a postdoc position in the group of Armin Biere at the Johannes Kepler University in Linz, Austria. Although the duration was relatively short (7 months), it was a very inspiring visit that influenced much of his later work. Armin's SAT solvers have been the strongest in the field for over a decade. Armin considers them "experimentally correct": i.e., they produce the right answer on a million small problems. However, Marijn suspected that one of the techniques in Armin's top-tier solver could accidentally produce wrong answers, which he demonstrated by providing a small example. This started a long discussion on how to make sure that automated reasoning tools are reliably correct.

Note that SAT solvers are used in industry and academia to show that hardware and software designs are correct. Also, an increasing number of long-standing math problems are tackled using SAT solvers. A bug in a SAT solver could therefore mean that such results are false. It is therefore important to ensure that we can have full confidence in the results.

One option would be to formally verify the solvers. However, that would be an enormous effort, which would substantially reduce their performance and make it harder to update the tools. The alternative that Marijn and co-authors explored produces a certificate of correctness of each result. Such certificates can be checked with a simpler tool, which ideally is verified in a trustworthy system.

After Linz, Marijn became a Research Fellow at the University of Texas at Austin in 2012, where he made various contributions to make his certificate vision a reality: He invented a single rule that captures all reasoning in SAT solvers. He showed that existing solvers can easily produce certificates based on this rule and he implemented an efficient checker for these certificates. Today, all top-tier SAT solvers support certificate logging, thereby boosting confidence in the correctness of the results. Marijn and others have also been using these certificates as proofs for a range of long-standing open math problems, including the Boolean Pythagorean Triples problem (which Tim Gowers called "the most disgusting proof ever"), the Boolean Erdős discrepancy problem, and Keller's conjecture.

During his time at UT Austin, Marijn's research also focused on how to exploit the enormous potential of massive parallel computation in the cloud. Important initial results on parallelizing SAT dated back to his time in Linz. The Texas Advanced Computing Center (TACC) provided him with an abundance of resources to apply and improve these techniques on hard math problems.

Parallel computation and certificates played a crucial role in finding smaller and smaller unit-distance (UD) graphs with chromatic number 5. In short, the method works as follows: given a UD graph, ask the SAT solver to produce a 4-coloring. If the chromatic number is 5 or larger, this is impossible. The solver can produce a certificate of that. Marijn developed techniques to minimize certificates. To extract a subgraph with chromatic number 5, one takes the induced subgraph consisting of all vertices mentioned in the certificate. This allowed him to improve Aubrey de Grey’s 1581-vertex result down to an 874-vertex result in a few days. By improving the certificate minimization techniques, Marijn was able to get it down to 510 vertices in a competition with Jaan Parts.

Currently, Marijn is an Associate Professor at Carnegie Mellon University (CMU), where he started in August 2019. One of his first results at CMU was the resolution of Keller’s conjecture. This problem was open since 1930. He finally solved it together with Joshua Brakensiek, JohnMackey, and David Narvaez using SAT solving technology. Together with his Ph.D. student Bernardo Subercaseaux, he recently determined that the packing chromatic number of the infinite grid is 15, a problem that was open for two decades. Again, SAT solving was essential to tackle this problem. With Emre Yolcu and Scott Aaronson, he is currently working on a moonshot project to solve the Collatz Conjecture using SAT.

Ron Graham presented Marijn Heule with a check, a prize for Marijn’s achievement in the Pythagorean Triples Problem. We will meet Marijn again later in this book, where we discuss 5-chromatic unit-distance graphs.

Oliver Kullmann is a German computer scientist, who is a lecturer in computer science at Swansea University, United Kingdom.

Victor W. Marek is an American computer scientist, born and educated in Poland. He is Professor-Emeritus in the Department of Computer Science at the University of Kentucky, where he served as a professor for well over three decades.

Let us roll back and look at related problems. Inspired by the old K. F. Roth’s conjecture (published by Erdős in 1961) [E61.22, problem 16, p. 230], Paul Erdős, András Sárkösy, and Vera T. Sós proved in 1989 a number of results and posed a number of conjectures [ESS]. I would like to present here one of each, see others in [ESS].

Erdős–Sárkösy–Sós’s Theorem 34.17 [ESS, theorem 3]. Any k-coloring of the positive integers, $k \leq 3$, contains monochromatic pairs x, y such that $x + y = z^2$, for infinitely many integers z.

The authors then pose a conjecture:

Erdős–Sárkösy–Sós’s Conjecture 34.18 [ESS, problem 2]. Let $f(x)$ be a polynomial with integer coefficients, such that $f(a)$ is even for some integer a. Is it true that for any k-coloring of positive integers, the equation $x + y = f(b)$ has a monochromatic solution with $x \neq y$ for some b (for infinitely many b)?

On the first reading, you may be surprised by the condition on $f(a)$ to be even for some integer a. You could, however, easily construct a counterexample to Erdős–Sárkösy–Sós’s Conjecture 34.18 if this condition were not satisfied. Indeed, let $f(x) = 2x^2 + 1$, and color the integers in two colors, one color for even integers and another for the odd ones. Obviously, there are no monochromatic solutions.

In 2006, Ayman Khalfalah, professor of engineering in Alexandria, Egypt, and Endre Szemerédi [KSz] generalized Theorem 34.17 to all k.

Kalfalah–Szemerédi’s Theorem 34.19 [KSz]. For any positive integer k, there exists $N(k)$, such that any k-coloring of the initial segment of positive integers $[N(k)]$ contains a monochromatic pair x, y such that $x + y = z^2$, for an integer z.

Khalfalah and Szemerédi also proved Conjecture 34.18.

Kalfalah–Szemerédi's Generalized Theorem 34.19 [KSz]. Given a positive integer k and a polynomial with integer coefficients $f(x)$ such that $f(a)$ is even for some a; there exists $N(k)$, such that any k-coloring of the initial segment of positive integers $[N(k)]$ contains a monochromatic pair x, y, $x \neq y$, such that $x + y = f(z)$, for some integer z.

Endre Szemerédi is a witty speaker, with humor reminiscent of Paul Erdős' – this was on display on 4 April 2007, when as a Visiting Fellow at Princeton-Math I attended his presentation of these results at the Discrete Mathematics Seminar.

Chapter 35
Ramsey Theory Before Ramsey: Van der Waerden Tells the Story of Creation

It is like picking apples from a tree. If one has got an apple and another is hanging a little higher, it may happen that one knows: with a little more effort one can get that one too.

– B. L. van der Waerden [Wae18]

A thing of beauty is a joy for ever.

– John Keats, *Endymion*

The third result in *Ramsey Theory before Ramsey* was proved by Bartel Leendert van der Waerden in 1926 and published a year later.

Monochromatic Arithmetic Progressions Theorem 35.1 (Van der Waerden, 1927, [Wae2]). For any k, l, there is $W = W(k, l)$ such that any k-coloring of the initial array of positive integers [W] contains a monochromatic arithmetic progression of length l.

B. L. Van der Waerden proved this pioneering result while at Hamburg University and presented it the following year [1927] at the meeting of *D.M.V.*, *Die Deutsche Mathematiker Vereinigung* (The German Mathematical Society) in Berlin. The result became popular in Göttingen, as the 1928 Russian visitor of Göttingen Alexander Y. Khinchin noticed and later reported [Khi1], but the result's original publication [Wae2] in an obscure Dutch journal hardly helped its popularity. Only Issai Schur and his two students Alfred Brauer and Richard Rado learned about it and improved upon Van der Waerden's result almost immediately (details in the next chapter); and somewhat later, in 1936, Paul Erdős and Paul Turán commenced density considerations related to Van der Waerden's result [ET] (more in the next chapter). Only after World War II, when Alexander Yakovlevich Khinchin's book *Three Pearls of Number Theory* came out in Russian in 1947 [Khi1] and again in Russian in 1948 [Khi2], in German in 1951 [Khi3], and in English in 1952 [Khi4], the result became a classic and has remained one of the most striking "pearls" of mathematics. N.G. de Bruijn kindly shared with me his correspondence with Van der Waerden (and with Erdős). In his April 5, 1977, reply to de Bruijn's compliment, Van der Waerden wrote: "Your praise 'A thing of beauty is a joy for ever' pleases me." Let me second de Bruijn: the praise, taken (without credit) from the great British poet John Keats, is well deserved!

Now that the success of Khinchin's booklet had made the result classic, the latter merited a special attention and commentary by its solver. Van der Waerden obligated and in 1954

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_35

published an essay *Der Beweis der Vermutung von Baudet* with a more expressive English title *How the Proof of the Baudet's Conjecture Was Found.* This essay has appeared four times in German: twice in 1954 [Wae13], [Wae14], in 1965 [Wae16], and posthumously in 1998 [Wae26]; and once in English in 1971 [Wae18]. It is not only invaluable as a historical document. The essay delivers a vibrant portrait of mathematical invention in the making. Van der Waerden presents all critical ideas of the proof in the most clear and engaging way. Thanks to the permission granted to me by Professor Bartel L. van der Waerden in his letter [Wae24] and the permission by Academic Press, London, I am able to bring this delightful, lively essay [Wae18] to you here instead of presenting a formal "dehydrated" proof of the result.

Enjoy! From here on, this chapter belongs to the Author; **Bartel Leendert van der Waerden recollects:**

Once in 1926, while lunching with Emil Artin and Otto Schreier, I told them about a conjecture of the Dutch mathematician Baudet:

If the sequence of integers $1, 2, 3, \ldots$ *is divided into two classes, at least one of the classes contains an arithmetic progression of* l *terms:*

$$a, a+b, \ldots a+(l-1)b,$$

no matter how large the length l *is.*

After lunch, we went into Artin's office in the Mathematics Department of the University of Hamburg and tried to find proof. We drew some diagrams on the blackboard. We had what the Germans call "*Einfälle*": sudden ideas that flash into one's mind. Several times such new ideas gave the discussion a new turn, and one of the ideas finally led to the solution.

One of the main difficulties in the psychology of invention is that most mathematicians publish their results with condensed proofs but do not tell us how they found them. In many cases, they do not even remember their original ideas. Moreover, it is difficult to explain our vague ideas and tentative attempts in such a way others can understand them.[1] To myself I am accustomed to talking in short hints which I alone can understand. Explaining these hints to others requires making them more precise and thus changing their nature.

In the case of our discussion of Baudet's conjecture, the situation was much more favorable for a psychological analysis. All ideas we formed in our minds were at once put into words and explained by little drawings on the blackboard. We represented the integers $1, 2, 3, \ldots$ in the two classes by means of vertical strokes on two parallel lines. Whatever one makes explicit and draws is much easier to remember and to reproduce than mere thoughts. Hence, this discussion between Artin, Schreier, and myself offers a unique opportunity for analyzing the process of mathematical thinking.

---

[1] And when mathematicians attempt to be subjective and include thoughts and emotions of the emergence of their results, most of journal editors, these priests of gloom and doom, would mercilessly cut manuscripts to bring them to an 'objective' and relentless theorem-proof style. – A.S.

It was clear to us from the very beginning that the case $l = 2$ is trivial. One need not even consider the infinite sequence of integers; it is sufficient to consider the three integers 1, 2, 3. If they are divided into two classes, one of the classes contains a pair of numbers (in arithmetic progression).

The next case we considered was $l = 3$. In this case, too, it is not necessary to consider all integers: it suffices to take the integers from 1 to 9. The numbers 1 to 8 can be divided, in several ways, into 2 classes without obtaining an arithmetic progression of 3 terms in one class, e.g. like this:

1 2		5 6		in the first class
	3 4		7 8	in the second class.

However, in any one of these cases, the number 9 cannot escape. If we put it into the first class, we have the progression 1 5 9, and if we put it into the second class, we get the progression 7 8 9. Just so in all other possible cases. I had observed this already the day before.

Next, Schreier asked if Baudet's conjecture is at all true for a certain value of l, is it always possible to find an integer $N(l)$ such that the conjecture holds already for the segment

$$1\,2\,3\ldots N(l),$$

in the sense that every division of this segment into two classes yields an arithmetic progression of length l in one class?

Schreier himself found the answer: it was Yes. If Baudet's conjecture holds for a fixed value of l, it is possible to find an N such that the conjecture holds already for the segment 1 2 ... N. This was proved by a well-known procedure from set theory, the "diagonal procedure." The argument is as follows.

If no such N existed, then for every N there would be a division D_N of the numbers from 1 to N into 2 classes such that no class contains an arithmetic progression of length l. Thus, one could obtain an infinite sequence

$$D_1\, D_2 \ldots$$

of such divisions. The number 1 lies, in every one of these divisions, in one of the two classes. Hence, it happens an infinity of times that 1 is in the same (first or second) class, and an infinite sequence $D'_1, D'_2, \ldots$ exists such that in all these divisions 1 is in the same class, say in class number i_1 ($i_1 = 1$ or 2).

In the divisions $D'_2, D'_3, \ldots$, the number 2 belongs to one of the two classes. Hence, by the same argument, an infinite subsequence $D''_2, D''_3, \ldots$ exists such that 2 is always in the same, i_2th class.

And so on. For every n, one obtains a subsequence of divisions

$$D_n^{(n)}, \qquad D_{n+1}^{(n)}, \qquad \ldots$$

such that in all these divisions, the integers 1, 2, ..., n are always in the same classes:

1 in class i_1
2 in class i_2
...
n in class i_n.

Next, one can form a “diagonal division” $D\,D$ of all integers 1, 2, 3, ... in which 1 lies in class i_1, 2 in class i_2, and so on. In this division, the number n lies in the same class as in the division $D_n^{(n)}$, hence the name “diagonal procedure.”

In this division $D\,D$, no arithmetic progression of length l could exist in which all terms belong to the same class. For if it existed, it would exist already in $D_n^{(n)}$, i.e., in one of the original divisions. But we have assumed Baudet’s conjecture to be true for the sequence of integers 1 2 3 ... and for this particular value of l. Thus, we obtain a contradiction.

From this point onward, we tried to prove the “strong conjecture,” as we called it, for a finite segment from 1 to $N(l)$, i.e., we tried to find a number $N(l)$ having the desired property. For $l = 2$ and $l = 3$, such numbers had been found already:

$$N(2) = 3,\ N(3) = 9.$$

So we tried to go from $l - 1$ to l. For this induction proof, the replacement of the original conjecture by a stronger one is a definite advantage, as Artin rightly remarked. If one can assume for $l - 1$ the existence of a finite bound $N(l - 1)$, one has a chance to find a proof for the next number l.

Next, Artin observed: If the strong conjecture is true for 2 classes and for all values of l, it must be true for an arbitrary number of classes, say for k classes. To prove this assertion, he first proposed k to be 4. The 4 classes can be grouped into 2 and 2. This gives us a rough division of the integers into 2 big classes, every big class consisting of 2 smaller classes. In one of the big classes, an arithmetic progression of $N(l)$ terms exists. The terms of this progression can be numbered from 1 to $N(l)$. These numbers are now divided into two smaller classes, and hence in one of the smaller classes, an arithmetic progression of length l exists. Thus, if the strong conjecture is true for 2 classes, it is also true for 4 classes. By the same argument, one finds that it also holds for 8 classes, etc., hence, for any number of classes $k = 2^n$. But if it holds for $k = 2^n$, it also holds for every $k \leq 2^n$ because we may always add a few empty classes. Hence, if Baudet’s conjecture holds for 2 classes, it also holds, even in the strong form, for an arbitrary number of classes.

We now tried to prove the “strong conjecture” for arbitrary k and l by induction form $l - 1$ to l. This means: we tried to find a bound $N = N(l, k)$ such that, if the integers from 1 to N are divided into k classes, one of the classes contains an arithmetic progression of length l.

Artin expected – and he proved right – that the generalization from 2 to k classes would be an advantage in the induction proof. For, he argued, we might now try to prove the strong conjecture for an arbitrary *fixed* value of k and for length l under the induction hypothesis that it holds for *all* k and for length $l - 1$. This means: we have a very strong induction hypothesis to start with, which is a definite advantage.

Following the line indicated by Artin, we now tried to prove Baudet’s conjecture for 2 classes and for progressions of length l, assuming the strong conjecture to hold for all k for progressions of length $(l - 1)$.

Next, Artin had another very good idea. If the integers 1, 2, . . . are divided into 2 classes, blocks of (say) 3 successive integers are automatically partitioned into $2^3 = 8$ classes. For each of the 3 numbers within, the block can lie in the first or second class and this gives us 8 possibilities for the whole block. Now the blocks of 3 successive integers can be numbered: block number n consists of the integers n, $n + 1$, $n + 2$. If the blocks are partitioned into 8 classes, their initial numbers n are also partitioned into 8 classes, and to this partition, we can apply the induction hypothesis. Thus, we obtain the following result: among sufficiently many successive blocks, we can find an arithmetic progression of $(l - 1)$ blocks all in the same class. The pattern of the distribution of integers over the two classes in the first block will be repeated, exactly as it is, in the other $(l - 2)$ blocks.

The same holds for blocks of arbitrary length m, each consisting of m successive numbers

$$n, n + 1, \ldots, n + m - 1.$$

The number of classes for those blocks is 2^m. Once more one can obtain arithmetic progressions of $(l - 1)$ blocks in the same class, with exact repetition of the pattern in the first block. Moreover, if the blocks are long enough, we can also find arithmetic progressions of $(l - 1)$ integers within each block.

In the simplest case $l = 2$, the conjecture is certainly true for all k, for if the integers from 1 to $k + 1$ are divided into k classes, there must be two integers in one of the classes. This is Dirichlet's "box principle"[2]: if $k + 1$, objects are in k boxes, one of the boxes contains two of them. A very useful principle in Number Theory.

Thus, starting with the obvious case $l = 2$, we tried to treat the case of 2 classes and $l = 3$ (although this case had been dealt with already by an enumeration of all possible cases). We represented the integers in the two classes by small vertical strokes on two parallel lines, as in Fig. 35.1.

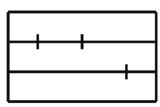


Fig. 35.1

Among three successive integers, there are always two in the same class, by the induction hypothesis, i.e., in this case, by the "box principle." Now consider a block of 5 successive integers. Among the first three, there are two in the same class; this gives us an arithmetic progression of length 2. The third term of this progression still lies within the block of 5. If it is in the same class as the first two terms, we have in this class a progression of length 3, as desired. Therefore, we may suppose that the third term lies in the other class, and we have, within every block of 5, a pattern like the one of Fig. 35.1.

I was drawing such blocks on the blackboard, and I thought: There are $2^5 = 32$ classes of blocks of 5; hence, among 33 successive blocks of 5, there must be 2 blocks in the same class. In the first of these blocks, a pattern like the one of Fig. 35.1 exists, and in the second block of 5, this pattern is exactly repeated (Fig. 35.2).

[2]In the USA, it is usually called the Pigeonhole Principle.

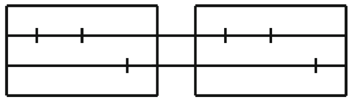


Fig. 35.2

What we wanted to construct were progressions of length 3. Hence, I drew one more block at the same distance from the second block as the second from the first, and I drew three strokes in the third block in the same position as the strokes in the first and second block (Fig. 35.3).

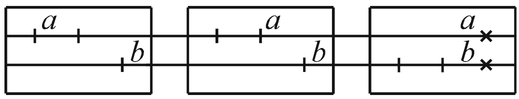


Fig. 35.3

The third of these strokes represents an integer, which may be in the first or second class. If it is in the first, we have in this class an arithmetic progression *a a a* (Fig. 35.3). If it is in the second class, we have in this class a progression *b b b*. Hence, we have in any case within the block of integers from 1 to $5 + 32 + 32 = 69$, an arithmetic progression of 3 terms in one class.

After having found this proof in the special case $k = 2$ and $l = 3$, I explained it to Artin and Schreier. I felt sure that the same proof would work in the general case. They did not believe it, and so I proceeded to present the proof in the next higher case $k = 3$, $l = 3$.

Instead of considering blocks of $3 + 2 = 5$, I now considered blocks of $4 + 3 = 7$ successive integers. Since the first four numbers of such a block are distributed among 3 classes, two of them must belong to the same class. The third term of the arithmetic progression starting with these two terms still belongs to the same block of 7. If the third term lies in the same class, we have a progression of length 3 in this class. Hence, we may suppose the third term to lie in another class. Thus, we obtain, in every block of seven, a pattern like the one in the first small block of Fig. 35.4.

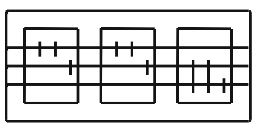


Fig. 35.4

The blocks of 7 are partitioned into 3^7 classes. Hence among $3^7 + 1$ successive blocks of 7, there are two belonging to the same class. In the first block, we have three integers in arithmetic progression, two of which belong to the same class, and this pattern repeats itself in the second block. If the second block is shifted once more over the same distance, one contains 3 blocks forming an arithmetic progression of blocks, as shown in Fig. 35.4.

In the third block, I drew 3 strokes in positions corresponding to the 3 strokes in the first or second block and I considered the possibilities for the third of these strokes. If it falls into the first or second class, we have an arithmetic progression of length 3 in the same class, by the same argument as before; but now the third stroke can escape into the third class. Thus, we obtain the pattern drawn in Fig. 35.4.

In every large block of $3^7 + 3^7 + 7 = h$ successive integers, we have such a pattern. Now the large blocks of h are divided into 3^h classes. Hence, among $3^h + 1$ successive large blocks, there are two belonging to the same class. Drawing the small blocks within the large ones, I obtained the picture of Fig. 35.5.

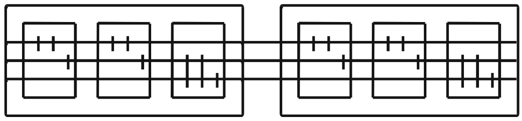


Fig. 35.5

Now shifting the second large block over the same distance and considering the third stroke in the third small block in the third large block, I showed that it cannot escape anymore. If it lies in the first class, there is a progression *a a a* in the first class. If it lies in the second class, there is a progression *b b b* in that class, and if in the third class, a progression *c c c* in that class. (Fig. 35.6).

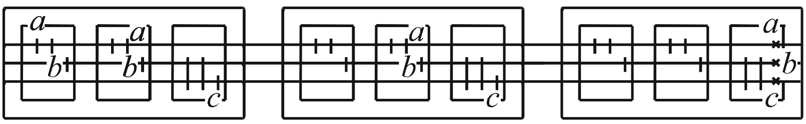


Fig. 35.6

After this, all of us agreed that the same kind of proof could be given for arbitrary k. However, Artin and Schreier still wanted to see the case $l = 4$.

As before, I first considered the case of 2 classes. For this case, I had already proved that among sufficiently many, say n, successive integers, there is a progression of 3 terms in the same class. We may suppose n to be odd. The distance between the first and the last term of the progression is $(n - 1)$ at most; hence, the difference between two successive terms is $\frac{1}{2}(n-1)$ at most. Now consider the fourth term of the same progression. All four terms lie within a block of

$$g = n + \frac{1}{2}(n-1)$$

successive integers. If the fourth term belongs to the same class as the other three, we are satisfied. Suppose it lies in the other class; then we have the pattern of Fig. 35.7.

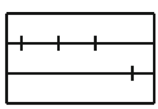


Fig. 35.7

In every block of g successive integers, such a pattern must occur. Now the blocks of g are divided into 2^g classes. Hence among sufficiently many, say $N(3, 2^g)$ blocks of length g, there are three blocks in arithmetical progression belonging to the same class. The pattern in the first block is exactly repeated in the second and third block (Fig. 35.8).

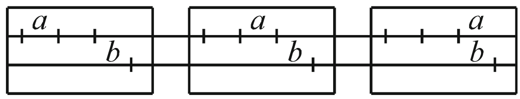


Fig. 35.8

Adding a fourth block to this progression, I easily obtained a progression *a a a a* in the first or *b b b b* in the second class.

Now, it was clear to every one of us that the induction proof from $(l - 1)$ to l works for arbitrary l and for any fixed value of k. Hence if Baudet's strong conjecture is true for length $(l - 1)$ and all k, it is also true for l and any k. Since it is true for $l = 2$, its truth follows quite generally.

Analyzing this record, one can clearly distinguish a succession of sudden ideas, which gave the discussion a new turn every time.

1. The first was Schreier's idea of restricting oneself to a finite segment from 1 to N. This idea was fundamental to the whole proof.
2. The second idea was to try an induction from $l - 1$ to l. This was quite a natural idea because the case $l = 2$ was obvious and the case $l = 3$ could be solved by enumerating all possible cases.
3. Artin proved: If the strong conjecture is true for 2 classes, it is also true for 4 classes. In his proof, another idea was implicit, viz.: If the conjecture is true for a segment of all integers from 1 to N, it is also true for any arithmetical progression of length N

$$a, a + b, \ldots, a + (N - 1)b$$

because the terms of this progression can be numbered by the integers 1 to N. This is also a central idea in the proof.

4. Next, Artin said: in an induction, it is always an advantage to have a strong induction hypothesis to start with. Therefore, let us start with the assumption that the conjecture holds for progressions of length $(l - 1)$ and for *all* k, and try to prove the conjecture for progressions of length l and for *one* value of k, say $k = 2$**.** Thus the plan for the proof was devised.
5. The next idea, which also came from Artin, was of decisive importance. He said: we can apply the induction hypothesis not only to single integers but also to blocks, for they too are divided into classes. Thus, we are sure that whole blocks are repeated $(l - 1)$ times.
6. After this, it was only natural to consider progressions of $(l - 1)$ integers within the blocks and to try to extend these progressions of length $(l - 1)$ to progressions of length l. The simplest non-trivial case is $l = 3$, and thus I was led, quite naturally, to consider patterns like the one of Fig. 35.2.

7. This pattern still does not contain a progression of length 3 in one class. Therefore, it was necessary to extend the progression of length 2 occurring in the second class in Fig. 35.2 to a progression of length 3. Hence, I extended the pattern of Fig. 35.2 by drawing the third block of Fig. 35.3, and I considered the third term of the progression *b b b*. As soon as attention was focused upon this term, it was clear that it cannot escape from forming an arithmetic progression of length 3 in the first or second class.

This final idea was accompanied by a feeling of complete certainty. I felt quite sure that this method of proof would work for arbitrary *k* and *l*. I cannot explain this feeling; I can only say that mathematicians often have such a conviction. When a decisive idea comes to our mind, we feel that we have the whole proof we are looking for: we have only to work it out in detail.

However, I can explain, to a certain extent, why Artin and Schreier did not feel so sure. They saw only the result: the presence of the progression *a a a* in the first class or *b b b* in the second one, but I had discovered a method for finding such progressions, and I was convinced that this method would work in higher cases as well.

It is like picking apples from a tree. If one has got an apple and another is hanging a little higher, it may happen that one knows: with a little more effort one can get that one too. The man standing next to me only sees that I have just got the first apple, and he is in doubt whether I can get the other too, but I myself have not only got the apple but I also have a feeling of the movement that enabled me to pick it.

The feeling that a method of proof can be carried over to the other cases is still sometimes deceptive. Often the higher cases offer additional difficulties. Still, feelings of this kind are extremely useful in mathematical research.

Finding the proof of Baudet's conjecture was a good example of teamwork. Each of the three of us contributed essential ideas. After the discussion with Artin and Schreier, I worked out the details of the proof and published it in *Nieuw Archief voor Wiskunde* **15**, p. 212 (1927). (Interesting applications and generalizations of the theorem proved in my paper were given by Richard Rado[3]).

A. J. Khinchin included the theorem among his "Three Pearls of the Theory of Numbers" (1952) and published a proof due to M. A. Lukomskaja, which is in all essentials the same as mine, the only difference being that in her proof the blocks are required to be non-overlapping. ■

* * *

Van der Waerden, assisted by Emil Artin and Otto Schreier, actually proved a "strong conjecture" as they called the result:

Van der Waerden's Theorem Strong Version 35.1 For all positive integers n and r, there exists an integer $W=W(n, r)$ such that if the initial set of integers $[W] = \{1, 2, \ldots, W\}$ is colored in r colors, then there exists a monochromatic n-term arithmetic progression.

It is natural to inquire whether the finality of n and of r is essential. Prove first that the finality of the length n of the guaranteed arithmetic progression is essential:

---

[3] R. Rado: Studien zur Kombinatorik, Ph.D Thesis Berlin 1931, *Math. Zeitschr.* 36, p. 424. Verallgemeinerung eines Satzes von van der Waerden, *Sitzungsber. preuss. Akad.,* Berlin 1933, p. 589. Note on Combinatorial Analysis, *Proc. London Math. Soc.* (2)48, p. 122.

Problem 35.2 Color the set of all positive integers in two colors in a way that forbids infinite monochromatic arithmetic progressions.

Of course, the finality of the number of colors is essential, for otherwise we can color each integer in its own color and thus exclude even length two arithmetic progressions. However, Paul Erdős and Ronald L. Graham proved a nice "consolation" result [EG]. We will say that a sequence is *representative* if each term is colored in a different color.

Theorem 35.3 (Erdős–Graham, [EG]). Any coloring of the positive integers in infinitely many colors contains arbitrarily long monochromatic or representative arithmetic progressions.

Hint Peek at the Szemerédi Theorem in Chapter 37 – and use it. ■

I have been unable to explain why the leader of the new field, Ramsey Theory, Paul Erdős almost universally quoted Van der Waerden's result as addressing only the case of two colors (see, for example, [E57.13], [E61.22], [E71.13], [E73.21], [E76.35], [E81.16], [E80.03], [E83.03], [E85.33], [E89.32], etc.). Is it because Van der Waerden's paper opens with Baudet's Conjecture for two colors, or because Erdős wanted, as he often did, to gain insight into the simplest case first (and then forgot about the general case)?

Besides Issai Schur and his former PhD students Alfred Brauer and Richard Rado, and also Erdős and Turán, practically nobody seemed to have appreciated and furthered Van der Waerden's proof during many years following Van der Waerden's publication. There was, however, one exception – a pair of mathematicians, who published on Van der Waerden's proof very shortly after its publication. Their paper, submitted for publication to the *Japanese Journal of Mathematics* on March 25, 1930, was somehow noticed and cited by Paul Erdős and Ronald L. Graham in their very fine and very hard-to-find 1980 problem book [EG]. I got my copy when Ron sent me one on the request of Paul Erdős.

Erdős and Graham characterized the Japanese paper as "an easy consequence of Van der Waerden's Theorem." In fact, it is much more!

Chapter 36
A Japanese Insight into Baudet–Schur–Van der Waerden's Theorem

Check for updates

The great surprise is, [KM] was published by the two Japanese mathematicians Sôichi Kakeya and Seigo Morimoto in 1930, much earlier than Erdős and Turán's 1936 paper. How did they get a hold of the little-read Dutch journal, where Van der Waerden published his result just 3 years earlier in German? The authors do misspell the name of Baudet everywhere, even in the title: *On a Theorem of M. Bandet [sic] and van der Waerden.* But they were first to recognize that credit is due to both mathematicians, Baudet for creating the conjecture, and to Van der Waerden for proving it. Without the conjecture, Van der Waerden would have had nothing to prove!

The authors of [KM] prove that in fact the statements of Theorem 35.1 and Theorem 36.1 are equivalent! In my opinion, Theorem 36.1 explains the essence of the celebrated Theorem 35.1 better than anything ever has.

The Kakeya–Morimoto Theorem 36.1 ([KM], 1930). If $A = \{a_1, a_2, \ldots\}$ is an increasing infinite sequence of integers with a bounded growth $a_{k+1} - a_k < d$, then A contains arbitrarily long arithmetic progressions.

Proof Let me present my version of the authors' proof; I think it is easier to read. The fact that differences $a_{k+1} - a_k$ are bounded by d suggests a d-coloring of the set of all positive integers in colors 0, 1, . . ., d–1 as follows: given a positive integer n, find the smallest term a in the sequence A such that $0 \leq a - n$. Obviously, $a - n < d$. We then color n in the color of $a - n$. By the Baudet–Schur–Van der Waerden's theorem, for any length l, there is a monochromatic arithmetic progression b_1, b_2,. . ., b_l of color, say, i. But then the progression b_1+i, b_2+i, . . ., b_l+i is both arithmetic and is entirely contained in A. ■

Kakeya and Morimoto then show that Baudet–Schur–Van der Waerden's theorem follows from the statement of their Theorem 36.1.

Theorem 36.2 (Kakeya–Morimoto, 1930, [KM]). If any increasing infinite sequence of integers $A = \{a_1, a_2, \ldots\}$ with $a_{k+1} - a_k < d$ for a fixed positive integer d, contains arithmetic progressions of l terms for any positive integer l, then the Baudet–Schur–Van der Waerden Theorem holds.

Proof I am presenting the authors' proof, in which I eliminated a number of typos in symbols and clarified the arguments. There is a number $n = n(l,d)$ such that any n consecutive terms of

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_36

A contain an arithmetic progression of l terms, for otherwise we can create a countable number of sequences:

$$a_{n1}, a_{n2}, \ldots, a_{nn}, \quad \ldots \tag{B}$$

such that $a_{n1} = 1$, $a_{n,\, i\,+\,1} - a_{n,\, i} < d$ for any n and i and none of them contains an arithmetic progression of l terms.

Since the number of possible values that $a_{1,2}$ can take is finite, there is an a_2 that appears infinite times as a_{n2}. Let us now keep only those arithmetic progressions from (B) that have a_2 as their second term. Continuing in the same fashion, we identify an a_3 that appears as the third term in infinitely many sequences (B), and so on. We end up with the infinite sequence:

$$1 = a_1, a_2, a_3, \ldots \ \text{ satisfying } \ a_{i+1} - a_i < d \tag{C}$$

whose first terms coincide with the terms of in the same position in some sequences (B). Thus, sequence (C) does not contain any arithmetic progression of l terms, which contradicts our assumption.

We can now prove Baudet–Schur–Van der Waerden's theorem by induction. Assume that it holds for k colors, i.e., the initial array of integers $1, 2, \ldots, n(k,l)$ colored in k colors contains a monochromatic arithmetic progression of l terms. Let the following array of integers be colored in $k + 1$ colors:

$$1, 2, 3, \ldots, v, \text{where } v = n(l, n(l,k)) \times n(l,k).$$

If the $(k + 1)$th color does not contain any sequence of $n(l, k)$ consecutive integers, then these integers are colored in the first k colors, and by the inductive assumption, one of these colors contains an arithmetic progression of l terms.

Otherwise, the $(k + 1)$th color has at least $h(l, n(l, k))$ terms with the difference between any two consecutive terms less than $n(l, k)$. Theorem 36.1 guarantees the existence of the desired monochromatic arithmetic progression in color $(k + 1)$. ■

Corollary 36.3 The Kakeya–Morimoto Theorem 36.1 is equivalent to the Baudet–Schur–Van der Waerden Theorem 35.1.

Kakeya and Morimoto also construct a lovely example, showing that in their Theorem 36.1, the words "arbitrarily long arithmetic progressions" cannot be replaced by "infinite arithmetic progressions."

Try to come up with a counterexample on your own. Then compare it to the following construction.

Counterexample 36.4 (Kakeya–Morimoto, 1930, [KM]). There is an increasing infinite sequence $A = \{a_1, a_2, \ldots\}$ of integers with $a_{k\,+\,1} - a_k \le 2$, such that A does not contain an infinite arithmetic progression.

Construction An infinite arithmetic progression P of integers is defined by an ordered pair (m, n) of integers, where m is the first term and positive n is the constant difference of P. Therefore, the set of all such progressions is countable, i.e., can be enumerated by positive integers to look like $P_1, P_2, \ldots, P_n, \ldots$

Now we construct a sequence S as follows. For the first term s_1 of S, we pick the first term of P_1. For the second term s_2 of S, we choose a term of P_2, which is greater than $s_1 + 1$, and so on. Now consider the increasing sequence A of all positive integers from which we removed all terms of the sequence S. Clearly, A does not contain any infinite arithmetic progression because it is missing a term from each arithmetic progression and satisfies the condition $a_{k+1} - a_k \leq 2$. ■

As you may know, I created and have been running for 39 years the Soifer (formerly Colorado) Mathematical Olympiad for middle and high school Olympians of my state. I like to extract Olympiad-style beautiful ideas from research mathematics and offer them to our young colleagues at a 5-problem 4-hour written annual competition. There is a reverse influence as well: some Olympiad problems inspire research.

The construction of Kakeya and Marimoto's Counterexample 36.4 had such an Olympiad flavor that I decided to use it as the hardest problem 5 in the 37th Soifer Mathematical Olympiad (SMO-37), October 2021. In the process, we found a much simpler counterexample than the one published by the two Japanese mathematicians. The only person, whom I show the problems in advance, has been Robert "Bob" Ewell, an ED and Air Force retired lieutenant colonel. In interaction with Bob, the problem grew to contain 3 parts: A, B, and C. I will show you problem 5 as evolution of ideas.

Second Proof of 36.4 (SMO-37, Problem 5A). Start the increasing sequence S with one positive odd integer s_1, followed by two consecutive evens starting with $s_1 + 1$, then by three consecutive odds starting with the previous even +1, etc. Assume S contains an infinite AP (arithmetic progression), call it S_1, of constant difference D. At some point, S will have more than $2D$ consecutive odd numbers, containing two consecutive terms of S_1, thus making all terms of S_1 odd. But further on S_1 will have an even integer: A contradiction. ■

Bob Ewell found the idea of this simple counterexample – so I had to "tighten the nuts" of the problem to disallow Bob's solution. Thus, Problem 5B was born.

Problem 5B (SMO-37). For an increasing sequence A of positive integers, A_n denotes the number of terms of A that do not exceed n. We say that the sequence's density $D(A) = 1$ if the ratio A_n/n becomes as close to 1 as we are pleased as n increases without bound. Is there a sequence A with $D(A) = 1$ that does not contain an infinite arithmetic progression?

Solution of 5B by Bob Ewell There is such a sequence.

Let A be the sequence of all positive integers except:

[1]. All of the integers between 1 and 10
[2]. The first ½ of the integers between 11 and 100
[3]. The first ¼ of the integers between 101 and 1000
[4]. The first $^1/_8$ of the integers between 1001 and 10,000
[5]. . . .
[6]. The first $1/2^k$ of the integers between 10^k and 10^{k+1}
[7]. . . .

Note that the number of integers removed at each power k of 10 (except the first 10) is $9 \times 10^k/2^k = 9 \times 5^k$. That is, the "holes" increase without bound. Therefore, no matter where an arithmetic sequence starts and no matter how big its constant difference is, the sequence will run into a hole too big to cross. It is easy to show that the density D(A) = 1. ■

Bob solved Problem 5B, thus "forcing" me to create Problem 5C to stop Bob's successes. :)

Problem 5C (Soifer, SMO-37). We call an increasing sequence A of positive integers *super dense* if for any positive integer n, A contains all integers from 1 through 10^n except at most n integers, and the differences between the consecutive integers excepted from A are strictly increasing. Is there a super dense sequence A that does not contain an infinite arithmetic progression?

Solution of 5B and 5C As in the construction of Counterexample 36.4, we enumerate all infinite arithmetic progressions of positive integers to look like $P_1, P_2, \ldots, P_n, \ldots$ and construct a sequence S as follows. For the first term s_1 of S, we pick the first term of P_1. For the second term s_2 of S, we pick the term of P_2 that is no less than $s_1 + 10$. For the third term s_3 of S, we choose a term of P_3, which is no less than $s_2 + 100$, and so on. Now consider the increasing sequence A of *all* positive integers from which we removed all the terms of the sequence S. Clearly, A does not contain any infinite arithmetic progression because it is missing a term from each of them. Its density $D(A)$ is the limit of $(10^n - n)/10^n = 1 - n/10^n$ as n increases without bound, which is obviously 1. ■

Notice *We can explicitly calculate the sequence* S if, for example, we use the following fantastic mapping:

$$f(a, b) = 2^{a-1}(2b - 1)$$

of ordered pairs of positive integers onto positive integers. Let now a be the first term and b the constant difference of an AP. Every positive integer can be uniquely expressed as a power of 2 times an odd integer; thus, each positive integer has a unique pair that maps into it. This inverse function f^{-1} maps an integer $2^{a-1}(2b - 1)$ into the pair (a, b), and we easily construct the terms of the sequence S:

$1 = 2^{1-1}(2{\cdot}1 - 1) \Rightarrow (1, 1)\ s_1 = 1$
$2 = 2^{2-1}(2{\cdot}1 - 1) \Rightarrow (2, 1)\ s_2 = 11$
$3 = 2^{1-1}(2{\cdot}2 - 1) \Rightarrow (1, 2)\ s_3 = 111$
$4 = 2^{3-1}(2{\cdot}1 - 1) \Rightarrow (3, 1)\ s_4 = 1111$
$5 = 2^{1-1}(2{\cdot}3 - 1) \Rightarrow (1, 3)\ s_5 = 11113$
$6 = 2^{2-1}(2{\cdot}2 - 1) \Rightarrow (2, 2)\ s_6 = 111130$
$7 = 2^{1-1}(2{\cdot}4 - 1) \Rightarrow (1, 4)\ s_7 = 1111301$
. . . .

■

Chapter 37
Whose Conjecture Did Van der Waerden Prove? Two Lives Between Two Wars: Issai Schur and Pierre Joseph Henry Baudet

As far as your advice to leave priority matter . . . alone, it is my opinion that the tiniest moral matter is more important than all of science, and that one can only maintain the moral quality of the world by standing up to any immoral project.
– L. E. J. Brouwer (From the February 24, 1929, *letter to H. Hahn, quoted from [Dal2], p. 651.)*

37.1 Prologue

Bartel L. van der Waerden credits "Baudet" [sic] with conjecturing the result about monochromatic arithmetic progressions. Decades later, Van der Waerden gives a most insightful story of the birth of his proof, which I have reproduced for you in Chapter 35. As I enumerated there, the "Story of Creation" appears four times in German: twice in 1954 [Wae13], [Wae14], in 1965 [Wae16], posthumously in 1998 [Wae26]; and once in English in 1971 [Wae18]. In these publications, Van der Waerden extends the credit for the conjecture to "the Dutch mathematician Baudet," still without the first name or even initials. Biographers of Van der Waerden faithfully follow him with crediting "Baudet" for the conjecture (see [Fre], [FTW], [Per], and [Bru1]).

On the other hand, Ronald L. Graham, Bruce L. Rothschild, and Joel H. Spencer in their definitive monograph [GRS1], [GRS2] cite Alfred Brauer [Bra2], [Bra3] in taking the credit for the conjecture away from Baudet and giving it to Issai Schur. Schur is also credited by Hillel Furstenberg in his pioneering paper [Fur1]. Consequently, practically all mathematicians have uncritically quoted or simply copied credit from [GRS1], [GRS2], or [Fur1].

False attributions are never pleasant. One may wonder, however, why the authorship of *this* conjecture is so extremely important that I have most thoroughly researched it and am dedicating this whole chapter to my findings. This is so because we have here, for the third time in the history of mathematics,[1] a totally new Ramseyan type question, quite uncommon in mathematics of the time: "if a system is partitioned arbitrarily into a finite number of

[1] First two being the Hilbert Theorem of 1892 [Hil] and the Schur Theorem of 1916 [Sch] – see Chapters 33 and 34.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_37

subsystems, then at least one subsystem possesses a certain specified property."[2] It was a major achievement indeed to envision and conjecture such a result, which allowed Ramsey Theory to be born. But whose achievement was it, Baudet's or Schur's? And who was "Baudet" anyway? My early investigative reports appeared in mid 1990s [Soi10], [Soi11], and [Soi12]. Let us look at the more complete evidence that I have been able to assemble to date.

37.2 Issai Schur

Germany has surely been one of the best countries in preserving documents through all the cataclysms that have befallen on and arguably were triggered by this land. Issai Schur's personnel file, and personnel forms it contains, is preserved in the Archive of the University Library of the Humboldt University at Berlin.[3] Let us make a good use of them.

[2] Leon Mirsky in [Mir], in reference to the Schur Theorem.

[3] Archive of Humboldt University at Berlin, documents UK–Sch 342, Bd. I, Bl. 1, 1R, and 2R, 3.

Young Issai Schur. (Courtesy of Hilde Abelin-Schur, his daughter)

Issai Schur was born on January 10, 1875, in the Russian city of Mogilyov (presently in Belarus) in the family of the merchant Moses Schur and Golde Landau. Being a Jew, Issai could not enroll in any Russian university. At 13, he went to live with his older sister in Libau, Russia (now Latvia), in order to attend the German language Nicolai-Gymnasium (1888–1894). That prepared him for entering a German university in 1894. In Berlin, on September 2, 1906, Issai Schur married Regina Malka Frumkin, born January 8, 1881, in Kowno (presently Kaunas, Lithuania), a medical doctor, also Jewish, and apparently an émigré from Russia. Issai Schur, who originally filled the personnel form in his hand, likely in 1916 (it was later updated, probably by clerks), on the line "Arian" promptly put "*nicht*" for himself and "*nicht*" for his wife. The happy and lasting marriage produced two children, Georg (named in honor of Schur's mentor, the celebrated algebraist F. Georg Frobenius), born on July 25, 1907 and Hilde, born on March 15, 1911.

Issai Schur. (Courtesy of his daughter, Hilde Abelin–Schur)

Issai Schur gave most of his life to the University of Berlin, first as a student (1894–1901; Ph.D. in Mathematics and Physics summa cum laude, November 27, 1901); then as a *Privatdozent* (1903–1909), *ausserordentlischer Professor* (equivalent to an associate professor, December 23, 1909 – April 21, 1913 and again April 1, 1916 – April 1, 1919); and *Ordinarius* (equivalent to a full professor, April 1, 1919 – September 30, 1935).[4] On April 1, 1921, Schur was appointed to the *Ordinarius* chair of Prof. Dr. Schottky with a very respectable compensation: 16200 marks base salary; plus local adjustment and family allowances; plus 5000 marks for lecturing the minimum of 8 hours a week. The only three years away from Berlin, 1913–1916, Schur spent at the University of Bonn. These years are important for our story, and we will thoroughly look at them in the next section.

Issai Schur was elected to a good number of academies of sciences. He was a legendary lecturer. Schur's student and friend Alfred Theodor Brauer (Ph.D. under Schur 1928) recalls [Bra2] that the number of students in Schur's elementary number theory courses often exceeded 400, and during the winter semester of 1930 even exceeded 500. Brauer would know, for as Schur's Assistant, he had to grade homeworks of all those students! Walter Ledermann, who estimates to have taken about 500 lectures from Schur, writes [Led1] that "Schur's lectures were exceedingly popular. I remember attending his algebra course which was held in a lecture theatre filled with about 400 students." Ledermann adds in his year 2000 interview [Led2]:

> I was absolutely captivated by Schur. I wrote about 300 lectures in fair copy in cloth bound book which I had until quite recently, running to something like 2000 pages of Schur's lectures.

Hitler's appointment as *Reichskanzler* by President von Hindenburg on January 30, 1933 changed this idyllic life. Schur's former student Menahem Max Schiffer recalls in his talk at the 4th Schur conference in May 1986 at Tel Aviv University, which was consequently published [Schi]:

> Now, the year 1933 was a decisive cut in the life of every German Jew. In April of that year [April 7, 1933 to be precise] all Jewish government officials were dismissed, a boycott of Jewish businesses was decreed, and anti-Semitic legislation was begun. When Schur's lectures were cancelled there was an outcry among the students and professors, for Schur was respected and very well liked. The next day Erhard Schmidt started his lecture with a protest against this dismissal and even Bieberbach, who later made himself a shameful reputation as a Nazi, came out in Schur's defense. Schur went on quietly with his work on algebra at home.

[4] Archive of Humboldt University at Berlin, documents UK–Sch 342, Bd. I, Bl. 4.

Issai Schur (left) and Edmund Landau. (Courtesy of Schur’s daughter, Hilde Abelin–Schur)

Ledermann shares with us vivid details [Led2]:

When Hitler finally came to power, all the Jewish faculty were dismissed instantly, including Schur who was not allowed to come even to the library anymore.

However Erhardt Schmidt, who was the decent sort of German, found that in the regulations of the Nazis there was a clause to say that these dismissals would not apply to two types of non-Aryan:

1) those who had fought in the First World War in the German army on the front, and
2) those who had during the First World War held a position making them German/ Prussian civil servants.

The first of these applied to Alfred Brauer who had been a soldier . . ., and yes, he was badly wounded, and the second applied to Schur because in 1916 he was an extraordinary professor at Bonn, so had effectively become a Prussian civil servant.

So, Schmidt applied this clause. He went to Goebbels and said, “You must abide by your own law and reinstate Schur for this reason,” and he was reinstated. He could then come to the University but he was not allowed to lecture. For supervision of my dissertation, I had to go to his house. It was nice to meet with him, he lived in a suburb of Berlin, to see him and his wife and talk not only about mathematics but also about the Jews. He said, “I can read the English Times which is still allowed,” all the other papers were taken over by the Nazis. I cannot bear this. And then the time came for me to have my exam, the oral, and he was allowed to come to take [conduct] this examination in mathematics for one hour. Also, a co-examiner was expected to come. They did not

> normally ask questions but would take a record, more like a secretary. This co-examiner was, unfortunately, none other than Bieberbach, who appeared in Nazi uniform, brown shirt and swastika. He came and sat down to take notes about what Schur was asking me. But I must say he was quite fair. He didn't interfere and I got a very good result.

Hindenburg negotiated with Hitler exemptions from the April 7, 1933, *Restoration of Professional Civil Service Law* for those Jews who fought for Germany in World War I, those who lost a father or a son in the war, and those who entered their civil service jobs (university professorships included) before the start of the war. Schur had held a civil service (university) appointment before the war and thus fell under the exemption. Nevertheless, by the order UI No. 6362 of the Prussian Minister for Science, Art and Public Education of on April 29, 1933, Schur and 18 other faculty were "relieved of their duties effective immediately" – yes, immediately, as was customary in the Nazi orders.[5] At that time, a representative of the English Jewish Emergency Council visited Schur. We are lucky to have had a witness at the meeting – Schiffer reports [Schi]:

> The lady asked Schur whether and where he wanted to go, because for a man of his reputation all doors would be open. But Schur responded that he did not intend to go; for he did not want to enable the Nazis to say, that many Jewish professors just left for better jobs. Besides, there were many younger colleagues which needed help much more urgently, and he would not take away their chances. He would stick it out in Berlin, for the craze of the Hitlerites could not last long.

I believe in Schur's incredible generosity and genuine care for his younger colleagues. Yet, there had to be more to his refusal to leave Germany early, when the Nazis came to power in 1933. In 1995, Schur's former student Walter Ledermann, Professor at the University of Sussex, UK, sent me his 1983 reprint [Led1], where he introduces additional reasons for the unfortunate Schur's decision to stay in Nazi Germany:

> When the storm broke in 1933, Schur was 58 years of age and, like many German Jews of his generation, he did not grasp the brutal character of the Nazi leaders and their followers. It is an ironic twist of fate that, until it was too late, many middle-aged Jews clung to the belief that Germany was the land of Beethoven, Goethe and Gauss rather than the country that was now being governed by Hitler, Himmler and Goebbels. Thus Schur declined the cordial invitations to continue his life and work in America or Britain. There was another reason for his reluctance to emigrate: he had already once before changed his language, and he could not see his way to undergoing this transformation a second time.
>
> So he endured six years of persecution and humiliation under the Nazis.

On October 7, 1933, the Prussian Minister for Science, Art and Public Education, by the order UI No. 8831, "canceled the suspension for Dr. Mittwoch and Dr. Schur, *Ordinarius* professors on the Philosophical Faculty," imposed by the previous order, effective – of course – immediately.[6] The legal exemption worked. Walter Ledermann, whom I quoted

---

[5] Archive of Humboldt University at Berlin, documents UK–Sch 342, Bd. I, Bl. 23 and Bl. 23R.

[6] Archive of Humboldt University at Berlin, document UK–Sch 342, Bd. I, Bl.24.

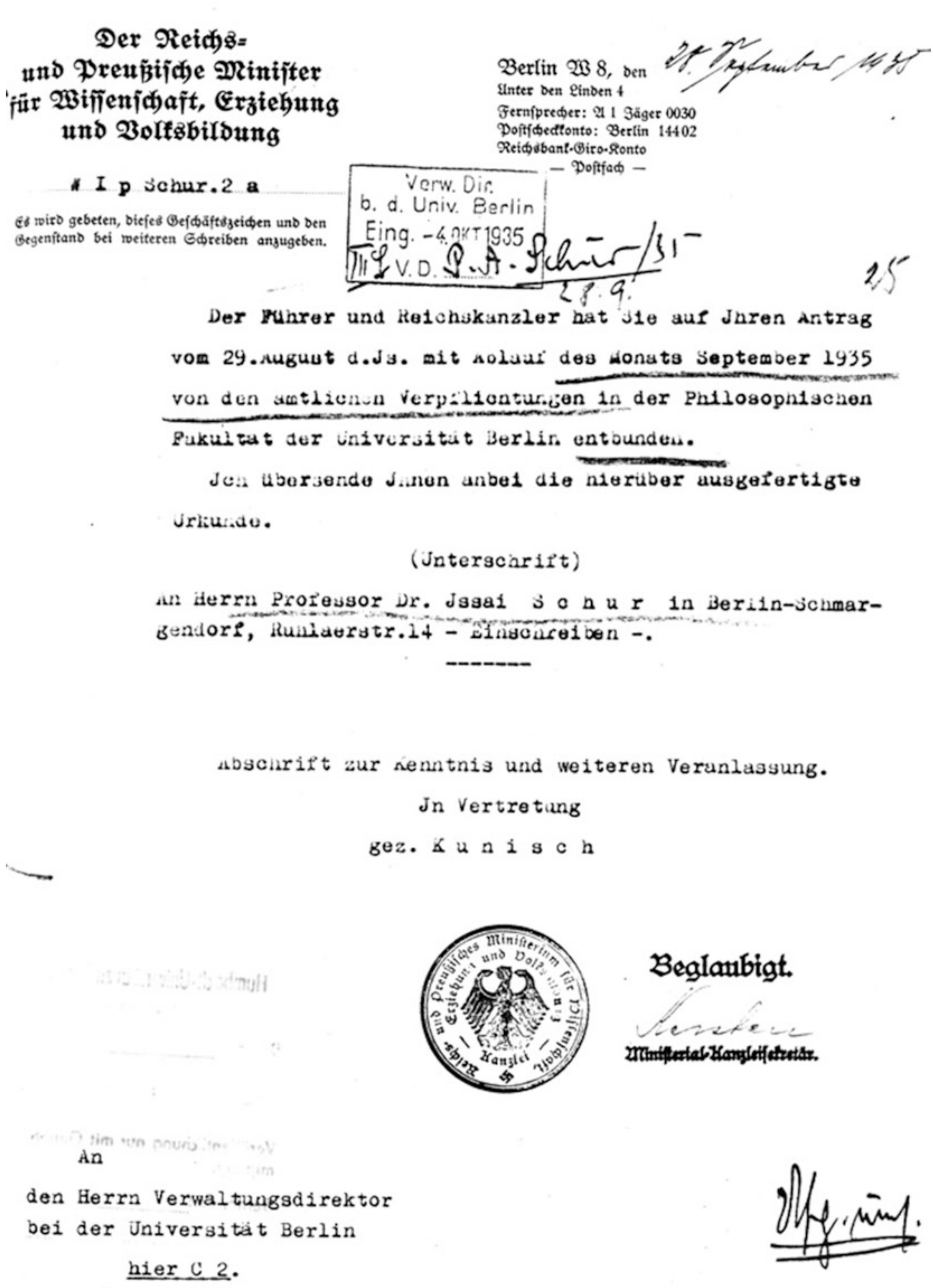
above, and Alfred Brauer [Bra2] credited Erhard Schmidt's efforts for the success. Consequently, Schur was able to carry out some of his duties but not all (no lecturing, for example) and not for long. Issai Schur was a famous professor, a pride of his university and of his profession. However, no achievement was high enough for a Jew in Nazi Germany. Following two years of pressure and humiliation, Schur, faced with imminent expulsion, "voluntarily" asked for resignation on August 29, 1935. On September 28, 1935, Reich's- and Prussian Minister of Science, Instruction and Public Education replied on behalf of "*Der Führer und Reichskanzler*," i.e., Adolf Hitler himself (see facsimile on one of the following pages)[7]:

> *Führer* and *Reichskanzler* has relieved you from your official duties in the Philosophical *Facultät* of the University of Berlin effective at the end of September 1935, in accordance with your August 29 of this year request.

As Henrik Hofer of the Humboldt University Library reports [Hof], Schur was the last Jewish professor to lose his job at the University of Berlin. Only a few of his closest friends had the courage to visit him, recalls Schiffer, and recollects one such visit, about which he learned from Schur himself [Schi, p. 180]:

> When he complained to [Erhard] Schmidt about the Nazi actions and Hitler, Schmidt defended the latter. He said, "Suppose we had to fight a war to rearm Germany, unite with Austria, liberate Saar and the German part of Czechoslovakia. Such a war would have cost us half a million young men. But everybody would have admired our victorious leader. Now, Hitler has sacrificed half a million of Jews and has achieved great things for Germany. I hope someday you will be recompensed but I am still grateful to Hitler." So spoke a great scientist, a decent man, and a loyal friend. Imagine the feelings of a German Jew at that time.

[7] Archive of Humboldt University at Berlin, document UK–Sch 342, Bd. I, Bl.25.



Der Reichs- und Preußische Minister für Wissenschaft, Erziehung und Volksbildung

W I p Schur.2 a

Es wird gebeten, dieses Geschäftszeichen und den Gegenstand bei weiteren Schreiben anzugeben.

Berlin W 8, den 28. September 1935
Unter den Linden 4
Fernsprecher: A 1 Jäger 0030
Postscheckkonto: Berlin 14402
Reichsbank-Giro-Konto
— Postfach —

Verw. Dir.
b. d. Univ. Berlin
Eing. -4. OKT 1935

Der Führer und Reichskanzler hat Sie auf Jhren Antrag vom 29. August d.Js. mit Ablauf des Monats September 1935 von den amtlichen Verpflichtungen in der Philosophischen Fakultät der Universität Berlin entbunden.

Jch übersende Jhnen anbei die hierüber ausgefertigte Urkunde.

(Unterschrift)

An Herrn Professor Dr. Jssai Schur in Berlin-Schmargendorf, Ruhlaerstr.14 – Einschreiben –.

Abschrift zur Kenntnis und weiteren Veranlassung.

Jn Vertretung

gez. Kunisch

Beglaubigt.

Ministerial-Kanzleisekretär.

An

den Herrn Verwaltungsdirektor
bei der Universität Berlin

hier C 2.

Letter relieving Issai Schur from his duties at the University of Berlin. (Courtesy of the Archive of the Humboldt University at Berlin)

Clearly, Erhard Schmidt, who, as we have seen, helped Schur after the latter's initial dismissal, held extreme nationalistic aspirations for Great Germany, *Deutschland über Alles*. Schmidt acknowledges and accepts the brutal sacrifice of half a million of Jews, including his friend Schur, and Schmidt is willing to sacrifice half a million of young German men for "great things for Germany." How low the morality fell in the Third Reich, if these were the views of "a decent man" (Schiffer's words), Erhardt Schmidt!

One very special 1936 visitor of Issai Schur, Paul Erdős recalls on the pages *Geombinatorics* [E95.32]:

> Schur was of the Russian Jewish origin. He always viewed himself as a German, and he was greatly attracted by the German culture. The horrible degeneration of Nazism was a great disappointment and a personal tragedy to him.

Menahem Schiffer lists Schur's numerous honors that were stripped away:

> He [Schur] was a member of many distinguished academies and learned societies; for example, the Prussian, Bavarian, Saxonian Academies of Science and many more. He had been ejected from each of them.

A document published in 1998 by the authors of [BFS] sheds light on one of these expulsions. Schur had been a member of the Prussian Academy of Sciences ever since his election in 1921. The Academy was going to publish works of Weierstrass – what can be political about that? The editorial board was to routinely sign off on the publication in a "*Zirkular.*" Let us look together at this document [BFS, p. 26]. The first two lines seem routine and are handwritten by Erhard Schmidt and Issai Schur, respectively (I am translating the lines from German here):

> Seen – 11.3.38 Erhard Schmidt
> Seen 12/3/38 Schur

Here comes Bieberbach, the founder of the racist doctrine of *German Mathematics* (that he opposes to *Jewish Mathematics*) and writes right below Schur, clearly hinting at Schur's presence:

> Bieberbach 29.3.38
> I am surprised that Jews still belong to the Academic Commissions. B.

In his turn Theodore Vahlen, a long-term Nazi and anti-Semite, a mathematician and an official in the Ministry of Education for University Affairs, in charge of hiring professors, agrees with Bieberbach:

> Seen Vahlen 30.3.38
> I request change. V.

The great Max Plank, a near 80-year-old icon of science, comes last and writes:

> Planck 3.4.38
> I will settle the affair. Planck

And settle Planck did. Just 4 days later, on 7 April 1938, Schur resigns from all Commissions of the Prussian Academy of Sciences. How does one assess Planck's role? Nazi collaboration, a pedantic fulfillment of his duties as the Secretary of the Academy, or a desire to dismiss Schur gentler than someone else, like Vahlen or Bieberbach would have done? We will never know for sure which one(s) of these motivations prompted Planck's actions. I for one deeply regret that, whatever the motive, Planck carried out the Nazi's dirty laundry. There were – had to be – other options. For example, Planck could have resigned from his Secretary position, or from the Academy itself. Meanwhile, the pressure on Issai Schur continues and later that year he resigns from the Academy itself.

On November 15, 1938, Issai Schur applies for a foreign passport, needed for leaving Germany. On January 14, 1939, the *Reichsminister* for Science, Instruction and Public Education states[8] that he "no longer objects to the issuance of a foreign passport for Dr. Schur" in view of "vulnerable health of Dr. Schur." He even approves paying Schur his emeritus remuneration through the date of Schur's departure.

On February 2, 1939, amid the Gestapo's "personal interest" in him,[9] depressed and sick, Schur leaves, I would say, runs away from Germany to Switzerland. Incredibly, the *Reichsminister* for Science, Instruction and Public Education believes that he could order Issai Schur where to live after Schur leaves Germany and, apparently, when to come back to the Third Reich, for on February 24, 1939, he issues the following order number W T Schur 4[10]:

> I hereby authorize the change of permanent address for the Emeritus Prof. Dr. Issai Schur of the University of Berlin, residing in Berlin–Schmargendorf, Ruhlaer Str. 4, first to Switzerland and thereafter to Palestine starting February 1st 1939 until the end of March 1941.[11]

Schur's wife of 33 years, Med. Dr. Regina Frumkin–Schur, joins him in Switzerland in March 1939. They stay in Bern for a few weeks with their daughter Hilde Abelin–Schur and her husband, Med. Dr. Chaim Abelin. Switzerland does not allow Schur to remain there permanently, so much for Swiss "neutrality." Broken mentally, physically, and financially, the Schurs move on to Palestine.

While in Palestine, without means, Schur has to sell his only valuables, scientific books and journals, to the Institute for Advanced Study, Princeton, where his former student and friend Alfred Brauer is Hermann Weyl's assistant and is charged with library acquisitions. This book transfer must have been painful for both Schur and Brauer. Schiffer recalls one 1939 episode that shows how infinitely professional Schur was:

> [Schur] agreed to give a lecture at the Hebrew University and this I will never forget. He spoke about an interesting inequality in polynomial theory with the customary clarity and elegance. Suddenly, in the middle of his talk he sat down, bent his head and was silent. We, in the audience, did not understand what was going on; we sat quietly and respectfully. After a few minutes he got up and finished his talk in his usual manner.
>
> I was sitting next to a physician from the Hadassa Hospital who had come to see this famous man. He was quite upset; after the lecture he told me that Schur had obviously

[8] Archive of Humboldt University at Berlin, document UK–Sch 342, Bd. I, Bl. 47 and Bl. 47R.
[9] [Bra2].
[10] Archive of Humboldt University at Berlin, document UK–Sch 342, Bd. I, Bl. 53 and Bl. 53R.
[11] Of all people, I should not be surprised, for when I was leaving another bastion of tyranny, the Soviet Union, in 1978, I too was told where to go and where to live.

> had a heart attack and he could not understand the self-discipline which had enabled Schur to finish his talk. That was the man Schur, for you!

Schiffer informs us that Schur eventually gets better, writes several research papers, supervises a number theoretic work of Theodore Motzkin, and "starts interacting with younger men at the Mathematics Institute." Issai Schur died from yet another heart attack in Tel Aviv on January 10, 1934, right on his 66th birthday.[12]

The list of Issai Schur's Ph.D. students, who became world-class mathematicians, is amazing. It includes Heinz Prüfer (1921), Richard Brauer (1925), Eberhard Hopf (1926), Alfred Brauer (1928), Bernhard Neumann (1932), Hans Rohrbach (1932), Wilhelm Specht (1932), Richard Rado (1933), and Helmut Wielandt (1935). The list of successful mathematicians, who were Schur's undergraduates or were influenced by him in other significant ways, is too numerous to be included here. Schur with his teacher and a student produced one of the most remarkable succession lines in the history of modern algebra: Ferdinand Georg Frobenius – Issai Schur – Richard Dagobert Brauer.

37.3 Argument for Schur's Authorship of the Conjecture

Issai Schur made major contributions to various areas of mathematics.[13] Our interest here lies in the result he obtained during 1913–1916 when he worked at the University of Bonn as the successor to the celebrated topologist Felix Hausdorff. There he writes his pioneering paper [Sch] containing, as he put it, "a very simple lemma, which belongs more to combinatorics than to number theory." We proved the Schur Theorem in Chapter 36. Here, I would like just to formulate it again for your convenience:

The Schur Theorem 37.1 (Schur, [Sch]). Let m be a positive integer and $N > m!e$. If the initial array [m] of positive integers is m-colored, then there is a monochromatic triple a, b, and c of the same color such that $a + b = c$.

The Schur Theorem gave birth to this novel way of thinking, a new direction in mathematics, today called the Ramsey Theory.

Leon Mirsky writes [Mir] on the centenary of Issai Schur's birth:

> We have here a statement of the type: "if a system is partitioned arbitrarily into a finite number of subsystems, then at least one subsystem possesses a certain specified property." To the best of my knowledge, there is no earlier result which bears even a remote resemblance to Schur's theorem. It is this element of novelty that impresses itself so forcibly on the mind of the reader.

Mirsky continues:

> After writing his paper, Schur never again touched on the problem discussed there; and this is in itself something of a mystery. For the strongest impression one receives on

[12] For more details see [Bra2], [Schi], [Led1] and [Soi10].

[13] For details see [Bra2] and [Led1].

scanning his publications is the almost compulsive striving for comprehensiveness. There are few isolated investigations; in algebra, in analysis, in the theory of numbers, Schur reverts again and again to his original questions and pursues them to the point of where one feels that the last word has been spoken ... Why, then, did he not investigate any of the numerous questions to which his Theorem points so compellingly? There is no evidence to enable us to solve the riddle. (Footnote: As will emerge from the discussion below, Professor Rado, if anyone, should be able to throw light on the mystery – and he tells me that he cannot.)

The latter Mirsky's statement, backed by Richard Rado, was echoed in the standard text on Ramsey Theory [GRS2, p. 70], thus becoming a universal view on this matter: "Schur never again touched on this problem."

I have solved the Mirsky's *"mystery,"* and my findings contradict Mirsky's, Rado's, and Graham–Rothschild–Spencer's conclusion. I will show in this section that the new Ramseyan mathematics, discovered by Issai Schur in his 1916 paper, remained dear to his heart for years to come. He thought about this new mathematics himself, and he passed his interest on to a number of his students: Hildegard Ille, Alfred Brauer, and Richard Rado.

As we have seen in Chapter 35, the third classic result of Ramsey Theory was published by B. L. van der Waerden in 1927, in which he presented "*Proof of a Baudet's Conjecture*" [Wae2]. The credit to Baudet for the conjecture remained unchallenged and unsubstantiated, until 1960, when Alfred Brauer (1894–1985) made his sensational revelations.

"I remember Alfred [Brauer]," told me over the phone Mrs. Hilde Abelin–Schur, the daughter of Issai Schur [Abe2], "he was Assistant of my father, and I was then a little girl." An Assistant, a doctoral student (Ph.D. in 1928), a colleague (*Privatdozent* at the University of Berlin), co-author, and a friend through the difficult years of the Nazi rule, Alfred Brauer had unique knowledge of Issai Schur. Away from Germany for over twenty years, he returned to Berlin in 1960 to pay tribute to his teacher. His moving talk about Issai Schur given at the Humboldt University of Berlin on November 8, 1960, appeared in print in 1973 as an introduction [Bra3] to the three-volume set of Schur's collected works that Brauer edited jointly with another former Schur's Ph.D. student Hans Rohrbach. This talk offered a wealth of information about Schur. In particular, it revealed that Issai Schur, inspired by E. Jacobsthal's results about quadratic residues,[14] came up with the following two conjectures:

Conjecture 37.2 For any positive integer k and any large enough prime p, there is a sequence of k consecutive quadratic residues modulo p.

Conjecture 37.3 For any positive integer k and any large enough prime p, there is a sequence of k consecutive quadratic non-residues modulo p.

As was the case with the Schur Theorem of 1916 [Sch], a search for a proof of number-theoretic conjectures 37.2 and 37.3 led Schur to conjecture a "helpful lemma":

[14] If the congruence $x^n \equiv a \pmod{m}$ has a solution for x, then a is called an *n-th power residue modulo m*. In particular, 2nd power residues are called *quadratic*.

Conjecture 37.4 (Issai Schur). For any positive integer *k*, there is $N = N(k)$ such that the set of whole rational numbers 1, 2, ..., *N*, partitioned into two classes, contains an arithmetic progression of length *k* in one of the classes.

Alfred Brauer describes the circumstances of Schur discovering that his conjecture 37.4 was proven:

> Many years passed, but neither Schur nor many other mathematicians, who were familiar with this conjecture, were able to prove it. One day in September of 1927 my brother [i.e., Richard Brauer, Ph.D. in 1925 under Schur] and I were visiting Schur, when [John] von Neumann came unexpectedly. He was participating in the meeting of the D.M.V.[15] and came to tell Schur that at the meeting Van der Waerden, using a suggestion by Artin, gave a proof of the combinatorial conjecture and was going to publish it under the title "*Beweis einer Baudetschen Vermutung*." Schur was very pleased with the news, but a few minutes later he became disappointed when he learned that his conjecture about sequences [i.e., conjectures 37.2 and 37.3 above] was not proven yet ... It would have made sense if Schur were to propose a change in the title of van der Waerden's publication or an addition of a footnote in order to indicate that this was an old conjecture of Schur. However, Schur was too modest for that.

Paul Erdős, a man of an incredible memory for events, told me that in everything concerned with Schur, Alfred Brauer was by far the most reliable source of information. Paul also shared with me a critical unpublished confirmation of Schur's authorship of the conjecture. During our long conversation,[16] that commenced at 7:30 PM on Tuesday 7 March 1995, in Boca Raton, Florida, during the traditional combinatorics conference's "Jungle Party," Paul told me that he heard about Schur's authorship of this conjecture from Alfred Brauer. Independently, he heard about it from Richard Brauer, a brilliant algebraist and the younger brother of Alfred. Finally, Schur's authorship was confirmed to Erdős by Erich Rothe, who obtained the information from his wife and Schur's former student Hildegard Rothe (born Hildegard Ille; Ph.D. in 1924 under Schur). As I am writing these lines, I am looking at a yellow lined sheet that Paul tore out of his notebook and next to his mathematical texts wrote for me "*Hildegard Ille*," so that I would remember her name when I get to write about her. Thank you, Paul, I remember!

I believe you will agree with me that I have produced as rigorous a proof as a historical endeavor allows that Issai Schur had the conjecture and created it independently from anyone else.

The historical research of this chapter shows for the first time that Issai Schur had been *the* most instrumental leader in the development of "*Ramsey Theory Before Ramsey.*" I did not know that myself until the completion of this research. Started with his 1916 theorem (Chapter 34), Schur's interest in not-yet-born Ramsey Theory continued with the conjecture on arbitrarily long monochromatic arithmetic progressions in finitely colored integers. Right

[15] *Deutsche Mathematiker–Vereinigung*, German Mathematical Society – the Annual September 18–24,1927 meeting took place in Bad Kissingen in Bavaria.

[16] Knowing the importance of Paul Erdős' information, I audio-recorded our conversation.

after Van der Waerden's publication, Issai Schur produced, as we will see in Chapter 38, *The Generalized Schur Theorem*, which generalized *at the same time* both the Schur and the Baudet–Schur–Van der Waerden theorems. With Schur's guidance, his former student Alfred Brauer proved a Ramseyan result of his own (Chapter 38). Schur offered Ramseyan type problems to his doctoral student Hildegard Ille. Under Schur's guidance, Richard Rado generalized the Schur and the Baudet–Schur–Van der Waerden theorems in his doctoral dissertation and important consequent publications. In fact, Rado contributed to Ramsey Theory, perhaps, more than anyone.

As proof on the Schur's pudding, we will observe in Chapter 38 that *Schur appears to have been first to raise the problem of arbitrarily long arithmetic progressions of primes in the 1920s–1930*, before Paul Erdős took the leading role in the development of Ramsey Theory. Erdős recalls on the pages of *Geombinatorics* [E95.32] on the occasion of Schur's 120th birthday in 1995:

> I first heard about Schur when I was a student of an old Hungarian algebraist Michael Bauer, who advised me to write to Schur about my results on prime numbers in arithmetic progressions. Schur was the first foreign mathematician with whom I corresponded. I wrote [to] him my elementary proofs on some of my results on prime numbers in arithmetic progressions, which Schur liked very much, the results were published in *Math. Zeitschrift* in 1935.

In fact, the Ramseyan baton from Schur to Erdős may have been passed at their 1936 meeting in Berlin. "I was told that Schur sometimes referred to me as the Sorcerer from Budapest," Paul recalled fondly in our conversations and in print [E95.32].[17] Amazingly, I found the eyewitness' reminiscences of this Erdős' visit of Berlin when Hilde Brauer, the widow of Schur's Assistant and close friend Alfred Brauer, gave me a gift of her wonderful *unpublished* memoirs [BraH]. She married Alfred on August 19, 1934, and as a "mathematical wife" from the Schur's circle, met Erdős during his Berlin visit:

> The latter [Paul Erdős], who was a child prodigy, surprised me at his first visit when he was barely twenty with curious interest for all details in bringing up a baby. He called all children epsilons but knew all the names of his friends' babies.

I have got to mention here one more Schur's activity, in which he, in a sense, predates Erdős. Schur's former student Richard Brauer (February 10, 1901–April 17, 1977) writes in the February 1977 introduction to his 3-volume collected papers [BraR] that appeared posthumously in 1980:

> He [Schur] conducted weekly problem hours, and almost every time he proposed a difficult problem. Some of the problems had already been used by his teacher Frobenius, and others originated with Schur. Occasionally he mentioned a problem he could not solve himself. One of the difficult problems was solved by Heinz Hopf and also by my brother Alfred and myself. We saw immediately that by combining our methods, we

[17] Paul Erdős was not only mathematically, but also personally attached to Issai Schur and his wife Regina. "I several times visited his widow. In 1965, I visited her in Tel-Aviv with my mother," writes Paul [E95.32].

> could go a step further than Schur. Our joint paper [BBH] in the list below originated this way.

So, Issai Schur had a great interest in creating problems and conjectures, and disseminating them on the regular basis, weekly, starting at least in 1920. Ramseyan-style problems and conjectures must have been part of this Schur's oeuvre. Paul Erdős, who took over the leadership of the Ramsey Theory, also had, as we all know, a great interest in problem posing. He created his first open problem in 1931. In 1957, Paul commenced his celebrated "some of my favorite unsolved problems" series of papers.

This inquiry into the life of Issai Schur was made possible by the invaluable help from Issai Schur's daughter Hilde Abelin–Schur; the widow of Alfred Brauer, Hilde Brauer; Schur's former student Walter Ledermann; Paul Erdős; Heiko Harborth; Henrik Hofer; and the Archive of Humboldt University of Berlin.

37.4 Enters Henry Baudet II

When in 1995, I presented my argument for Issai Schur's credit in an essay [Soi10] written on the occasion of his120th birthday, I specifically included a historically significant disclaimer:

> Nothing presented here excludes the possibility that Baudet created the conjecture independently from Schur. N. G. de Bruijn [Bru3], clearly understanding the rarity of Ramseyan ideas at the time, hypothesizes that Baudet was inspired by the 1916 paper [Sch] of Schur to independently create the conjecture. Perhaps, in the future historians would shed light on the question whether Baudet was an independent from Schur author of the second [counting Hilbert–1892, third] conjecture in the history of Ramsey Theory. Until then the conjecture ought to be rightfully called Schur's.

When my essay [Soi10] appeared, I learned from N. G. de Bruijn about the existence of P. J. H. Baudet's son, Henry Baudet, or as he sometimes called himself Henry Baudet II, and mailed him a copy of my paper. I sowed an essay and harvested a fury! The young Henry Baudet (his full name Ernest Henri Philippe Baudet born on January 29, 1919, in Scheveningen; he was 76 at the time) replied in style all his own:

> I write to you in my own English, which is far from good, but it might be better than your own French or Dutch.

He then offered a counterexample[18] to Schur's 1916 theorem and questioned Brauer's assessment of Schur:

> "Too modest" seems hardly possible and hardly believable, considering the revolutionary essence of the theorem or the conjecture.

Henry was clearly upset with my putting in doubt his father's credit. In my August 30, 1995 letter, I admitted that indeed "my French and Dutch are far inferior" to his fine English and offered Henry to publish in my quarterly *Geombinatorics* his essay challenging

---

[18] Schur's Theorem survived; Henry simply misunderstood it.

my proof of Schur's authorship of the conjecture, if he so desired. I also offered Henry Baudet II to join me in the investigation of whether his father Henry Baudet I created the conjecture independently from Issai Schur.

Henry offered to help with documents upon his return to Holland from his summer home in Bourgogne, France. In addition to being a History Professor, Henry was The Historian of the Delft Technical University, and the last PhD student of the legendary Johan Huizinga of *The Waning of the Middle Ages* fame. From letter to letter, I was promoted from "Professor Soifer" to "Alexander," to "Sasha." Our correspondence for the ensuing year was very intense: we exchanged some 30 letters (letters, not e-mails). My family and I then paid a 5-day visit to Henry and his wife Senta Govers Baudet in their centuries-old stone house in the medieval village *Corpoyer-la Chapelle*, population 26, in Bourgogne, France.[19] As I am writing these lines, I am holding in front of me a copy of Henry's book *Mon Village en France* warmly inscribed to my (then) wife Maya and me by Henry on August 1, 1995. Later that year we also visited the Baudet family in their Dutch house in the town of Oegstgeest, population ca. 24,000, nestled on the outskirts of Delft.

I learned much about Henry and Senta helping Jews in the Netherlands occupied by Nazi Germany for the long five years 1940–1945. Henry recalls [Bau5]:

> I myself, finally, started studying history at Leiden University but this was interrupted when the Germans, during the war, closed the University. Somehow, nevertheless, I could remain in touch with my professors, at least in the beginning. Of course, the German occupation made life extremely difficult, and this every year more and more. Resistance was a new activity we had to learn; hiding Jews was a daily concern and hiding ourselves was another. We lost many friends but somehow or other I got through myself (though my wife, then my girlfriend, then 17 years [old], got temporarily into jail for helping Jewish classmate to escape – she (I mean: her Jewish girlfriend) lives in Dallas now and we see each other and call each other by telephone).

In fact, Senta Govers Baudet's name is inscribed in Yad Vashem – The World Holocaust Remembrance Center in Israel as she was awarded the high title of a *Righteous Among the Nations*, granted to non-Jews who risked their lives to save Jews during the Holocaust. Senta helped her Jewish friend Liny L. Yollick escape from the Netherlands by lending her Senta's identification card. The escape was successful, but silly Liny sent the card back with a boy who was caught by the Germans. On June 27, 1942, Senta was imprisoned by the Germans and spent a week in jail, interrogated daily and nightly. Only her consistent denial of loaning the card to Liny, had finally convinced the jailers.[20] This was but one episode of the young family's participation in the resistance. In fact, Henry and Senta risked their lives, on numerous occasions by helping Jews hide or escape. They themselves had to hide from the Germans, who came to look for them on occasion.

[19] Both Henry Baudet II and his son Remy Baudet, a wonderful violinist (music, even more than mathematics and chess, was a family tradition for generations), looked so Gascogne, that they could play Alexandre Dumas Père's D'Artagnan without any make-up.

[20] I thank Yad Vashem, The Holocaust Martyrs' and Heroes' Remembrance Authority, for sharing with me copies of the relevant documents substantiating Senta' high recognition.

With Henry's help, I was able to successfully investigate the question of whether Baudet I earned the credit that Van der Waerden so nonchalantly had given him.

My dear friend Henry Baudet II was one of the most charming people I have met in my life. He passed away on December 16, 1998. In 2003, Delft Technical University created the *Henry Baudet Institute* dedicated to the history of design, one of his many interests.

37.5 Pierre Joseph Henry Baudet

B. L. van der Waerden gave Baudet credit in his 1927 paper [Wae2], which in fact was called *Beweis einer Baudetschen Vermutung* (i.e., *Proof of a Baudet Conjecture*). We do not find Baudet's initials in Van der Waerden's paper. Indeed, Van der Waerden did not even know that at the time his publication came out, Baudet had been dead for six years. As is often the case with young and brilliant mathematicians, Van der Waerden was probably not interested in the history of the problem he solved and in the identity of the author of the conjecture. In reply to my questions, Van der Waerden answers on April 24, 1995 [Wae25]:

1. I heard of "Baudet's Conjecture" in 1926.
2. I never met Baudet.
4. I never met Schur.
5. I never heard about Schur's [1916] result.

By the time Van der Waerden publishes a *detailed* story of the emergence of his proof in German in 1954 and in English in 1971 (presented in Chapter 35), he is not only a celebrated mathematician but also a famous historian of science, author of the well-known book *Science Awakening* [Wae15] and numerous historical articles. Sometimes he is deservedly harsh toward other historians [Wae15]:

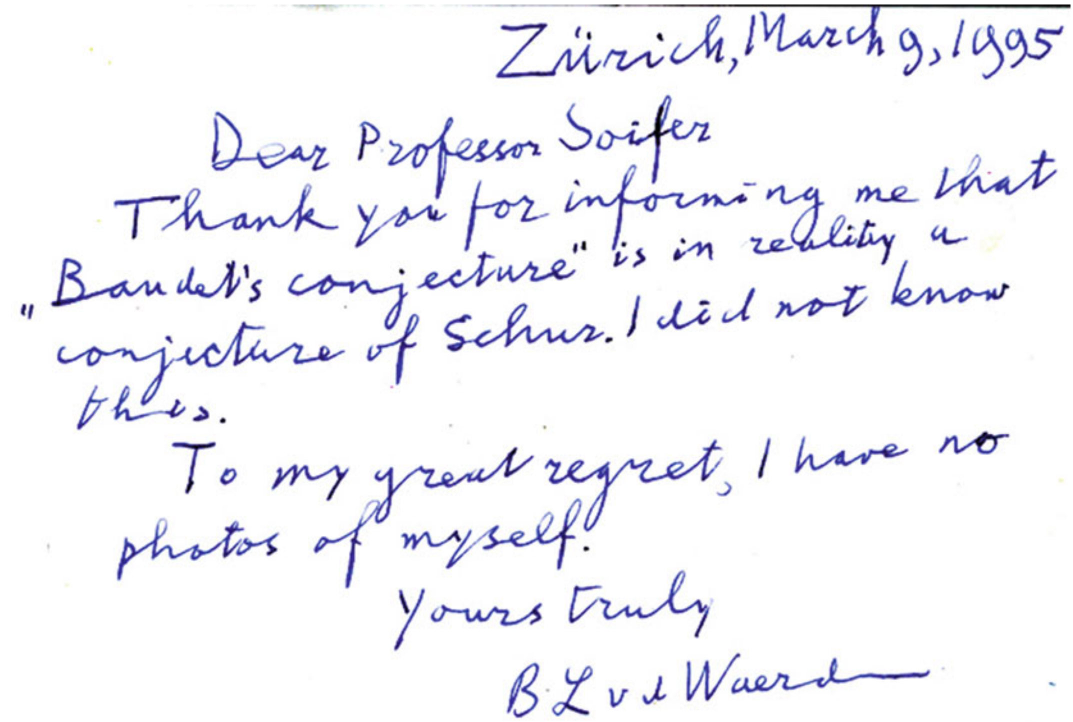
> How frequently it happens that books on the history of mathematics copy their assertions uncritically from other books, without consulting the sources! How many fairy tales circulate as "universally known truths"!

Yet Van der Waerden-historian does not investigate the authorship of the conjecture that became *his* classic theorem. Biographers of Van der Waerden faithfully follow the Master and credit "Baudet" with the conjecture, ignoring – or being ignorant of – Brauer's reminiscences [Bra3], and providing no independent historical analysis (see [Fre], [FTW], [Per], etc.).

I thought that in all likelihood someone, sometime during the long years between 1927 and 1971, must have mentioned to Van der Waerden Brauer's assertion that Van der Waerden proved Schur's conjecture. Nobody, apparently, has until now, as you can see from Van der Waerden's March 9, and April 4, 1995 replies [Wae23], [Wae24] to my inquiry[21]:

> *Dear Professor Soifer: Thank you for informing me that 'Baudet's conjecture' is in reality a conjecture of Schur. I did not know this.*

[21] See the facsimile of [Wae23] in this chapter.



Zürich, March 9, 1995

Dear Professor Soifer

Thank you for informing me that „Baudet's conjecture" is in reality a conjecture of Schur. I did not know this.

To my great regret, I have no photos of myself.

Yours truly

B L v d Waerden

Van der Waerden, March 9, 1995 letter to Alexander Soifer

While Van der Waerden's acceptance of my argument for Schur's credit was important, it contributed nothing to the question whether Baudet created the conjecture independently of Schur. As I wrote [Soi10], "Perhaps, in the future, historians will shed light on the question whether Baudet was an author, independent of Schur." This future has arrived: let us look at Baudet's role in our saga.

It appears that Alfred Brauer is first to speak in print about Baudet on 8 November 1960 [Bra3] (see also its later English translation [LN]) ever since Baudet's obituaries appeared in 1921 [Schuh] and 1922 [Arr]. Since Brauer knows firsthand that Schur has created the conjecture (and, I gather, assumes it to be unlikely that two people could independently come up with such a revolutionary conjecture), he attempts to "*prove*" that Baudet did *not* create the conjecture independently by showing how the conjecture got from Schur to Baudet:

> Baudet at that time was an unknown student at Göttingen, who has later made no mathematical discoveries. On the other hand, at this time Schur's friend Landau was a professor at Göttingen, who obviously knew the conjecture, and used to offer unsolved conjectures as exercises to every mathematician he met. It was therefore highly probable that Baudet learned the conjecture directly or indirectly [from Landau].

Brauer repeats his assertions in English in print in 1969 [Bra2]:

> It seems that the title of van der Waerden's paper "Beweis einer Baudetschen Vermutung" [Wae2] is not justified. Certainly [sic] van der Waerden heard about the conjecture from Baudet, a student at Goettingen.

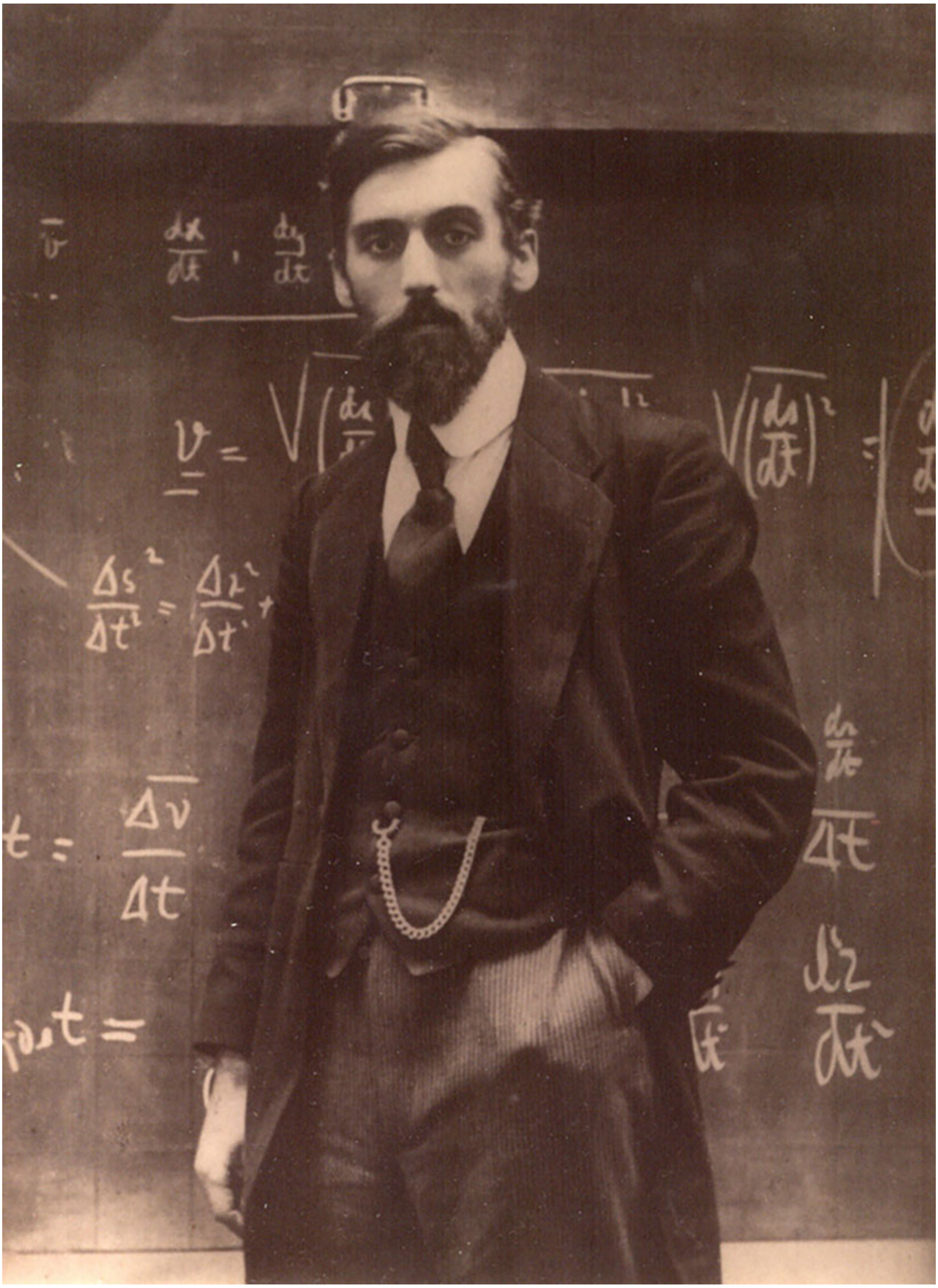
When Alfred Brauer speaks about Baudet (I wish he did not!), he enters the area not personally known to him. Consequently, Brauer presents his hypotheses as if they are true facts. In fact, I find Brauer's hypotheses to be dramatically false. Baudet "at that time" was not "an unknown student at *Göttingen*," but instead a brilliant young Ph.D. from *Groningen*. Brauer's allegation that Baudet "later made no mathematical discoveries" was as gratuitous as it was incorrect: in addition to publishing his doctoral thesis [Bau1] and the inaugural speech [Bau2], Baudet published three papers [Bau3], [Bau4], and [Bau5] that appeared in *Christiaan Huygens* – not bad for someone who left this world untimely at the age of 30. Baudet became a Full Professor at Delft University at the tender age of 27 – can this be said of many mathematicians, then or now?

Alfred Brauer's valuable testimony about Schur's creation of the conjecture, as well as his regrettable misrepresentations about Baudet, are repeated by Graham–Rothschild–Spencer in their standard texts on Ramsey Theory [GRS1], [GRS2] and from there are copied by a good number of publications. It is time, therefore, to set the record straight, and convey to the world how great a man the world lost in Pierre Joseph Henry Baudet.

The following account of Baudet's life was possible only due to the indispensable assistance of Henry Baudet II, the son of the mathematician Pierre Joseph Henry Baudet. Unless otherwise credited, the following information, slightly edited, comes from Henry Baudet II's letters to me [BII1 – BII13] and my personal interviews with him in his Medieval house in Bourgogne, France.

> My father was born on January 22, 1891, in Baarn (province of Utrecht, The Netherlands) in nothing less than a psychiatric clinic, where my grandfather – a neurologist – was medical superintendent. A few years later my grandparents moved to The Hague, where my grandfather started a private practice. So it was in The Hague that my father grew up, attended the elementary school and then the Gymnasium from which he graduated in 1908. He was a dedicated chess player and cellist. (In this, he followed the family tradition: we all are musicians and chess players, though not on his level).
>
> In September 1908 my father enrolled as a student of mathematics at Leiden University, where he studied under Kluyver. I know next to nothing about his study in Leiden, except the fact of his early fame as a chess player, a musician, and a future mathematician. He obtained his master's degree in 1914, as far as I know just on the eve of the World War I and became a mathematician at the same Gymnasium in The Hague where my father had been a pupil**.** He stayed there until his 1919 appointment as a Professor at the University of Technology at Delft (then still named the Technical High School).
>
> As a student at Leiden, he met my mother [Ernestine van Heemskerck] who studied in the Faculty of Arts, and my parents got married on April 7th of 1914 ... My sister (also a mathematician) was born in 1915 on the 31st of January. I myself arrived four years later on January 29th, 1919. So, all of us are Aquarius.
>
> How my father and Schuh[22] met, I don't know, probably in the Society of Mathematicians. They were, however, close friends since 1914 or 1915 ... With Schuh as supervisor, my father began to work on his thesis, but he could not take his doctor's degree with him, as Delft had no doctorate in mathematics. And [Johan A.] Barrau [1874–1953, a professor of mathematics at Groningen University] ultimately took over Schuh's job.

[22] Frederick Schuh, 1875–1966, Ph. D. under Diederik Korteweg, as was L. E. J. Brouwer after him, a very versatile mathematician, with numerous publications in analysis, geometry, number theory, statistics, recreational mathematics, teaching of mathematics, etc.



Pierre Joseph Henry Baudet (1891–1921). (Courtesy of Henry Baudet II)

The eulogy “In Memoriam Prof. P. J. H. Baudet” [Arr] by Dr. E. Arrias appeared on January 28, 1922. The author, who has known Baudet for 15 years, reports astonishing talents of Pierre Joseph Henry:

> At 15, Baudet was known for virtually never losing a game [in chess] and playing several simultaneous games blindfold ... But all these achievements were outshone by the miraculous things he has done with the Laskagame, invented by Dr. E. Lasker [a mathematician and the legendary world chess champion during the incredibly long period 1894–1921]. Before Lasker had his new game published, he submitted it to Baudet for evaluation. With his characteristic tempestuous application Baudet mastered this game; it was as if he finally had found something that could fully satisfy his wits. This exceptionally intricate game with its discs in four different colours, its capricious, almost incalculable combinations, suited his mathematical brain exactly. It is, therefore, not surprising that having studied the game for half a year, he could scarcely be beaten by Lasker himself... Thanks to his enthusiasm a Lasca society was founded in The Hague, and even a first national tournament organized, but after everything had taken shape, he died one day before the tournament, to which he had been looking forward like an eager child (for in spite of his scientific greatness he was a child in joy) ...[23]
>
> As proficient as he was at board games, as high was his reputation as a musician ... Being an extraordinarily sensitive cellist, he completely mastered the technique of this instrument. Many were the times that he contributed to the success of concerts by his impassioned playing. And all this without score; a feat only very few people are so privileged. With him it was not a matter of learning notes, but he absorbed the complete picture of the composition, and even when he had not seen the composition for ten years, he was able to conjure it up clearly and to play it from memory, when only hearing the piano part ... He was excellent at reading scores and he conducted already during his grammar school period. He was fully familiar with theory and counterpoint. Only recently, he could prove this when the vice-chancellor of Delft University asked him to orchestrate the Don Juan for the students’ string orchestra. Next to all his excessively many occupations, this task could be added without any problem. He finished it just before death overtook him
>
> . . .
>
> It was pure scientific curiosity that had made him master this as well as everything he did: ... learning Hebrew and the four Slavic languages simultaneously was no trouble at all, since he was learning anyway – and in fact this was far more interesting – that comparative linguistics ... Stacks of work are lying in his study; constantly new ideas suggested themselves to him which he noted down only in lapidary form. He did not get around to publish much, but his confrere friends will need years of hard work to sort out and work on his sketchy notes.

[23] As I learned from Professor N. G. de Bruijn [Bru1], “In his *Brettspiele der Völker* (Berlin 1931) Lasker describes a game of ‘Laska’ he lost to Baudet at a tournament in The Hague 1920. (‘Laska’ was Lasker’s own invention, which he tried to promote at a time he thought that eventually all serious chess games would lead to a draw.)”

> On the birthday of Jesus this highly gifted man with his magnificent Christ like features parted from this earthly life, at the same age, as his greatest master. But in our thoughts he will rise again and stay alive for us as long as we keep breathing!

Baudet defended his doctoral thesis cum laude in 1918 at Groningen University and became Professor in Pure and Applied Mathematics and Mechanics at Delft Technical University in 1919. He was 27. Pneumonia brought his life to an untimely end on the Christmas Day of 1921. His first obituary [Schuh] was written by his friend and teacher Frederik Schuh.

Pierre Joseph Henry Baudet was an extraordinary man indeed. But did he create the conjecture?

The credit given Baudet by Van der Waerden [Wae2] is insufficient since he "never met Baudet" and "heard of 'Baudet's Conjecture in 1926", i.e., years after Baudet's passing [Wae26]. However, by backtracking the link from Van der Waerden to Baudet, we reach firmer grounds.

37.6 Argument for Baudet's Authorship of the Conjecture[24]

The search, in fact, was started by Henry Baudet II. Born on January 29, 1919, Henry lost his father at the age of not quite two and always wanted to find out more about him. In 1962–1963, professor of tax law and an amateur mathematician Tj. S. Visser gave a talk *Attack on Sequences of Natural Numbers* attended by Henry Baudet and his 15–16-year-old son Rémy. Unbelievably, the four-page brochure (in Dutch) of this talk survived in the family papers and was shared with me by Henry. Thus, we are granted an attendance to Visser's lucid and informed talk:

> My story is about the most beautiful statement of number theory, The Theorem of Baudet. The pearl of Baudet . . .
>
> Baudet is the early departed in the beginning of this century Delft's Professor of Mathematics, born in Nenegouw . . .
>
> His pearl of the theory of numbers is: If one divides the natural numbers 1, 2, 3, 4... *ad infinitum* into a random number of boxes, then there is nevertheless always at least one box which contains an arithmetic progression of arbitrary length . . .
>
> This proposition was formulated by Prof. Dr. P. J. H. Baudet in 1921. He died shortly after, leaving a wife and a baby. Many celebrities tried to find a proof of this theorem. The young, also Dutch mathematician succeeded. His name was B. L. van der Waerden. He published his proof in 1927 in *Het Nieuw Archief* under the title *Beweis einer Baudetschen Vermutung*.
>
> It takes five pages, uses no higher mathematics but is very heavy. He seems to have found it during a holiday session at Göttingen where his astuteness rightly won large admiration. Bartel van der Waerden is a son of the engineer-teacher Theo, doctor in technical sciences, a very prominent person elected to Parliament from the S.D.A.P.,

[24]This is an expanded version of my tribute [Soi11] to P. J. H. Baudet that was published on the occasion of his 105th birthday.

> known as '*rooie Theo*' [Red Theo]. The young Bartel became professor at Groningen, was later oil-mathematician, is now at Zurich director of the Mathematical Institute, and is world-renowned.
>
> After 1927, the statement and its proof fell asleep.

Tj. Visser then conveyed how the Russian mathematician Aleksandr Yakovlevich Khinchin brought the theorem back to life by publishing it, with a slightly different proof found by his student M. A. Lukomskaja, as one of the pearls in his book *Three Pearls of Number Theory*, which appeared in Russian, German, and English.

As an amateur mathematician, Henry-the-son was fascinated by the conjecture. "Could I write to him [i.e, to Van der Waerden]?" he asked the family friend and his father's mentor Frederick Schuh (February 7, 1875–January 6, 1966). "Of course," Schuh replied. Henry recalls:

> In the context and the fact that I proposed to Schuh to give me the address of Van der Waerden, it was clear that Schuh considered it [the conjecture] to be an important affair. He agreed that I should write.

I asked Henry (we had long interviews in his centuries-old Bourgogne stone house):

> Does it appear that Schuh was in total agreement that it was Baudet's conjecture that Van der Waerden proved?

"Absolutely, absolutely yes, absolutely," replied Henry. And so, on September 1, 1965, Henry Baudet II wrote to Van der Waerden in style already known to you from Henry's first letter to me:

> I am the son of my father. It is always the case, but you understand what meaning this introduction has in this case. Somehow from afar I was following your publications, and thus I was able to get into my hands your work in the *Abhandlungen aus dem Mathematischen Seminar Hamburg* [Wae16]. For me this is not a completely closed book. Having at one point started in mathematics, I have become a historian in the end, and it is something entirely different.
>
> In this letter to you, a fairly remarkable fact is taking place. It is a fact that I cannot say anything special, but nevertheless I wanted very much to establish a contact with you. Of course, I would like to ask you whether you have a reprint of your publication of 1926, in which you present a well-known proof; possibly also other publications, if such exist related to my father, especially to the abovementioned work in *Hamburgsche Abhandlungen*.
>
> Last year in Zürich I tried to find your name in the telephone book. Unfortunately, I was unable to find there your name. I also tried to contact you at the University of Zürich, but also without result.
>
> As far as I can follow number theory, I find it exciting. If I were to become a mathematician, my inclinations would have certainly led me in this direction – in the direction of numerical mathematics and number theory. In my free time I continue to deal with Fermat and Mersenne; although "in general" with the history of mathematics. I would appreciate it very much if I could hear something from you and possibly you could send me one or several copies of your works of those where you have written about my father.

On October 20, 1965, Van der Waerden replied to Henry:

> It was very nice to receive a letter from you. I have not known your father and have never written anything about him. I heard about his [!] conjecture which he had posed at *Het Wiskundig Genootschap* (Mathematical Society) in Amsterdam.
>
> I am sending you a *overdrukje* (reprint) of my work from *Hamburger Abh.* and on loan a photocopy of my work in *Het Archief* from 1926. I will further ask the publisher Birkhäuser to send you a copy of my psychological research "Einfall und Überlegung" in which the history of the solution of this problem is also considered.

Thus, Van der Waerden stated that *P. J. H. Baudet posed his conjecture at the Mathematical Society in Amsterdam.* Van der Waerden attached to his letter copies of his original proof [Wae2] and his just published reminiscences [Wae16]. Henry Baudet II discussed this correspondence with Frederick Schuh, a major figure in the Amsterdam mathematical circles in the 1920s. This is why the following Henry Baudet's May 27, 1996 reply to my inquiry is the crux of the matter [BII12]:

> When I told Schuh about my correspondence with Van der Waerden, he would have definitely told me that the conjecture was not my father's, if it had been not his.

Schuh did not correct Henry Baudet, most likely because for him P. J. H. Baudet's authorship of the conjecture was a long-known fact.

After Henry Baudet the son, the next person who showed active interest in the authorship of the conjecture was N. G. de Bruijn. *Wiskundig Genootshcap* (Mathematical Society) decided to publish a 2-volume edition entitled *Two Decades of Mathematics in the Netherlands*: *1920–1940. A retrospection on the occasion of the bicentennial of the Wiskundig Genootschap.* The books were to reproduce short works of the leading Dutch mathematicians of the period, such as Van der Corput, Van der Waerden, Van Danzig, each followed by a commentary. Van der Waerden was to be represented by *Beweis einer Baudetschen Vermutung* [Wae2], with a commentary by de Bruijn, who in his March 29, 1977 letter posed to Van der Waerden several questions about the history of the conjecture. The latter replied on April 5, 1977 [Wae19]. I thank Nicolaas G. de Bruijn for sharing with me this very important Van der Waerden's letter an translating it from the original Dutch:

> I will happily answer your questions.
>
> 1. I am quite sure that I heard about the conjecture for the first time in 1926, around the time I got my Ph.D. in Amsterdam. I probably picked it up at one of the monthly meetings of the *Wiskundig Genootschap*, where Schuh appeared regularly. I do not know if it was Schuh himself or someone else who made me aware of this.
>
> 2. Yes, the entire affair happened on a single afternoon. Only the cases k=2, k=3, I had already figured out before.
>
> 3. I think I only later heard of I. Schur's proposition.
>
> 4. No, I do not know anything about Baudet. I have a vague memory that he was a friend or pupil of Schuh.
>
> 5. My biography: I have studied mathematics, physics, astronomy, and chemistry. Mathematics mostly with Mannoury, Hendrik de Vries and Brouwer. Astronomy with the excellent Pannekoek. In 1972 I retired in Zürich. Not "emeritus" because that does not exist in Switzerland.

> Included is the Bibliography with a few corrections. Furthermore, I have nothing to add to your piece. You praise "A thing of beauty is a joy forever,"[25] pleases me.

Thus, Van der Waerden got the conjecture in 1926 directly or indirectly from Frederick Schuh, Baudet's mentor and close friend, and the authorship of Baudet came to Van der Waerden with the conjecture. Van der Waerden has even "a vague" but correct memory that Baudet was Schuh's "friend and/or pupil." Thus, we have traced the way the conjecture traveled from Baudet to Van der Waerden via Schuh. However, one question remains open: Did Baudet independently create the conjecture or received it indirectly from Schur (try not to mix up here Schur and Schuh)? This is the question I was unable to address until December 18, 1995, when Henry Baudet II, the son and historian, came up with what he humorously named "*A Second Conjecture of Baudet*" [BII4]:

> It seems reasonable to suppose that neither Professor F. Schuh nor my father were informed of Schur's work. Though Germany was 'next door,' the World War broke nearly all contacts, which were only slowly restored in the course of the '20s.

And Henry Baudet II found convincing evidence to prove his conjecture. The first major mathematical event after World War I was unquestionably *Congrés International des Mathématiciens*, which took place during September 22–30, 1920 in Strasbourg, France. The whole world was represented there, with the notable exception of the German mathematicians, even though Strasbourg was located right by the French border with Germany. The wounds of World War I were still very painful. On the French initiative, the Germans were banned from the 1920 and the consequent 1924 International Congress of Mathematicians. It was not until the congress of 1928 that they were allowed to rejoin the world of mathematics.

Both J. A. Barrau and P. J. H. Baudet attended the Strasbourg's 1920 Congress. Baudet mailed to his wife *daily accounts* of his meetings at the Congress, and these letters have survived the long years and another war that followed. The letters report the meetings with a most impressive group of mathematicians: Denjoy, Fréchet, Valiron, Châtelet, Dickson, Eisenhart, Le Roux, Typpa, Lebesgue, Larmon, Young, De Vallée Poussin, Deruyts, Jordan, Montel, and Volterra [Bau3]. The letters also capture Baudet's impressions and emotions of days long gone [Bau5]:

> I am in nearly permanent contact with the Americans here. They are after all the nicest people here. And this is not only my opinion but also Barrau's. The nicest of all is Eisenhart. [Letter of September 29, 1920]
>
> At 11 P.M. all the cafés here are closed. You understand that this is not our cup of tea. It will be much better at our next Congress. That will be in the U.S. in 1924. The Americans here are really very nice people. Dickson and Eisenhart are their principal representatives, Eisenhart brought his wife who is quite a nice person. We talked a lot in these days, and she definitely expects you [Ernestine Baudet] too in the U.S. next time. You see: nothing can change it, you must join me next time. Barrau told Dickson about the critical review he [Barrau] had written and has modified after my severe critical

[25] The text in quotation marks is in English in Van der Waerden's otherwise Dutch letter.

> comments. The consequence of the discussion was that Dickson asked me to write him about the matter, as Barrau and I had here no copy of our controversial texts. [Letter of September 23, 1920]

Thus, Baudet and Barrau met Princeton's Professor and future Dean and Mathematics Chair Luther Pfahler Eisenhart. Do not forget his name: we will meet Eisenhart again later in this book when he will invite Van der Waerden to come and work at Princeton.

Baudet and Barrau also met and had discussions with the famous American number theorist Leonard Dickson. The meeting with Dickson attracted my attention in particular because Dickson's result inspired Issai Schur to come up with the Schur Theorem of 1916. However, this route only confirmed Baudet II's conjecture. Right before the Congress, in April 1920, Dickson had completed volume 2 of his monumental 3-volume *History of the Theory of Numbers* [Dic2]. He did cite there (p. 774) Schur's 1916 paper [Sch]: "* I. Schur gave a simpler proof of Dickson's theorem." But in the Preface Dickson explained that "the symbol * before the authors' names" signified "that the papers were not available for review," i.e., even Leonard Dickson, the most informed number theorist of his time, had not himself seen Schur's 1916 paper before the Congress.

Geographically speaking, Baudet and Schur had one chance to meet in August 1921, when Henry and Ernestine Baudet with their daughter Puck visited their friend and the legendary world chess champion Emanuel Lasker and had a short stay in his Berlin house. Puck "still has clear recollection of their stay at the Laskers, particularly when their rowing boat on the Wannsee[26] was wrecked,"[27] because neither Lasker nor Baudet could swim and had to be rescued. We are fortunate to have a photograph from this visit. Puck, however, does not remember visiting the University.

[26] You would recognize the name of *this* lake. The Lasker–Baudet humorous episode took place at the site where on January 20, 1942, fifteen high-ranking civil servants and SS officers discussed "The Final Solution" of the Jewish question in Europe. They decided to deport European Jews to the East and murder them all.

[27] [BII7].

Seated Ernestine, Puck and P. J. H. Baudet; standing (from the left) Emanuel Lasker and a Gymnasium Rektor, Lasker's house, August 1921, Berlin. (Courtesy of Henry Baudet II)

The family correspondence has survived, and it does not indicate that any new acquaintances were made during this Berlin trip, which took place just a few months before the untimely passing of P. J. H. Baudet.

Thus, it is plausible to conclude that Baudet and Schur never met and that P. J. H. Baudet discovered the conjecture independently from Issai Schur.

My investigation into the life of Pierre Joseph Henry Baudet was made possible by the invaluable and enjoyable help of Henry Baudet II and Nicolaas G. de Bruijn, to whom I extend my deepest gratitude.

37.7 Summing Up

The evidence presented here clearly shows that two brilliant men, Issai Schur and Pierre Joseph Henry Baudet, independently created the third conjecture of *Ramsey Theory before Ramsey*. From now on, let it be known as *The Baudet–Schur Conjecture*. What can be a happier conclusion to historical research!

Obviously, without the conjecture, no proof would have been possible. To conjecture such a pioneering result was surely as great a contribution as its proof by Bartel L. van der

Waerden. It is therefore fitting to call the monochromatic arithmetic progressions theorem after all three contributors: *The Baudet–Schur–Van der Waerden Theorem.*

At the time when Alfred Brauer's work [Bra1] proving two original Schur's conjectures appeared in 1928, Frank Plumpton Ramsey was working on his pioneering work [Ram2], that he submitted for publication later that year. A few years later, in 1932, Issai Schur's student Richard Rado defended his doctorate dissertation [Rad1], which was Rado's first fundamental contribution to Ramsey Theory. During the winter of 1932–1933 Paul Erdős and George Szekeres wrote their first Ramseyan paper [ES1]. Since then, Paul Erdős inspired many mathematicians around the world to enter the field. A new era of a maturing Ramsey Theory began.

I ought to point out amazing ways in which the lives of the players of this story are interwoven. Mentor and friend of Baudet, Frederik Schuh was instrumental in Van der Waerden getting to know the Baudet–Schur conjecture. Baudet's Ph. D. thesis *Promotor* (supervisor) was the very same Johan Antony Barrau, who in 1928, while moving to Utrecht, offered Van der Waerden his chair at Groningen, and again in 1942 proposed Van der Waerden for his chair at Utrecht. Read much more about it all in the following chapters, dedicated to vast generalizations of the Baudet–Schur–Van der Waerden Theorem and to my search for Van der Waerden the man.

The brutal war separated the authors of the Baudet–Schur–Van der Waerden Theorem and their families. As we have seen here, Baudet's son Henry Baudet II and his girlfriend Senta worked in the Dutch underground saving lives of Jews. Issai Schur was thrown out of the University of Berlin and following years of humiliation escaped to Palestine; his tired heart soon gave up. Being Dutch, Van der Waerden served as a Professor in Germany the entire Nazi time. Van der Waerden's life and fate prompted more controversy among scholars who care about historical truth than even presidential elections in the U.S. I had to get to the truth since no one before me had done it. Read the chapters dedicated to it later in this foliant.

Chapter 38
Monochromatic Arithmetic Progressions or Life After Van der Waerden' Proof

38.1 Schur's Generalization

And God said, "Let there be light."
– Genesis

And there was light: Issai Schur – who else – produced the first spark, a generalization of the Baudet–Schur–Van der Waerden Theorem. In fact, his result generalizes both, the Schur and the Baudet–Schur–Van der Waerden Theorems. With all the search engines of today's Internet, one would be hard pressed to find it, for it does not appear in any Schur's paper: This most modest man gave it to his former student Alfred Brauer to publish!

Alfred Brauer writes [Bra2] that a few days after his and Richard Brauer's 1927 visit of Schur, he proved Schur's conjecture about quadratic residues (conjecture 37.2), with the use of the Baudet–Schur–Van der Waerden Theorem. Schur then noticed that Brauer's method of proof can be used for obtaining a result about sequences of n-th power residues. Soon Issai Schur found a short, Olympiad-like, brilliant way to prove the following result that generalized *both* theorems.

The Generalized Schur Theorem 38.1 (I. Schur, [Bra1], [Bra2]). For any k and l, there is $S(k, l)$ such that any k-coloring of the initial array of positive integers $[S(k, l)]$ contains a monochromatic arithmetic progression of length l together with its constant difference.

Proof For 1 color, we define $S(1, l) = l$ and the statement is true.

Assume the theorem is true for k colors. We define

$$S(k+1, l) = W(k+1, (l-1)S(k, l) + 1),$$

where $W(k, l)$ is as defined in theorem 35.1. Let the initial array of integers $[S(k + 1, l)]$ be colored in $k + 1$ colors. Then by theorem 35.1 (see the right side of the equality above), there is a $(l - 1)S(k, l) + 1$ term monochromatic arithmetic progression

$$a, a+d, \ldots, a+(l-1)S(k, l)d.$$

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_38

For every $x = 1, 2, \ldots, S(k, l)$, this long monochromatic arithmetic progression contains the following l-term arithmetic progression:

$$a, a + xd, \ \ldots, a + (l - 1)xd.$$

If for one of the values of x, the constant difference xd is assigned the same color as the progression above, we have concluded the proof of the inductive step. Otherwise, the sequence

$$d, 2d \ldots, S(k, l)d$$

is colored in only k colors, and we can apply to it the inductive assumption to draw the required conclusion. ■

A great proof, don't you think! It is interesting to note here that unlike the Baudet–Schur–Van der Waerden Theorem, the Generalized Schur Theorem does not have a Szemerédi-style density generalization – see more about it later in this chapter.

Schur wanted Alfred Brauer to include this theorem (as well as the one about n-th power residues) in Brauer's paper because Schur believed to have used Brauer's method in these proofs. Schur did not want to take away any credit from his student. The student had to oblige but he "always called it Schur's result"[1] and gave Schur credit everywhere it was due in his paper [Bra1] that appeared in 1928. A few weeks later Brauer also proved Schur's conjecture about quadratic non-residues (conjecture 37.3), which appeared in the same wonderful, yet mostly overlooked paper [Bra1].[2]

Schur's ingenious contributions to *Ramsey Theory before Ramsey* do not end here. We will come back to them later in this chapter. For now, I wish to speak about density results.

38.2 Density and Arithmetic Progressions

Let us look at how this flourishing field has evolved. We will start with the key definition from the Erdős–Turán 1936 paper [ET]. Denote by $r_l(N)$ the maximum number of integers not exceeding N such that no l of them forms an arithmetic progression. Paul Erdős and Paul Turán proved a number of results about $r_3(N)$ and conjectured that

$$r_3(N) = o(N).$$

This conjecture was proved in 1953 by Klaus F. Roth [Rot]. The only conjecture about the general function $r_l(N)$ in Erdős–Turán paper was attributed to their friend George Szekeres and was later proven false. Sixteen years have passed before Endre Szemerédi proved in 1969 [Sz1] that

[1] [Bra2].

[2] I say 'overlooked' because the leading Ramsey Theory book [GRS2] contains almost identical result (theorem 2, p.70) without reference or credit to Schur.

$$r_4(N) = o(N).$$

In a 1973 paper, Paul Erdős [E73.21, pp. 118–119] remarks: "[this] very complicated proof is a masterpiece of combinatorial reasoning." A very surprising paragraph follows (ibid.):

> Recently, Roth [1970] obtained a more analytical proof of $r_4(n) = o(n)$. $r_5(n) = o(n)$ remains undecided. Very recently, Szemerédi proved $r_5(n) = o(n)$.

Clearly, Erdős added the last sentence in the last moment and should have removed the next to last sentence. The latter result has never been published, probably because Endre Szemerédi was already busy trying to finish the proof of the general case. On 4 April 2007, right after his talk at the Princeton Discrete Mathematics Seminar, I asked Szemerédi whether he had that proof for 5-term arithmetic progressions and what came of it. Endre replied:

> Hmm, it was so close to finding the proof of the general case, maybe two months before, that I did not check all the details for 5. It was more difficult than the general case.

Indeed, in 1974, he submitted and in 1975 published [Sz2] a proof of the general case, i.e., for any positive integer l

$$r_l(N) = o(N).$$

This work in one stroke earned Szemerédi the reputation of a wizard of combinatorics. Since then the terminology changed and I wish to present here a more contemporary formulation than the one used in Szemerédi [Sz2]. We will make use of the notion of "proportional length," known as *density*, in the sequence of positive integers $N = \{1, 2, \ldots, n, \ldots\}$. The *density* is one way to measure how large a subset of N is. Its role is analogous to the one played by length in the case of sets on the line R of reals.

Let A be a subset of N; define $A(n) = A \cap \{1, 2, \ldots, n\}$. Then *density* $d(A)$ of A is naturally defined as the following limit if one exists:

$$d(A) = \lim_{n\to\infty} \frac{|A(n)|}{n}.$$

The *upper density* $\bar{d}(A)$ of A is analogously defined as

$$\bar{d}(A) = \lim_{n\to\infty} \sup \frac{|A(n)|}{n}.$$

Now we are ready to look at a classically simple formulation of Szemerédi's result.

The Szemerédi Theorem 38.2 Any subset of N of positive upper density contains arbitrarily long arithmetic progressions.

In various problem papers, Erdős gives the date of Szemerédi's accomplishment and Erdős' payment as 1972 (sometimes 1973, and once even 1974). The following statement appears most precise as Erdős made it very shortly after the discovery at the 3–15 September 1973, International Colloquium in Rome [E76.35] and places Szemerédi's proof around September 1972:

> About a year ago Szemerédi proved $r_k(n) = o(n)$, his paper will appear in "Acta Arithmetica"...

Erdős was delighted with Szemerédi's result and awarded him $1000 in late 1972 to 1973 [E85.33]:

> In fact denote by $r_k(n)$ the smallest integer for which every sequence $1 \leq a_1 < a_2 < \ldots < a_l \leq n$, $l = r_k(n)$ contains an arithmetic progression of k terms. We conjectured
>
> $$(15) \qquad \lim r_k(n)/n = 0.$$
>
> I offered $1000 for (15) and late in 1972 Szemerédi found a brilliant but very difficult proof of (15). I feel that never was a 1000 dollars more deserved. In fact several colleagues remarked that my offer violated the minimum wage act.

On 4 April 2007, in a personal conversation, Szemerédi confirmed my historical deductions:

> I proved [the] general case in fall 1972 and received Erdős' prize in 1973.

I refer the interested reader to the original paper for the proof, which is brilliant and hard. Partial results are proved in [GRS2] (it is remarkable that even this standard text in the field did not include Szemerédi's complete proof). We do not know whether Szemerédi alone could have prepared his truly hard proof for publication: Ron Graham tells me that he spread numerous sheets around his house while trying to figure out and write up a proof one could understand.

While Szemerédi's Theorem is a very strong generalization of the Baudet–Schur–Van der Waerden Theorem, Paul Erdős, and Ronald L. Graham observe in their 1980 problem book [EG, p. 19] that the analogue of the Szemerédi Theorem does not hold for the Generalized Schur Theorem 38.1. Can you think of a counterexample before reading the one below?

Observation 38.3 (Erdős–Graham, 1980). Szemerédi-like generalization does not hold for the Generalized Schur Theorem.

Proof The set of odd integers of density ½ cannot contain even a 2-term arithmetic progression and its difference! ■

38.3 Who and When Conjectured What Szemerédi Proved?

As I mentioned earlier, throughout this book (and my life), I have given credit for a result to both the creator of the conjecture and the author of the first proof. Truly, without good conjectures, we would not have many good results. Moreover, pioneering conjectures, such as the Baudet–Schur, played a major role in paving the way for new mathematics. Our question here naturally is: Who and when conjectured what Szemerédi proved?

No one would expect a mystery here – just look at Szemerédi's 1975 paper, in which he presents the history of advances in good detail. He starts with giving credit for conjecturing his theorem to Paul Erdős and Paul Turán in their 1936 paper [ET]. And so, I am looking at this short important paper – without finding the conjecture, except for the case of 3-term arithmetic progressions. This incorrect credit is then repeated in the standard Ramsey Theory texts by Graham–Rothschild-Spencer [GRS1] and [GRS2] in 1980 and 1990 respectively, and from there on everywhere else, until 2002 when leaders of the field Ronald L. Graham and Jaroslav Nešetřil notice the discrepancy and explain it in the following way [GN, p. 356]:

> Although they [Erdős and Turán] do not ask explicitly whether $r_l(N) = o(N)$ (as Erdős did many times since), this is clearly on their mind as they list consequences of a good upper bound for $r_l(N)$: long arithmetic progressions formed by primes and a better bound for the van der Waerden numbers.

Clearly, my friends, Ron and Jarik, and I agree that the conjecture does not appear in the 1936 paper [ET]. Their argument that the young Erdős and Turán had the conjecture "clearly on their mind" could be viewed more as an eloquent homage to the two great mathematicians than a historical truth. Besides, mind reading is not a part of mathematics nor history. We therefore must research further.

In his 1957 first-ever open-problem paper [E57.13], Paul Erdős indicates that before him and Turán, Issai Schur (!) called on studying longest arithmetic-progression-free opening arrays of positive integers. Erdős writes:

> The problem itself seems to be much older (it seems likely that Schur gave it to Hildegard Ille, in the 1920's).

Erdős returns to Issai Schur's contribution in his second 1961 open-problem paper [E61.22], which in 1963 also appears in Russian [E63.21][3]:

> The problem may be older, but I cannot definitely trace it. Schur gave it to Hildegard Ille around 1930.

Paul told me that he "met Issai Schur once in mid 1930s," more precisely in 1936 in Berlin. They shared a mutual admiration (as we have seen in section 37.3). Undoubtedly, they discussed prime numbers, but likely not arithmetic progressions. Erdős learned about Schur's interest in arithmetic progressions and early Ramsey-like conjectures and results from Hildegard Ille (1899–1942). Now, this requires a bit of explanation because they probably had never met!

Erich H. Rothe (1895–1988), *Dr. phil. Universität Berlin* 1926 under the eminent Erhard Schmidt and Richard Mises, in 1928 married a fellow student Hildegard Ille (1899 – Iowa, 1942), *Dr. phil. Universität Berlin* 1924 under Issai Schur. They taught at *Universität Breslau*, Germany (later and earlier Wrocław, Poland) until, as the Jews, they were forced to flee Nazi Germany in 1937 and came to the United States. In 1942, Hildegard passed away at the young age of 43. The accomplished mathematician Erich Rothe held a professorship at the University of Michigan at Ann Arbor from 1941 until his retirement in 1964. His eulogy (*Notices of Amer. Math. Soc.*, 1988, 544) quotes Chair of the University of Michigan Math. Department D. J. Lewis saying that "Rothe was a scholar of the old school. He was very broadly educated . . . He was a wise and judicious man of much wit. His companionship was very much in demand."

Erich Rothe was Paul Erdős' source of reliable information on problems and conjectures in number theory that Issai Schur shared with Rothe's wife Hildegard (Ille) Rothe. From Erich

[3] This Russian publication does not appear in any of Paul Erdős' bibliographies.

Rothe, Erdős learned about Schur's authorship of the monochromatic arithmetic progressions conjecture, proved later by Van der Waerden (Chapter 37). From Erich Rothe, Erdős learned that Issai Schur yet again contributed to number theory and Ramsey theory when he asked his graduate student Hildegard to investigate arithmetic progression-free arrays of positive integers. To my surprise, *no one* before acknowledged credit Erdős gave to Schur in his first open-problem papers [E57.13], [E61.22] and [E63.21].

I believe, however, that Erdős learned about Schur being first to investigate this subject after Erdős and Turán independently rediscovered it: their paper [ET] was published in 1936, while Erich and Hildegard Rothe came to the Unites States in 1937; moreover, Erdős–Rothe conversations took place after Hildegard's passing in 1942. Paul was certainly correct when in both his 1957 and again 1961 open-problem papers he wrote, "The first publication on the function $r_k(n)$ is due to Turán and myself." This was an important paper, and Paul knew that. Yet, it contained the "density" conjecture only for 3-term arithmetic progressions. Graham and Nešetřil are correct when they write [GN] that "Erdős did [pose the general case conjecture] many times," but the real question is: when did he pose the conjecture for the first time?

I am reading yet again Erdős' first 1957 open-problem paper. Paul writes:

> In [ET] we stated our conjecture that $\lim r_3(n)/n = 0 \ldots$ Roth [Rot] proved that $r_3(n) = o(n) \ldots$ The true order of magnitude of $r_3(n)$ and, more generally, of $r_k(n)$, remains unknown.

Paul discusses the general function $r_k(n)$, but the conjecture in the general case is not here. If the conjecture were to exist consciously in his mind, he would have included it in this open-problem article, I am almost certain of it. Paul had not, and this, in my opinion, is a reliable indicator that the general conjecture did not exist yet in 1957.

In the second 1961 open-problem paper, Paul publishes the general conjecture explicitly for the first time:

> For $k > 3$ the plausible conjecture $r_k(n) = o(n)$ is still open.

This "*still* open" indicates that Erdős created the problem well before he submitted this paper, which was "Received October 5, 1960." This suggests the birth of the general conjecture in 1957–1959.

During his 23 December 1991 "favorite problems" lecture at the University of Colorado at Colorado Springs, Paul indicated the time when he offered for the first time the high prize of $1000 for *this* conjecture:

> Twenty-five years ago I offered $1000 for it.

This places the $1000 offer in 1966 or so. In early January 1992, in Colorado Springs, Paul confirmed to me in person that this was the highest prize he has ever paid:

> The maximum amount of money I paid [was] $1000 to Szemerédi in 1972. This was a conjecture of Turán and myself. If you have a sequence of positive density, then it contains arbitrarily long arithmetic progression.

Paul also told me then that "Turán and I posed this problem in the early 1930s." I hope, however, that my argument, presented here, indicates that it took time for the plot to thicken, that it was a long pregnancy, and from the early seeds in the 1930s the great conjecture had grown inside Paul Erdős' head and was born in 1957–1959.

Even after Szemerédi, Erdős was not quite happy with the state of knowledge in this field. In 1979, he offered an extravagant prize for the discovery of the asymptotic behavior (published in 1981 [E81.16]):

> It would be desirable to improve [lower and upper bounds] and if possible to obtain an asymptotic formula for $r_3(n)$ and more generally for $r_k(n)$. This problem is probably enormously difficult and I offer $10,000 for such an asymptotic formula.

Erdős' $10,000 Open Problem 38.4 Find an asymptotic formula for $r_3(n)$ and more generally for $r_k(n)$.

We have already witnessed Erdős directing research on the chromatic number of the plane and creating a good number of related problems. Here too Erdős is in the driver's seat (well, actually, Paul did not drive), following a prophetic start by Issai Schur.

Endre Szemerédi (born 21 August 1941, in Budapest; Ph.D. 1970, Moscow State University under Israel M. Gelfand) is the State of New Jersey Professor of computer science at Rutgers University and researcher in combinatorics and discrete mathematics division of Alfréd Rényi Institute of Mathematics in Budapest. In 1989, he was elected to the membership in the Hungarian Academy of Sciences. In 2012, Endre received the highest award a mathematician can win, the Abel Prize.

38.4 Paul Erdős' Favorite Conjecture

During our Colorado Springs joint work on (not yet finished) book *Problems of pgom Erdős*, between December 24, 1991 and January 9, 1992, I asked Paul which of his open problems were his favorite. Paul gave me a list of a few favorites. He started it with this problem [Soi29]:

> One of the most interesting problems is this: If you have a sequence the sum of whose reciprocals diverges, then for every r, there are r terms that form an arithmetic progression.

On another occasion during these two working weeks, Paul told me that he offered, not surprisingly, the highest prize for the same problem:

> The largest amount of money, which I offered really is: if you have a sequence of integers the sum of whose reciprocals diverges, then it contains arbitrarily long arithmetic progressions. This would imply in particular that the primes contain arbitrarily long arithmetic progressions. That is $3000.[4]

[4] In fact, already in 1980, the prize reached $3000. I read the footnote [EG, p. 11]: "One of the authors (P. E.) currently offers US $3000 for the resolution of this problem.

The Erdős \$3000 Conjecture 38.5 A set $A = \{a_1, a_2, \ldots, a_n, \ldots\}$ of positive integers, where $a_i < a_{i+1}$ for all i, with the divergent sum $\sum_{n \in N} \frac{1}{a_n}$, contains arbitrarily long arithmetic progressions.

What brought Paul Erdős to this conjecture? On September 15, 1979, in the problem paper [E81.16] submitted to the premier issue of *Combiniatorica*, Paul writes:

> In this connection I conjecture that if $\sum_{r=1}^{\infty} \frac{1}{a_r} = \infty$ then for every k there are k a_r's in arithmetic progression. Since Euler proved that the sum of the reciprocals of the primes diverges, our conjecture would settle the conjecture of primes . . . I offer 3000 dollars for the proof or disproof of the conjecture.

It appears that Paul Erdős offered for the first time his (then) largest prize, \$3000 in his 1976 talk "To the memory of my lifelong friend and collaborator Paul Turán" at the University of Manitoba, Canada Conference [E77.28]. (In the paper [E77.26] submitted the previous year, 1975, I see the prize of \$2500.) The highest prize and high frequency of including this conjecture in talks and papers indicate that this was one of Erdős' favorite conjectures. During his second talk at the University of Colorado at Colorado Springs on March 17, 1989, referring to this conjecture, Paul said [E89.61]:

> I should leave some money for it in case I leave. "Leave" means, of course, get cured of the incurable decease of life.[5]

The prize stood at \$3000 for nearly two decades, when in one of his last problem articles [E97.18], written in 1996 and published posthumously in 1997, Paul raised the prize to \$5000:

> I offer \$5000 for a proof (or disproof) of this [problem]. Neither Szemerédi nor Furstenberg's methods are able to settle this but perhaps the next century will see its resolution.

Since, as Paul believed, it may be a while before this conjecture is proven, we ought to record it with the new, highest ever Erdős' (serious) prize:

The Erdős \$5000 Conjecture 38.5′ A set $A = \{a_1, a_2, \ldots, a_n, \ldots\}$ of positive integers, where $a_i < a_{i+1}$ for all i, with the divergent sum $\sum_{n \in N} \frac{1}{a_n}$, contains arbitrarily long arithmetic progressions.

One question remains: when did Erdős first pose this problem? I searched for evidence in the ocean of his writings and found three indicators. First, in a paper submitted on 7 September 1982 to *Mathematical Chronicle* (now called *New Zealand Journal of Mathematics*), that appeared the following year [E83.03], Paul writes:

> This I conjectured more than forty years ago.

In the same year, 1982, Paul spoke at the Conference on Topics in Analytic Number Theory in Austin, Texas. I read in the proceedings (published in 1985 [E85.34], p. 60):

[5] Quoted first in [Soi14]. Earlier Paul mentioned leaving some money for this conjecture in some of his papers, e.g., [E77.28].

> I conjectured more than 40 years ago that if $a_1 < a_2 < \ldots$ is a sequence of integers for which $\sum_{i=1}^{\infty} \frac{1}{a_i} = \infty$ then the a_i's contain arbitrarily long arithmetic progressions.

Thus, both of these publications indicate that the conjecture was posed before 1942. On the other hand, in the 1986 Jinan, China, Conference proceedings (published in 1989 [E89.35]) Paul writes (p. 142):

> About 30 years ago I conjectured that if $\sum_{n=1}^{\infty} \frac{1}{a_n} = \infty$, then the a's contain arbitrarily long arithmetic progressions.

This would date the birth of the conjecture at about 1956. This information allows us to conclude only that this important conjecture is old and was born somewhere between very early 1940s and mid 1950s. The conjecture is obviously hard, for in spite of all assaults, it remains open. Moreover, even its weakest $250 version has not been conquered:

Paul Erdős' $250 Conjecture 38.6 A set $A = \{a_1, a_2, \ldots, a_n, \ldots\}$ of positive integers, where $a_i < a_{i+1}$ for all i, with the divergent sum $\sum_{n \in N} \frac{1}{a_n}$, contains a 3-term arithmetic progression.

In his 1983 survey, Ronald L. Graham proposes a "related perhaps easier conjecture." This beautiful conjecture is still open today, which is a good indicator that it is not so easy as it may seem. Z^2 denotes the set of points in the plane (i, j) with integral coordinates i, j.

Graham's Conjecture 38.7 [Gra3]. If A is a subset of Z^2 and $\sum_{(i,j) \in A} \frac{1}{i^2 + j^2} = \infty$, then A contains a square.[6]

The Erdős $5000 Conjecture 38.5 is still open. However, the existence of arbitrarily long progressions of primes has been proved by two brilliant young mathematicians, Ben Green and Terence Tao [GT] (they first submitted their proof on April 8, 2004; the 6th revision is dated September 23, 2007). Quite expectedly, their result is an existence proof and does not help to construct long arithmetic progressions of primes. In the first edition of this book, the credit for the longest actually constructed example consisting of 24 terms, went to Jarosław Wróblewski, a mathematician from the Wroclaw University, Poland. The present World Record is AP27, an arithmetic progression of 27 primes, was constructed by Rob Gahan of Ireland on 23 September 2019[7]:

$$224584605939537911 + 81292139 \cdot 23\# \cdot n,$$

where $n = 0, \ldots, 26$ and $p\#$, called "p primordial," stands for the product of all primes not exceeding p (in particular, $23\# = 2 \times 3 \times 5 \times 7 \times 11 \times 13 \times 17 \times 19 \times 23 = 223092870$).

On 15 April 2010, I received an email from the celebrated Norwegian mathematician Helge Arnulf Tverberg (1935–2020). I admired him and his 1966 theorem, but never communicated with him before:

> Dear Professor Soifer,

6 In our convention, a square is a set of its 4 vertices.

7 https://www.primegrid.com/forum_thread.php?id=7012&nowrap=true#133172.

Reading your excellent colouring book, I suddenly recalled an episode from Canberra in 1988.There P.E. gave a problem lecture, and he described the (then) $3000 conjecture (38.5). After a short break he said that he would also pay $3000 for a counterexample. To this the famous group theorist B.H. Neumann commented immediately: "and $6000 for both?" That was so immediate that B.N. laughed as loud as everyone else, apparently one part of the brain had produced this joke before some other part had taken it in.

I must add that unfortunately I have always been extremely interested in much of the material in your book, but, although trying, never been able to prove any result in those fields. I once found an amusing trifle: an alternative proof of [Esther Klein's] tool 31.3. Points A, ..., E in the plane, in general position, define a natural drawing of K_5 and therefore, by Kuratowski's theorem, the segment AB, say, intersects CD, which makes quadrilateral ACBD convex.

Best regards from
Helge Tverberg

38.5 Hillel "Harry" Furstenberg

Two years after Szemerédi's combinatorial proof was published, which incidentally used the Baudet–Schur–Van der Waerden Theorem, in 1977 Harry Furstenberg published a totally different proof [Fur1], using tools of Ergodic Theory. In fact, in doing so, Furstenberg created a new field, *Ergodic Ramsey Theory*. "Both results are beyond the scope of this book," write the authors of the standard text [GRS2] about Szemerédi's and Furstenberg's proofs – they are beyond the scope of this book too, for my goal is to introduce ideas and the excitement of mathematics of coloring and "meet" the people behind these ideas.

I first met this remarkable mathematician in Keszthely on Lake Balaton in July 1993, where we celebrated Paul Erdős' 80th birthday with a fitting conference, attended by who-is-who in Erdősian mathematics. Hillel (Harry) Furstenberg looked exactly the way I imagined Moses (in different clothing, of course). In fact, he looked much more like the Prophet than Charleston Heston ever had, Hollywood make-up trickery notwithstanding. We then met at Princeton when Harry was an invited speaker. Yakov "Yasha" Sinai, the host of Harry's visit, invited me to a reception. I sat next to Harry and asked him to write his autobiography for this book, which he kindly agreed to do. Harry was born right when Adolf Hitler fired Issai Schur from his professorship, in the same city of Berlin. Harry, you have the podium, and I am joining your audience!

I was born in Berlin on 29/9/35. I have few recollections of Berlin of the time. I remember my sister (older than myself by 3 years) pointing out a boarded-up bakery, saying this was Hitler's bakery. Apparently, she (over)heard that because of Hitler this Jewish establishment had been closed off, I remember some visits to a synagogue. We actually lived next to one (33 Brunnenstrasse) which today is a perfume factory, with only a lintel giving evidence of the one time use as synagogue, because the words "This is the gate to the Lord, the righteous shall pass through" appear on it.

Already before Krystallnacht (8/11/38) some of my parents' Jewish friends had received expulsion orders from the Nazis. Our own expulsion order came soon after Krystallnacht and my parents frantically searched for shelter. One of my early recollections is that of the morning after Krystallnacht when the four members of my family lined up underneath the broken windows of our basement apartment viewing the damage. I was old enough to realize the seriousness of the occasion.

From letters I later found, I discovered that my parents had sent a request to the Australian government for asylum and were refused. I have no idea to how many other places we applied. Fortunately, an aunt of mine was able to deposit 1,000 pounds sterling with the bank of England, thereby obtaining for us temporary asylum. We arrived in England sometime in 1939, shortly before the Blitzkrieg over London. I remember the shelters in London, the women knitting, I remember the skies at night criss-crossed with searchlights, and I remember my mother, sister and myself being sent to Norfolk, out of the London danger. My father hoped very much at that time to come to America and join my mother's brother who had recently bought a poultry farm in East Brunswick, NJ. He had a health problem (a thyroid condition), and knowing the Americans were strict, he underwent what was at that time risky surgery to rectify the problem. He did not survive the surgery and my widowed mother took her two children to the U.S. where we arrived shortly before the outbreak of WW II. We stayed at my uncle's poultry farm for over a year, and I attended McGinnis Elementary School. Kindergarten, first and second grades were in one room. Two years in the room were enough for me, so that when after two years we moved to Manhattan, I found myself in third grade in PS 169, near 168th street where we lived. The Rabbi of the nearby synagogue that we attended convinced my mother that I should go to a Jewish Day School, and that she needn't worry that I'd become a Rabbi myself. I attended Yeshiva Rabbi Moses Soloveitchik through eighth grade and got the rudiments of a traditional orthodox education. I graduated that institution in 1948 and, again, with some persuasion by the Rabbi, continued at a Jewish High School called Talmudical Academy, now known as Yeshiva University High School. Spending some summers in summer-school I finished high School in 1951 and continued at Yeshiva University, Since the college and high school were located in the same building I had already in high school come under the influence of Professor Jekuthiel Ginsburg, editor of Scripta Mathematica, a journal devoted to historical and recreational aspects of Mathematics, and from whom I first heard of Paul Erdős, and believed even then that he must be very old. (Shlomo Sternberg was also a student at the high school at the time and we both had found our own proofs of the famous problem of showing that if two angle bisectors of a triangle are equal then the triangle is isosceles, and we went to share our discoveries with Professor Ginsburg. Thenceforth he would regularly give us problems to solve.) Prof. Ginsburg realized that for me to devote myself to mathematics, I would need an income, which he obtained for me by having me do editorial work for Scripta. I learned to draw diagrams that were used in the magazine, and I sharpened my mathematical German and French by translating papers sent to the journal in those languages. I don't recall now if any of those translations were ever actually used. In the early fifties, Ginsburg took advantage of his friendship with various prominent mathematicians and set up a graduate school in mathematics at Yeshiva University. The first staff members traveled to Y. U. from their home institutions: Eilenberg and Kolchin from Columbia, Jesse Douglas from City College, Gelbart from Syracuse. I graduated in 1955 receiving both a B.A. and an M.Sc. degrees.

I continued at Princeton, having made the decision not to pursue a rabbinic career at Y. U. and quickly came under the influence of Salomon Bochner who took an interest in me because of his own religious background, and I imagine he found in me someone with whom he could share ideas in a long abandoned area of his past experience. (His father was an accomplished Jewish scholar, and Bochner kept in his office a portion of his father's library which with its annotated volumes attested to his father's scholarship.)

I received my Ph.D. in 1958 and two weeks later was married to Rochelle Cohen from Chicago, whose grandparents had immigrated to the U.S. from Poland. I spent one year as instructor at Princeton, followed by two years at MIT as C.E. Moore Instructor. Following a path taken by Eugenio Calabi (Bochner student – MIT and University of Minnesota) we moved to Minneapolis where we lived from '61 to '65 except for one year I spent as a visitor to Princeton ['63–'64]. During this time I was negotiating taking a position in Israel at the Hebrew University and in the summer of 1965 we made our move, spending first several months in Paris with a Sloan fellowship which provided our income during the year of our move. I also took a half-time position at Bar-Ilan University, and I'm proud particularly of Alex Lubotzky who was my Ph.D. student at Bar-Ilan, and is now my colleague at the Hebrew University.

* * *

Harry was professor at the Hebrew University's Einstein Institute of Mathematics from 1965 until his retirement in 2003. At the time of our meeting at Princeton, he split his time between the Hebrew University and Yale. He won major awards: Israel Prize (1993) and Eolf Prize in Mathematics (2006/2007). Furstenberg shared the highest award, Abel Prize, with Grigory Margulis (2020) "for pioneering the use of methods from probability and dynamics in group theory, number theory and combinatorics."

Furstenberg both created a new field of mathematics, the Ergodic Ramsey Theory, and founded a school in this new field. This manifested itself in 1996, when Furstenberg's scientific son and grandson joined together in generalizing Furstenberg's result. Vitali Bergelson (Ph.D. under Furstenberg 1984 at the Hebrew University, born in Kiev in 1950) and Alexander Leibman (Ph.D. under Bergelson in 1995 at the Technion, born in Moscow in 1960), both presently at Ohio State University, obtained [BL] what is often called the *Polynomial Szemerédi Theorem*. The authors in their paper give several versions of their result. Here is one, most relevant to our theme (it is the authors' Theorem B_0 for $l = 1$):

The Bergelson–Leibman Theorem 38.8 [BL]. Let $p_i(x)$, $i = 1, \ldots k$, be polynomials with rational coefficients taking on integer values on integers and with the zero last coefficients, i.e., $p_i(0) = 0$. Then any subset of positive integers N of positive upper density contains for any array of integers $v_1, v_2, \ldots, v_n$ a set of the form

$$\{a + p_1(x)v_1, a + p_2(x)v_2, \ldots, a + p_n(x)v_n\}$$

for some $a, x \in N$.

In particular,

The Bergelson–Leibman Theorem, Version II, 38.9 Let $p_i(x)$ be polynomials with integer coefficients with the zero last coefficients, i. e., $p_i(0) = 0$. Then any subset of N of positive upper density contains a set of the form

$$\{a + p_1(x), a + p_2(x), \ldots, a + p_n(x)\}$$

for some $x \in N$.

You can easily observe the validity of the following corollary that Paul O'Donnell will use in his result later in the book:

BLT's Corollary 38.10 For any positive integers *m, r,* any *r*-colorings of the set *N* of positive integers contain arbitrarily long monochromatic arithmetic progressions whose constant difference is an *m*-th power of a positive integer.

Presently, new exciting developments came from the pen of Vitali Bergelson. I have got to share those with you however briefly.

38.6 Bergelson's AG Arrays

In 2005, Vitaly Bergelson [Ber] extended the Ramseyan hunt for arithmetic progressions to *geoarithmetic progressions.* The following two easy exercises highlight the setting better than any words.

Proposition 38.11 Any coloring of positive integers *N* in finitely many colors contains arbitrarily long monochromatic *geometric* progressions.

Proof Given *m*-coloring *C* of the set *N,* and a positive integer *k*. Pick an integer t, $t > 1$. The coloring *C* of the whole set *N*, of course, assigns colors to all elements of the subset $\{t^n : n \in N\}$. Now we get the new coloring C' of the set *N* by assigning the color of t^n to *n*. For the coloring C', the Baudet–Schur–Van der Waerden theorem guarantees the existence of an *n*-term monochromatic arithmetic progression a, $a + d$, ..., $a + (k - 1)d$. The numbers t^a, t^{a+d}, ..., $t^{a+(k-1)d}$ form a *geometric* progression, and under the original coloring *C,* they are assigned the same color. ■

This proposition shows that we need to look for the existence of something more sophisticated than geometric progressions. Bergelson looked for an appropriate new term: he used *AG set*, then *geoarithmetic progression*. I propose a term *array* as more descriptive, as we really have here a square array of numbers.

Geoarithmetic array – or for short *AG array* – *of rank k* is a set of the form

$$\left\{r^j(a + id); i,j \in \{0, 1, \ldots, k\}\right\}.$$

Observe: an *AG* array contains lots of arithmetic and geometric progressions, and more.

Proposition 38.12 There is a set of positive (additive) density that contains no 3-term *geometric* progressions.

Proof Just pick the set of square-free positive integers. ■

This proposition shows that we need here a different notion of density, a sort of geometric density. In his introduction, Bergelson offers an example of what this means.

A set $A \subseteq N$ is *multiplicatively large* if for some sequence of positive integers a_1, $a_2, \ldots, a_n, \ldots$

$$\lim_{n \to \infty} \sup \frac{|A \cap a_n F_n|}{|a_n F|} > 0,$$

where $F_n = \left\{ p_1^{i_1} p_2^{i_2} \ldots p_n^{i_n} : 0 \leq i_j \leq n, 1 \leq j \leq n \right\}$ and where $\{p_i\}$ is the sequence of primes in some arbitrarily preassigned order.

We are ready now to look at a special case of Bergelson's result.

The Bergelson Theorem 38.13 Let $A \subseteq N$ be a multiplicatively large set. Then A contains AG arrays of arbitrarily large rank.

Observe: for any coloring of N in r colors, at least one of the monochromatic sets is multiplicatively large and thus contains AG arrays of arbitrarily large rank. It is clear that Vitaly Bergelson and his coauthors are up to vast generalizations of the celebrated results of Ramsey Theory. I wish them much success.

38.7 Van der Waerden's Numbers

Through Issai Schur, Richard Rado was aware of the Baudet–Schur–Van der Waerden theorem from the beginning, and generalized it, but in his early years he did not seem to be much interested in numerical bounds. On the other hand, already in his 1935 celebrated joint paper with George Szekeres, Paul Erdős showed interest in numerical bounds of combinatorial functions. So, when the leaders of Ramsey Theory Erdős and Rado got together in 1951, the result was the paper ([ER] (read November 15, 1951; published 1952), that poineered quantitative evaluation of Van der Waerden's numbers. Having addressed the Ramsey Theorem, Erdős and Rado created *Van der Waerden's function*, and therefore *Van der Waerden's numbers* (they do not use the word "numbers" per se, but what are Van der Waerden's numbers if not values of Van der Waerden's function?), and introduce a natural notation $W(k,l)$ for both:

> The last example of the paper is not concerned with Ramsey's theorem but with the following theorem due to van der Waerden [Wae2]. Given positive integers k and l, there is a positive integer m such that, if the set $\{1, 2, ..., m\}$ is divided into k classes, at least one class contains l+1 numbers which form an arithmetic progression. The least number m possessing this property is denoted by $W(k,l)$, (*van der Waerden's function*).

> Our final example yields what seems to be the first non trivial, no doubt, extremely weak, lower estimate of W, namely $W(k, l) > ck^{(1/2)}l^{(1/2)}$. An upper estimate of W, at any rate one which is easily expressible explicitly in terms of the fundamental algebraic operations, seems to be beyond the reach of methods available at present.

The Erdős–Rado notation $W(k,l)$, in today's conventions, would stand for $W(k,l + 1)$. The second variable, as used today (and in theorem 35.1), stands for the number of terms in the arithmetic progression. When the number of colors is $k = 2$, we simply omit the first variable: $W(l) = W(2,l)$.

Observe, the Erdős–Rado's interpretation of the notation simplifies statements of some results. For example, best lower bound, due to Elvyn R. Berlekamp, is simpler in the Erdős–Rado notation, than in the one he used in the original paper [Berl]:

Lower Bound 38.14 (Berlekamp, 1969). $W(k) > k2^k$ if k is a prime (Erdős–Rado's understanding of the notation is used).

In today's standard notation (where the variable k stands for the number of terms in AP), the result reads as $W(k + 1) > k2^k$.

Surprisingly, Berlekamp's result remains the best known for primes after five decades. In 1990, Zoltán Szabó, using Lovász' Local Lemma, found the best known lower bound for all n [Sza].

Lower Bound 38.15 (Szabó, 1990). For any $\varepsilon > 0$, $W(k) \geq \frac{2^n}{n^\varepsilon}$ for large enough n.

The upper bound has withstood all assaults for decades. Erdős writes in 1957 (I have simply changed the notation to the one used today), [E57.13]:

> All known functions $W(k)$ increase so rapidly that they do not even satisfy the condition
>
> $W(k) = k^{k^{\cdot^{\cdot^{k}}}}$ (k exponents).

The problem was that all known proofs of the Baudet–Schur–Van der Waerden Theorem used double induction. This prompted doubts even in such a mathematical optimist as Paul Erdős, who wrote in 1979 [E81.16]:

> Until recently nearly everybody was sure that $W(k)$ increases much slower than Ackermann's function. I first heard doubt expressed by Solovay which I more or less dismissed as a regrettable aberration of an otherwise great mind. After the surprising results of Paris and Harrington [PH] Solovay's opinion seems much more reasonable, and certainly should be investigated as much and as soon as possible.

Yet, Ron Graham persisted with optimism and bet \$1000 on it in his 1983 survey [Gra2]:

> There is currently no known upper bound for $W(k)$ which is primitive recursive.[8] This is because all available proofs leading to upper bounds involve at some point a (perhaps intrinsic) *double* induction, with k as one of the variables. This leads naturally to rapidly growing functions like the Ackermann function which may help to explain the enormous gap in our knowledge here. The possibility that $W(k)$ might in fact actually have

[8] See [Soa] for definitions and comparison of rapidly growing functions.

> this Ackermann-like growth has been strengthened by the work of Paris and Harrington [PH], Ketonen and Solovay [KS], and more recently Friedman [Fri], who show that some natural combinatorial questions do indeed have *lower* bounds which grow this rapidly (and even much more rapidly ...). In spite of this potential evidence to the contrary, I am willing to make the following [conjecture].

Graham then formulated the conjecture for first proof (or disproof) of which he was offering $1000 ever since the late 1970s:

Graham's $1000 Van der Waerden's Numbers Conjecture 38.16 [Gra2].

$$W(k) < 2^{2^{\cdot^{\cdot^{2}}}}$$

for $k \geq 1$, where the number of 2's is k.

Paul Erdős asked for less, just for a primitive recursive upper bound, in the 1984 conference talk in Japan, published the following year [E85.33; p. 75]:

> I give 100 dollars for a proof that f(n) is primitive recursive and 500 dollars for a proof that it is not.

Ron's and Paul's expectations were soon rewarded. Saharon Shelah proved exactly what the doctor ordered (I mean Doctor Erdős): Shelah's *Primitive recursive bounds for van der Waerden numbers* [She1] was published in 1988 "with a beautifully transparent proof," as Gowers commented later [Gow, p. 466].

Shelah's Upper Bound 38.17 [She1]. Van der Waerden's numbers are primitive recursive.

Ron Graham described this event in the December 29, 2006 e-mail to me:

> I gave Shelah the check [a consolation $500 prize for conjecture 38.16] when he was lecturing at Rutgers (as you know, he visits there for 2 months each year). It was shortly after he proved his bound, which was somewhat before it was published. Incidentally, the original title of his paper was quite different from what appeared!

Erdős too gave Shelah the highest praise in many talks. Here, for example, is a quotation from Erdős' 1988 talk at the 7th Fischland Colloquium in Wustrow, Germany [E89.27]:

> This was certainly a sensational triumph.

Shelah's result inspired Paul Erdős to pose a new, most challenging conjecture. In [E94.21], first submitted on January 25, 1993, and published a year later, Paul Erdős wrote:

> It was a great achievement when a few years ago Shelah gave a primitive recursive bound for $W(k)$. Probably, this bound was still much too large, perhaps $W(k) < 2^{2^k}$.

We thus get Paul Erdős' conjecture, which he repeated in 1996 (posthumously published in 1997 [E97.18]):

Paul Erdős' 1993 Van der Waerden's Numbers Conjecture 38.18 [E94.21].

$$W(k) < 2^{2^k}.$$

In 1998, Timothy Gowers announced and in 2001 published his incredible 124-page *A New Proof of Szemerédi's Theorem.* His upper bound for Van der Waerden's numbers appears on the next to last page as "corollary 18.7":

Gowers' Upper Bound 38.19 [Gow]. Let *k* be a positive integer and let $N \geq 2^{2^{2^{2^{2^{k+9}}}}}$. Then however set $\{1, 2, \ldots, N\}$ is colored with two colors, there will be a monochromatic arithmetic progression of length *k*.

In other words,

$$W(k) \leq 2^{2^{2^{2^{2^{k+9}}}}}.$$

In answering my inquiry, Ron Graham wrote to me in the 28 December 2006 e-mail:

> Regarding the payment to Gowers, I gave him the check during a talk I gave in Hungary (again in connection with celebrating Erdős' mathematics but I'm not sure of the exact year). I attach a photograph showing the actual presentation. I interrupted my talk and came down into the audience to give him the check!

In one of my talks, I used this photograph that Ron kindly provided. One day, I received a communication from Tom C. Brown, a professor at Simon Fraser University in Vancouver, Canada. Imagine, he was the photographer of Graham–Gowers photo! Tom sent me copies of his correspondence with Paul Erdos, his reprints, and a fine quality photo, with all the details of the event:

Date: July 07, 1999
Place: Hungarian Academy of Sciences, Budapest
Subject: Ron Graham gives $1000 check to Tim Gowers
Photographer: Tom Brown
(The time was either late morning or early afternoon.)

Ron Graham presenting the $1000 check to Tim Gowers; Photo by Tom C. Brown. (Courtesy of Ron Graham and Tom C. Brown)

Tim Gowers [Gow, p. 586] seemed to question whether he fully deserved the $1000 reward:

> Ron Graham has conjectured in several places (see e.g. [GRS2]) that the function $W(k)$ is bounded above by a tower of twos of height k. Corollary 18.7 [i.e., result 38.19 above] proves this conjecture for $k \geq 9$, and indeed gives a much stronger bound. It looks as though more would be needed to prove it for $k = 7$ (for example) than merely tidying up our proof. For $k \leq 5$, the exact values of $W(k)$ are known and satisfy the conjecture.

Gowers should not worry. Graham's $1500 ($500 to Shelah and $1000 to Gowers) is clearly the money best ever spent in the encouragement and support of mathematical research.

As to Ron Graham, as soon as he paid Tim Gowers, he offered another $1000 conjecture [Gra7], [Gra8]. Prefacing this conjecture, Graham wrote [Gra7], [Gra8]:

> In particular, this [Gow] settled a long-standing conjecture I had made on the size of $W(n)$ . . ., and as a result, left me $1000 poorer (but much happier). Undaunted, I now propose the following:

Graham's 2007 \$1000 Van der Waerden's Numbers Conjecture 38.20 [Gra6], [Gra7]. For all *k,*

$$W(k) < 2^{k^2}.$$

Observe, that for $k > 3$, we have $2^{k^2} < 2^{2^k}$, thus, Graham's 2007 conjecture is harder – if true – than Paul Erdős' 2003 conjecture 38.18. Which one is "better"? Only time will tell – a very long time, I believe.

We have discussed here the asymptotic behavior of the function $W(k)$. So little is known about its exact values for small k that in their 1980 monograph [EG] Erdős and Graham exclaimed "It would be very desirable to know the truth here." A few values were found in 1969 by Vašek Chvátal [Chv] (first three) and in 1978 by R. S. Stevens and R. Shanturam [StSh] (the last one):

$$W(2) = 3$$
$$W(3) = 9$$
$$W(4) = 35$$
$$W(5) = 178$$

In spite of all dramatic improvements in computers, no further values have been computed in three decades that followed, when in 2008 Michal Kouril and Jerome L. Paul determined the next Van der Waerden's number [KoP]:

$$W(6) = 1132.$$

For other Van der Waerden's numbers (cases when more than 2 colors are used, or settings are non-symmetric), please, consult section 2.3 of the impressive 2004 monograph [LR] by Bruce M. Landman and Aaron Robertson.

It is time to say a few words about our genius record holders.

38.8 Saharon Shelah

The time stamp – late 1974; the place – Moscow. I went to Anna Petrovna Mishina's Abelian Group Seminar at Moscow State University. She told us that the young Israeli mathematician Saharon Shelah had just published a solution of the Whitehead problem.[9] This was a sensational news, for everyone who was somebody in Abelian Group Theory tried to solve this problem – and failed. Better yet, the answer was not a yes or a no, as we all expected, but "it depends" – depends upon the system of axioms for set theory!

[9] Must an Abelian group G with $\mathrm{Ext}(G, Z) = 0$ be free?

Nine years later. Roll forward to the spring of 1984. As an American, attending the Abelian Groups and Modules Conference in Udine, Italy, dedicated to László Fuchs's 60th birthday, I was introduced to Saharon Shelah at a dinner the night before the conference's opening. Answering his question "what are you working on?" I shared with Saharon my problems and conjectures. The following morning Saharon invited me to his hotel room and, to my surprise and delight, offered to collaborate on solving my problems. Right there he handed in to me a page with a *finite* lemma, which was the only element I was missing for settling one of my conjectures dealing with *uncountable* abelian groups! His question "Why do people attend conferences?" – I answered quite traditionally, "To show their latest results, to learn about achievements of others, and to socialize." "None of this makes any sense," Saharon replied, and added "People should attend conferences in order to solve together problems they could not solve on their own." And so, I missed many talks, was not allowed by my new coauthor to drink wine (and that is in Italy!), but by the end of the week, we solved all my problems and proved all my conjectures – this became the subject of our two joint papers in the *Journal of Algebra*, recommended for publication by the wonderful gentleman and mathematician David Buchsbaum of Brandeis University. (David was one of my first American references in 1978–79, and we met at Brandeis and also in my Boston apartment.) It was a special, inspirational experience to work with Saharon. It also required a full concentration, for he was an amazingly quick learner and thinker. I was impressed by his ca. 400 publication; however, Saharon observed that it was a far cry from Paul Erdős' 1500 works. On the conference's excursion day, I was sharing a bench on the bus with László Fuchs, the honoree of the conference. "I am working with Saharon, and he is a genius," I told László. "But of course," replied he, as if it was something obvious.

A good fortune preserves the photo of the Honoree Laszlo Fuchs with Saharon Shelah and me.

From the left: Alexander Soifer, Laszlo Fuchs, and Saharon Shelah, Udine, Italy, April 12, 1984

Nine years later. The night before the opening of Paul Erdős' 80th birthday conference in the summer of 1993, Saharon arrived very late in Keszthely on Lake Balaton, Hungary, and invited me to join him right away for an 11 p.m. supper. During the meal, I told Sharon all I knew about the chromatic number of the plane problem. He was excited and after supper left to sleep on it. The next morning Saharon simply said, "I have not seen the light." He has a philosophical view on choosing his battles, which he shared once with me: "Nobody cares how many problems I cannot solve – people care only how many I can."

Nine years later. In September 2002, for the first time, we met in the United States. Saharon invited me to his Rutgers University in Piscataway, New Jersey, for a week of fun of the mathematical kind. This was a productive week. To our own surprise, we showed that the chromatic number of the plane may depend upon the system of axioms we choose for set theory. We constructed a distance graph on the line, whose chromatic number was 2 in the standard **ZFC** system of axioms for set theory, and uncountable in **ZFS**. I will share with you what **ZFS** stands for and many details and results of this meeting and its results in one of the last chapters of this book. Saharon worked in such a complete concentration that I noticed him wearing one blue and one brown sock. The next day the color coordination remained unchanged. On the third day (like in fairytales), he wore a matching pair of brown socks. This is how I was able to conjecture that his wife Yael arrived from Israel and joined Saharon in New Jersey – she did. We met again in the fall 2003 and extended our construction from the line to the plane.

Saharon Shelah was born in Jerusalem, Israel, on 3 July 1945. He is the Abraham Robinson Professor of Mathematical Logic at the Einstein Institute of Mathematics of the Hebrew University, Jerusalem, and the Distinguished Visiting Professor at Rutgers University, Piscataway, New Jersey, where he spends every September and October. He is one of the great problem solvers of all time, who has won numerous awards, including the first Erdős Prize (1977), Karp Prize of the Association for Symbolic Logic (1983), George Pólya Prize (1992), Israel Prize for Mathematics (1998), János Bolyai Prize (2000), Wolf Prize (2001), The EMET Prize for Art, Science and Culture (2011), Leroy P. Steele Prize for Seminal Contribution to Research (2013), Hausdorff Medal of the European Set Theory Society (2017), Schock Prize in Logic and Philosophy of the Royal Swedish Academy of Sciences (2018), and the Honorary Doctorate from the Technische Universität Wien (2019). The count of his papers has now surpassed 1300. Saharon has also authored 9 major books. Since he has had over 200 coauthors, we can initiate the Shelah number not unlike the Erdős number.

38.9 Timothy Gowers

William Timothy Gowers, born on November 20, 1963 in Wiltshire, received his doctorate at the University of Cambridge under the famed Hungarian combinatorialist Béla Bollobás. Following productive years 1991–1995 at the University College London, he has been a Fellow of Trinity College and the Rouse Ball Professor of Mathematics at the University of Cambridge.

In 1998, Gowers won the Fields Medal and a year later was elected Fellow of the Royal Society. Having attended his talks at Princeton-Math, I can attest to the elegance and lucidity of Tim's presentations of his great combinatorial results. He is an expositor of mathematics as well, with *Mathematics: A Very Short Introduction* to his credit, and a much-much longer "introduction" to mathematics: *The Princeton Companion to Mathematics*, 2008.

Chapter 39
In Search of Van der Waerden: The Early Life

I can tell you that I am very much impressed with the thoroughness and integrity whereby it is written. I was also amazed that you have been able to collect so many facts, letters and data from that period. So much work! Your description is very objective but humane and it is most interesting how out of all these facts slowly one gets an image of a real person of flesh and blood behind these facts.

– Dorith van der Waerden

Thank you for sending me your triptych, which I read with great interest! This history is so complex, but you got so much information, I was astounded. Reading was very compelling – my greatest compliment for the study you made.

– Theo van der Waerden

39.1 What You Will Find in This and the Following Three Chapters

My distinguished colleagues John J. Watkins and Robin J. Wilson called *The Mathematical Coloring Book*, 2009 [Soi44] "labor of love." However, in a friendly way, they opined that I wrote too much about Van der Waerden, too much labor or too much love. Their remark is well taken, but . . . What am I to do? To simply state that my 20-year-long archival research proved that Van der Waerden was not "a Nazi collaborator" nor "a strong anti-Nazi"? and instead of proof declare Trump's "believe me"? In my entire life, I've never asked anyone to believe me – this is *your* choice. So, I have to present as rigorous a proof as the historical genre allows in this complicated case. Moreover, that is not my only goal here.

I believe that moral principles lie not outside of the profession, but rather are a critical part of its foundation. And my research into the life of Van der Waerden and to a degree his friend Werner Heisenberg is a rich lesson on the moral foundations of mathematics and science and on the value of human life on our planet. So, if you are not too curious about these problems, you could just read the sections "Van der Waerden and Van der Corput: A Dialog in Letters," Heisenberg's "On Active and Passive Opposition in the Third Reich," "The *Het Parool* Affair," and "From Rolf Nevanlinna Prize to Abacus Medal: A Noteworthy History of an IMU Prize."

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_39

This and the following three chapters are not the same that appeared in the first 2009 edition of *The Mathematical Coloring Book* [Soi44]. Instead, they are a *very condensed version* of my 2015 definitive 500-page foliant *The Scholar and the State: In Search of Van der Waerden* [Soi47]. Here, I will omit almost all footnotes and references to archival documents – this would make your reading "smoother" and perhaps more enjoyable. I am omitting revealing pages of secret recordings of the ten distinguished German physicists during their 1945 detention in Farm Hall near Cambridge, England, and the Swiss years 1952–1996 of Van der Waerden. If you are interested in a complete report of my 20+ year historical research, with many hundreds of supporting documents and many rare photographs, you would benefit from reading the complete book [Soi47]. On the other hand, I am adding here an important new material that has surfaced after *The Scholar and the State* was published in 2015: the renaming by the International Mathematics Union of the Rolf Nevanlinna Prize after its 40-year-long existence, an unwritten agreement in post-World War II Germany to forget everything and forgive almost everyone; the Russian war on Ukraine that started in 2014 with the support of many Russian cultural celebrities, including some mathematicians, etc. The Past repeating itself in the Present comprises three new sections of this book: Today I, Today II, and Today III.

39.2 Why Van der Waerden and Why Me?

Good 'history' is possible when historians take the initiative to undertake their own investigations of what has been accepted as 'fact.'

– Harriet Sepinwall

It is hard to be a historian. It is difficult if you have not lived in the time you write about, and if you have, it is even worse.

– Nicolaas G. de Bruijn
e-mail to A. Soifer, 6/1/2004

In 1990, I started research for the first edition of *The Mathematical Coloring Book* [Soi44]. In addition to presenting *mathematics of coloring as an evolution of ideas*, I chose to include biographies of major creators of these ideas. Van der Waerden proved a critical theorem about monochromatic arithmetic progressions in finitely colored integers; thus, I had to include his biography. I first thought that I could simply quote biographies, if not from *Encyclopedia Britannica*, at least from scholarly books and articles. I immediately ran into irreconcilable contradictions in published texts. For example, many authors stated that Van der Waerden was a strong anti-Nazi, while others alleged him to be a Nazi collaborator. It soon became clear that I had to do my own historical research if I wanted to present the truth. Of my many biographies in this book, Van der Waerden's was the most complex and controversial. Consequently, it took the longest – over 20 years – to research his life in many archives of Germany, Holland, Switzerland, and the United States, and interview senior, reliable informants.

Bartel Leendert van der Waerden was a distinguished mathematician and historian of science. In 1927, he published the theorem on monochromatic arithmetic progressions in finitely colored integers [Wae2] that later became a classic. Together with the 1916 Schur Theorem on monochromatic solutions of the equation $x + y = z$ in finitely colored integers [Sch], the Baudet–Schur–Van der Waerden Theorem gave birth to *Ramsey Theory*, before Ramsey's posthumous 1930 publication [Ram2].

Van der Waerden made major contributions to algebraic geometry, abstract algebra, group theory, number theory, combinatorics, analysis, and statistics. In addition to mathematics, he contributed to quantum mechanics, wrote on the psychology of discovery, and published liberally on the history of mathematics, astronomy, and natural sciences in antiquity. Among the many books, Van der Waerden wrote the two-volume *Moderne Algebra* [Wae3], one of the most influential mathematical books ever written. This 1930–1931 book is still in print today, nearly a century later!

Clearly, Van der Waerden deserved a book-length biography. It is surprising that when I commenced my research in 1990, no such book existed. "Why is that?" I asked N.G. de Bruijn, who in 1952 accepted Van der Waerden's chair at the University of Amsterdam. In his reply, De Bruijn shared with me what I would call "Theory of Matters Biographical" [Bru7]:

> My advice to scientists who would like to have books about them after their death is (apart from obvious things like doing important work and having lots of students):
>
> 1. Stay in your country.
> 2. Stay in a single subject.
> 3. Don't get old.
> 4. And, if you do happen to get old: try to write an autobiography.
>
> Van der Waerden missed points 1, 2, and 3 and was too modest to write an autobiography.

I discovered an enormous volume of documents – ca. several thousand – many not introduced earlier in historical scholarship, for example, Princeton and John Hopkins universities' job offers to Van der Waerden, the latter one accepted by him. I dug up contemporaneous newspapers not used before, which provided me with vivid snapshots of the day of their issue. I discovered that most witnesses of the Nazi era were reluctant to recollect those painful times. Some of my informants possessed vital information, and for the first time conveyed it to me. This book would not penetrate the subject as deeply without input from such eyewitnesses as N.G. de Bruijn, H.J.A. Duparc, Beno Eckmann, Paul Erdős, Bartel's niece Dorith van der Waerden and nephew Theo van der Waerden, and Bartel's son Hans van der Waerden.

While none of his three homelands, Holland, Germany, and Switzerland produced any books on the life of Van der Waerden – there were numerous biographical articles that missed or ignored the Dutch and American sources. Most authors penned celebratory articles aimed to fabricate Van der Waerden's image as a hero, moreover, a German hero. An *anonymous* referee of my 2015 biography [Soi47] complained about my "scratching at the star's brilliant image, creating thus a portrait full of antipodes. The reasons for this kind of portrait are hard to understand." I can only answer, *c'est la vie*, that's life, life is full of contradictions, and Van der Waerden had his share of them. Whoever says the truth shall die!

The Biblical wisdom agrees with me, "The waters wear the stones" (Job 14:19). Stonewalling the truth will sooner or later collapse. *The truth like water will find its way* *out.* Yes, my anonymous referee, I know that many of my colleagues believe that a scholar should be evaluated based on scholarly achievements alone, with no regard to moral bearings. However, I am with Albert Einstein, who on November 20, 1950, wrote, in English:

> The most important human endeavor is the striving for morality in our actions. Our inner balance and even our very existence depend on it. Only morality in our actions can give beauty and dignity to life.

Many of my predecessors apparently believed that a personal acquaintance with Van der Waerden automatically made them experts on his life. Their repetition of Van der Waerden's own words, such as "everyone knew ... that he was a strong anti-Nazi," uncritical copying from each other, mixed with "cheerleading," hardly added up to history.

Professor Miles Reid's approach in his 1988 Cambridge (!) University Press book [Rei] did not contribute to history either when he wrote:

> Rigorous foundations of algebraic geometry were laid in the 1920s and 1930s by Van der Waerden, Zariski and Weil (Van der Waerden's contribution is often suppressed because a number of mathematicians of the immediate post-war period, including some of the leading algebraic geometers, considered him a Nazi collaborator).

Even if "leading algebraic geometers," presumably Oscar Zariski and André Weil, had such an opinion, their fine mathematical achievements did not automatically make them custodians of the historical truth. It was very unfortunate that such a heavy accusation was leveled against Van der Waerden without any substantiation at all. In fact, Van der Waerden publicly criticized the Nazi regime from its inception until May 1935, at which time he was warned by the Leipzig University administrators that meddling in German political affairs could cost him his German professorship.

In these four chapters, I choose to write not about Van der Waerden's mathematics – we have looked at some of it already –but about the life and fate of a scholar, a fine person from a distinguished Socialist family, who finds himself in the Nazi tyranny and accepts certain compromises that in my opinion erode his moral grounds.

Van der Waerden was important to Holland. He was one of the two best Dutch mathematicians of the XX century (together with Brouwer). Bartel belonged to the family of the beloved Congressman Dr. Theo van der Waerden. Why was Van der Waerden so important to Germany that some German authors tried to make a *German hero* out of him? He was one of the most brilliant young mathematicians of Europe and authored the famous *Moderne Algebra* [Wae3]. Most of all, German authors adored Van der Waerden for sticking with Germany to the bitter end.

My approach to historical writing differs from most of the historical research literature in a number of essential ways. I unapologetically open my "kitchen" to you so that you can join me in my research, ride with me on the trains of thought, and feel the adrenalin of Dr. Watson when he joins Sherlock Holmes. I try to use the present tense as much as possible so that you and I can "*live with*" the personages of my narrative and not merely *read about* them. I often quote long documents generously to give you a flavor of the person and the epoch and to give the players in my drama greater roles while reserving a lesser part for myself. I try to stay close to documents and eyewitnesses, not going further than one step away from the evidence. I may disagree with the personages of my book and on occasion argue with them, but I treasure the life and work of B.L. van der Waerden, Johannes G. van der Corput, Werner

Heisenberg, Carl Friedrich von Weizsäcker, Peter Debye, Niels Bohr, Max Planck, Albert Einstein, Erich Hecke, Issai Schur, P.J.H. Baudet, Henry Baudet II, N.G. de Bruijn, Beno Eckmann, and others. I realize that in writing about them I open my own integrity to your judgment, and that is fair.

Little did I know when I commenced this research how passionately people feel still today about the Third Reich and World War II, European suffering, and the Holocaust. In a sense, *writing a book on these topics is akin to crossing a mine field: one wrong word – and you are history.* As an illustration, it suffices to recall overzealous attacks on Dr. Daniel Goldhagen for his book "Hitler's Willing Executioners: Ordinary Germans and the Holocaust."

I view history – as I do written word in any field, mathematics included – to be a genre of literary art. A book ought to be written by one human for another human, yet it often feels as if scholarly books, including history, are written by a robot for another robot. History is inevitably subjective. In fact, I believe that everything in this world is subjective, that a claim of objectivity is a lie, or at best a noble but unachievable goal. Even in writing my documentary prose, I must choose hundreds of documents out of thousands that I have assembled. My book is a means of my self-expression, and so I feel compelled to express my views on some important issues that the personages of my book are facing. While trying to be fair, I realize that some of you may disagree with my views and lessons I have learned from history. I hope they will publish their views, documents, and arguments, for in a *substantive constructive debate* we get closer to the ever-elusive truth.

I am most interested in history when observations of the past elucidate problems of the present and help us solve them. Analyzing Van der Waerden's life under the Nazis allows us an insight into problems of a scholar in a totalitarian state. You will readily realize that problems touched on here are with us today and merit our discussion here, not elsewhere. This super-objective of keeping in mind problems of today should not in any way affect the thoroughness of research into Van der Waerden's life. I pledge to be a historian first and a scholar concerned about current affairs second.

39.3 The Family

Those of us fortunate to grow up in an inspiring family know how profound the family's influence is. It may not be clear in the early years, but with age, this influence becomes more apparent. Mark Twain put it best:

> When I was a boy of 14, my father was so ignorant I could hardly stand to have the old man around. But when I got to be 21, I was astonished at how much the old man had learned in seven years.

Bart's father, Dr. Theo van der Waerden, and his younger brother Jan studied civil engineering at the Delft Technical University, where they joined first student-socialists of Holland. Upon graduation c. 1910, Theo taught mathematics and mechanics in Leeuwarden, Dordrecht, and finally for 20 years, 1902–1922, in Amsterdam. In 1911, he earned the degree of Doctor of Technical Sciences by defending his thesis entitled *Education and Technology* (*Geschooldheid en Techniek*).

A year earlier, on June 28, 1910, Dr. Theo was elected a representative of *SDAP* (*Sociaal-Democratische Arbeiderspartij*) to the Provincial government of North Holland, where he remained until 1919. Theo was editor of *The Socialist Guide* (*De Socialistische Gids*), where after 1916 he started publishing articles on economic issues. From September 17, 1918, and until his passing on June 12, 1940, he was a *SDAP*'s universally admired member of the House of Representatives (*Tweede Kamer*) of the Dutch Parliament. Theo-the-grandson informs:

> [Dr. Theo van der Waerden] was beloved and important in Dutch history. In 1939, he was to be the first Socialist Cabinet minister, but he was already ill.

Published on the day of his passing, Dr. Theo's moving eulogy[1] was entitled "A worker with a warm heart and a sober mind" ("*Een werker met een warm hart en een nuchtere geest*"):

> The working class loses in him one of the pioneers of the socialism in the Netherlands, who has not saved himself, a man, who always gave the best he can offer to the people.
>
> We remember him in gratitude and respect.

Bart's mother, Dorothea van der Waerden, was very much loved by her three sons. Let me share with you two amazing family photos, taken in 1916 and 1925. The room is the same, all family members occupy the same seats, and so we could virtually feel the passage of the 9 years.

Dr. Theo, Bart, Dorothea, Ben and Coen van der Waerden, 1916. (Courtesy of Dorith van der Waerden)

[1] *Het Volk*, June 12, 1940.

Dr. Theo, Bart, Dorothea, Ben and Coen van der Waerden, 1925. (Courtesy of Dorith van der Waerden)

When the sons left the family house in the latter 1920s, Dr. Theo and Dorothea built a magnificent house "Breidablik" at *Verlengde Engweg* 10, in Laren.

Bart's middle brother Coen van der Waerden (December 29, 1904–December 24, 1982), became a member of the Senate (*Eerste Kamer*) of the Dutch Parliament from *PvdA* (*Partij van de Arbeid*), for the total of 10 years (1957–1966 and 1970–1971) and was one of the leaders of his party. Coen was a spokesman on economic issues and a member of the union wing of *PvdA*.

I learned much about Bart's youngest brother Benno (Ben) van der Waerden and his heroic conduct during the Nazi occupation of the Netherlands from his daughter Dorith van der Waerden, a psychologist:

> My father, Benno, born 2 October 1909, died 9 of May 1987. My mother's name was Rosa Eva Louise Weijl – here comes the Jewish root – born 26 July 1909. She died 4 years ago. They met in 1939 and married 4 months later in the same year. . . He studied law [University of Amsterdam, 1927–1932] and became a lawyer. He became appointed in 1949 [to a judge of the City of Amsterdam]. . .
>
> The fact that my father married a Jewish woman was no coincidence I believe. In the thirties my father was active in helping German Jews to escape from Germany to Holland. During the occupation, he made false identity cards for Jews and helped

them to change identity. I do not know much more about it as this period was never spoken about in our family as in most families.

I am the only one who is again politically active in local politics for a green leftish party *GroenLinks*.

The Netherlands was overrun by the German invaders over the course of five short days of 1940: May 10–15. The Socialist-Democrat Dr. Theo van der Waerden would have likely been on an early list of the Dutch sent to a concentration camp. Records show that he denied the German invaders that pleasure by succumbing to cancer at 8 in the morning on June 12, 1940. After Dr. Theo's passing, his wife Dorothea lived in the Laren house together with her sister. Unable to cope with depression caused by the German occupation, Dorothea drowned herself in a nearby lake on November 14, 1942.

Bartel was understandably proud to belong to this distinguished family of public servants. In the difficult postwar times, he will invoke his father and brothers as high arbiters of his character and integrity.

From the left: Camilla, Bartel, Theodorus, Coenraad, Dorothea and Benno van der Waerden; 30th Anniversary of Theo & Do's marriage, Circa August 28, 1931, Freudenstadt, Southern Germany. (Courtesy of Coenraad's son Theo van der Waerden)

39.4 Van der Waerden at Amsterdam

In 1919, Bartel entered the University of Amsterdam very early – he was 16 (as was L. E. J. Brouwer before him when the latter started at Amsterdam). Following doctoral examination at Amsterdam, Van der Waerden is awarded a Rockefeller fellowship at Göttingen University for 7 months (1925–1926) for studying abstract algebra under Emmy Noether. Bartel impresses not only Noether but also Göttingen's leaders David Hilbert and Richard Courant, who will write letters of recommendation for the young Dutchman. Van der Waerden defends his doctorate at Amsterdam under Hendrick de Vries.

39.5 Van der Waerden at Hamburg

In 1975, Van der Waerden commences to tell the Story of Hamburg [Wae20]:

> [In 1926] I went to Hamburg as a Rockefeller fellow to study with Hecke, Artin and Schreier.

He confirms it to the interviewer on May 4, 1993 [Dol1]:

> After one semester at Göttingen, Courant started to take notice of me. He procured for me, on the recommendation of Emmy Noether, a Rockefeller grant for one year. With this I studied another semester at Göttingen and one semester at Hamburg with Artin.

In his 1930 *Moderne Algebra* [Wae3], Van der Waerden enumerates his Hamburg studies when he lists the sources of this book:

> A lecture [course] by E. Artin on Algebra (Hamburg, Summer session 1926).
> A seminar on Theory of Ideals, conducted by E. Artin, W. Blaschke, O. Schreier, and the author [i.e., Van der Waerden] (Hamburg, Winter 1926/27).

Van der Waerden is engaged with Hamburg leaders of algebra; the time there is one of the most important in Bartel's mathematical life [Wae20]:

> I met Artin and Schreier nearly every day for two or three semesters.

The Hamburg time also allows an insight into the views and personality of Van der Waerden. During the January 15, 1927, interview with Van der Waerden, the Rockefeller official Wilbur Earle Tisdale, the assistant to Augustus Trowbridge, the head of the Paris Office of the International Education Board, Tisdale notes Van der Waerden's predilection for categorical opinions:

> While he [Van der Waerden] is young, he has very clear and definite opinions – perhaps too much so. I talked to him concerning Kloosterman and, in his frank way, he told me he considered Kloosterman to be lazy, an average straight forward worker, but temperamental and requiring conditions to be just right before he can work... His feeling is that [Edmund] Landau, at Göttingen, is a man without particular vision.

In spite of being too critical of his colleagues, Van der Waerden leaves a positive overall impression on Tisdale:

> Van der Waerden appeals to me as a very intense, gifted and enthusiastic individual. He has the unfortunate defect of stammering, especially in his more intense moments, but he is so agreeable to talk to that the defect is rather minimized. I explained to him how the seriousness of such fellows as himself might be influential in justifying the appointment of future fellows, to which he reacted most enthusiastically and agreeably.

39.6 The Story of the Book

Emil Artin (1898–1962), a framer of abstract algebra, promised Richard Courant to write a book on abstract algebra for the Courant-edited *Yellow Series* of Springer-Verlag. During the summer of 1926, he gave a course on abstract algebra attended by Van der Waerden who took meticulous notes. Artin agreed to share the writing of *his* book, based on *his* lectures, with the 23-year-young Dutchman. However, as we all know, The Book appeared a few years later under one name, that of the Student and without the Master.

What happened is a question of enormous importance, for The Book became one of the most famous books in the history of mathematics. Yet, I found *no research* published on this subject. Van der Waerden told *his* Story of The Book, *his* interviewers and *his* former Ph.D. students repeated it, and most historians and mathematicians uncritically accepted thus invented fairytale. I invite you to join me in taking a close look at the documents. It is most appropriate first to give the podium to Professor Van der Waerden, who in 1975, after Artin's passing, tells us how enormous Artin's contributions to The Book really were [Wae20]:

> Artin gave a course on algebra in the summer of 1926. He had promised to write a book on algebra for the "Yellow Series" of Springer. We decided that I should take lecture notes and that we should write the book together. Courant, the editor of the series, agreed. Artin's lectures were marvelous. I worked out my notes and showed Artin one chapter after another. He was perfectly satisfied and said, "Why don't you write the whole book?"
>
> The main subjects in Artin's lectures were fields and Galois Theory. In the theory of fields Artin mainly followed Steinitz, and I just worked out my notes. Just so in Galois Theory: the presentation given in my book is Artin's.
>
> Of course, Artin had to explain, right at the beginning of his course, fundamental notions such as group, normal divisor, factor group, ring, ideal, field, and polynomial, and to prove theorems such as the *Homomorphiesatz* and the unique factorization theorems for integers and polynomials. These things were generally known. In most cases I just reproduced Artin's proofs from my notes.
>
> I met Artin and Schreier nearly every day for two or three semesters. I had the great pleasure of seeing how they discovered the theory of "real fields," and how Artin proved his famous theorem on the representation of definite functions as sums of squares. I included all this in my book (Chapter 10). My sources were, of course, the two papers of Artin and Schreier in *Abhandlungen aus dem mathematischen Seminar Hamburg5* (1926), p. 83 and 100.

Van der Waerden gives further credits to Artin (ibid.):

> In chapter 5 (*Körpertheorie*) I mainly followed Artin and Steiniz...
>
> Chapter 7 on Galois Theory was based on Artin's course of lectures...
>
> In chapter 10 ... (a) the Artin–Schreier theory of real fields and representation of positive rational functions as sums of squares ... In treating subject (a) I closely followed the papers of Artin and Schreier.

Van der Waerden repeats his story during the 1993 interview, and the interviewer-historian, Professor Dold-Samplonius publishes it [Dol1]:

> Artin was supposed to write a book and wanted to write it with me. Having finished the first chapter, I showed it to Artin. Then I sent him the second and asked him about the progress of his part of the book. He hadn't yet done anything. Then he gave up the idea of writing the book with me. Nevertheless, the book is based on lectures of Artin and Noether.

The idyllic picture is further embellished by Dold-Samplonius in her 1997 eulogy of Van der Waerden [Dol2]:

> Artin gave a course on algebra that summer, and, based on Van der Waerden's lecture notes, the two planned to coauthor a book on algebra for Springer-Verlag's"Yellow Series." As Van der Waerden worked out his notes and showed Artin one chapter after another, Artin was so satisfied that he said "Why don't you write the whole book?"

"Artin was so satisfied," Van der Waerden and Dold-Samplonius lead us to believe. In fact, Artin was so dissatisfied that he obviously *refused* to write the book together with Van der Waerden. I read – in disbelief – the revealing Richard Courant's August 6, [192]7, letter to Van der Waerden:

> Dear Herr v.d. Waerden!
>
> Herr Artin has sent me a copy of the enclosed letter about which I am somewhat astonished and concerned. Do you understand Artin's attitude? I don't. Is there any personal sensitivity behind this or are these differences of an objective nature? In any case, one cannot force Artin. But I would like to hear your opinion before I answer him.
>
> Hopefully, you haven't become angry. – I wish you a good recovery and a good vacation, and remain with friendly greetings
>
> Your [Courant]

Clearly, Artin refused to write The Book with Van der Waerden, and thus "astonished" Courant. Artin must have felt offended by Van der Waerden, but how?

Let us look at the surviving shreds of evidence. The skies are cloudless on November 29, 1926, as we glance at Courant's letter to Van der Waerden:

> Dear Herr Van der Waerden!
>
> What about this admission of your *Habilitation*. It would be very good to get this thing moving.
>
> How are you doing otherwise? How is the book by Artin and you coming along?

We see the first clouds in Van der Waerden's December 2, 1926 reply to Courant:

> The Yellow Book is making progress; I have finished writing a large part; I have half-finished other parts, and the plan for the whole is becoming more precise in details

through the conversations with Artin, the only thing is that Artin himself writes very little.

So, Artin has given his course on which the book is to be based, Artin is making his material "more precise in details through the conversations," but "Artin himself writes very little," or – as Rudyard Kipling would have put it (see "How the Camel Got His Hump" in *Just So Stories* [Kip]) – Artin does not "fetch and carry like the rest of us." Two months later, on February 2, 1927, we observe the skies becoming overcast as the Student is dissatisfied with the Master:

> My coexistence with Artin is still very fruitful. He forever digs up nice things that will also have to come into the book, and from our conversations many details emerge by which the proofs are simplified or new contexts are uncovered. Even if he does not work on the book directly, it is still coming forward.

So, Artin has not only provided a well-thought-out lecture course, ready for notetaking, but he further contributes to the joint book: "he forever digs up nice things," "many details emerge by which the proofs are simplified or new contexts are uncovered." But Artin won't "plough like the rest of us" (Kipling again), and the Student is upset that the Master "does not work on the book directly" and, just as in his letters to Courant, probably accuses the Master to his face of *not writing down his fair share* of "nice things." As Van der Waerden recalls, "He [Artin] hadn't yet done anything [sic]." That would explain Artin's explosion and refusal to write *his* book with this Student. Now we can better understand the quoted above 1993 interview [Dol1]. In fact, Van der Waerden tells us the truth, but not all the truth and without the context behind it, the context that would have allowed us to understand what happened. Let us revisit it, now that we have established the context and thus are able to understand Van der Waerden's words:

> [I] asked him [Artin] about the progress of his part of the book. He hadn't yet done anything [!]. Then he gave up the idea of writing the book with me.

But never mind the Master: the Student has gotten everything he needs, and can now publish The Book by himself, with the blessing of his mentor and the "Yellow Series" founder and editor Richard Courant.

I have coauthored works with others. It never mattered to us who would write down joint ideas and proofs. Such great mathematicians as Israel M. Gelfand, Paul Erdős, and Saharon Shelah often left the writing of joint works to their coauthors. I know that first-hand, for Erdős and Shelah have been my co-authors. I am surprised by Van der Waerden's narrow view of coauthorship. Producing a book requires not merely writing it down, but first of all discovering and assembling numerous ideas, theorems, proofs, trains of thought, giving the whole material a structure and style. In all these tasks Artin's contributions were overwhelming, and to publish The Book of Artin's ideas and proofs without Artin at least as a coauthor was unfair, in my opinion.

On the title page of The Book – what an unusual place for acknowledgments – Van der Waerden gives credit to Artin's lectures (and Noether's lectures) as being "used" in the book – but is that enough? Numerous theorems, proofs, and ideas contributed by Artin are not credited to Artin.

Van der Waerden publishes the two volumes in 1930 and 1931 in the *Yellow Series*. The great book has a great success. It excites and inspires generations of mathematicians (me included) and brings B. L. van der Waerden worldwide fame.

Unquestionably, Van der Waerden deserves credit for writing down and editing the book. How much credit depends upon how close the book is to Artin's lectures and how publishable Artin's lectures were. Those who attended Artin's summer 1926 lectures are no longer with us and thus cannot help us answer this question. But during my long 2002–2004 and 2006–2007 work at Princeton University, I found among the present Princeton professors a good number of Artin's students from his Princeton's 1946–1958 years: Gerard Washnitzer (who took all of Artin's courses 1947–1952), Harold W. Kuhn, Robert C. Gunning, Hale F. Trotter, Joseph J. Kohn, and Simon B. Kochen. Independently interviewed, they were amazingly unanimous in their assessments of Artin's lectures, unanimous even in epithets they used to describe the lectures. Tall, slender, handsome, with a cigarette in one hand and chalk in the other, without ever any notes (except, sometimes a small piece of paper extracted for a second from a jacket pocket), Artin delivered elegant, smooth, well-thought-out lectures, so much so, that notes, carefully taken, could be quite close to a finished book. Harold W. Kuhn, who took Artin's 1947 course, recalls:

> Artin's lectures were composed like a piece of music, with introduction, exposition, development, recapitulation, and coda.

Van der Waerden took notes of Artin lectures in his generation; Serge Lang did so in his. In his book [Lan1, p. vi], Lang calls Van der Waerden's book "Artin–Noether–Van der Waerden" – fair enough – but then he should have called his own book "Artin–Lang," *n'est-ce pas*?

There was another way to credit and honor the teacher. Van der Waerden gave a noble example of it when he had not "nostrified" somebody else's lecture notes. But of course, this was a special case of his admired mentor, *Fräulein* Emmy Noether [Wae20]:

> I took notes of the latter [Emmy Noether's] course, and these notes formed the basis of Emmy Noether's [!] publication in *Mathematische Zeitschrift* 30 (1929) p. 641.

The Book is prominently mentioned by Van der Waerden in his 1982 Oxford, England, talk, in which he quotes Hermann Weyl's Memorial Address for Emmy Noether:

> A large part of what is contained in the second volume of Van der Waerden's "Modern Algebra" must be considered her [i.e., Noether's] property.

Van der Waerden then responds to Weyl's remark with modesty and admiration for Noether:

> I gladly admit that this is perfectly true.

39.7 The Theorem on Arithmetic Progressions

> *Now again about the respectability of combinatorics. Even in 1926, when Van der Waerden proved the conjecture, the subject was not mainstream.*
>
> – Nicolaas G. de Bruijn

At the Bad Kissingen September 1927 annual meeting of the *Deutsche Mathematiker-Vereinigung* (*DMV* for short, the German Mathematical Society), Bartel L. van der Waerden announced a proof of the following theorem [Wae2]:

> For any k, l, there is $N = N(k,l)$ such that the set of positive integers 1, 2, ... , N, partitioned into k classes, contains an arithmetic progression of length l in one of the classes.

The Dutch Professor Wouter Peremans, Ph.D. 1949 under Van der Waerden, writes [Per, p. 135] that this "result ... made him [Van der Waerden] at one stroke famous in the mathematical world." I love this result, this is why I became interested in Van der Waerden's life in the first place. However, the original appearance of this result could not have possibly made Van der Waerden "famous." It took time for this theorem to be noticed and taste for such new Ramsey-type ideas to develop. Initially, Van der Waerden himself must not have thought highly of the value of this result and did not expect others to appreciate it, for he published it in "a second order" Dutch journal *Nieuw Archief voor Wiskunde*, whereas his algebraic geometry papers that he considered important, he published in the prestigious journal *Mathematische Annalen*. Nicolaas G. de Bruijn, who knows best, explains [Bru3, p. 116]:

> Old and respectable as the "*Wiskundig Genootschap*" may be, it has never been more than a small country's mathematical society. Accordingly, it is not surprising that the society's home journal, the "*Nieuw Archief voor Wiskunde*," has a relatively small circulation, and, as a second order effect, the *Nieuw Archief* does not get more than a small part of the more important contributions of the Dutch to mathematics.

De Bruijn elaborates on Van der Waerden's paper and the obscurity of combinatorics at the time in his January 15, 2004, e-mail to me [Bru5]:

> Now again about the respectability of combinatorics. Even in 1926, when Van der Waerden proved the conjecture, the subject was not mainstream. Van der Waerden did not send his paper to one of the leading mathematical journals, like the *Mathematische Zeitschrift*, but to the *Nieuw Archief*, home journal of the Dutch Mathematical Society, a journal that was unavailable in many libraries.

From Van der Waerden's captivating reminiscences of *How the Proof of Baudet's Conjecture Was Found* [Wae13, Wae14, Wae18, Wae26], we learn that the proof was obtained as the result of collaboration of three mathematicians, Emil Artin, Otto Schreier, and Bartel L. van der Waerden.

As you already know, Van der Waerden in fact proved the conjecture created independently by Pierre Joseph Henry Baudet and Issai Schur. As Van der Waerden informed me, he had never met either of his coauthors of what I equitably named [Soi3] the Baudet–Schur–Van der Waerden Theorem.

39.8 From Göttingen to Groningen

In the waning days of February 1927, Van der Waerden passes his *Habilitation* at Göttingen University under Richard Courant and soon becomes Courant's *Assistent* and *Privatdozent* at Göttingen. The following year Professor J. A. Barrau decides to vacate his Groningen position and move to Utrecht and recommends Van der Waerden for his place. Following several exchanges between the Curators and the Cabinet, on August 7, 1928, Queen Wilhelmina of the Netherlands assents to the appointment. And thus, *Professor Bartel Leendert van der Waerden* is born at the tender age of 25!

In the middle of his Groningen years, in 1929,Van der Waerden accepts a particularly productive visiting appointment at Göttingen: in July, he meets there his future wife. Beautiful Austrian Camilla Rellich, two years Bartel junior (born September 10, 1905), is the sister of Franz Rellich, who in the same year (1929) defends his PhD dissertation under Richard Courant. Already on September 27, 1929, Bartel and Camilla unite in a marriage that will last a lifetime. Their first child, Helga, is born in Groningen on July 26, 1930. Their other two children will be born in Germany: Ilse on October 16, 1934, and Hans Erik on December 7, 1937.

Groningen seems to have been a stepping stone for a number of fine mathematicians. Van der Corput was there too, and Van der Waerden recalls learning much of mathematics from him. At Groningen, Van der Waerden finished The Book.

39.9 Transformations of the Book

The Book was the main outcome of Van der Waerden's years at Groningen. Everyone who has written a book would agree that Van der Waerden proved to be a great expositor of the new abstract view of algebra. He writes in the preface of the 1930 first edition of volume 1 that The Book, started as Artin's lecture notes, has substantially changed, and by the time of its release, it was difficult to find Artin's lectures in it. I know of no way to verify this statement today. Granted, Van der Waerden's contribution must have grown significantly from 1927 to 1930. However, it is also clear that an unusually large contribution of the non-author Artin remained insufficiently credited in The Book, as we have seen when we cited Van der Waerden's own 1975 words. The Book became an instant classic, enjoyed by generations of mathematicians. I remember reading during my freshman university year (1966–1967) the early Russian translation (Vol. 1, 1934; Vol. 2, 1937) with delight and profit. The book was so rare that I was not allowed to take it home and had to read it in my university library.

Unlike his mentors Brouwer and Hilbert, Van der Waerden apparently did not have firm principles related to the foundations of mathematics that he was willing to fight for, as the story of changing – and changing back – his *Moderne Algebra* book shows. It is surprising that the quick learner Van der Waerden has seemingly failed to see the importance of the "Battle over the Foundations" that raged for decades and to take a firm position on it. The leading historian of the Axiom of Choice Gregory Moore writes in his wonderful book [Moo]:

> In 1930, Van der Waerden published his *Modern Algebra*, detailing the exciting new applications of the axiom [of choice]... Van der Waerden's Dutch colleagues persuaded

> him to abandon the axiom in the second edition of 1937. He did so . . . [which] brought such a strong protest from his fellow algebraists that he was moved to reinstate the axiom [of choice] and all its consequences in the third edition of 1950.

Indeed, in January 1937, in the Preface to the second edition of volume 1, Van der Waerden discloses the surprising transformation of The Book [Wae6]:

> I have tried to avoid as much as possible any questionable [sic] set-theoretical reasoning in algebra. Unfortunately, a completely finite presentation of algebra, avoiding all non-constructive existence proofs, is not possible without great sacrifices. Essential parts of algebra would have to be eliminated, or the theorems would have to be formulated with so many restrictions that the text would become unpalatable and certainly useless for a beginner. . .
>
> With the abovementioned aim in mind, I completely omitted those parts of the field theory which rest on the axiom of choice and the well-ordering theorem. Other reasons for this omission were the fact that, by the well-ordering principle, an extraneous [sic] element is introduced into algebra and, furthermore the consideration that in virtually all applications the special case of countable fields, in which the counting replaces the well-ordering, is wholly sufficient. The beauty of the basic ideas of Steinitz' classical treatise on the algebraic theory of fields is plainly exhibited in the countable case.
>
> By omitting the well-ordering principle, it was possible to retain nearly the original size of the book.

Then, in the July 1, 1950, Preface to the third edition of volume 1, I read with puzzlement Van der Waerden's justification of the reversal [Wae11]:

> In response to many requests, I once again included sections about well-ordering and transfinite induction, which were omitted in the second edition, and on this foundation, I presented the theory of fields developed by Steinitz in all its generality.

It appears as if the victory of Brouwer's intuitionism, which manifested itself in the second edition, was short lived. In the end, Hilbert's set theoretic foundation of mathematics triumphed in The Book.

On March 15, 1977, Dirk van Dalen, the biographer of L.E.J. Brouwer, interviewed Van der Waerden and has kindly shared with me that never completely published interview, and so we can "hear" Van der Waerden himself commenting on the transformations of The Book, on his commitment to good pedagogy, and his "always philosophically. . . fluctuating" views:

> **Van Dalen**: In your book on algebra, you took different positions on constructivism, where at one time the well-ordering theorem was included and another time not. You have a paper on effective factorization of polynomials. Was that under the influence of Brouwer?
>
> **Van der Waerden**: Yes, of course. That varying position in different editions was not a change of fundamental position, philosophically I have always been fluctuating, but that was for pedagogical reasons.
>
> If you look at the factorization in two factors, then I think it may be good pedagogy to show it constructively. Later I thought to do it as I used to.

39.10 On to Germany

Ever since his student years, Van der Waerden aspired to a job in Germany, perhaps the place-to-be at the time. The leading German mathematicians had a very high opinion of him: it suffices to observe that Van der Waerden was ranked 3rd on the list of the all-important David Hilbert's succession at Göttingen. Documents in the National Archive of the Netherlands show that on June 27, 1930, Leipzig University became officially interested in considering Van der Waerden for a position of *ordinarius*, the approximate German equivalent of an American full professor. The attempts to keep Van der Waerden in Holland failed, and he succeeds Otto Hölder at Leipzig. On May 1, 1931, twenty-eight years of age, Van der Waerden starts as an *ordinarius* at the *Universität Leipzig*.

Germany in 1931 was the center of the mathematical world, and Leipzig, although not a match to Göttingen and Berlin, was a very fine university, with a world-class program in physics. Once at Leipzig, Van der Waerden joins the seminar conducted by the physicists Werner Heisenberg, who will soon win the Nobel Prize "for the creation of quantum mechanics...", and Friedrich Hund. Heisenberg held fond memories of Niels Bohr's famous seminar in Copenhagen which he attended in the 1920s. He tried to reconstruct the spirit of that seminar at Leipzig. Heisenberg's seminar was the powerhouse of thinkers on matters physical. His assistants and guests included Felix Bloch (Nobel Prize 1952), the Russian genius Lev Landau (Nobel Prize 1962), the future American hydrogen bomb's leading creator Edward Teller, the future member of the Manhattan project Victor F. Weisskopf, the Heisenberg–Hund–Bohr student Carl-Friedrich Baron von Weizsäcker, the future Princeton professor Ariel Wintner, and many other outstanding minds.

Van der Waerden was an extremely quick learner. He picked up physics from them as he had earlier learned algebra from Noether and Artin. Already the following year, in 1932, Van der Waerden publishes a book on applications of group theory to quantum mechanics in the Springer *Yellow Series* [Wae4].

Van der Waerden becomes a friend of young Carl-Friedrich von Weizsäcker. On February 12, 2011, Carl-Friedrich's son, Professor Ernst Ulrich von Weizsäcker, shared with me a story he heard from his father:

> Dear Alexander,
>
> It so happened that I was in touch with Thomas Goernitz [one of the closest colleagues of Carl-Friedrich von Weizsäcker] recently who brought back to my memory that my father was extremely thankful to Bartel Leendert van der Waerden after the latter had served as the examiner in physics at the Ph.D. exams. My father was extremely young at the time, 21 years old only, and felt he was very inexperienced in experimental physics. But Van der Waerden was fascinated, so it seems, with what my father knew and explained in theoretical physics, so he let him speak and speak and the time was over before they could turn to experimental physics. And the whole thing ended in a top rating for my father. That was in 1933, one of the darkest years for Germany and the world, as you know.

For Heisenberg, his former student and colleague Carl Friedrich von Weizsäcker became the closest confidant. In October 1934, he writes to his mother:

> Only the friendship with Carl Friedrich, who struggles in his own serious way with the world around us, leaves open to me a small entry into that otherwise foreign territory.

Hitler's ascent to power at the dawn of 1933 found Bartel van der Waerden contemplating his second Rockefeller (IEB) fellowship.

Chapter 40
In Search of Van der Waerden: The Nazi Leipzig, 1933–1945

40.1 The Dawn of the Nazi Era

From 1933 till 1940 I considered it my most important duty to help defend the European culture, and most especially science, against the culture-destroying National Socialism.

– Bartel L. van der Waerden

The compromises you will have to make will later be held against you, and quite rightly so... But in the ghastly situation in which Germany now finds herself, no one can act decently.

– Max Planck to Werner Heisenberg

The Russian thinker and exiled revolutionary Leon Trotsky insightfully describes the situation in Germany and points out the complacency of academics in his early, June 10, 1933, article [Tro]:

> The immense poverty of National Socialist philosophy did not, of course, hinder the academic sciences from entering Hitler's wake with all sails unfurled, once his victory was sufficiently plain. For the majority of the professorial rabble, the years of the Weimar regime were periods of riot and alarm. Historians, economists, jurists, and philosophers were lost in guesswork as to which of the contending criteria of truth was right, that is, which of the camps would turn out in the end the master of the situation. The fascist dictatorship eliminates the doubts of the Fausts and the vacillations of the

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_40

> Hamlets of the university rostrums. Coming out of the twilight of parliamentary relativity, knowledge once again enters into the kingdom of absolutes. Einstein has been obligated to pitch his tent outside of the boundaries of Germany.
>
> On the plane of politics, racism is a vapid and bombastic variety of chauvinism in alliance with phrenology. As the ruined nobility sought solace in the gentility of its blood, so the pauperized petty bourgeoisie befuddled itself with fairy tales concerning the special superiorities of its race.

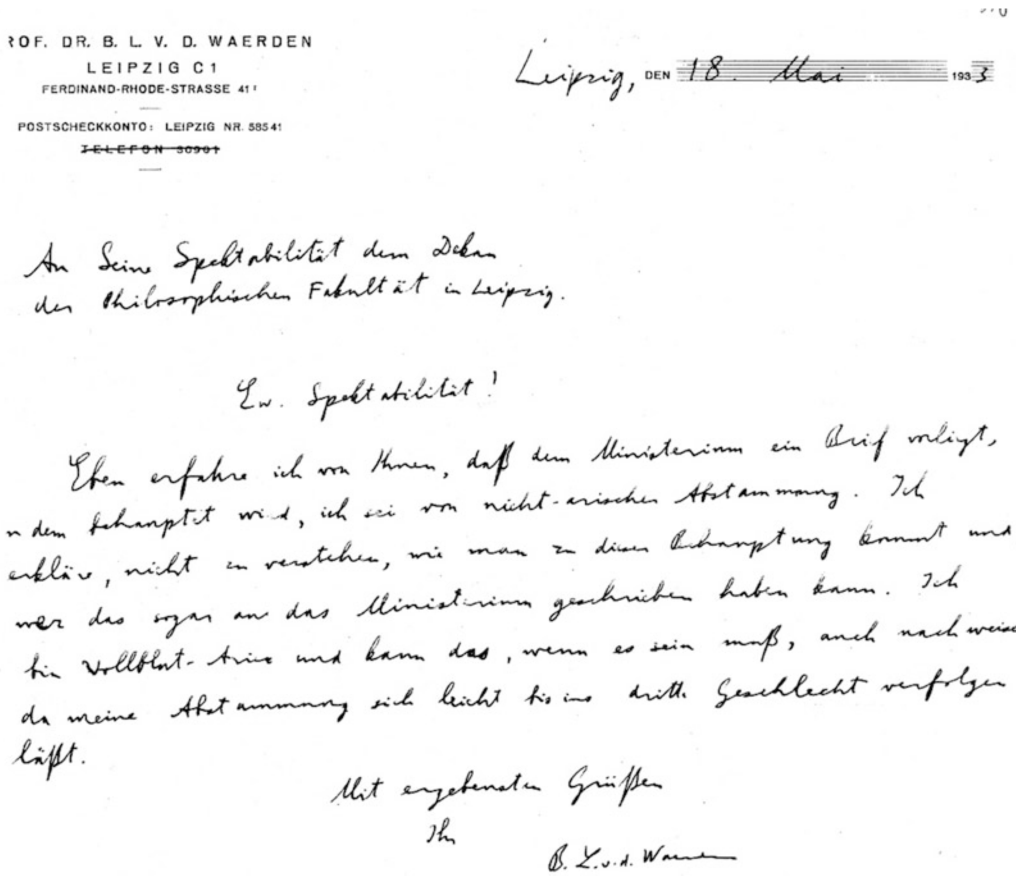
The April 7, 1933, "Law for the Restoration of the Professional Civil Service" (*Gesetz zur Wiederherstellung des Berufsbeamtentums*) was signed and put into an immediate effect by Reich Chancellor Adolf Hitler, Reich Minister of the Interior Wilhelm Frick, and Reich Minister of Finance Johann Ludwig (Lutz) Graf Schwerin von Krosigk. The law rid German universities of all Jewish (by Nazi definition) professors, except civil servants in office prior to August 1, 1914, those who fought at the Front for the German Reich or its Allies in the World War, and those whose fathers or sons fell in the World War.

Dekan Weickmann of Leipzig's Philosophical Faculty did not wish to fall behind the swiftly rolling Nazi avalanche and immediately expressed his limitless support for "the efforts of the government directed at the limitation of Jewish influence at German universities" and inquired from Dresden what they should do with the foreigner Van der Waerden and the Jew Felix Bloch.

By some accounts, Leipzig University alone lost 35 academics to dismissal, resignation, forced retirement, and death. Heisenberg's Ph.D. 1928, brilliant assistant and a companion in hiking and skiing outings, Felix Bloch was among those dismissed for being Jewish. Bloch asked for and received help from Heisenberg's mentor, coauthor, and friend, Physics Nobel Prize Laureate 1922 Niels Bohr. In June 30, 1933, letter to Bohr, Heisenberg is grateful "for . . . your efforts on behalf of our young physicists, whose well-being lies in all our hearts" and apologizes for the new Third Reich, "for all of that which is now happening in this country." A year later Bloch will accept a job at Stanford University and, in 1952, win the Nobel Prize.

The 1933 firings include Van der Waerden's teachers and mentors at Göttingen, Emmy Noether and Richard Courant. These perturbations briefly affect Van der Waerden, who is alleged to be a foreigner (correctly) and a Jew (incorrectly). Friedrich, the leader of the mathematics students' organization (*Führer der mathematischen Fachschaft*), argues that as a foreigner Van der Waerden is not fit to be the Director of the Mathematics Institute. In his defense against Friedrich's accusations, Van der Waerden writes the following letter to Dekan Ludwig Weickmann on May 18, 1933 [see facsimile]:

> Your Magnificence!
>
> I have just learned from you that the Ministry possesses a letter in which it is claimed that I am of a non-Aryan descent. I declare that I do not know how that conclusion was reached and who could have written this to the Ministry. I am a full-blooded Aryan and I can prove that if necessary, because my ancestry can be tracked for three generations.
>
> With loyal regards,
>
> Yours
>
> B. L. v. d. Waerden


ROF. DR. B. L. V. D. WAERDEN
LEIPZIG C1
FERDINAND-RHODE-STRASSE 41
POSTSCHECKKONTO: LEIPZIG NR. 58541
~~TELEFON 30901~~

Leipzig, den 18. Mai 1933

An Seine Spektabilität den Dekan
der Philosophischen Fakultät in Leipzig.

Ew. Spektabilität!

Eben erfahre ich von Ihnen, daß dem Ministerium ein Brief vorliegt, in dem behauptet wird, ich sei von nicht-arischer Abstammung. Ich erkläre, nicht zu verstehen, wie man zu dieser Behauptung kommt und wer das sogar an das Ministerium geschrieben haben kann. Ich bin Vollblut-Arier und kann das, wenn es sein muß, auch nachweisen, da meine Abstammung sich leicht bis ins dritte Geschlecht verfolgen läßt.

Mit ergebenen Grüßen
Ihr
B. L. v. d. Waerden

B. L. van der Waerden claims his "full-blooded" Aryanness. (Courtesy of Leipzig University)

The number of Aryan generations in Van der Waerden's ancestry quickly grows, for the next day, on May 19, 1933, Leipzig's Rektor Achelis informs Minister Hartnacke of Saxony that the accusation of Van der Waerden being non-Aryan is incorrect, that Vander Waerden has a proof that *five* [sic] generations of his ancestors have been Christians, and thus Van der Waerden should be able to retain his directorship.

Meanwhile, even those Jews, who were exempted from firing under the April 7, 1933, law, found themselves under an immense pressure to resign. Nazi students boycotted and disrupted classes of Jewish professors, one of whom was the Göttingen number theorist Edmund Landau. Van der Waerden mentions his actions against Landau's boycott in "The Defense," a document he will write for the de-Nazification Boards of Utrecht and Amsterdam Universities after the war: "In 1933 I traveled to Berlin and Göttingen to protest the boycott of [Edmund] Landau's classes by Göttingen Nazi students." In June 1933, the great physicists Max Planck and Werner Heisenberg, the latter by now Van der Waerden's close friend, circulate a petition in support of Van der Waerden's Göttingen mentor Richard Courant, who fights his unlawful dismissal as a veteran of World War I.

Not everyone immediately understood how dangerous the Nazi regime promised to be. The United States' official early posture was to order a cup of coffee and view the confrontation between Nazism and Socialism. Some Americans, e.g., members of the *Emergency Committee in Aid of Displaced Foreign Scholars*, the *U.S. Emergency Rescue Committee*, and the *Unitarian Service Committee*, were rescuing children and great minds of Europe, such as Albert Einstein, Emmy Noether, Marc Chagall, Max Ernst, Erich Maria Remarque, Lion Feuchtwanger, Thomas Mann, Heinrich Mann, and Berthold Brecht. Others, such as FBI Director J. Edgar Hoover and his agents, were spying on the rescued refugees and even trying to get some of them, the famous writer Lion Feuchtwanger included, deported out of the United States (read more in Alexander Stephan's excellent monograph *Communazis* [Ste]). To my disbelief, I learned that even the founder and first director of the Institute for Advanced Study, Princeton, Abraham Flexner, with full support of one of the leading Princeton-Math professors and future chair Solomon Lefschetz, ridiculed Einstein for being an outspoken anti-Nazi, as you can see from Flexner's September 28, 1933, letter to Felix M. Warburg of New York City:[1]

> Dear Mr. Warburg:
>
> In reply to Miss Emanuel's note containing the cables from you and Lockar Lampson I am writing to you as follows:
>
> "Suggest you cable Lockar Lampson as follows signing your name opinion here in academic and official circles strongly to effect that Professor Einstein should not participate in Albert Hall meeting regardless of subject of his discussion Please give him my former telegram as well as this Unquote Am writing you."
>
> I may add that last night Professor Lefschetz, who holds the highest professorship in mathematics in Princeton University and is himself a Russian Jew, came to see me and asked me if I could not in some way shut Einstein up, that he was doing the Jewish cause in Germany nothing but harm and that he is also seriously damaging his own reputation as a scientist and doing the Jewish situation in America no good.
>
> I may add for your private information that I am seriously concerned as to whether it is going to be possible to keep him and his wife in this country. I have been pleading with them all summer to show the elements of common sense, and their replies have been vain and foolish beyond belief. You have doubtless noticed in the morning paper that the German government has retracted in part its attitude toward Jewish merchants. Einstein is simply making it as hard as possible for the German government to climb down. Scores of individuals in New York and in Princeton have spoken to me about him, his wife, and their conduct, and without a single exception in thorough condemnation, despite the fact they are all bitterly opposed to the present German regime. Though he is of course not a Communist, he is now only partially a Pacifist. The clipping, which Miss Emanuel sends, is correct in maintaining that his presence on the platform will do no good to anybody. The case is very different with a man like Austen Chamberlain, who has been Foreign Secretary and is a Christian gentleman, and in his hands it ought to be left.
>
> To cap the climax, Einstein has made practically no sacrifice whatsoever. He and his wife are better taken care of today than they have ever been in their life if they will only

[1] I am grateful to the Shelby White and Leon Levy Archives Center, Historical Studies-Social Science Library Archive of the Institute for Advanced Study Princeton for this letter and permission to reproduce it. A good part of this letter was first quoted in [Sie3].

> behave themselves. Other German Jewish scholars like Frank and Haber, both Nobel Prize medalists, have actually given up their posts either voluntarily or through suppression and allowed the world to judge, with the result that they are more highly esteemed than ever and their dignity has hurt the German Government a good deal more than Einstein's everlasting publicity.
> With all good wishes,
> Sincerely yours,
> Abraham Flexner

Bringing Albert Einstein to the Institute for Advanced Study was the greatest luck of Abraham Flexner's life. How many people would know today of the Institute for Advanced Study if Einstein did not work there? And yet, like the worst kind of appeasers of Nazi Germany, Flexner and Lefschetz are looking for "some way to shut Einstein up," to prevent him from speaking against Nazism! Perhaps, Flexner and Lefschetz merit a little break: after all they are American Jews, removed by the Atlantic Ocean from the horrors of Nazi Germany, and not wise enough to comprehend Nazism at its early stage. Surely Richard Courant, himself a Jewish refugee from Nazi Germany, appreciates Einstein using his acclaim and reputation to warn the world about the dangers of Nazism? Sadly, Shakespeare comes to mind: "*Et tu, Brute? You too, Brutus?*" Constance Reid reports:

> Einstein, who had been in America for the past few months, had been making a number of widely publicized statements deploring "brutal acts of violence and oppression against persons of liberal opinion and Jews . . . in Germany [which] have aroused the conscience of all countries remaining faithful to ideals of humanity and political liberties." On March 29 [1933] . . . the government in Berlin had announced that Einstein had inquired about taking steps to renounce his Prussian citizenship.
>
> "Even though Einstein does not consider himself a German," Courant wrote, "he has received so many benefits from Germany that it is no more than his duty to help dispel the disturbance he has caused . . . What hurts me particularly is that the renewed wave of anti-Semitism is . . . directed indiscriminately against every person of Jewish ancestry, no matter how truly German he may feel within himself, no matter how he and his family have bled during the war and how much he himself has contributed to the general community. I can't believe that such injustice can prevail much longer – in particular, since it depends so much on the leaders, especially Hitler, whose last speech made quite a positive impression on me."

So much for the acclaimed cleverness of Courant: Einstein is ungrateful, and Hitler leaves "quite a positive impression" on Courant! Even the well-known anti-Nazi, Physics Nobel Laureate and Einstein's friend Max von Laue urges Einstein to abstain from politics:

> Here they are making nearly the entirety of German academics responsible when you do something political.

It sounds as if the German academics do not wish to do or say much and be responsible for anything. Einstein sums up his position in his reply to von Laue. His words call on scholars to leave the ivory tower and assume responsibility for world affairs, to be counted in the struggle for truth and justice:

> I do not share your view that the scientist should observe silence in political matters, i.e., human affairs in the broader sense ... Does not such restraint signify a lack of responsibility? Where would we be had men like Giordano Bruno, Spinoza, Voltaire, and Humboldt thought and behaved in such a fashion? I do not regret one word of what I have said and am of the belief that my actions have served mankind.

What about the great old man of physics Max Planck? Surely, he could understand Einstein's sacrifice in service of the world? Not so. Dominic Bonfiglio observes:

> Later that May[1933], as Nazis and their sympathizers were preparing to burn 20,000 books in Berlin, Max Planck was recorded in the Academy's minutes as saying that "through his political behavior himself [Einstein] rendered his continued membership in the Academy impossible." There were few things that surprised Einstein more about Hitler's rise to power than the way the majority of German academics responded to it. In August, Einstein told a colleague that he probably wouldn't see his country of birth again.

Einstein never did. The prominent Einstein's critics should recall words of the XVIII century French playwright Molière and own the responsibility for their inaction:

> It is not only what we do, but also what we do not do, for which we are accountable.

The 1930s American government's official policy of appeasement toward Nazi Germany is regrettable, to put it mildly. However, it reflects the position of the majority of the American population. The leading wire service Associated Press (AP) allows us to clearly see this. On March 7, 1934, AP reports from New York City:

> Twenty-two speakers presented the "Case of Civilization against Hitler" at a mass meeting in Madison Square Garden, New York, March 7. Edward J. Neary, Executive Committee member of the American Legion, is shown [this text was accompanied by a photo of E.J. Neary, for which I do not have copyrights] as he presented the case of war veterans against Hitler. The audience was composed of liberals, Jews, and other anti-Nazis. [I own this AP photo, but do not have copyrights to share it with you.]

I am shocked to read the last sentence of this AP report. Hitler has been in power for over 14 months, yet AP and the American people do not get it. The report insinuates that only fringe elements of the American society are against Nazi Germany's crimes: "liberals, Jews, and other anti-Nazis"! However, closer to the start of the World War II, the American public opinion will slowly shift against Nazi Germany. Mass demonstrations will follow. One such very large "Stop Hitler Parade" will take place in Manhattan on March 25, 1939 [I own this AP photo, but do not have copyrights to share with you this human river flowing along a major street].

40.2 The Princeton Job Offer

For 20 months, 2003–2004, I worked at Princeton University as a "Visiting Fellow." It does not mean "Visiting Dude" or "Guy" – in translation from the British, this title means a "Visiting Researcher." And so, I researched math and history, sometimes alone, other times jointly with John H. Conway or with the Israeli genius Saharon Shelah at Rutgers University. I was constantly thinking about Van der Waerden and his fate and discussed my findings with the grateful and valuable audience of Princeton Math colleagues during the daily coffee hours. From the grapevine I heard that once upon a time Van der Waerden was offered a job here, but no evidence has ever been published. In the spring of 2003, I asked the departmental administrator Scott Kinney for any relevant documents. He checked in the secretive file room and told me there was no record, "maybe because he never actually came to work here." My 2003 inquiries into the Princeton University Archive and the Institute for Advanced Study Archive penned nothing. There are countless dead ends in the maze of historical research; was I at one?

A year later, when I was about to leave Princeton and go back to my gorgeous Colorado mountains, I decided to try and see whether there existed any trace of a Princeton faculty discussion about inviting Van der Waerden. In my June 3, 2004, e-mail I queried the Mathematics Department Chair, Nick Katz:

> As you probably know, I am writing a book on Ramsey Theory together with the history of its early creators. You will provide my historical research a very essential help if you allow me to read minutes/notes of Princeton math department faculty meetings for 1933–1934 (or better yet 1933–1945). Best wishes! Alexander

Imagine my delight when the following day I received Nick's reply:

> I have left both the minutes you requested, and also the minutes of what seems to have been a university-wide "research committee," with Scott, for you to look at. We only ask that these materials, which are irreplaceable, stay in the building. Good luck with your book. Best, Nick

Irreplaceable? You bet! This was a treasure trove, unclaimed and unread by anyone in three-quarters of a century! Chairs at Princeton Math usually rotate every three years, and these two old priceless folders, holding the concise documentary history of Princeton, quietly sat in a drawer of the chair's desk. One of the two old-fashioned folders was entitled "Department of Mathematics, Minutes of Department Meetings, September 29, 1931–March 29, 1949." The other untitled folder contained Minutes of the Research Committee, later called Scientific Research Committee of Princeton University, together with various financial documents, dating from January 23, 1926, to 1949.

The minutes of the Mathematics Department did not mention Van der Waerden. However, the minutes of the Research Committee recorded an official job offer Princeton University made to him!

Princeton University was offering Bartel to get out of the ugly young Nazi state and come to Princeton as a visiting professor. I read the yellowed pages with the greatest interest:

> A meeting of the Research Committee was held on Tuesday, May 9, 1933, in Dean [of the Faculty Luther Pfahler] Eisenhart's office, Fine Hall, at 12:00 noon. Present: Dean Eisenhart, Professors [Edwin Grant] Conklin [Biology], [Rudolph] Ladenburg [Physics], [Solomon] Lefschetz [Mathematics], [Henry Norris] Russell [Astronomy] and [Sir Hugh] Taylor [Chemistry]."

Section 2 of these minutes is of our prime interest:

> Dean Eisenhart reported the desire of the Department of Mathematics to secure Professor van der Waerden of Leipzig on the Mathematics funds for the first term of 1933–34 at a salary of $3500. Dean Eisenhart reported that in this case also there would be a delay on account of uncertain conditions in Germany.

By this "also" Dean Eisenhart referred to section 1 of these minutes, which is important to us as well and reads as follows:

> Dean Eisenhart reported the inability of Professor Heisenberg to give a definite answer to the offer of an eight weeks engagement at a salary of $3000 at the present time owing to the conditions in Germany. Professor Heisenberg suggested that he might be able to give a definite reply at the end of the year. Dean Eisenhart has written to Heisenberg on the assumption that his letter meant the end of the academic year and suggested that decision by July would be acceptable.

Section 7 of the minutes is relevant too:

> Professor Lefschetz raised the question of alternatives to Professor Heisenberg in case it was found impossible to secure his services. After discussion, it was decided that the matter be left in abeyance until further reports were available concerning the German situation.

The Princeton Research Committee choices are good predictors of Nobel Prize winners. At another 1933 meeting, the Committee identifies Erwin Schrödinger, who would win a Nobel Prize in Physics for 1933, as a backup for Heisenberg, the soon to be the Nobel Prize winner for 1932.

Let us return to the Third Reich, year 1933. Heisenberg wants to wait and see how the German situation develops. He does not like to lose best Jewish German physicists, including his assistant Felix Bloch, because it will be bad for physics in Germany. But he is excited about the Nazi promise of the German national revival. On October 6, 1933, unbelievably, Heisenberg writes to his mother about "much good" in the Nazi intentions:

> Much that is good is now also being tried, and one should recognize good intentions.

As the XI century abbot Saint Bernard of Clairvaux observed, "The road to hell is paved with good intentions." And Nazi Germany has certainly been en route to that destination.

Following the wishes of the Research Committee and Mathematics Department, Princeton University offers Professor Van der Waerden a Visiting Professorship for the September 15, 1933–February 15, 1934, semester. On June 27, 1933, Van der Waerden asks Dekan Weickmann for the approval of his Princeton visit:

> To His Magnificence Dekan of the Philosophical *Facultät* at Leipzig.
>
> I would like to inform Your Magnificence that I received a prestigious invitation to give invited lectures at the University of Princeton (America) in the winter term 1933/34. As it becomes clear from the attached letters, Princeton offers optimal conditions for scientific research and inspiration by interaction with other mathematicians. For that reason I intend to accept the invitation if a leave of absence is approved for September 15 to February 15, and an appropriate replacement can be found.

> I therefore ask the *Facultät* to forward my application for the leave of absence to the Government. The directors of the Mathematics Institute will contact you with suggestions regarding my replacement.
> Respectfully submitted,
> B.L. v.d. Waerden

The wheels of the young Nazi bureaucracy move surprisingly swiftly in this case. The following day, on June 28, 1933, a letter supporting the leave, is sent to the *Facultät* by the three codirectors of the Mathematics Institute: Professors Van der Waerden, Paul Koebe, and Leon Lichtenstein. On June 30, 1933, Dekan Weickmann throws his support in a letter to the Saxon Ministry of People's Education in Dresden. On July 15, 1933, Van der Waerden sends a letter to Dekan Weickmann inquiring whether the Dekan has any news from the Ministry, and on the very same day Dekanin turn sends his inquiry to Councilor Seydewitz of the Ministry. On July 18, 1933, Seydewitz sends two letters: one to Dekan Weickmann, approving the leave without pay (as is requested by Van der Waerden, who is to be paid well by Princeton); and to *Privatdozent* Dr. Friedrich Karl Schmidt of Erlangen University, inquiring whether the latter would accept a replacement position at Leipzig. On July 24, 1933, Schmidt accepts the replacement job. Thus, everything – all approvals and the replacement – is ready for the cross-Atlantic voyage of the Van der Waerden family, when five days later, on July 29, 1933, Van der Waerden's letter shocks everyone (even me as I read these documents):

> To the Saxon Ministry of People's Education
> Attention Councilor Seydewitz
>
> Since in my opinion (and also in the opinion of the directors of Mathematics Institute) my presence at the Mathematics Institute this coming winter is urgently necessary, I respectfully ask the Ministry to revoke the leave that has already been approved. I will inform my replacement Dr. F. K. Schmidt as well as the Philosophical *Facultät* about my decision.
> Yours respectfully,
> B.L.v.d. Waerden

Thus, Van der Waerden has jumped through all bureaucratic Nazi hoops but in the end rejects the Princeton job! The Bard would have summarized the Princeton story as *Much Ado about Nothing*.

This was the first major junction in the life of Van der Waerden: had he come to Princeton, as a fine and young mathematician Van der Waerden would have most likely received further, more permanent offers from Princeton University or from the recently founded Institute for Advanced Study. His life – and the history of algebraic geometry – would have been different. But Van der Waerden chooses to remain in Nazi Germany, as does his friend Werner Heisenberg, who during several prewar years has not accepted job offers from Princeton, Harvard, Yale, Columbia, University of Michigan, and other fine American universities. Heisenberg's devotion to doing physics in Germany and his nationalism as reasons for staying in the Third Reich have been well established. Van der Waerden's Princeton opportunity has not been known, let alone explained, until my unearthing of the old dusty Princeton folder.

Van der Waerden's most surprising rejection of the Princeton offer begs a natural question: why did he do it? He explains it in the August 12, 1933, letter to Oswald Veblen, the first permanent mathematics professor at the Institute for Advanced Study Princeton:

> Like you, I am very sorry that we will not meet in Princeton next winter, but it was really impossible for me to leave Leipzig at this time.

As we have learned from the Leipzig University archive, all permissions have been granted. It therefore appears that Van der Waerden chooses courtesy over the truth in his letter to Veblen. But what is the truth?

Van der Waerden asks his mentor Richard Courant to help him receive a second Rockefeller (IEB) Fellowship, this time for work on algebraic geometry in Italy, primarily under Federigo Enriques and Francesco Severi in Rome. On March 2, 1933, Courant, still at Göttingen, pays Van der Waerden the highest praise and "informally and personally" asks Dr. W. E. Tisdale, the Rockefeller Official in Paris, whether support for Van der Waerden is possible.

Tisdale receives the letter on March 6, 1933, and *the same day* replies to Courant, asking to have Van der Waerden provide more details, which Van der Waerden does in his March 12, 1933, two-page letter (received in Paris on March 31, 1933). This letter, written in English, provides an insight into Van der Waerden's view of the state of algebraic geometry:

> Algebraic geometry, originated in Germany in the work of Clebsch, [Emmy Noether's father Max] Noether and others, has been continued during the last 30 years nearly exclusively by Italian mathematicians: Enriques, Castelnuovo, Severi, and others. They have developed methods and theorems, which are of extremely high interest both for algebra and geometry, but which are still awaiting an exact algebraic foundation: The contact between Italian geometry and German algebra is missing. I think this is a typical case in which your Foundation can help. I know the algebraic methods which can serve as a base for algebraic geometry very well, perhaps best of all German mathematicians.

Thus, Van der Waerden considers himself to be *the best* German mathematician for the job of putting algebraic geometry on the foundation of abstract algebra, and he may be correct. Moreover, for the first time in written records that I have unearthed, Van der Waerden casts himself here as a *German mathematician.* A successful Rockefeller (IEB) fellow the first time around, Van der Waerden expects an easy approval of his second fellowship. So, has Van der Waerden simply preferred Rome over Princeton? Indeed, I found a proof of it in his own words – even before he jumped through the Leipzig bureaucratic hoops – in an (undated, but definitely written in May or else June of 1933) letter to Richard Courant:

> I still thank you many times for your efforts at Rockefeller. I only got a reply from Tisdale that now there are sufficient documents to discuss the case with his colleagues in Paris …
>
> I have an offer from Princeton University, with a stipend, to spend the coming winter semester (Sept.–Jan.) there. This offer came already at the beginning of April [1933].

> But it does not tempt me as much as the Rome trip; I also do not know whether the regime will allow this much of a leave of absence.

As we know, at some point – more precisely, on July 24, 1933 – Van der Waerden has learned that “the regime will allow this much of a leave of absence.” He may have hoped even as late as in late July 1933 to get the Rockefeller money for Rome. Is this why Van der Waerden cancels the visit to Princeton? Perhaps, but there could have been another important reason for not going to Princeton *or* to Rome: Van der Waerden does not really wish to leave Germany for the first winter of the Third Reich (ibid.):

> I cannot judge yet whether it is not smarter [sic] to spend this winter in Leipzig.

What is so smart about staying in Nazi Germany during the winter of 1933–1934? We will never know for sure, but a plausible question is in order: Did Van der Waerden not wish to raise suspicion of the young and already cruel Nazi regime? Now that Van der Waerden is not going to go to Princeton anyway, it is easy for him to be conscientious (ibid.):

> I believe I will suggest to the Americans that this time they could spend their money better than to get me out because I still have a position that I can keep.

It appears likely that the Rockefeller people, once they learned of the Princeton offer to Van der Waerden, have chosen to use their funds to support those mathematicians who depended solely upon Rockefeller money and thus decided not to fund Van der Waerden’s second fellowship. In fact, already on March 29, 1933, the Rockefeller official Dr. W. E. Tisdale shows a complete knowledge of Van der Waerden’s situation in his diary:

> Van der Waerden, past fellow now at Leipzig is excellent. As a matter of fact, Princeton wants to get him in the faculty to replace shifts due to Flexner’s activity [i.e., the creation of the Institute for Advanced Study]. They will probably ask him to come for a semester in which they could have a mutual exchange of views.

Yes, the Princeton position would have likely become permanent for Van der Waerden. It seems clear that Princeton mathematicians were unhappy with Van der Waerden’s “smart” choice to stay in Nazi Germany when they offered him a great opportunity to get out. As we will see later, they will remember this rejection after the war, when Van der Waerden will become eager to come to Princeton from war-devastated Holland.

40.3 Eulogy for the Beloved Teacher

Fired from Göttingen University for being Jewish, Emmy Noether got a job at Bryn Mawr College near Philadelphia in the United States. This was not a good match for research-oriented Noether, but it was a job in a safe place at the difficult time of emerging Nazism. On April 14, 1935, she passed away. World-renown scholars wrote touching eulogies: Albert Einstein and Hermann Weyl in the USA; Pavel Aleksandrov in the Soviet Union, where Noether was planning to visit later that same year. Nazi Germany was another matter. A eulogy for a Jew and a liberal would not be appreciated by the Nazi authorities. Nevertheless, this is exactly what Van der Waerden did. He published in the *Mathematische Annalen* a heartfelt *Obituary of Emmy Noether* ([Wae5], translated into English in [Dick]). Let us pause in our narrative and pay homage to Emmy Noether and her favorite pupil Bartel L. van der Waerden's bravery:

> Our science has suffered a tragic loss. On April 14, 1935, Emmy Noether, our devoted collaborator at the *Mathematische Annalen* for many years, a highly unique person, and a scientist of great importance, died following a surgical operation…
>
> The maxim by which Emmy Noether was guided throughout her work might be formulated as follows: "Any relationships between numbers, functions, and operations only become transparent, generally applicable, and fully productive after they have been isolated from their particular objects and been formulated as universally valid concepts" …
>
> During her last eight years in Göttingen, prominent mathematicians from all over Germany as well as abroad came to consult with her and attend her lectures. In 1932, together with E. Artin, she received the Ackermann-Teubner memorial award for arithmetic and algebra. And today, carried by the strength of her thought, modern algebra appears to be well on its way to victory in every part of the civilized world.

40.4 One Faculty Meeting at Leipzig

> *In Germany itself this situation was aggravated by the isolation of the individual. Communication became increasingly difficult – only the most intimate friends dared to speak their minds to one another.*
>
> – Werner Heisenberg

Leipzig Faculty, including some major players of the May 8, 1935, Faculty Meeting. From the left, first row: Friedrich Klinger, Werner Heisenberg; second row: Bernhard Schweitzer, Joachim Wach; third row: Hermann Heimpel, Theodor Hetzer, Konstantin Reichardt, and *Dekan* Helmut Berve. April 1935. (Courtesy of Leipzig University)

May 1935 commenced with the Governor (*Reichsstatthalter*) of Saxony Martin Mutschmann dismissing five remaining Jewish professors from Leipzig University – Dr. of Medicine Bettmann, and four Philosophical *Facultät* professors: Joachim Wach (theology), Benno Landsberger (Semitic and Eastern philology), Friedrich Wilhelm Daniel Levi (mathematics), and Fritz Weigert (photo chemistry), all veterans of World War I, and as such exempted from the dismissal under the April 7, 1933 Law. On Friday, May 2, 1935, Leipzig's new *Rektor*, the psychologist Felix Emil Krueger (1874–1948), appointed just in April 1935, discussed these firings with the *Staatssekretär* Theodor Vahlen (coincidentally a mathematician), who was in charge of the Third Reich's university appointments in the *Reichserziehungsministerium* and reported directly to the *Reichsminister* Bernhard Rust.

Rektor Krueger announced these firings on Wednesday, May 8, 1935, in the afternoon at the faculty meeting of the Philosophical *Facultät*. He wanted to merely test the faculty's sentiments and not have a full-blown discussion. However, five professors bravely questioned the legality and morality of the firings and forcefully spoke in support of their fired Jewish colleagues. They were Bartel L. van der Waerden, who led the fight; physicists Werner Heisenberg and Friedrich Hund, whom you have already met in this book, classical archeologist Bernhard Schweitzer (1892–1966), who later earned the honor of being the first post-World War II Rektor of Leipzig University (May 1945–December 1945), and Russian-born German and Nordic philologist Konstantin Reinhardt (1904, St. Petersburg, Russia–1976, New Haven, USA), who in three years would leave Germany for the United States where in 1947 will become a professor of German philology at Yale University.

The discussion during this faculty meeting was passionate. News about it outraged the Ministry. Short-tempered (as is often the case with bureaucrats in tyranny) Nazi officials demanded an "immediate report." The Saxon Ministry of People's Education issued an urgent demand ("tomorrow by 1 P.M.") for the "precise" stenography of the meeting. The recording secretary Junker reconstructed the meeting's stenography on May 21, 1935, based on the detailed notes he had taken during the meeting.

Let me translate for you the *entire* reconstructed stenography, which is so cinematographic that we can "hear" voices of the participants and "see" their actions. In the first edition of *The Mathematical Coloring Book*, I presented a good part of it, leaving Van der Waerden's son Hans van der Waerden not satisfied. I agree with him, and now present here the complete text.

<u>Transcript.</u>
Ministry of People's Education — Dresden-N 6, May 17, 1935
To the *Rektor* of the University — <u>Leipzig.</u>

It has been alleged that the following happened at the faculty meeting of the Philosophical *Fakultät* on Wednesday afternoon: It has been asserted that Professor v.d. Waerden openly protested against the actions of the Governor (*Reichsstatthalter*). He pointed out that Wach had been a combatant in the war and the law explicitly stated that veterans of non-Aryan descent were exempt from the dismissal. So this would be abuse of the law and he himself [Van der Waerden] would feel ashamed if a man who gave his blood for him were now treated in such a way. He asked the *Fakultät* to make a unanimous resolution opposing the [dismissal] decision.

It is asserted that nobody objected, but Professor Golf[2] forbade v.d. Waerden to speak in the tone he was using and emphasized that insults of this kind were not usual at German universities. I stated that Professor Hund had not exactly approved of the actions of the Governor (*Reichsstatthalter*).

The Ministry asks for a detailed report.

/Signed for/ Geyer

- - - -

Leipzig, [May] 20, [19]35
The Rektor asks Herr Dekan Berve[3] for an immediate report.
In Leipzig
Signed Krueger
Rektor.
Transcript
5.21.1935

Dear Herr Dekan!

Herr Rosenberg has just informed me that you wish to see the exact transcript of the meeting of 5. 8.1935 by tomorrow at 1 P.M.

I assume that you do not care about the whole transcript but rather only the account of the Discussion of the Dismissal of the four colleagues.

From here on I write for you what I took down as a stenographer. I noted word for word the phrases that the particular gentlemen used. In the transcript of the *Facultät* [meeting] I used these phrases only as the basis of my formulations. The statement by Herr v.d. Waerden drew the warning from Herr Golf, a statement which I wanted to hand you at the time when Golf burst out (I enclose the note), and which I have omitted from the official transcript, as something regarded as irrelevant and "resolved" by Herr Golf and because it does not accord with conventions of the *Facultät* to record distractions.

- - - - - - -

The Dekan said that the Governor (*Reichsstatthalter*) [of Saxony], upon the request of the [Saxon] Ministry [of Science and Culture], dismissed 4 people. They are Mr.'s Wach, Landsberger, Levi, and Weigert. (Regarding this, it is noticed in the stenographic original: "§6. Teaching arrangements withdrawn—put in retirement.") Professor von Weigert is kw, and his position cannot be refilled.[4]

(Afterwards there were other issues and finally: The issue of the withdrawal of their titles of doctors).

---

[2] Professor of Agriculture Arthur Golf (1877–1941), Rektor of Leipzig University (October 1933–March 1935, and again October 1936–March 1937), member of *NSDAP* (*Nationalsozialistische Deutsche Arbeiterpartei*, known as the Nazi Party) since 1932, the author of *Nationalsozialismus und Universität. Rektoratsrede* (Leipzig, 1933).

[3] Helmut Berve (1896–1979), classicist and historian, member of the Nazi Party since 1933.

[4] Siegmund-Schultze advises me that "kw" likely means "*kann wegfallen*" = can disappear, which is a note that even today is attached to positions which the administration intends to eliminate.

Herr Reichardt asks: The dismissals are based on §6 of the law. Would it be possible to learn something about legal issues related to this question? After all, they fought at the front and were combatants in the war. And among the students this caused a considerable uproar.

Dekan called upon the Rektor who was present.

Rektor: I cannot tell you everything I discussed in Berlin. Everything is still in flux. When I returned from Berlin, I called the Ministry. The next day I was asked over the phone to submit a report. Now the report is before me. I want to tell you the significant things in it. I have reported to the Deputy Secretary Vahlen who directly reports to the *Reichsminister* [Rust], about the current situation in Leipzig and the recent dismissals. That happened on May 2. At that time the *Dekan* only informed me that 4 *Dozenten* [app. associate professors] at the Philosophical *Fakultät* were affected by the dismissals. Meanwhile I have learned that Dr. of Medicine Bettmann was also affected. He has also been dismissed. I was asked in the presence of the General Counsel, Count Rantzau, to characterize the instructors affected by this action, and their military service. In addition to which I suggested to discuss their relations abroad and depict the consequences of their dismissal. I did it as well as I could. I mentioned the reputation of the professors. I emphasized that Mr. Landsberger was regarded as a leader in his field, that he had relations to England. Levi had an offer from Tehran. Wach, whom I have known since his *habilitation* at Leipzig, had just received a one-year leave for a visiting position in America. Weigert had severe problems with his ears. And he had participated in war-related scientific investigations during the war.

Regarding the consequences in Leipzig, there is a certain uproar among the students of those affected, which I discussed in more details. Among the instructors too. Mainly because dismissals were based on the §6. Several instructors had asked me whether this paragraph can be used in their own fields and whether the *Fakultät* that is responsible for the completeness of the course offerings, had been consulted. Most of the colleagues had expressed the opinion that §6 could not be applied to veterans of the war. The opinion of the lawyers was that there was inconsistency between these actions and the prerogatives of the Minister, who alone has the right to dismiss. Also in the case of Landsberger suggestions should be made for an immediate successor. But a position that was canceled based on §6 cannot be re-occupied. This is a contradiction but the people in Berlin told me that it is not an obstacle that could not be overcome. In many other cases a similar procedure has been followed against non-Aryan professors. They filled the position some months later. In that case the position must be included in the budget again. It has turned out that the position is indispensable.

The Rektor has summarized his thoughts as follows: I am not familiar enough with the legal situation to respond appropriately and therefore I have asked for a full clarification of the legal situation.

Dekan: I will be in Dresden tomorrow and I feel it is my duty to point out to the Ministry that the *Facultät* has not been consulted. [I] have also received letters from foreign students.

v.d. Waerden: Can't the Rektor say anything about the official reasons?

Rektor: I can't. In Berlin they did not even know the names of these people.

v.d. Waerden: And how about Dresden? After all, it is natural to suspect that it is against the Jews and there are no [other] official reasons.

Dekan: The dismissals were done “in the interest of the service” (“*im Interesse des Dienstes*”). It is not our responsibility to go further into that.

Heisenberg: This action has caused dismay among many of us because they [we] felt that it did not satisfy the meaning of the law. This is: combatants belong to the people’s community! It is our duty to help them in every respect, especially because their students have already stood up for them. It is necessary that the *Facultät* says that it is about people who have put their life at risk for us.

Golf: These are concerns that are justified. But please do not continue the discussion and do not ask questions. The report has now been sent to Dresden. The reply will come. The Dekan travels to Dresden tomorrow. Any further discussion today is therefore superfluous. We hope that we will be informed about the reply.

Hund: I believe that I cannot refrain from expressing the sentiment among the group of colleagues. If these actions become a fact, this would show that a meaning of the exemption in the law, that men who have fought on the frontlines could not be expelled, would be violated. For us that would be a serious disappointment in the Government. Many of us, who have not been to the frontlines, including myself, would have to be ashamed before these men.

v.d. Waerden: It would be useful if an unambiguous decision could be reached regarding the rights of the combatants and the meaning of the law, which is obviously disregarded.

Dekan: I may remark that I allow this discussion only so that I can report in Dresden about the sentiment among the *Facultät* committees.

Golf: I feel satisfied with what the Rektor has told us. But I want to advise (in a louder voice) Herr v.d. Waerden to be more cautious. He said: a paragraph of the law has been violated. He obviously did not keep in mind that this amounts to saying that the Governor has violated the law. We don’t know his reasons and it is not up to us to make a judgment. So, please, be more careful, be more cautious with your comments.

v.d. Waerden: (in a loud whisper directed at Golf) Thank you!

Golf: (across the table, loudly): The matter is thus closed!

Schweitzer: We have learned in part about the legal basis of the matter, and in part we have been promised a complete clarification. But there is also an aspect of decency to the matter. Among the non-tenured faculty members the revocation of the teaching permits is tantamount to an indefinite dismissal. Under the law this is only possible in case of a disciplinary action. Maybe it is possible to inquire in Dresden whether or not an indefinite dismissal is justified in this case. Even the most junior assistants are protected against such a dismissal.

Rektor: I haven’t restricted myself in Berlin to the legal side of the matter, but I have also mentioned its extraordinary severity.

Dekan: We now discuss point 4 on the agenda. . .

\- - - - - -

These are my notes of the debate. I still have the original stenography. I did not make any further notes, but I accept responsibility for the correctness of what I have noted.

I greet you with Heil Hitler!

Yours,

Signed Hch. Junker

As we can see from this incredible stenography, the five protesting professors use moral and legal arguments in opposing the dismissals of their Jewish colleagues, the draconian dismissals for cause (§6) without the right to ever work in the profession. Van der Waerden makes a legal argument based on the exemption for Jewish veterans of World War I provided in the April 7, 1933, law. Of course, he knows that Nazi Germany lives not by the law but by the latest word of the Nazi leaders. Yet Van der Waerden demands from the Nazi State to live by its own laws:

> It would be useful if an unambiguous decision could be reached regarding the rights of the combatants and the meaning of the law, which is obviously disregarded.

Van der Waerden's son, Hans van der Waerden, observes [WaH2]:

> He [B.L. van der Waerden] decided, whatever happened, to stay aloof of German politics, put a bridle on his personal anti-fascist feelings (without denying them), and never to speak overtly neither in opposition to Nazi ideology nor in favor of it.

Yes, I agree, in general. However, during this faculty meeting, Van der Waerden goes beyond his typical judicial approach to the Nazi regime and attacks one of the pillars of Nazi ideology, its anti-Semitism:

> It is natural to suspect that it is against the Jews and there are no [other] official reasons.

Heisenberg and Hund too address both legal and moral aspects of the dismissal:

> This action has caused dismay among many of us because they [we] felt that it did not satisfy the meaning of the law ... It is necessary that the *Facultät* says that it is about people who have put their life at risk for us. (Heisenberg)
>
> If these actions become a fact, this would show that a meaning of the exemption in the law, that men who have fought on the frontlines could not be expelled, would be violated. For us that would be a serious disappointment in the Government. Many of us, who have not been to the frontlines, including myself, would have to be ashamed before these men. (Hund)

A public protest against the firing of Jewish professors in 1935 was a rare and brave act. As I reported in 2004 [Soi4], the stenography of the meeting left on me an impression that Heisenberg, Hund, and Van der Waerden, the three professors who protested the strongest, were co-conspirators, who discussed between themselves not only physics but also politics. Having now read Heisenberg's 1971 memoirs [Hei2], I find there a confirmation of my conjecture. Thirty-six years later, Heisenberg remembers affairs of year 1935 and shares them with us:

> Political interference in university life became more and more intolerable. One of my faculty colleagues, the mathematician Levy, who, by law, should have enjoyed immunity because of his distinguished war record, was suddenly relieved of his post. The indignation of some of the younger members of the staff – I am thinking particularly of Friedrich Hund, Karl Friedrich Bonhoeffer, and the mathematician B. L. van der Waerden – was so great that we thought of tendering our resignations and of persuading other colleagues to follow suit.

In 1935, a mass resignation of some of the leading professors, including the Nobel Laureate Heisenberg, could have shaken up even the unshakeable Nazi state – if it were to become widely known. This was an incredibly daring plan, which would have cost all the participants their professorships and careers in the Third Reich, and possibly more. But the plan has not been implemented. Heisenberg explains (ibid.):

> Before taking this grave step, I decided to discuss the whole question with an older man, who enjoyed our full confidence. I accordingly asked Max Planck for an interview and then paid a visit to his home in the Grünewald section of Berlin . . .
>
> I told him about the latest developments in Leipzig and about the plan of some of the younger staff members to resign. But Planck was convinced that all such protests had become utterly futile.
>
> "I am glad to see that you are still optimistic enough to believe you can stop the rot by such actions. Unfortunately, you greatly overestimate the influence of the university or of academicians. The public would hear next to nothing about your resignation. The papers would either fail to report it or else treat your protests as the actions of misguided and unpatriotic cranks . . .
>
> In these circumstances, your resignation would have no effect at the present time other than to ruin your career – I know you are prepared to pay that price. But as far as Germany is concerned, your actions will only begin to matter again after the end of the present catastrophic phase. It is to the future that all of us must now look. If you resign, then, at best, you may be able to get a job abroad. What might happen at worst, I would rather not say. But abroad you will be one of countless emigrants in need of a job, and who knows but that you would deprive another, in much greater need than yourself? No doubt, you would be able to work in peace, you would be out of danger, and after the catastrophe you could always return to Germany – with a clear conscience and the happy knowledge that you never compromised with Germany's gravedigger . . .
>
> If you do not resign and stay on, you will have the task of quite a different kind. You cannot stop the catastrophe, and in order to survive you will be forced to make compromise after compromise . . . I think that all of us who have a job to do and who are not absolutely forced to emigrate for racial or other reasons must try to stay on and lay the foundation for a better life once the present nightmare is over. To do so will certainly be extremely difficult and dangerous, and the compromises you will have to make will later be held against you, and quite rightly so. I cannot blame anyone who decides differently, who finds life in Germany intolerable, who cannot remain while injustices are committed that he can do nothing to prevent. But in the ghastly situation in which Germany now finds herself, no one can act decently. Every decision we make involves us in injustices of one kind or another. In the final analysis, all of us are left to our own devices . . ."
>
> And that is how we left it. On the train journey back to Leipzig, the conversation kept going round and round in my head. I almost envied those of my friends whose life in Germany had been made so impossible that they simply had to leave. They had been the victims of injustice and would have to suffer great material hardship, but at least they had been spared the agonizing choice of whether or not they ought to stay on . . . And what precisely were the compromises Planck had hinted at? At the beginning of each lecture, you had to raise your hand and give the Nazi salute. But hadn't I raised my hand

> to wave at acquaintances even before the advent of Hitler? Was that really a dishonorable compromise? And then you had to sign all official letters with "Heil Hitler." That was much less pleasant, but luckily I, for one, didn't have to write all that many official letters, and when I did, the new salutation invariably meant "I don't want to have close contact with you." We were expected to attend celebrations and marches, but I felt it ought to be possible to get out of quite a few. A compromise here, a compromise there, and where did you draw the line? Had William Tell been right to refuse homage to Gessler's hat, thus endangering the life of his own child? Ought he to have compromised? And if the answer was no, ought we to compromise with our own Gesslers?
>
> Conversely, if one decided to emigrate . . . might you not be simply leaving the field to those madmen, those spiritually unhinged creatures whose demented plans were driving Germany headlong into disaster?

Observe that these are 1971 Heisenberg's recollections of his 1935 thoughts. By 1971, he knows that the Third Reich ended up killing tens of millions of innocent people; some 6,000,000 million of Jews alone. How could he reduce his compromises with the Nazi state to merely the salutation "Heil Hitler"? What about lending the prestige of one of the world leading physicists to the Third Reich? Worse yet, what about working on an atomic bomb and an atomic reactor under Hitler, for Hitler? Heisenberg continues (ibid.):

> Planck had said that we might be faced with alternatives that would be equally unjust. Were such situations possible? I tried to think up an extreme situation which, though it had not occurred in reality, was not too far-fetched, not quite obviously beyond a humane solution. This was the example I finally hit upon: A dictatorial government has jailed ten of its opponents and has decided to kill at least the most important of the prisoners. At the same time, the government is terribly anxious to justify this murder before the rest of the world. Accordingly, it makes an offer to another of its opponents, say, a jurist who has been left at liberty because of his high international renown: if he can produce and sign a legal justification for the murder of the most important prisoner, then the other nine will be released and allowed to emigrate. If he refuses, all ten prisoners will be killed. The jurist is left in no doubt that the dictator is in earnest. What is he to do? Is it clear conscience, a "white waistcoat," as we used to call it cynically, worth more than the lives of nine friends? Even his suicide would be no solution; it would merely lead to the immediate slaying of the innocent ten.
>
> Thinking along these lines, I remembered a conversation with Niels Bohr, during which he referred to the fact that justice and love were complementary concepts. Although both are essential components of our behavior toward others, they are, in fact, mutually exclusive. Justice would force the juror to withhold his signature, the more so as the political consequences of his signing might be such as to destroy more innocent people than the nine friends. But would love refuse the cry for help sent by the desperate families of the nine friends?
>
> After a while, I realized how extremely childish it was to go on playing such absurd mental games. What mattered was to decide here and now whether I ought to emigrate or to stay in Germany. "Think of the time after the catastrophe," Planck had said, and I felt he was right. We would have to form islands, gather young people round us and help them to live through it all, to build a new and better world after the holocaust

> [Heisenberg uses a small "h"]. And this was bound to involve compromises, for which we would rightly be held to account – and perhaps even worse . . . By the time the train pulled into Leipzig, I had made up my mind: I would stay on in Germany, at least for a time, continue working at the university, and, for the rest, do my bit as best as I possibly could.

I am compelled to reply to these three great physicists.

"Absurd mental games," you say, Professor Heisenberg? How often does one remember the 1935 thoughts in 1971 – and prominently insert them in his book? Clearly, this train of thought mattered a great deal to you. Moreover, between 1935 and 1971, you included a very similar kill-one-save-ten situation in your unpublished 1947 document "On Active and Passive Opposition in the Third Reich" [Hei1]. I wish to test your morality *theory* by my *experiment*:

Dr. Heisenberg, would *you* sign a death sentence for "the most important" innocent person in order to save others? Would *you* sign a death sentence for "the most important" protester of the May 1935 faculty meeting, Bartel van der Waerden, in order to save Carl-Friedrich von Weizsäcker and Friedrich Hund? I believe that you were the most loyal friend of people in your close circle, and thus you would have never signed such a death sentence. Thus, your clever theory, praising the *morality of collaboration with the Nazi regime* in killing an innocent person, does not pass the ultimate test by experiment.

Signing a death warrant to an innocent would make you an accomplice of the criminal Nazi regime.

Dr. Niels Bohr, I deeply admire you as a scholar and man. Do you really believe, as Heisenberg reports, that "justice and love were complementary concepts"? I'd say that the complement of love is indifference, while justice is synonymous with impartiality (recall the image of Lady Justice, a blindfolded woman holding a scale). So, by your logic indifference and impartiality are synonyms – and I submit, they are not. The indifferent juror would sign a death verdict for the innocent person – what does he care – while the impartial juror will not.

Dr. Max Planck, I share some of your views, which I learned only on November 12, 2010, when I read them quoted in Heisenberg's book [Hei2]. You warned Werner: "If you do not resign and stay on . . . in order to survive you will be forced to make compromise after compromise . . . and the compromises you will have to make will later be held against you, and quite rightly so." I agree with you, and for this very reason, I would have advised Heisenberg to leave Nazi Germany rather than stay on and thus support the criminal state by his nuclear research and by his high worldwide reputation.

In the summer of 1939, just before the start of World War II, physicist (Nobel Prize 1938) Enrico Fermi warns his friend Heisenberg about inevitable compromises and responsibility for them, very much like Max Planck. However, while Planck drew a conclusion of staying in Nazi Germany, Fermi urges his friend to leave:

> Whatever makes you stay on in Germany? You can't possibly prevent the war, and you will have to do, and take the responsibility for, things which you will hate to do or to be responsible for.

There was no shortage of advice. In Heisenberg's May 12, 1935, letter, he briefs his mother that the Leipzig University Rektor pressures Heisenberg to enter the German Army as a reserve officer in order to remedy his part in the faculty meeting protest and to demonstrate his loyalty to the Third Reich. Heisenberg does follow RektorKrueger's advice and serves as a reserve officer in the Army of Nazi Germany.

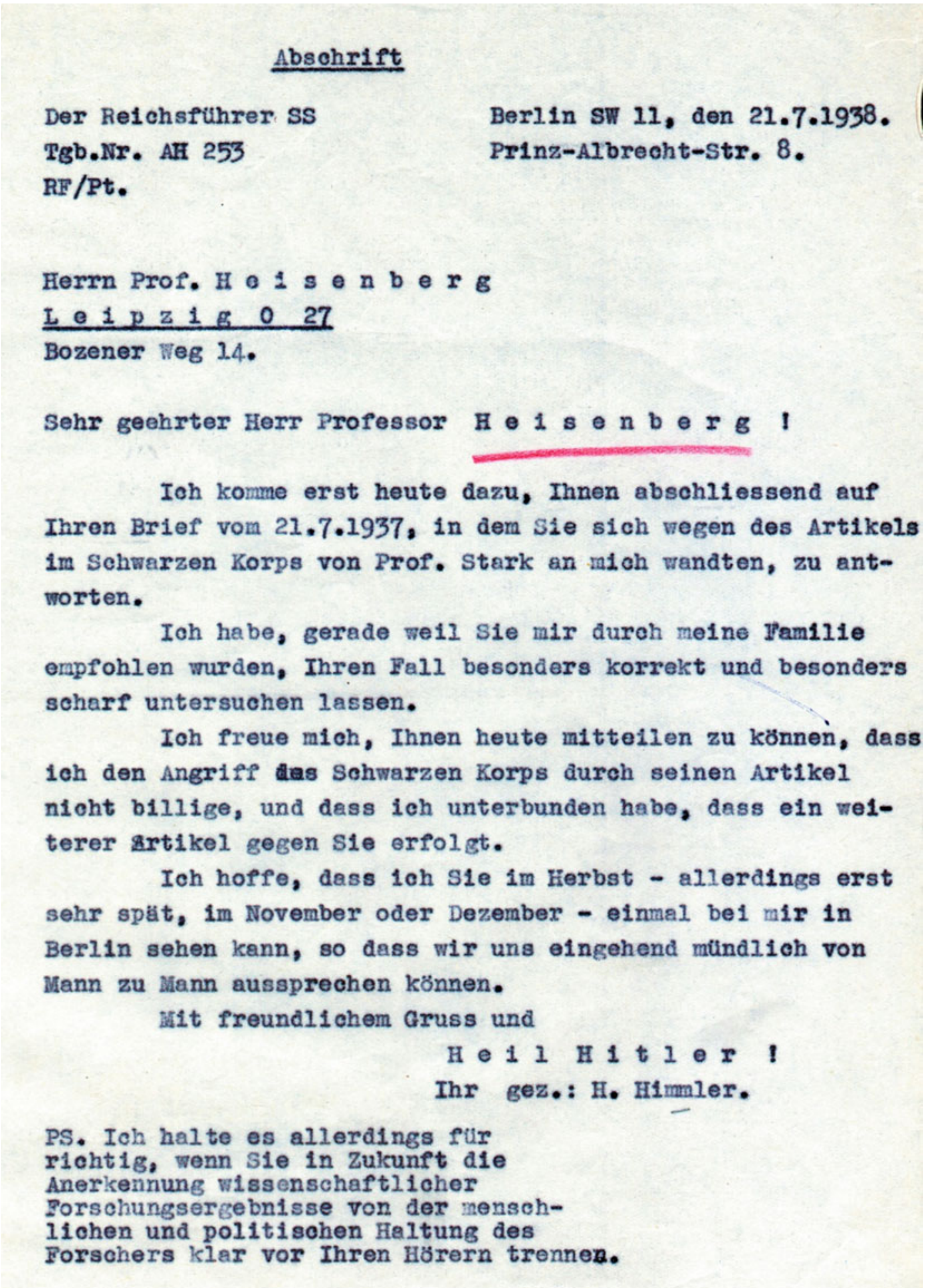
Ever since the late 1920s, Philipp Lenard (Nobel Laureate 1905) and Johannes Stark (Nobel Laureate 1919) had promoted the notorious notion of "Aryan Physics" contrasted with "Jewish Physics" of Einstein and others. On July 15, 1937, Stark called Werner Heisenberg a "White Jew" in the *SS* newspaper *Das Schwarze Korps* (*The Black Corps*). Heisenberg was outraged, as Van der Waerden would remember even a decade later. And so just two years after the heroics of the May 1935 faculty meeting and pledge to "do my bit as best as I possibly could," Heisenberg allows himself a shocking compromise with the Nazi regime by entering in a "contract with the devil." An old proverb warns, be careful what you wish for: you just might get it. Just six days after Stark's article, in the July 21, 1937, letter, Heisenberg asks none other than the *SS Reichsführer* Heinrich Himmler for protection.

In one year to the day (!), the desired protection is granted by Himmler, who on July 21, 1938 writes about it to his subordinate, *Gestapo* chief, *SS*-Lt. General Reinhard Heydrich, *SS-Obergruppenführer*, Chief of the Reich Main Security Office, including the *SD*, *Gestapo* and *Kripo* (Heydrich was the one who presided over the January 20, 1942, Wannsee Conference, dedicated to the "Final Solution," plans for the deportation and extermination of all Jews in German-occupied territories) [Gou, 116–119]:

> Dear Heydrich,
>
> I have received the good and very objective report on Professor Werner Heisenberg, Leipzig. I enclose herewith a very proper letter of Professor Prandtl, Göttingen, with which I agree. I also enclose a copy of my letter to Heisenberg for your information . . .
>
> I believe that Heisenberg is a decent person and that we cannot afford to lose or to silence this man, who is still young and can still produce a rising generation in science.

One would think that "a decent person" is a high compliment. However, here it comes from one of the Nazis' top mass murderers, someone whose taste in morality we must question. The same day Himmler promises protection in a letter to Heisenberg personally (see a photocopy of the letter in this chapter) [ibid]:

> Only today can I answer your letter of July 21, 1937, in which you direct yourself to me because of the article of Professor Stark in "*Das Schwarze Korps*."
>
> Because you were recommended by my family, I have had your case investigated with special care and precision.
>
> I am glad that I can now inform you that I do not approve of the attack in "*Das Schwarze Korps*" and that I have taken measures against any further attack against you.
>
> I hope that I shall see you in Berlin in the fall, in November or December, so that we may talk things over thoroughly man to man.
>
> With friendly greetings.
>
> Heil Hitler!
>
> Your,
>
> H. Himmler
>
> P.S. However, I consider it best if in the future you make a distinction for your audience between the results of scientific research and the personal and political attitude of the scientists involved.



Abschrift

Der Reichsführer SS
Tgb.Nr. AH 253
RF/Pt.

Berlin SW 11, den 21.7.1938.
Prinz-Albrecht-Str. 8.

Herrn Prof. Heisenberg
Leipzig O 27
Bozener Weg 14.

Sehr geehrter Herr Professor Heisenberg !

Ich komme erst heute dazu, Ihnen abschliessend auf Ihren Brief vom 21.7.1937, in dem Sie sich wegen des Artikels im Schwarzen Korps von Prof. Stark an mich wandten, zu antworten.

Ich habe, gerade weil Sie mir durch meine Familie empfohlen wurden, Ihren Fall besonders korrekt und besonders scharf untersuchen lassen.

Ich freue mich, Ihnen heute mitteilen zu können, dass ich den Angriff des Schwarzen Korps durch seinen Artikel nicht billige, und dass ich unterbunden habe, dass ein weiterer Artikel gegen Sie erfolgt.

Ich hoffe, dass ich Sie im Herbst - allerdings erst sehr spät, im November oder Dezember - einmal bei mir in Berlin sehen kann, so dass wir uns eingehend mündlich von Mann zu Mann aussprechen können.

Mit freundlichem Gruss und

Heil Hitler !

Ihr gez.: H. Himmler.

PS. Ich halte es allerdings für richtig, wenn Sie in Zukunft die Anerkennung wissenschaftlicher Forschungsergebnisse von der menschlichen und politischen Haltung des Forschers klar vor Ihren Hörern trennen.

Copy of Himmler's Letter to Heisenberg, July 21, 1938. (Courtesy of Leipzig University)

And thus, Heisenberg receives Himmler's high approval to speak about relativity theory, under the condition that he makes no mention of its creator Albert Einstein. It is hard to believe that such a brilliant mind, Werner Heisenberg, would ask one of the most brutal Nazi murderers, Heinrich Himmler, for favors. However, Goudsmit leaves no doubts about it by including facsimiles of both Himmler's letters, to Heydrich and Heisenberg, in his book *Alsos* ([Gou], pp. 116 and 119).

In the midst of the Nazi regime crimes against humanity, Heisenberg's defense of a theory – the relativity theory as it were – seems insignificant, while his demand for restoring his personal "honor" appears petty. In my eyes, Heisenberg's appeal to Himmler and Himmler's grant of protection fare among the darkest stains on Werner Heisenberg's reputation. The contract that Leipzig's Dr. Heisenberg reached with *SS Reichsführer* Himmler eerily reminds me Johann Wolfgang von Goethe's classic book about another scientist, Dr. Faust, entering in a contract with the Devil. In fact, Goethe spent his early years in Leipzig, studying at Leipzig University. Leipzig's 15th century *Auerbachs Keller* restaurant with its legend of Dr. Johann Georg Faust's barrel ride became the only real location in Part One of Goethe's "Faust."

Heisenberg paid a high price for his high *SS* protection. This protection ended forever the days when Heisenberg could publicly criticize any actions of the Nazi regime, even if he were so inclined, for Heisenberg became a highly protected asset of this criminal regime. Heisenberg had countless opportunities to emigrate, for right before the war commenced, he was in the United States and received offers from a number of leading American universities. However, Heisenberg chose to stay in and to serve Germany – Nazi Germany, as was the case.

Let us return to the Third Reich, year 1935. Shortly after the Leipzig faculty meeting, the entire Van der Waerden family, Bartel, Camilla, and their daughters Helga and Ilse are spending their summer vacation in Bartel parent's magnificent house in Laren, near Amsterdam. On August 10, 1935, Bartel writes a letter to Richard Courant, who is already living in New York:

> Personally, we are all doing very well. Our oldest daughter Helga had her appendix removed yesterday. The operation seems to have been successful. We are here in Holland for two months and rest up our souls from the constant tensions, hostilities, orders and paperwork . . . Ministries examine who has not yet been completely forced into line [of National Socialism], who is a friend of Jews, who has a Jewish wife, etc., as long as they themselves are not torn apart by their fight for power.

This paragraph truly opened my eyes to Van der Waerden's mid-1935 perception of his situation. He is not a prisoner of the "Ivory Tower": he is acutely aware of life around him. But Van der Waerden views life in the Nazi state not as a *tragedy* but as a *farce* and writes about it with amusement. *The entire family is abroad* in Bart's Homeland, Holland, yet he does not seem to give any thought about the whole family remaining in Holland!

40.5 Germany Treacherously Invades Holland

German thunder . . . will come and when you hear crushing, as it has never crashed before in all of world history, you will know, German thunder has finally reached its goal. With this sound, eagles will fall dead from the sky, and lions in the most distant desert in Africa will put their tails between their legs and crawl into their royal caves . . . And the hour will come.

– Heinrich Heine, 1834

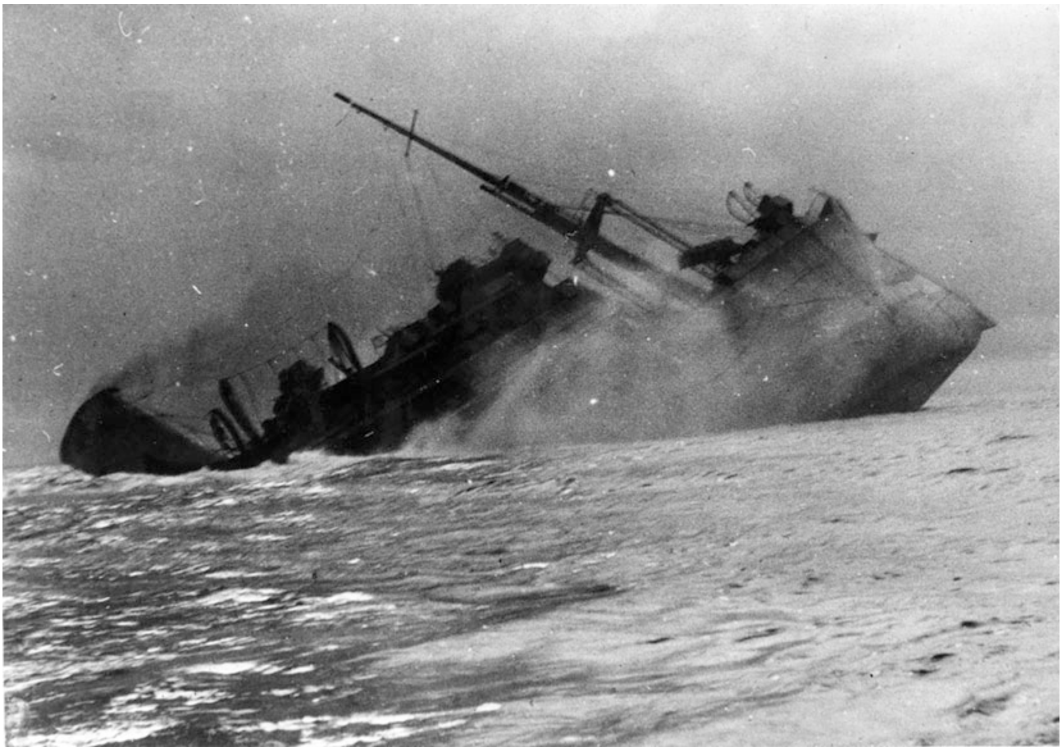
What I should explain to the Dutch people is, however, not my actions before 1940, but those after the Netherlands had been attacked by Germany . . . I have never given a class or worked on things that could be used for military purposes.

– Bartel L. van der Waerden

The Netherlands safely lived in neutrality through World War I. It hoped to repeat it in World War II. However, the Dutch plan of neutrality crumbles when on May 10, 1940, Germany treacherously attacks the Netherlands, as well as Luxembourg, Belgium, and France (Norway and Denmark were attacked earlier, on April 9, 1940). Queen Wilhelmina of the Netherlands and her government flee to London. Her daughter, the future Queen Juliana, and her family go into exile in Canada.

Destruction of Rotterdam, May 14, 1940; Bundesarchiv_Bild_146-2005-0003, Rotterdam, Zerstörungen. (Wikipedia)

The Dutch fight against the overwhelming advantage of the Third Reich Navy and Air Force. German bombers set the entire inner city of Rotterdam ablaze. The devastation of Rotterdam, serving as a threat to do the same to Utrecht and Helder, Amsterdam and Den Haag, forces the Netherlands to surrender the following day, May 15, 1940.



Dutch Ship being torpedoed by a German submarine, October 1945. (Photo ANEFO; Archive of Alexander Soifer)

Right on its first page, The New York Times reports the reaction of President Franklin Delano Roosevelt, who “condemns” the invasion but is determined to keep American “neutrality” [sic] toward Hitler:

> WASHINGTON, May 10--President Roosevelt twice today condemned Germany’s invasion of Belgium, Holland and Luxembourg as an unwarranted aggression on neutral countries and as threatening the cultural and scientific civilization of the world...
>
> On both occasions the President impressed his determination to keep America at peace and safeguard the nation’s neutrality.

Some condemnation! Roosevelt is prepared to pay for his “neutrality” by throwing Holland, Belgium, and Luxembourg to the hungry Nazis! As to Hitler, if he cherished the plans to create *Großgermanisches Reich Deutscher Nation* (Greater Germanic Reich of the German Nation) that would include the Netherlands, Flemish Belgium, Luxembourg, Norway, and Denmark, the brutal invasion was a ridiculous way to go about it.

The main personage of our story Van der Waerden finds himself in an awkward situation: he is a professor and civil servant of Nazi Germany that has waged the brutal unprovoked war against his Homeland, Holland. How is he affected by this course of events? What is his take on the matter? Upon Nazi Germany's invasion of the Netherlands, many Dutch citizens inside Germany are at first treated as enemies and interned. In fact, right on May 15, 1940, the day Holland capitulated, Van der Waerden is suspended by the Rektor from teaching at Leipzig University:

> I already asked you yesterday over the phone to refrain from any teaching activity until further notice. I herewith repeat this order in writing and ask you to discontinue your administrative activity as Director of Mathematical Seminars and Mathematical Institute.
>
> Meanwhile I have asked the Ministry for a decision whether in view of you being an official and your oath to the *Führer* my order regarding your activity as a Professor and Director of the Institute should continue.

Van der Waerden was detained but soon released. Permission to lecture again was granted to him via telephone by the senior servant Dames on 11 June 1940. His reaction to this brief suspension allows us an unexpected insight into his views of Germany and Holland. He understands from the beginning that the suspension is likely to be short-lived, but that as a condition for reinstatement as a professor at Leipzig he may be asked to accept Nazi Germany's citizenship. The day following the suspension, on May 16, 1940, Van der Waerden writes about his dilemma to a trusted friend, Editor of the *Mathematische Annalen* Erich Hecke [Wae8]:

> For the time being I am not allowed to teach courses. But the Rektor has already written to Berlin and asked for an authorization to allow me to carry on my office. The Dekan predicts that this would be smoothly approved; maybe I would be asked to become a German citizen. You will understand that I would be uncomfortable with that at this time. In principle I have no objections against German citizenship, but at this moment when Germany has occupied my homeland, I really do not want to abandon my neutrality and take the German side.

Thus, "*in principle*" Van der Waerden has "*no objections against German citizenship.*" He merely "does not want to abandon his *neutrality*" between the brutal invader, Nazi Germany, and his victimized Homeland. How does one explain such insensitivity toward the Homeland? Could it be that Van der Waerden by now believes that he belongs to Germany, German culture in general, and to German science and mathematics in particular? If so, this would explain his neutrality and reluctance to leave Germany when in the middle of the war he receives a job offer from Utrecht University.

40.6 A Dream of Göttingen

Before Hitler's ascent to power, Germany arguably occupied the highest mathematical ground in the world, and Göttingen University was its greatest peak. From Felix Klein to David Hilbert, the Göttingen mathematicians created an unparalleled school. In 1928, Richard Courant and the Rockefeller Foundation created in Göttingen the Mathematical Institute populated by some of the finest scholars. At this time, even young brilliant Americans, such as Saunders Mac Lane, were attracted by Göttingen. Mac Lane recollects [Mac]:

> The Mathematical Institute in Göttingen in 1931 had an outstanding tradition: Gauss, Riemann, Dirichlet, Felix Klein, Minkowski and Hilbert. It was located in a new and ample building (thanks to the Rockefeller Foundation, which had also provided such a building for mathematics at Paris).

Van der Waerden spent his happy young years at Göttingen. He was the favorite student of Emmy Noether, habilitated under Richard Courant, served at Göttingen as Courant's *Assistent* and *Privatdozent*. Fond memories of the great Göttingen must have inspired a dream to work there again. It is only natural that in late 1943–early 1944, Van der Waerden tries to convert his Dream of Göttingen into reality. The choice of people he asks for help in obtaining a Göttingen professorship is surprising for someone who thought of himself as a "strong opponent of the Nazi regime." There is an Old Russian proverb, "Tell me who your friends are, and I will tell you who you are."[5] As all universal declarations, it does not fit all cases. And yet, there is a grain of truth in this folk wisdom. Let me introduce to you the two Van der Waerden's helpers (more information about them can be found in [Rem], [Sie3], [Seg], [Geo], and other sources).

The first helper, Wilhelm Süss, a professor of mathematics and Rektor of Albert Ludwig University of Freiburg, a 1934–1937 member of the *SA* (Storm Troopers), joined the Nazi Party (*NSDAP*) in 1937, and the *Nationalsozialistischer Deutscher Dozentenbund* (Nazi Lecturers Confederation) in 1938. During 1937–1945, Süss was the *Führer* of the *Deutsche Mathematiker-Vereinigung* (*DMV*). He distinguished himself by enthusiastically initiating the expulsion of Jews from the *DMV* membership rolls right after becoming its president, even before he was ordered to do so by his Nazi patrons. "Jews were not merely excluded from *DMV*; the Nazis attempted to eliminate them from the history of the *DMV*, as if they had never existed" [Seg]. In 1938, Süss also initiated the expulsion of Jews from editorial boards. Consequently, he got such a clout with high Nazi officials that on August 3, 1944, Hermann Göring himself approved the creation of the *Mathematisches Forschungsinstitut Oberwolfach* on the hills of the Black Forest. Naturally, Süss served as Oberwolfach's first director. Van der Waerden was friendly with Süss, gave a talk at Süss' invitation at Freiburg University in 1944, and corresponded with Süss until the latter's passing away in 1958.

In August 1985, I spent a delightful week at Oberwolfach. Then I was not a historian and did not know that this scenic mathematical retreat was authorized by Hermann Göring and paid for by Nazi money. The *Mathematisches Forschungsinstitut Oberwolfach* has been providing a valuable service to the international mathematical community. And yet, it would be hard for me now to stay there again, for ghosts of the past would spoil the serenity of the rolling hills and the delight of scientific exchange.

I hear you asking me: What can *Mathematisches Forschungsinstitut Oberwolfach* do today about its Nazi past? To begin with, Oberwolfach must stop lying about its history. Oberwolfach Director, 2002–2013, Prof. Dr. Dr. h.c. Gert-Martin Greuel certainly knows the history of the institution he has led for 11 years. Yet, Greuel conceals the Nazi roots and

[5] In 2014, a new version of this proverb was born in light of the Russian annexation of Ukrainian Crimea: "Tell me whose Crimea is, and I will tell you who you are."

starts the Oberwolfach Institute history in 1946 in his chapter "Mathematics Between Research, Application, and Communication" in the book *Raising Public Awareness of Mathematics*, Behrends, E., et al. (eds), Springer, ISBN 978-3-642-25710-0.[6]

I feel affinity to the refugee from Nazi Germany Professor Max Dehn, who as a Jew was expelled from the *DMV* in 1935. When invited to rejoin it in 1948, Dehn replies [Sie3, p. 393]:

> I cannot rejoin the *Deutsche Mathematiker-Vereinigung*, I have lost confidence that such an association would act differently in the future than in 1935 . . . I am not afraid that the new *DMV* will again expel Jews, but maybe next time it will be so-called communists, anarchists or "colored people."

The second helper, Helmut Hasse, "a good German and National Socialist" as Van der Waerden describes him in his June 13, 1942, letter, was an excellent algebraist, a major contributor to class field theory. He was a member of the anti-Semitic *Deutschnationale Volkpartei*, led from 1928 by the eventual member of Hitler's first cabinet Alfred Hugenberg. Segal argues [Seg] that "Hasse was no anti-Semite, and, for example [sic], remained friendly with [Hasse's 1921 Ph.D. thesis advisor Kurt] Hensel until his death in 1941." How can one example – or two, Hasse was friendly with his coauthor Emmy Noether – prove that Hasse was not an anti-Semite? Isn't it typical for an anti-Semite to hate all Jews except for a few personal friends?

Princeton Mathematics Professor Willy Feller told then young Gerard Washnitzer (Professor Emeritus during my years at Princeton University) that Feller was present at Hasse's lecture at the Oslo International Congress of Mathematicians (July 13–17, 1936). While giving a lecture on number theory and emphasizing great significance of class field theory, Hasse mimicked a Yiddish accent while uttering "*Satz – beweis – satz – beweis.*"[7]

Hasse viewed "Hitler as a national hero" and on October 29, 1937, applied for membership in the Nazi Party [Seg, 124–167]. The fanatical Nazis required from its members not to have a "full-Jewish" ancestor living after 1800, whereas Hasse was a "1/16 Jew as a consequence of a baptized great-great-grandmother" (ibid.). Hasse appealed the rejection to Hitler himself (who did grant a few exceptions). As Hasse was a *Korvet-Kapitän* (Corvette Captain, equivalent to the US Lieutenant-Commander) serving in Nazi Germany's War Navy starting in 1939 (and through the end of the war in 1945), the decision on his Nazi Party membership was postponed until after the war (ibid.). This put Hasse in a most opportune situation, and he took a full advantage of it: he was a card-carrying member of the Nazi Party during the Nazi era and claimed not being a Nazi after Nazi Germany lost the war. "Normal heroes" love to always be on top!

Hasse expressed the most hateful attitudes toward people of other races and ethnicities. Let me share with you several vivid examples, some of which I published for the first time in [Soi47] and others appear for the first time here.

Jacopo Barsotti told Princeton's Gerard Washnitzer, that as a graduate student, Barsotti attended Hasse's talk in Pisa after the start of World War II and before Italy's collapse. During

[6] https://www.researchgate.net/publication/301171280_Mathematics_Between_Research_Application_and_Communication

[7] Recorded interview with Professor Gerard Washnitzer, Commons Room, Fine Hall, Princeton University, March 2004.

the talk, Leonida Tonelli asked Hasse about the fate of the Polish mathematicians, and in particular about Juliusz Schauder. Hasse replied,

> Poles should not do mathematics. They should work in coal mines and agricultural labor.[8]

This event was independently confirmed to Washnitzer by other Italian mathematicians during the 1950 Cambridge (USA) International Congress of Mathematicians.[9]

In March 15, 1939, letter to Harvard Professor Marshall Stone, Hasse urged the exclusion of the German refugees to the United States from serving as reviewers for the *Zentralblatt für Mathematik*:

> Looking at the situation from a practical point of view, one must submit that there is a state of war between the Germans and the Jews . . .

"The state of war," Herr Hasse? The state of war between armed to the teeth Nazis and unarmed innocent victims? Fortunately, there were American mathematicians (far from all) who understood the nature of the real war. As C. R. Adams reports (ibid.):

> Mr. Veblen insists that there is a war by the Germans against *civilization.*

It is amazing that even many years after the war ended, during which the world learned so much about the crimes of Nazism, Hasse did not change his racist views. Segal, who presents much material on Hasse [Seg], describes how in the 1960s at Ohio State University, USA, Hasse claimed that "slavery in America had been a good institution for blacks."

I must quote here a letter [Lan2] published in Germany and the USA by Serge Lang, which graphically portrays Hasse's views and behavior during the war and the Nazi occupation of France and Norway:

> I take this opportunity to put in the record some information concerning Hasse's behavior after France's defeat in 1940. In the fall of 1940, Hasse went to meet Elie Cartan at his home in Paris. Hasse was dressed in a German uniform. The only other person present was Elie Cartan's son, Henri Cartan, whom I heard personally report the

---

[8] Hasse is here in a complete accord with the Nazi policies toward the Polish population. Richard C. Lukas writes [Luka]: "The German campaign against the Poles focused largely but not exclusively upon the elimination of anyone with the least political or cultural prominence. Years before their invasion of Poland, the Germans drew up lists of prominent Poles slated for execution or imprisonment . . . The Nazi determination to obliterate the Polish intelligentsia resulted in wiping out forty-five percent of Polish physicians and dentists, forty percent of professors, fifty-seven percent of attorneys, thirty percent of technicians, and a majority of leading journalists." [Let us add the Katyn Massacre, where the Soviet NKVD murdered ca. 21,000 Polish officers and intellectuals.]

The famous French mathematician Jean Dieudonné [Die, p. 16] addresses specifically the fate of mathematicians: "In Poland the mathematical schools were physically annihilated, since half the mathematicians were massacred by the Nazis. They did not recover their standing until after 1970."

[9] Recorded interview with Professor Gerard Washnitzer, March 26, 2004; 3:30–5:30 P.M., Commons Room, Fine Hall, Princeton University. Confirmed by Washnitzer during Sunday, December 3, 2006, 12:45–2:30 interview, Commons Room, Fine Hall, Princeton University.

> encounter publicly in the late fifties, as follows. Hasse acted in a very friendly way, and proposed to Elie Cartan that French and German mathematicians should cooperate, independently of the circumstances which were otherwise occurring. Elie Cartan answered in an equally friendly fashion that it was an excellent idea, but that the Poles should also take part. Hasse then answered no, that the Polish people were a separate people with whom it was not possible to collaborate. Elie Cartan then answered that under these conditions, it was impossible to start a French–German mathematical cooperation.
>
> Some 40 years later, in 2000–2001, at the *Max-Planck Institut* in Bonn, I heard for the first time an account from the Norwegian mathematician Arnfinn Laudal, of a similar visit that Hasse made to Thoralf Skolem in Oslo. Laudal got the story from Skolem himself, and the story was confirmed recently by Skolem's children. Hasse had shown up at Skolem's home dressed in a German Navy [*Korvette Kapitän*] uniform, but was refused entrance by Skolem, on the doorsteps. Hasse had come with a proposition like the one he had made to Elie Cartan.
>
> There occurred a vigorous and high-voiced exchange between Skolem and Hasse. Thus, Hasse's visit to Elie Cartan was not an isolated event.

Helmut Hasse was not content to merely do mathematics in the Ivory Tower and believe in "*Mathematik über alles*." No, Hasse took a full advantage of his status of a distinguished mathematician to spread the racist venom for decades, from the Congress of 1936 to the American visit in the 1960s.

Yet, Peter J. Roquette (Ph.D. under Hasse, 1951) and Günther Frei (Ph.D. under Van der Waerden, 1968) portray Hasse as a man of the highest moral standing. How can one believe Roquette–Frei when they contradict the accounts by such universally admired scholars as Cartan, Skolem, Veblen, and Siegel? The examples of Hasse's behavior and his bigotry I introduced here have been omitted by Roquette and Frei. Moreover, I read in Frei in disbelief [Fre, p. 65]:

> Fighting against politically-minded and fanatical students and striving for the conservation of the scientific importance of the famous institute took most of Hasse's time and energy . . . In Hasse's seminar with the young and gifted students Witt and Teichmüller – Siegel did participate later on – important articles on congruence function fields were written.

"Fighting against politically-minded and fanatical students," you allege? Fanatics were precisely the students Hasse and obviously Frei favor. Gifted as they may have been, Ernst Witt and Oswald Teichmüller *were* storm troopers, members of the notorious *Sturmabteilung*, the Assault Division, "Brownshirts."

With no disagreement from me, Frei calls Carl Ludwig Siegel "the most eminent mathematician in Germany" [Fre, p. 65]. But then Frei omits or is ignorant of Siegel's assessment of Hasse. Let me help my colleague Frei. On March 22, 1939, Siegel, having returned to Germany from the Institute for Advanced Study Princeton, wrote to Oswald Veblen [Seg, p. 165]:

> After the November pogrom, when I returned from a trip to Frankfurt, full of nausea and anger at the bestialities in the name of the higher honor of Germany, I saw Hasse for the first time wearing Nazi-party insignia! It is incomprehensible to me how an intelligent

> and conscientious man can do such a thing. I then learned that the foreign-policy occurrences of recent years had made Hasse into a convinced follower of Hitler. He really believes that these acts of violence will result in a blessing for the German people.

Frei and Roquette with the collaboration of Franz Lemmermeyer edited in 2014 a new edition in English [FLR] of the correspondence between Artin and Hasse which they first published in 2008 in German [FR]. They did a fine job of mathematical commentary. However, an Old Russian proverb warns, "A spoon of tar can spoil a barrel of honey," and the authors added a spoonful of tar to their commentary. Have Frei and Roquette addressed Hasse's application to the membership in the Nazi Party, his strong support for Hitler, service as a *Korvette Kapitän* in the *Obercomando der Kriegsmarine* (The Supreme Command of the War Navy), instances of Hasse's racism and anti-Semitism, etc.? Nothing of the kind is mentioned in the 2014 book.

Frei apparently thinks that the best defense of Hasse accused of anti-Semitism is to flash a positive quote from a Jew. And so, he does precisely that, in the quote that refers to *very early pre-Nazi times* [FLR, p. 29]:

> Abraham Adolf Fraenkel, who like Hasse received his Ph.D. in Marburg under the supervision of Hensel, who was Hasse's colleague in Kiel, and who later was rector of the Hebraic University in Jerusalem, writes in his book [Fra67, p. 153]:
>
> *Personally,* my experiences with Hasse were positive throughout, and I always found him to have a flawless character.

This "*persönlich*" in Frei's quote, by all logic of style begs "*aber*" ("however") in the next sentence. And so, I order Fraenkel's memoirs [Fra67] to check my conjecture, and *voila*: "*aber*" does open the very next sentence, and the paragraph ends in Fraenkel's "dismay" (!) over Hasse's Nazi period conduct:

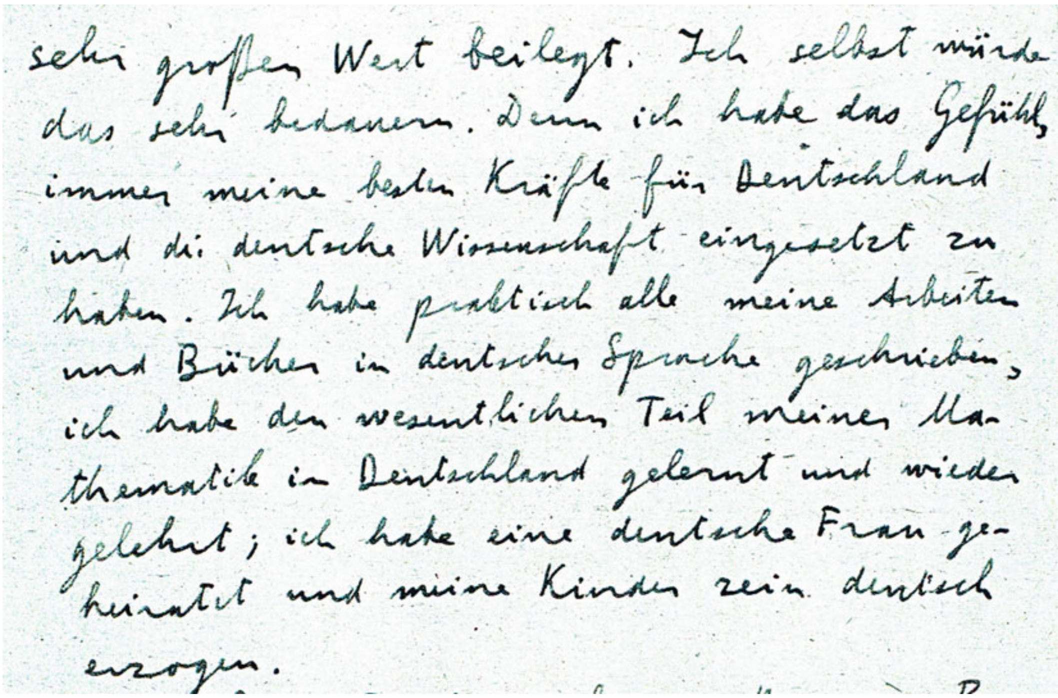
> *However*, some years later, after he [Hasse] had become a professor at Göttingen, a crisis shook his life: one of his opponents found out that he had a Jewish [great-] great-grandfather. Although the German racial laws only reached as far as the grandparents and besides, in his appearance and bearing he made a completely "Aryan" impression, he felt he was in an unbearable situation. He appealed to Hitler, who named him an honorary full Aryan along with some other outstanding, not purely Aryan scholars. Then, he joined the National Socialist Party, but after the war did not crave an alibi, in contrast to the majority of opportunistic careerists. In June 1946, when I met the most important British mathematician, G.H. Hardy and *to my dismay heard these details about Hasse*, Hardy was busy writing a letter to the British occupation authorities in Göttingen, demanding that he be restored to his position in view of his scholarly importance, after he had been dismissed from the University due to his party membership.

So, why do Frei and Roquette go to such a great extent in fabricating a myth of loveable Hasse? Is it because for them *Mathematik über alles* and all moral concerns are negligible? Or is it because there was a severe shortage of heroic mathematicians in Nazi Germany? You want a hero, write about Erich Hecke. *There is an eternal dispute whether mathematics is discovered or invented. There is no dispute – history ought not to be invented, gentlemen!*

Now that I have introduced the helpers, whose Nazi affiliation Van der Waerden knew well, we are ready to return to Van der Waerden himself. In his March 14, 1944, letter, Van der Waerden asks the *DMV* President Süss whether he should accept Utrecht's offer. The Utrecht offer is apparently used in this letter by Van der Waerden as leverage for obtaining another position. Van der Waerden really longs for a professorship at Göttingen:

> Dear Herr Colleague!
>
> Please, allow me the liberty to approach you with the following personal matter. In the last few years I have repeatedly been subjected to difficulties that hurt me very much. I have repeatedly been invited to give many presentations abroad, the first time already before this war, but permission has every time been denied to me. I have been considered for an appointment in Munich, but the appointment did not come off. Now the *Facultät* in Göttingen has nominated me; but the actual appointment seems to miscarry again. I have just [sic][10] received an offer from Utrecht. Faced with the necessity to decide for or against accepting this call, the question arises whether the described above opposition is not an indication of the fact that from the authorities' side my work in Germany is not wanted or at least not a great deal of worth is placed in it.
>
> I would certainly personally strongly regret that, because I spent my best energies for Germany, which I applied to the German Science [*die deutsche Wissenschaft*]. I have written practically all my works and books in the German language, I have learned and also taught a major portion of my mathematics in Germany; I have a German wife, and my children were raised pure Germans [see facsimile].
>
> As a sign that I should not give in to my fear, I hope that I would really receive a call to Göttingen, on which I personally place a great deal of value.
>
> If you in your position as a head of the *DMV,* can take a stand in my question, I would ask you to get in contact with Herr Hasse (*Blu-Wannsee, Am Sandwerder* 7), with whom I have spoken about this call to Göttingen and to whom I am also sending a copy of this letter.
>
> With my best greetings and thanks
> Your very devoted
> B.L.v.d. Waerden

[10] Van der Waerden puts a smokescreen here. He informed his Dekan and Rektor about the Utrecht offer on January 4, 1943, i.e., over 14 months earlier. Moreover, on February 25, 1944, or18 days prior to this letter, Van der Waerden informed his Dekan, Rektor, and the Minister of his final decision *not* to accept the Utrecht offer and stay in Nazi Germany until the end of the war.



sehr grossen Wert beilegt. Ich selbst würde
das sehr bedauern. Denn ich habe das Gefühl
immer meine besten Kräfte für Deutschland
und die deutsche Wissenschaft eingesetzt zu
haben. Ich habe praktisch alle meine Arbeiten
und Bücher in deutscher Sprache geschrieben,
ich habe den wesentlichen Teil meiner Ma-
thematik in Deutschland gelernt und wieder
gelehrt; ich habe eine deutsche Frau ge-
heiratet und meine Kinder rein deutsch
erzogen.

B. L. van der Waerden's "Germanness," a facsimile of a fragment of the letter to Süss. (Courtesy of ETH)

This letter suggests that perhaps Van der Waerden does not perceive himself as Dutch any more but instead belongs to the German culture with all his heart and soul, with his "best energies for Germany" applied "to the German Science," with writing practically all his "works and books in the German language," with teaching "a major portion" of his "mathematics in Germany," with "a German wife," and with "children raised pure Germans."

His son, Hans van der Waerden, shows a great insight in his comments about this transformation in his September 10, 2010, letter to me [WaH2]:

> Another of your key documents is my father's declaration of his *attachment to Germany* (German mathematics, German culture, Germany as a whole). Indeed, by this time, it seems that my father, without becoming a nationalist like Heisenberg, had come to feel like a German citizen, losing much of his attachment to his Dutch origin. Becoming something like an average non-fascist German, his feelings in the years 1943/44, when the outcome of the war was uncertain – and he was pondering over the Utrecht offer – might be summarized as follows: "Let us be patient, things will change, the war will be over some day, maybe by some treaty acceptable to both sides, when they are sufficiently exhausted and disgusted by mutual mass-murdering; and probably after some serious defeats this horrible Nazi regime will be overthrown and Germany – my Germany – can become again a decent member of the international community." This, at least, was what thousands of intellectuals were silently hoping [for], as can be proved by numerous documents produced after the war. No reason to believe that my father differed from them.

Back to the letter; Van der Waerden asks Süss to use his influence with the Nazi authorities to help Van der Waerden materialize his Göttingen dream and in particular to contact the other helper Helmut Hasse at Göttingen, to whom Van der Waerden has already written. On March 31, 1944, *Rektor* Süss promises help "not only in a personal, human sense, but as *Führer* of the *DMV*":

> Very esteemed Colleague,
>
> Your letter from 14 March, which I found waiting here yesterday after two weeks' absence, in the meantime is forcing me continually to reflect a good deal and is giving me a lot to think about. At least I would like to express this right away, so you do not believe that I have little regard for your concerns or do not feel them myself. Fundamentally I can assure you now that I will try to help you in the limited way that is possible for me to do so, not only in a personal, human sense, but as *Führer* of the *DMV*. Mr. Hasse has just written to me too about the entire matter after he spoke with you. I will need a few days to find a quiet moment I need to think through the situation before I dare to say anything more precise.

Five weeks later, on May 19, 1944, Süss comes again:

> Very esteemed, dear colleague,
>
> Weeks ago I gave a brief answer to your letter from the middle of March. In the meantime I have repeatedly thought things over and, also prompted by a letter from Mr. Hasse and other considerations, have had a cause to reflect about that. It would likely be best if we could speak about all the issues. This is one reason why I would like to be permitted to invite you to a lecture in our little colloquium in Freiburg. Then afterwards we could find time to consult with one another, as I have in mind.

Thus, Süss leaves specifics of his help to a personal meeting with Van der Waerden, and thus out of our historical reach. From his next letter we only learn that on Monday, July 10, 1944, Van der Waerden is to give a talk "Babylonian and Greek Algebra" at Süss' Albert Ludwig University of Freiburg.

What about Helmut Hasse, who corresponded with both Van der Waerden and Süss regarding the Dream of Göttingen? I have been able to find two of his letters to Van der Waerden. In the letters, the sender is stamped as "Korv-Kap (*Korvette Kapitän*) Prof. Dr. Hasse, *Obercomando der Kriegsmarine* (The Supreme Command of the War Navy), Berlin-Wannsee."[11] During 1939–1945, Hasse has been the Commander of the department FEP III of the German Navy Ordnance (*Marinewaffenamt*). On June 23, 1944, Hasse writes to Van der Waerden on the Military Postcard with a round seal of *Obercomando der Kriegsmarine*, in a handwritten beautiful Gothic style, known as "*Sütterlin*." He offers Van der Waerden to "harness" himself in Nazi Germany's war research and has already arranged such a war research position with the people who can make it happen for Van der Waerden (underlines are Hasse's):

> Dear Herr van der Waerden,
>
> I am very happy that you have had such a tremendous success. Right away I let Dr. Fränz know by word of mouth and arranged with him that you should be given an

[11] Somebody must look into Hasse's active work on torpedoes for the Nazi War Navy.

> official research commission from the office in charge (*BHF* = *Bevollmächtigter der Hochfrequenzforschung*) [The Command of High Frequency Research]. I hope that is all right with you. To me it seems in other regards favorable for you to let yourself be "harnessed" in this way into the current research projects. I was also a while ago in Freiburg and spoke with Süss among others about you. You will hear from him how things are in G. [Göttingen]. A decisive change in the situation there has not happened since our last conversation.
> With fond regards and best wishes,
> Your H. Hasse

Thus, Nazi War Navy Captain Hasse from the Supreme Command of the War Navy has arranged a Nazi military research position for Van der Waerden. In his "Defense" after the war, Van der Waerden will write that he has never taken part in military research – and I trust him. However, a Nazi war-related job has been created for him by his Nazi helper Hasse. "A strong opponent of the Nazi regime" ought not to ask the Nazis for favors.

Van der Waerden deserved a professorship at Göttingen, but his Dream of Göttingen never materialized. His friend Werner Heisenberg, who did not particularly dream of Göttingen, easily landed there after the war and the six-month Farm Hall detention. Nobel Prize has its privileges.

The end of the war finds the Van der Waerden family – Bartel, Camilla, and their children – in the Austrian countryside at Tauplitz, near Graz, in the house of Camilla's mother [Dol1]. Bartel does not wish to return to Leipzig; we will discover his reason later. He and his family allow the American liberators to transport them, as displaced persons, from Austria to Holland, where Bartel thinks he still has that job offer from Utrecht University. After all, in the two and a half years of Utrecht's courting him, he has never said "no" – to them! Let us follow Van der Waerden and his family to Holland.

Chapter 41
In Search of Van der Waerden: Amsterdam, Year 1945

41.1 Home, Bittersweet Home

Following the war's last "three months, distant from all culture and barbarism"[1] in the Austrian Alps, the Van der Waerdens are liberated by the American Armed Forces. Bartel is not thrilled about the hardships of their liberation, as he describes it on July 1, 1945, in a letter to Otto Neugebauer[2] from the camp for displaced persons at the town of Sittard in the southernmost Dutch province of Limburg[3]:

> When the Americans had liberated us, we were like cows pushed together in cattle wagons and transported to Holland, my wife, 3 children and I. The transport lasted 16 days, it was horrible. The children were of course sick but then recovered here in the camp.

Months later, in November 1945, Van der Waerden is still angry at the Americans, whose "friendly offer" turned into a distasteful experience, as he writes to Richard Courant of New York[4]:

> ~~When the Americans came, and we were given a friendly offer to get a direct trip to Holland, the misery began. Three weeks we spent in hard freight cars [*Güterwagen*] and in dirty unsanitary camps with poorly prepared and hard to digest food.~~[5]

[1] Van der Waerden, July 1, 1945, letter in German to Otto Neugebauer; Library of Congress, Manuscript Division; possibly from the Veblen Papers.

[2] Otto E. Neugebauer (1899–1990), a historian of mathematics, an anti-Nazi, the founder of *Zentralblatt für Mathematik* (1931) and of *Mathematical Reviews* (1940).

[3] Van der Waerden, July 1, 1945, letter to Otto Neugebauer; Library of Congress, Manuscript Division; possibly from the Veblen Papers.

[4] November 11, 1945, Van der Waerden's letter in German to Richard Courant; ETH, Hs 652: 10649 (unfinished and unsent, 2 pages survive). The complete 3-page letter was sent on November 20, 1945; New York University Archives, Courant Papers.

[5] Throughout the book, strikethrough text represents words lightly crossed out in the original as if for the purpose to remain easily readable by Van der Waerden and consequently by us.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_41

Van der Waerden knows, however, that by comparison with many other survivors, he has done all right, or perhaps, he does not wish to appear as a whiner to his friend Richard Courant, and so he crosses out the above description and replaces it with a moderated one:

> The repatriation was less than attractive. Three weeks in freight wagons and camps, but of course one can survive that.

On July 1, 1945, Bartel van der Waerden is about to become a free man. He expects to get a ride from the camp to Laren very soon, for in writing on that day from the Sittard camp to his American colleagues Lefschetz, Veblen, and Neugebauer, he gives the *Breidablik* return address. Indeed, *Breidablik* is ready to provide the roof over the heads of Bartel and Camilla and their children. In a few days, the Van der Waerdens make it to this magnificent house. Now they need to find bread for their table.

The Van der Waerdens have had it much easier in Germany during the war than the people of the occupied Netherlands. After the five years of occupation and a devastating last winter, the so-called *Hongerwinter* ("The Hunger Winter"), when some 30,000 people died of starvation and malnutrition, life in the Netherlands immediately after the war is no bed of roses. Bartel assesses it on July 1, 1945:

> Holland is freed from oppression, but it is – like Germany and Austria – in a desolate state. Food supply is sufficient, but all other necessities of life are lacking.

Postwar life in Holland must have been even harsher on the Van der Waerdens, who arrived in Holland with practically nothing. Even half a year later, they are so short of bare necessities that Bartel has to step on his (considerable) pride and on December 29, 1945, ask Richard Courant in New York for help:

> I thank you very much for sending me the two volumes of Courant–Hilbert. Your kindness gives me courage to utter another wish. We are so short of underwear and warm clothes for the children. Helga is 15, Ilse 11, Hans 8 years old. My father's house is extremely cold. Perhaps your wife has got some wool or things the children don't wear anymore? They can be as old and ugly as they may: my wife can change nearly anything into anything. And further: Would it be possible to send a sheet (for a bed)? We have only 4 sheets for 5 beds, and it is quite impossible to get any here.
>
> I hope that you and your wife will not be angry with me for asking so much. If it is difficult for you, or if your people need the things more than I, please don't send anything.

Bartel gets help from his large family. His numerous aunts send him apples and things. On December 29, 1945, the younger brother of Bart's father, Uncle Herman van der Waerden, offers to make shoes for Bart's son Hans, who without shoes cannot even go out. Hans van der Waerden responds to my question about his postwar years in Holland [WaH1]:

> Concerning my life as a boy in Laren, which is within the period you are interested in, is the only time I clearly remember. For me, far away from the burden of political past, it was a wonderful time, that makes me feel homesick ever since, as soon as I cross the border to the Netherlands or hear someone talk my beloved childhood language.

41.2 The New World or Old?

> *I do not mind his remaining a German Professor until the end –*
> *I do mind his remaining a German Professor at the beginning!*
>
> – Otto Neugebauer

After the war, Van der Waerden could have returned to Leipzig University. There he would have been given a hero's welcome, for he stayed with Germany to the end of the war. Why did he not return to Leipzig?

This question occupied me for many years, until unexpectedly I found the answer in Van der Waerden's letter to the new Princeton mathematics chair Solomon Lefschetz. Even Lefschetz never learned the answer, for it was contained *only* in the handwritten copy Van der Waerden kept to himself, in which the answer was written and then lightly crossed out so that Van der Waerden – and consequently I – can read it! I learn here – and nowhere else – that Van der Waerden does not wish to go back to Leipzig because Leipzig is now in the Soviet zone of occupation, and he has no desire to live under the Russian rule. As someone who has lived under the Soviet rule, revoked Soviet citizenship, and started life all over as a refugee in America, I can relate to Van der Waerden's – and his friend Heisenberg's – distaste for the Russian tyranny. However, was the Nazi tyranny, which they both accepted, any better?

Van der Waerden does not wish to stay in Holland, Austria, or Germany due to their "desolate state." He believes he could get a position in Holland, likely referring to his old never accepted Utrecht's offer but prefers to come to America. Unlike in 1933, Van der Waerden is now very interested in Princeton, for he writes this letter in English to Lefschetz right upon his return to Holland, while still in the Sittard camp for displaced persons, on July 1, 1945[6]:

> Dear Professor Lefschetz!
>
> Peace at last, thank God! By the help of our mighty allies, Holland is freed from oppression, but it is – like Germany and Austria – in a desolate state. Food supply is sufficient, but all other necessities of life are lacking: not even railways are going. Scientific work and international contact are practically impossible.
>
> In March, my home in Leipzig being destroyed by bombs, I could escape with my family from the bomb hell to Austria. From there we have just been repatriated to Holland. ~~Returning to Leipzig, which belongs now to the Russian zone of occupation, seems impossible and, even if possible, not advisable.~~ I can get a position in Holland ~~probably~~ but Holland is in a heavy political and economic crisis, as I said before. For all these reasons I should like to go ~~temporarily or definitively~~ to America.

In particular, Van der Waerden wishes to be invited to Princeton again:

> Several years ago, you encouraged me to write to you if I wanted to be invited to America. In the year 1939 [actually in 1933] I was invited to come to Princeton as a guest for half a year. Do you think that this invitation could be repeated? I should enjoy

---

[6] Van der Waerden to Lefschetz, July 1, 1945; handwritten letter in English; ETH, Hs 652:11346.

very much getting into contact with the American mathematicians again, especially with those of Princeton. I shall accept with joy any invitation of this kind…

With best greetings to Veblen, [von] Neumann and the other Princetonians.
Yours very sincerely
B.L.v.d. Waerden

The same day, July 1, 1945, Van der Waerden writes a nearly identical letter to Oswald Veblen at the Institute for Advanced Study Princeton. The only difference is in the justification for the desire to come to America: in addition to "a desolate state" of Holland, Germany, and Austria, Van der Waerden pays a high praise to mathematics in the United States:

I have been cut off from international mathematics, whose heart pulses in America, for five years, and I want to regain contact as soon as possible.

The third July 1, 1945, letter Van der Waerden sends to Otto Neugebauer. The first reply, the August 20, 1945, letter from Lefschetz, is not promising:

Dear Dr. van der Waerden:

Your letter of July 1st reached me in due time. I was very sorry to hear about your losing your home in Leipzig and can well understand your desire to come to the United States (who does not feel the same way in Europe just now?). However, we are in a complete state of flux here and the time does not seem very propitious for bringing in scientists from the outside, especially professors in former German universities. I have transmitted copies of your letter to some mathematicians that know you, in particular to the members of the Institute for Advanced Study, for the pre-war invitation that you mention can only have come from them. They have informed me that there is nothing available at the present time. One of them did express the hope that you would accept the position at Utrecht since, no doubt, you are very badly needed there. I confess that I agree a little bit with him.

Yours sincerely,
S. Lefschetz

Van der Waerden could not have found Lefschetz's letter encouraging. No doubt he senses a thinly concealed irony behind Lefschetz's rhetorical question: "Who does not feel the same way in Europe just now?" Lefschetz is even blunter when he acknowledges that the time is not "very propitious for bringing in scientists from the outside, especially professors in former German universities." Lefschetz seems to imply that Van der Waerden made a wrong choice by staying in Nazi Germany and now has to pay the price for being on the wrong side of the divide during the war. In Lefschetz's "defense," one should note that he treated sarcastically the vast majority of humans around him.

I must add that in his reply Lefschetz is factually wrong: not only did the 1933 invitation come from Princeton University and not from the Institute for Advanced Study, but Lefschetz himself attended the meeting of the Princeton's Research Committee that decided to invite Van der Waerden.

A few months later Princeton starts looking for an algebraist, but Lefschetz does not even inform Van der Waerden, for he has someone else in mind; he is willing to even curb his usual sarcasm and charm that someone. On Wednesday, October 17, 1945, Lefschetz writes to the algebraist of his choice, who at that time is at Indiana University, Bloomington[7]:

> Dear Artin,
>
> Owing to recent losses in our department, to which now must be added Wedderburn's retirement (soon to be official), I feel very strongly that we should add a major scientist to our staff. You are the first person of whom I thought in this connection and, if possible, I would just as soon not go further in my search. Your achievements as a mathematician, together with your well-known sympathetic influence on the younger men, do indeed make you the man of the hour.

Two days later, on October 21, 1945, Emil Artin happily responds[8]:

> Dear Lefschetz:
>
> It is with very great joy that I received your letter and I feel deeply honored that you are thinking of me. I would not be a mathematician if I would not feel greatly interested and attracted by a chance to go to Princeton. Princeton is now after all the center of all mathematics.

As if especially for the sake of my book, Artin then asks:

> How did the case of Van der Waerden go on after his letter? I am here so isolated that I get the news only after long detours. I[s] something specific known of the German mathematicians?

Artin's question shows that Lefschetz widely circulated Van der Waerden's July 1, 1945, letter asking for a Princeton job, likely together with Lefschetz's sarcastic reply. On October 27, 1945, Lefschetz informs Artin that Van der Waerden has not been invited to Princeton[9]:

> Nothing has been done regarding Van der Waerden – nothing, at least from this side.

Surprisingly, Lefschetz then shows knowledge of the *secret detention* in Farm Hall, England, of Heisenberg and other leading German physicists, who during the war were involved in research on atomic bomb and reactor:

> We have no information about German mathematicians whatsoever. I did learn two days ago that Heisenberg and all the nuclear physicists are being detained though well treated. Some more of "*maladie du siècle*" [disease of the century].

Before replies from America could arrive, Van der Waerden writes two letters to his good friend Heinz Hopf, a (Jewish) German mathematician, now a Swiss citizen and professor at the ETH in Zurich. I have been unable to locate these letters, but according to Hopf's August 3, 1945, reply, they were written on July 19 and 21, 1945. Hopf opens his letter with praising Switzerland and its neutrality:

> Here in Switzerland one is of course less fanatical, exactly this in my opinion, a particularly important and fortunate consequence of our neutrality...

---

[7] Typed letter in English; Personnel File of Emil Artin, Princeton University.

[8] Handwritten letter in English; Personnel File of Emil Artin, Princeton University.

[9] Typed letter in English, Personnel File of Emil Artin, Princeton University.

At the end of the letter, Hopf adds his personal tribute to Switzerland:

> My wife and I are doing fine… we are happy that we are Swiss.

It is plausible that this praise of the Swiss neutrality and Hopf's happiness with Swiss citizenship plant in Van der Waerden a seed of interest in living in Switzerland. Hopf is unhappy that the Swiss consider – as they should in my opinion – "Hitlerism" to be a part of the German culture:

> I beg you, by the way, not to misunderstand the above comment about neutrality, the open opinions here are completely unified against Germany, the bitterness about the Nazis is gigantic, but the boundaries between Hitlerism and the German culture are not always observed here either.

As the author of this narrative, I am compelled to ask: Professor Hopf, and what are the alleged boundaries? Wasn't Nazism (to a great regret of many) a product and part of the German culture every bit as Marxism or music of J.S. Bach and Beethoven were? Of course, there is high culture and low culture, but both of them are parts of culture in a broader sense of the word, and who can – or should – split them apart?

Hopf understands the liability of Van der Waerden's spending the entire Nazi era in Germany and offers Van der Waerden a line of defense:

> One would perhaps argue this way: he has worked as a professor in Germany even during a period of abuse of his homeland by Germany because he believed that he could thus contribute somewhat to the saving of the culture in Europe; we respect that; but he must be consistent and extend this attempt to salvaging culture in Germany. I believe it would be very difficult to argue against this argument.

Finally, Hopf scolds the Dutch for not immediately jumping on the opportunity to hire Van der Waerden:

> When the Dutch, whom you can approach with clean conscience and offer them your services, do not want you, then in my opinion they hurt themselves, and that is their business. I consider it certain that in a few years, when the waves calm down a bit, somewhere in the world you will work again in the profession – assuming naturally that you with your family can economically survive until then, which I am not sure about.

Van der Waerden will quote these lines to the Dutch almost immediately. In 1945, Switzerland did not allow even a brief visit to the former Nazi Germany Professor Van der Waerden. As we will see, the Swiss will drop their "neutrality façade" the very next year.

Sometime in July–August 1945, Hopf writes about Van der Waerden's plight to his friend and famous German historian of mathematics Otto Neugebauer, who now lives in the United States and edits *Mathematical Reviews* that he created in 1940 after Springer-Verlag put pressure on Neugebauer to Nazify *Zentralblatt für Mathematik*. On August 15, 1945, Neugebauer replies to Heinz Hopf in English as follows:

> I have heard directly from Van der Waerden. I do not mind his remaining a German Professor until the end – I do mind his remaining a German Professor at the beginning! However, I feel very differently than the Lord and [thus] I do not intend to do anything positive or negative.

On November 11, 1945, Van der Waerden writes to his mentor and friend Richard Courant in New York about the bombings of the late months of the war, his tough repatriation, and his new job at *Royal Dutch Oil*, also known as *Royal Dutch Shell*, or simply *Shell*. On December 13, 1945, Courant sends a guarded reply in English. Before deciding whether to renew their old friendship, Courant desires to know why Van der Waerden has chosen to stay in Nazi Germany:

> I wish very much that there were an opportunity of talking to you personally and for that matter to other old friends who have been in Germany during the war. Of course, so much has happened in the meantime that in many cases much will have to be explained before one can resume where one left off. Your friends in America, for example, could not understand why you as a Dutchman chose to stay with the Nazis.

Moreover, Courant makes his request for an explanation public: at the top of the letter, I see a handwritten inscription:

> cc. sent to: Reinhold Baer, U. of Ill. Urbana
> Herman Weyl – Inst. for Advanced Study Princeton
> Veblen

Courant's papers include both Van der Waerden's November 20, 1945, handwritten letter and its typewritten copy, which suggests that Courant had it typed and copies sent to the same addresses as his reply. As Lefschetz before him, Courant too apparently believes that Van der Waerden made the wrong choice. On December 20, 2004, I had an opportunity to ask over the phone Ernest Courant, the elder son of Richard Courant and a prominent nuclear physicist in his own rights, a natural question: "What did your father think about Van der Waerden?" He replied as follows, as I jotted down his words:

> He [Richard Courant] considered him [Van der Waerden] a great mathematician and was a bit critical of him for being perhaps too comfortable in Nazi Germany.

Thus, America and Switzerland have to wait. Beggars could not be choosers, and so Dr. Van der Waerden is now – finally – willing to seriously entertain a professorship in his "desolate" (his word) Homeland. Van der Waerden is up for big surprises, as we will see in the next few sections. He has returned to his homeland as if an alien, not understanding the psyche and the mood of the Dutch people, who experienced horrific five years of occupation. As the historian Louis de Jong sums up [Jon],

> The Germans succeeded by and large in exploiting the economic potential of the Netherlands, and they succeeded in deporting most of the country's Jews.

I should add, some 80% of the Dutch Jews did not survive the war and the Holocaust. De Jong continues,

> Their [Germans'] attempt at Nazification, however, failed miserably,[10] and they were totally unable to prevent the growth of a flourishing underground movement, whose

[10] De Jong specifies elsewhere in his book (p. 33), "The Dutch Nazi movement never won the support of more than 1.5 percent of the Dutch population."

three main achievements were to keep up people's morale (principally through the underground press); to care for some hundreds of thousands [!] of people who were living in hiding; and to provide the Allies with vital military information . . .

Nations of heroes do not exist. But there were among the Dutch tens of thousands of ordinary human beings, men and women, who did save the country's soul.

41.3 "The Defense"

Some of the stories are difficult to believe. Part of all this is the way people always talk about their past. The reasons they give for their behaviour in the past may be just inventions, colored by how history took its course.

– Nicolaas G. de Bruijn

Van der Waerden expects that the Utrecht chair, first offered to him in December of 1942, is still waiting for him. He also does not mind a chair at Amsterdam. However, following the liberation, the *Militair Gezag* (Military Authority) installed *Commissie van Herstel* at each of the five Dutch universities, which gradually became known as *College van Herstel* (Recovery Board, or Restoration Board), formed to advise the Military Authority on how to act against collaborators and other pro-German professors and staff members, and when the university could be reopened. It was expected that all suspect staff would be removed in a few months' time. In fact, the removal took much longer. I am grateful to Dr. Peter Jan Knegtmans, *The University Historian* at the University of Amsterdam, for the information on *College van Herstel* and the workings of the City of Amsterdam, contained in his e-mails [Kne4] and [Kne5] to me. The Dutch postwar educational and governmental systems were a "jungle," and it has been invaluable to have such a uniquely qualified jungle guide.

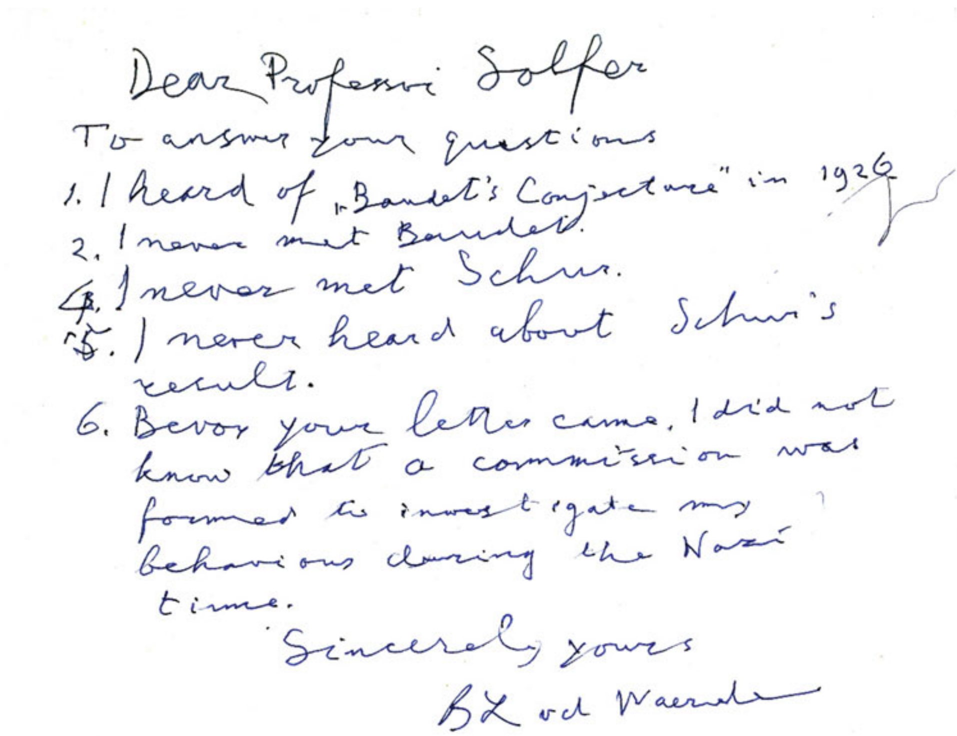
Utrecht University's *College van Herstel en Zuivering* (Board of Recovery and Purification), as it was called there, was installed on 18 June 1945, while the University of Amsterdam's *College van Herstel* (Board of Recovery) was installed on June 8th, 1945. At the time the University of Amsterdam belonged to the City. Yet *B. en W.*, the Executive, consisting of the *Burgemeester en Wethouders* (mayor and at the time 6 aldermen), could not appoint professors; only the city council that numbered 45 could appoint them. Moreover, an appointment of a professor needed a Royal assent. The Queen could not give her assent if the government did not submit to her a request for assent. On the other hand, the government would not submit a request for assent if there was even a slight chance that the Queen would refuse it, as she had a few times during those postwar years.

Originally Dutch, Professor of History of Mathematics at the Massachusetts Institute of Technology Dirk J. Struik (1894–2000) maintained close ties with the leading Dutch colleagues and based the following 1995 statement to me [Str] on a letter he had received from Jan A. Schouten in 1945–46:

Though he [Van der Waerden] stayed at Leipzig University during the Hitler days, he was able to protect Jewish and left wing students.[11] This was brought out after the war when his behavior in Leipzig was scrutinized by a commission of his peers in the Netherlands. He was entirely exonerated.

On April 12, 1995, I quoted this statement in my letter to Professor Van der Waerden and asked him to describe for me in detail this "commission of his peers," its membership and charge. On April 24, 1995, Van der Waerden mailed his reply [Wae26] (see the facsimile of his letter in this section):

Before your letter came, I did not know that a commission was formed to investigate my behaviour during the Nazi times.



Dear Professor Soifer

To answer your questions

1. I heard of "Baudet's Conjecture" in 1926
2. I never met Baudet.
4. I never met Schur.
5. I never heard about Schur's result.
6. Bevor your letter came, I did not know that a commission was formed to investigate my behaviour during the Nazi time.

Sincerely yours

BL vd Waerden

B.L. van der Waerden, April 24, 1995, letter to Alexander Soifer

[11] As we have seen, Van der Waerden spoke against firing of Leipzig's Jewish professors in May 1935 and published papers of Jewish authors in the *Annalen* until 1940. I have found no evidence of him protecting "Jewish and left wing students," and Van der Waerden never claimed it himself.

Many years later I discovered that the University of Amsterdam's *College van Herstel* (*CvH*) *did* investigate Van der Waerden, and the City executive board, *B.&W.*, wrote about Van der Waerden to *CvH*, a de-Nazification board.[12] Van der Waerden *knew* about the investigations, for on July 20, 1945, just a few weeks after he returned to Holland, he wrote in his own hand his "Defense" and forwarded it to the Amsterdam's *College van Herstel*, and also to the Utrecht's *College van Herstel en Zuivering*. This Van der Waerden's defense of his reasons for staying in Nazi Germany and his activities in the Third Reich is a most important testimony, never discussed in historical scholarship before 2004 [Soi6]. I feel compelled to include the translation of this Dutch handwritten document in its entirety, with my commentaries. You can see its facsimile in my 2015 book [Soi47].

41.3.1 Defense

Since 1931 I have been a Professor at Leipzig University. The following serves as an explanation as to why I stayed there until 1945:

1) From 1933 till 1940 I considered it to be my most important duty to help defend the European culture, and most especially science, against the culture-destroying National Socialism. That is why in 1933 I traveled to Berlin and Göttingen to protest the boycott of Landau's classes by Göttingen Nazi students. In 1934[1935] Heisenberg and I strongly protested against the dismissal of 4 Jews in a faculty meeting at Leipzig. Because of that I got a reprimand from the Saxon Government (*Untschmann*) and an admonition that as a foreigner I should not interfere in German politics. What my wife and I have personally done to help Jewish friends with their emigration is not relevant here, but what is, is that as [an] editor of the *Math. Annalen* I accepted until 1942 articles of Jews and "*Jüdische Mischlinge*" (Nazi term for people of Jewish and Aryan mixed blood), furthermore that in the *Gelbe Sammlung* [Yellow Series] of Springer which I was partially responsible for, an important work by a Jewish author appeared in 1937 (Courant-Hilbert, *Methoden der Mathematischen Physik II*), and that in 1941 I was the Ph.D. advisor of a non-Aryan. In 1936 [1935], when my esteemed teacher Emmy Noether died, I pointed out the great merits of this Jewish woman.

I could not have known in advance that all this would be like "punching a brick wall" [*vechten tegen de bierkaai*] and that the Nazis would drag the entire German culture with them into their destruction. I still hoped that the German people would finally see reason and would put an end to the gangster regime. Meanwhile my work was not altogether for nothing because my students, such as [Herbert] Seifert, Hans Richter, Wei-Liang Chow, Li En-Po, Wintgen, etc., whose dissertations were accepted in the *Math. Annalen*, have done an excellent work at Leipzig. If I had not been in Germany, these [students] would likely not have encountered the problems that I have given them.

[12] Dr Knegtmans [Kne2] refers to the April 17, 1946, letter from *B. en W.* of Amsterdam to *CvH*, *Archief Curatoren* nr 369, which says that "the [Van der Waerden's] appointment did not go through also because the Minister had told the City Council beforehand that he would not ratify it."

As Hopf advised, Van der Waerden justifies his staying in Nazi Germany by stating that it was his "most important duty to help defend the European culture, and most especially science, against the culture destroying National Socialism." However, as is evident from Hopf's reflection, many of Van der Waerden's contemporaries found it difficult to separate "German culture" from "Hitlerism." Given Van der Waerden's scruples regarding "the gangster-regime" (his words), his fellow scientists – then and now – considered his willingness to serve that regime naïve at best and hypocritical at worst. Van der Waerden continues his "Defense" with part 2, dedicated to the five years of the German occupation of Holland:

> This all may serve for closer understanding of my attitude towards the Nazis. What I should explain to the Dutch people is, however, not my actions before 1940, but those after the Netherlands had been attacked by Germany.
>
> 2) From 1940 to 1945. After the breakout of the war with the Netherlands, I was first locked up and then released on the condition that I do not leave Germany. So I was practically in the same position as those who were forced laborers in Germany.
>
> If I had given up my position, then I would have probably been forced to work in an ammunition factory.

To say that a university full professor was "in the same position as those who were forced laborers in Germany," was a dramatic exaggeration, and it likely appeared as such to the Dutch who read the Defense.

> I have never worked for the *Wehrmacht* [the German Army], I have never given a class or worked on things that could be used for military purposes.

While we have already learned from the June 23, 1944, postcard, that German War Navy *Korvette Kapitän* Prof. Dr. Helmut Hasse had arranged a war-related job for Van der Waerden in the Command of High Frequency Research, I have no reason to think that Van der Waerden accepted that war-related job.

However, Van der Waerden has taught students, many of whom may have served the *Wehrmacht* and some definitely "worked on things that could be used for military purposes." For example, Professor of History of Mathematics at Frankfurt University Moritz Epple informs us in his report on my manuscript of [Soi47] that Herbert Seifert, Ph.D. 1932 under Van der Waerden, volunteered for war work at the *Institut für Gasdynamik*, which was "a part of *Luftfahrtforschungsanstalt Hermann Göring* at Braunschweig, one of the major facilities of aviation research in Nazi Germany, built between 1936 and 1938 ... It was one of the most important places in Nazi Germany for developing knowledge about supersonic aircraft."

Besides, by working in Nazi Germany's Civil Service, Van der Waerden contributed to "the gangster regime" and lent his credibility and acclaim as a distinguished scientist to that of the Third Reich.

> In 1943[December 1942] the *Faculteit* of Physics and Mathematics at Utrecht asked me whether I would accept an appointment as a Professor there. I asked them to postpone the matter if possible until after the war, because I did not want to be appointed by the Van Dam department.

It suffices to point out here that coming home at the Utrecht *Faculty request*, even with the approval by the Nazi-collaborating Minister Jan van Dam, would have been much better for

Van der Waerden's reputation in his Homeland than continuing to serve the Third Reich to the end.

> I do not need to add to this that I have never been a member of any *NS* [National Socialist] organization or have sympathized with them, because that is self-evident for a decent thinking human being. It was commonly known in Germany that I was not a Nazi and because of that the government distrusted me and did not give me permission to go to the Volta Congress in Rome in 1939, and to give lectures in Hungary or to French prisoners of war, or to partake in the Congress of Mathematicians in Rome.

This is true, however, the Nazi government did allow Professor Van der Waerden to travel inside and outside Germany: for example, to travel to Holland in 1933, 1935, 1938, 1939, 1940, 1942, and possibly in 1944. Moreover, on some of these trips, for example in 1935 and 1939, he was accompanied by his entire family and could have remained in Holland.

> The Faculty at Munich suggested me as a successor to Carathéodory, but the party authorities declared me "*untragbar*" [intolerable], and the appointment did not happen.
>
> Also, my wife, who is Austrian, has been strongly opposed to the Nazi regime from the very beginning.
>
> Laren, N-H [North-Holland], 20 July 1945 B.L.v.d. Waerden

Indeed, in the Munich deliberations Van der Waerden was perceived as a philo-Semite, and this must have cost him the Munich job. We will attempt to gain some insight into Mrs. Camilla van der Waerden's views in the next section.

With the "Defense" submitted, Van der Waerden hoped to get a professorship at Utrecht or Amsterdam. Van der Corput was the key man to this end.

41.4 Van der Waerden and Van der Corput: A Dialog in Letters

Why would I go to Holland where the oppression became so intolerable and where every fruitful scientific research was impossible?

– Bartel L. van der Waerden

It was not at all fitting for a Dutchman to make mathematics in Germany flourish in those years when Germany was preparing for war and was kicking Jews from every position and place.

– Johannes G. van der Corput

Johannes Gualtherus van der Corput (1890–1975) was a professor of mathematics at Groningen (1923–1946) and Amsterdam (1946–1954). During the war and the German occupation of Holland, he took an active part in the Dutch underground and in 1945 spent a week in a Nazi jail for hiding people from the occupiers in his house. According to Dr. Knegtmans (June 10, 2004, e-mail to me, [Kne7]), "Van der Corput belonged to a small group of Groningen professors that had developed some ideas about the post-war university in the sense that it had to become a moral [!] community that would be able to withstand any authoritarian threat or defiance. Van der Leeuw, the first post-war Minister of Education, had belonged to the same group."

Prof. Dr. Gerardus J. van der Leeuw, Minister (1945–1946) of Education, Culture and Sciences (*Onderwijs, Kunsten en Wetenschappen*) appointed Van der Corput to be the Chair of the Committee for the Coordination and Reorganization of Higher Education in Mathematics in The Netherlands (*De Commissie tot Coördinatie van het Hooger Onderwijs in de Wiskunde in Nederland*). The Committee became known as "The Van der Corput Committee." In 1946, Van der Corput will become one of the founders and the first director of the *Mathematisch Centrum* (Mathematics Center) in Amsterdam.

Van der Corput knew Van der Waerden from their 1928–1931 years working together at Groningen, where young Bartel learned quite a bit of mathematics from him [Dol1]. Van der Corput hosted Van der Waerden's October 10–14, 1938, visit for giving talks at Groningen University. The colleagues corresponded even during the war and the German occupation of Holland. In early 1944, Van der Corput recommended the book about the history of sciences in antiquity, which Van der Waerden had been writing, to the Dutch publisher J. Noorduijn en Zoon N.V. – Gorinchem. Eventually, in 1954, this book was published in Dutch and in 1961 in English in an expanded beautiful edition as *Science Awakening* [Wae15].

Right after the war, the friends lived in an absolute sense not far from each other, Van der Corput in Groningen, and Van der Waerden in Laren near Amsterdam, but on the Dutch scale the trip from Laren (Amsterdam) to Groningen was a major journey. And so, to our good historical fortune, their preferred means of communication were letters. Van der Waerden saved handwritten copies of his own letters (the first plain paper copier, Xerox 914, was invented only in 1959!) and Van der Corput's original letters; they are now preserved at the ETH Archive in Zurich.

A voluminous file of their 1945 correspondence, lying in front of me as I am writing these lines, is an invaluable resource for understanding their views on moral standards of scholars during the Nazi era and the occupation of Holland, and, more generally, eternal moral dilemmas posed by the war and its aftermath. I will let the correspondents do most of the talking. A number of different handwritten versions of some of these letters exists. Some copies were sent to third parties, such as Van der Waerden's close friend and fellow mathematician Hans Freudenthal (1905–1990). All this indicates that Van der Waerden took this exchange extremely seriously, as did Van der Corput.

On July 29, 1945, Van der Corput sends Van der Waerden a letter in which he conveys his new leading role in the mathematical higher education of the Netherlands:

> I have been appointed chairman of a commission to reorganize higher education in mathematics in the Netherlands, which will have as its primary duty to offer advice for the filling of vacancies in mathematics.

Van der Corput realizes that his new authority to advise Minister van der Leeuw, calls for a new responsibility, and so he continues with probing questions:

> Your letter made me do a lot of thinking. I never understood why you stayed in Germany between 1933 and 1940, and also why after 10 May 1940 [the day Nazi Germany attacked the Netherlands] you did not return to the Netherlands as so many succeeded in doing, if need be to go into hiding here ["some hundreds of thousands of people . . . were living in hiding"[Jon]]. Rumors went around about you that you were

not on our side anymore, at least not entirely. That could have been slander. I would find it important if you could explain to me the situation completely and in all honesty.

Van der Corput concludes by sharing his own resistance activities:

People were in hiding in my house throughout the entire war, 23 in total, of which 5 were Jews; I was a representative at Groningen of the Professors Resistance Group. When I was arrested in February 1945, they found two people in hiding in my house, of which one was Jewish. I was suffering from angina and was released from prison after a week. My house and all my furniture were impounded [by the authorities] but we moved back on the day of liberation . . . I was on the Board of *Vrij-Nederland* [Free Netherlands][13] and was arrested for disseminating illegal literature.

Van der Waerden replies on July 31, 1945. He expresses delight with his friend being in charge of all Dutch university appointments in mathematics, including Van der Waerden's own appointment – perhaps, too much of a delight – but then, understandably, carefully crosses most of the delight out:

I am very happy to be able to direct my defense to the right address against the things that have been blamed on me completely unexpectedly from all sides. ~~So you are chairman of the commission which will decide on the future occupation of the professorships of mathematics, perfect! An illegal work of the highest order and what is more, benefitting me. Delightful!~~

From the following lines, we discover how the writing of the "Defense" has come about. We also learn that Van der Waerden has attached a copy of the "Defense" to this letter:

When I spoke with Freudenthal about it [professorship at Amsterdam] and told him that I was looking forward to possible collaboration with him, he firstly pointed out the difficulties, especially from students' circles, that could be expected, and for the aspersions that would be cast upon me because of my stay in Germany after 1933. He advised me to write down my defense [!], which I had presented to him verbally. I have done it, and after conversations with others, I have added a few more things . . . In this situation you now come forward and ask for my justification. *Voila*! I hereby include a copy of the piece.

Van der Waerden then explains why he did not return to the Netherlands when Nazi Germany waged an unprovoked war against his Homeland:

I truly did not come to the idea of returning to the Netherlands after 1940 and going into hiding here. At the end of 1942 I had come to Holland and spoke with all sorts of people (honestly no *NSB*-ers[14] because those do not belong to my circle of friends) but there was nobody who gave me [such] advice; the concept of going into hiding, furthermore, did not exist at that time.

[13] *Vrij Nederland*, an underground newspaper.

[14] *Het Nationaal Socialistische Beweging* (National Socialist Movement, a Nazi party in the Netherlands).

Van der Waerden is incorrect when he alleges that he "truly did not come to the idea to return to the Netherlands." Starting in December 1942, he had discussed the idea of coming back to a professorship at Utrecht with Barrau and Carathéodory. However, most troubling is the next statement: "the concept of going into hiding, furthermore, did not exist at that time [end of 1942]." In fact, hiding commenced immediately after the invasion of Holland in May 1940, and hundreds of thousands Dutch people, wanted by the Nazis, went into hiding. Van der Waerden knew about it very well at least since his late 1942 visit of Holland and wrote about it to Hecke on 6 April 1943: "Maybe he [Blumenthal] is in hiding like thousands of others." Van der Waerden then spells out what could be *the real reason why he did not wish to come home to Holland during the war*:

> Why would I go to Holland where the oppression became so intolerable and where every fruitful scientific research was impossible?

These words make me think that Van der Waerden has never seriously considered going back to Holland during the German occupation of his Homeland. It seems to me that Van der Waerden feels no responsibility for the "intolerable oppression" that his new country, Nazi Germany, imposed on his Homeland. In a statement that Van der Corput must have found particularly disingenuous, Van der Waerden claims that his "struggle" for the German culture and science has been as noble as Van der Corput's underground activities in Holland, and it is the people in Holland who are guilty of not understanding his "struggle" "against the Nazis":

> For your struggle of which I have heard with great delay and only in part, I had great admiration and undivided sympathy, but I could not partake in it from that distance, because I did not have enough contact with you. Since 1933, I waged another struggle, together with other reasonable people such as Hecke, Cara[théodory], and Perron against the Nazis and for the defense of culture and sciences. That I was on the good side of that struggle was, as I thought, universally known. I did not expect that people here in Holland would have so little understanding of it.

Van der Corput is unhappy with some of the answers, He shows Van der Waerden's letter to some of his trusted colleagues, Marcel Gilles Jozef Minnaert (1893–1970), Professor of Astronomy at Utrecht University, and Balthasar van der Pol (1889–1959), Professor of Theoretical Electricity at the Technical University of Delft. Finally, on August 20, 1945, Van der Corput makes his displeasure known to Van der Waerden and asks him a key question, whether Van der Waerden is demanding a full and unconditional exoneration or is pleading difficult circumstances:

> Your letter has not completely satisfied me. You complain that we here in Holland lack sufficient understanding of your troubles, but after reading your letter I wonder whether you have a sufficient understanding of troubles which we had to deal with here and of what was to be expected of a Dutchman in these years. It is not clear to me from your letter whether you consider your attitude in the past faultless or whether you plead mitigating circumstances.

Van der Corput refuses to condone Van der Waerden's actions during the war, comparing them unfavorably to his own unambiguous rejection of Nazism from the beginning of Hitler's rein:

> Concerning me personally, in January 1939, I turned down [Erich] Hecke's invitation, passed on to me by [Harald] Bohr, to give one or more lectures, because I refused to come to Germany as long as Hitler was in power. Consequently, I have not been in Germany after 1932. In connection with this position of mine that was shared by many of us, I do not understand how you can so easily gloss over those years between 1933 and 1939. Indeed, it was not at all fitting for a Dutchman to make mathematics in Germany flourish in those years when Germany was preparing for war and was kicking Jews from every position and place.

These are powerful words, let us read them again: "*It was not at all fitting for a Dutchman to make mathematics in Germany flourish in those years when Germany was preparing for war and was kicking Jews from every position and place*." Van der Corput then cites the 1939 incident that, apparently, still bothers him and directly asks whether Van der Waerden and his wife were Nazi sympathizers:

> Furthermore, I remember that after a lecture at Groningen, in the *Doelenkelder*[15] you spoke with appreciation of the regime in Germany, and more especially of Göring,[16] upon which I advised you better to stop this because this was not well received by the students of Groningen. I have to add that I do not know whether or not you were being serious at that time, but it made a strange impression on us, who considered Hitler a grave danger for humanity. Furthermore, I was informed from various sides that your wife was pro-Hitler, and that when she was supposed to come to stay in Holland, she even stated as a condition that no bad could be spoken about Adolf. I say this because you write that your wife was always against the regime. It is better that these things are discussed in the open, because then you can defend yourself.

In spite of his serious reservations, Van der Corput clearly wants to help Van der Waerden and by doing so help Dutch mathematics:

> I myself think that the Netherlands should care for its intellect and especially one like yours. I have always regretted that you went to Germany, and I will look forward to it if you can be won back completely for the Netherlands . . .
>
> I would want nothing better than for everything to be all right. Because there isno Dutch mathematician with whom I would like working more than with you. I would find it fantastic if we could work on mathematics at the same university again. Then, I think, we could found a mathematical center.

Van der Corput holds significant power and appropriately assumes a commensurate responsibility, and this is the reason for his asking these tough questions:

> I hope that you will not just excuse me for these questions but understand them. Before the government can appoint someone, it will conduct a very detailed investigation, and it is to be expected that it will also ask for my advice. It is therefore necessary for me to be well informed.

---

[15] The steakhouse *De Doelenkelder* still exists in Groningen: call 050-3189586 for reservations!

[16] Hermann Göring, Commander-in-Chief of *Luftwaffe* (German Air Force), President of the Reichstag, Prime Minister of Prussia, and Hitler's designated successor.

Perhaps to Van der Corput's surprise, Van der Waerden remains nonchalant in his immediate four-page reply. He proudly asserts his complete innocence and "demands a complete exoneration." Van der Waerden then quotes the letter he received from Hopf just about two weeks prior, in which Hopf blames the Dutch for conducting the de-Nazification of the Netherlands:

> You ask whether I want to plead mitigating circumstances. Absolutely not! I demand a complete exoneration because I do not think that I can be blamed for anything. And I am also convinced that when my case ~~now or after a few years when the understandable commotion and confusion caused by the German terror has calmed down~~ is looked at objectively, that this exoneration will be given me. This conviction I shared with Hopf at Zurich who (following a conversation with Kloosterman about me) writes: "When the Dutch, whom you can approach with clean conscience and offer them your services, do not want you, then in my opinion they hurt themselves, and that is their business. ~~I consider it certain that in a few years, when the waves have calmed down a bit, somewhere in the world you will work again in the profession.~~"
>
> ~~Also the English and the Americans, and above all the Russians, make a distinction between the Nazis, whom they want to destroy, and the German culture, which they want to help resuscitate. Should we not try to make this objective way of judgment acceptable also in the Netherlands again?~~[17]

Van der Waerden continues by presenting, again, his (and Hopf's) opinion that one must differentiate between "the Hitler regime" and "the German culture":

> Your most important accusation, I assume, is the words "It was not at all fitting for a Dutchman to make mathematics in Germany flourish in those years when Germany was preparing for war and was kicking Jews from every position and place."
>
> In this sentence two things are identified with each other that I see as the strongest opposites: the Hitler regime and the German culture. What was preparing for the war and was throwing out the Jews was the Hitler regime; what I was trying to make flourish or rather to protect against annihilation was the German culture. I considered and still consider this culture to be a thing of value, something that must be protected against destruction as much as possible, and Hitler to be the worst enemy of that culture. Science is international, but there are such things as nerve cells and cell nuclei in science from which impulses are emitted, that cannot be cut out without damage to the whole. And I mean that this standpoint is principally defensible even for a Dutchman, and I should not be in the least ashamed for having taken this position.
>
> Of course, it is understandable that people here in Holland today do not want to know, to see a difference between the Nazis and Germany or the German culture. Germany attacked the Netherlands and shamefully abused it, and the whole German people are also responsible for that. For the duration of the war this position is completely true, but one must not use this as measure to assess events that happened before the war.

[17] This paragraph is thinly crossed out in this version but was not crossed out in another, unfinished version in my possession.

> By the way, nobody at the time thought to condemn my actions. In 1934 or 1935 the Dutch Government itself officially allowed me to continue my activities in Leipzig. The student organization invited me in 1938 for a series of talks, among other places at Groningen, a certain Van der Corput asked me in 1943 to write a book for his "*Wetenschappelijke Reeks*" [Scientific Series], and I could name a lot of other things like that.

As we know, the lives and fates of the two Dutch scholars, Debye and Van der Waerden, have diverged. In January 1940, Debye leaves Germany for the United States, while Van der Waerden stays in Germany to the end of the Nazi era. Van der Waerden explains to Van der Corput:

> Debye too stayed in Germany until the end of 1939, when the Germans gave him a choice: either leave or assume the leadership of war research (*Kriegsforschungen*). Had they given me this choice I would have left.

Would Van der Waerden have left? Why would a Nazi ultimatum be necessary for a Dutch citizen, Van der Waerden, to leave the Third Reich? Couldn't he have simply accepted the Utrecht job when it was first offered or not returned back to Nazi Germany from one of his many visits of Holland, including 1935 and 1939 visits with his whole family? Van der Waerden then explains his complimentary statement about Herman Göring made during his 1939 visit of the Netherlands:

> This is what concerns the official part of the affair. Now the personal part. You seem to remember that I spoke appreciatively in the *Doelenkelder* about the regime in Germany and more specifically about Göring. You must therefore consider me as somebody without an elementary sense of right and wrong; because Göring is, as everybody knows, a clever crook, whose henchmen burned the Reichstag and who used that to abolish socialist parties. An unprecedented deception of the people that was used to destroy the democracy and the parties to which I, because of tradition, friendship, and because of my own father, was connected. And I would have defended that criminal? And moreover, the Hitler regime? And now I would twist around like a weathervane and contend that I was always against Hitler? In other words, that makes me a deceiver, a cunning liar! Nevertheless, you always willingly offer me your mediation, not only with words but also with deeds, with Noordhoff, present my *defense* to Minnaert, and write that you do not like to work with anybody more than with me. I do not understand that attitude. Or rather I can only give one explanation to it, namely that deep in your innermost a voice tells you: no, I know that man from before as decent and truth-loving, let me give him an opportunity to defend himself.
>
> Well, I can guarantee you that what you write about the *Doelenkelder* must be a misunderstanding. I have never uttered a word of defense of the Nazi regime to anybody. The question which we spoke about in the *Doelenkelder* was, if I am not mistaken, not whether this regime was defensible, but how can people cope in Germany in spite of this regime. How is science under these circumstances possible? Then I may have mentioned a few facts from which it was apparent that at Leipzig especially and more importantly in mathematics, the pressure from above was not as claustrophobic as people imagined it here. I may have mentioned in connection with something or other that Göring was not an anti-Semite and even appointed Jews in his ministry, or I have

> told how popular he was with the people and with his subordinates or something like that. But to defend Hitler or Göring? Impossible!

I take Van der Waerden at his word; he himself was not an anti-Semite and not a Nazi sympathizer. However, he is now asserting that the Nazi pressure at Leipzig was not too bad, no big deal, only all Jewish professors fired, the Jewish professors he had unsuccessfully tried to defend in 1935. He then declares that the second man of Nazi Germany, Herman Göring, was not an anti-Semite. Really? Van der Waerden then goes on to explain his wife Camilla's demand that no criticism of Germans be made in her presence should she visit the Netherlands:

> Now about my wife supposedly being a Nazi. Would you believe that this is the third time that I hear this spiteful slander? I cannot figure out where this slander is coming from. We, my wife and I, have avoided any contact with the Nazis in Leipzig like the black plague. Our acquaintances were only people who shared our horror for the Nazi regime. And then, when she stayed in Holland, she asked that nothing bad be said about Adolf? Do you honestly believe that my father, when we stayed with him in 1939, would have accepted such a condition, or whether my brothers would have been content with it? The truth is that my wife could not tolerate it when bad was spoken about the Germans. Indeed, German is her mother's tongue, and she knew so many kind people in Germany. If you do not want to believe all of this [based] on my word, then please write a letter to Frau Lotte Schoenheim, Hotel Stadt Elberfeld, Amsterdam. From 1932 up until her emigration to the Netherlands in 1938, she has been frequently in conversation with my wife and me, and after that in Holland has stayed in contact with my family. She knows our opinion not only from words but also from deeds.

Again, I take Van der Waerden at his word. *According to him*, Camilla "could not tolerate it when bad was spoken about the Germans." Were all Germans in Nazi Germany above criticism in 1939? Did not Van der Waerden himself write above in this very letter that "Germany attacked the Netherlands and shamefully abused it, and the whole [!] German people are also responsible for that"?

And one more question: doesn't Van der Waerden feel *some* personal responsibility for German crimes against humanity? Does he really imagine that being off the German citizenship rolls frees him from any responsibility for the horrific actions of the country he has lived in and worked in Civil Service for 14 long Nazi years?

This handwritten letter is particularly important to Van der Waerden: he encloses a large handwritten part of it, entitled "From a letter to Prof. J. G. van der Corput," in his January 22, 1946, letter to Hans Freudenthal together with "The Defense," which has earlier been submitted to the Amsterdam's *College van Herstel* and Utrecht's *College van Herstelen Zuivering*.

In his immediate, August 28, 1945, reply, Van der Corput soft pedals on his probing questions and assures Van der Waerden of his support:

> Am I mistaken if I have an impression that you wrote your letter in a somewhat irritated state? I believe that I have consistently acted in your interest; also during a conversation with the Minister I pointed out that the Netherlands should be very careful not to lose a man like you. I even said that the Netherlands should rejoice if we get you back for

> good. But there are general rules, and it needs to be determined how much those apply to you.
>
> I have always considered it impossible that you are a "weathervane, a hypocrite, and a cunning liar," and I still consider it impossible. With my remark I wanted to show that you in my opinion did not sufficiently realize how we thought of the Hitler regime even then. It was all joking, and I never attached much significance to it, but when afterwards remarks were made indicating doubt, I thought it was important for you that I mention this in my letter. I would be very sorry if I hurt you by it, but it is still better to bring these things out in the open and to give you an opportunity to rebut them. To my great pleasure I found out today that it was said that at the Mathematical Congress in Oslo [1936] you were known as a strong anti-National Socialist.
>
> Immediately after receiving your letter, I made sure that this week Friday night or Saturday morning there will be a meeting between me and the Minister of Education about this matter. The Minister has already told me in the first conversation that the cabinet has spoken about general rules concerning the persons who were in German service during the war. Those rules were to be finalized then. Whether or not this has happened since then I will find out this week.

Van der Corput leaves the last two points of Van der Waerden's letter (presumably Bartel's praise of Herman Göring and Camilla's defense of all Germans in the Third Reich) to a confidential in-person conversation, and thus, to my regret, out of reach of historical scholarship. These points are so important that Van der Corput is willing to travel early in the morning from Groningen to Laren for a person-to-person discussion:

> About the various other points of your letter, I would like to speak with you in person next week. Tuesday September 4, I hope to get to Laren for this before 9 o'clock in the morning.

But not to worry anyway:

> Be assured that it is my sincere desire to keep you for the Fatherland and for higher education.

Soon success seems to be around the corner. Van der Corput communicates the first hopeful signs on September 11, 1945:

> I have discussed your case with Oranje and Borst, leaders of the Professors' Resistance. After my explanation neither one of them saw any problem with your appointment at one of the Dutch universities. They of course cannot decide anything, but as is evident to me, it is much easier for the minister and his department if they know that there is no opposition from that particular side. I have the impression that things will be all right and that after a few months we will be able to collaborate again ...
>
> P.S.: . . . During my absence Van der Leeuw has called to tell me that both parts of my most recent letter were "good." One of the parts concerned my statement that we do not need to fear any opposition from Borst and Oranje . . . It will all work out, that is my opinion.

Five days later Van der Corput is ready to celebrate 'mission accomplished' (the phrase made famous by the US President George W. Bush):

> I have just received a written confirmation from Van der Leeuw . . . He writes: "As far as van der Waerden is concerned, we will just count on it that it is all right."
>
> This means that he is prepared to appoint you.
>
> I am very much pleased with this, both personally and in the interest of the country.
>
> P.S.: I am passing the message on to Utrecht right now.

On September 22, 1945, Van der Waerden optimistically describes the state of his job hunting to his confidant Hans Freudenthal:

> Minister Van der Leeuw told Van der Corput that now that Van der Corput and Borst and Oranje of the Professors Resistance Group consider me as sufficiently "pure," he also considers the affair "OK." My appointment at Utrecht is therefore very close.

On September 29, 1945, Van der Corput informs Van der Waerden by a telegram that *College van Herstel en Zuivering* of Utrecht University got on Van der Waerden's board as well:

> Minnaert[18] signals *College van Herstel* considers Van der Waerden sufficiently politically reliable and desires appointment at Utrecht
>
> Van der Corput

However, about a month later, unexpectedly, skies over the two friends become cloudy. Van der Corput informs Van der Waerden about it in his October 24, 1945, letter:

> Indeed, difficulties concerning your appointments arise now again. As there is someone in higher education, who works against you and among other things maintains that you had to use – and did regularly use – the Hitler salute at the inception of your classes in Germany. Be so kind to give very clear answer to this question, so that I can contradict it if this slander comes about again.
>
> . . . This week I received an invitation from the Faculty of Natural Sciences at Amsterdam to become Weitzenböck's replacement. This shows that the opposition against your nomination in Amsterdam is too strong. I do not know what I am going to do. Personally, I like Utrecht better, but maybe I can do more for mathematics in Amsterdam . . .
>
> I am not happy about the turn that the problems in mathematics [appointments] have taken. I would be particularly sorry if certain illegal circles [*illegale kringen* – he probably means former underground circles] will successfully delay your appointment at a Dutch university.

Van der Waerden answers right away, on October 26, 1945. He does not give a "very clear answer to this question" of the Hitler salute, or any answer for that matter. He shares Van der Corput's pessimism about his academic prospects in the immediate future and blames the students and Minister of Education Van der Leeuw for it:

---

[18] Marcel Gilles Jozef Minnaert, a member of the "Van der Corput Committee." Documents in the archive of Utrecht University show that Minnaert – in a sense – represented Van der Waerden to the Utrecht's *College van Herstel en Zuivering*, which most likely had never met with Van der Waerden in person. This was a very beneficial representation for Van der Waerden, because as an outspoken critic of Nazism Minnaert spent nearly two years in a Nazi prison, from May 1942 to April 1944.

After what I have read in the *Vrij Katholiek*[19] about the radical demands of the students and the willingness of Van der Leeuw to listen [to them], I think it will take some time before I can get a position at Utrecht. I have something else now, as of October 1, 1945, I am working for *Bataafsche*.[20]

Van der Corput's reply comes a full month later, on November 26, 1945. He opens his letter with the good news:

I very much want you to have a position in higher education. The Committee for Mathematics [*Wiskundecommissie*] intends to create the Center for Pure and Applied Mathematics, most likely in *A'dam* [Amsterdam], and if the Center comes into being, I want you to work there.

Then there come the bad news:

Butthere are problems, and I hate time and again asking you these questions and asking you for clarifications, but I have to do this. In order to support you I need the answers to these questions.

It now centers around three clearly indicated points.

The first. Your father and your uncle repeatedly and with a lot of emphasis insisted before [!] the war that you should leave Germany. They felt it was your duty to leave but you refused, and they considered it as neglect of your duties.

The second: some people are certain that your wife is an anti-Semite, others believe that this is too strong a statement, but she did not want to have anything to do with Jews.

The third. During the war there was an opportunity for you to go to America, but you refused, for you [argued that you] needed to stay because you could do a good work for your students, some of whom were Jews. If this is true that even during the war, when you had a chance to go to the United States, you still did not want to leave, this will create definite difficulties for you.

Apparently, without receiving a reply for twelve days, Van der Corput writes again on December 8, 1945, this time quite apologetically:

I am not asking you these things for myself . . . I want to collaborate [with you] as much as I can . . . It would be very unpleasant if these questions would somehow cause the deterioration of my relations with you or your wife. Please, understand I only need it for the government.

Now Van der Waerden replies immediately, on December 10, 1945. He first reassures Van der Corput of their friendship:

It would be a pity if our cordial relationship should become the victim of our correspondence. But I see no risk of that.

When there was now and then a ring of annoyance, this was in fact directed against the people who disseminate such gossip against me, but not against you, of whom I know that you are tirelessly active in my interest and that of the Dutch science!

---

[19] *DeVrij Katholiek* (The Free Catholic) monthly of the Free Catholic Church in the Netherlands.

[20] *Bataafsche Petroleum Maatschappij* (*B.P.M.*), today known as the Royal Dutch Shell.

Van der Waerden then spells out his fundamental "democratic" principles:

> On the other hand, I also cannot imagine that you are incensed by the fundamental democratic anti-Fascist position that I have adopted in my letter. My viewpoint is that where appointments are concerned, only capacities of the appointee should be taken into account, and not – as is usual with Fascist regimes – the person's character, past, and political trustworthiness.

"Only capacities"? As the author of this narration, I am compelled to ask: Bartel, don't you agree that mathematicians do not live in a vacuum, and thus their "character," their "past," and their moral fabric matter? Would *you*, for example, hire Bieberbach, a decent mathematician and indecent man, an anti-Semite and the leader of the notorious race-based *Deutsche Mathematik*? Haven't you seen enough examples of the Nazis using "capacities" of their scientists, and "the power of German engineering" (as today's Volkswagen commercials delight in stating) to evil ends? How about the uses of U-boat submarines, Messerschmitt aircraft, Wernher von Braun rockets, gas chambers and crematoria for an efficient extermination of human beings, medical experiments on prisoners in the camps? *Mathematik über alles? Mathematics above all? Mathematics above morality?*

Van der Waerden then invokes his father Dr. Theo as the influence of his life:

> I have been raised under the apprenticeship of my father who was a man of democratic principles; subsequently I have been under Hitler's control, and I have seen to which terrible consequences the opposite view leads.
>
> You too [sic] have actively opposed Nazism and fought for democracy and freedom of our nation. Therefore, I cannot imagine that you would hold my viewpoint against me, even though you do not share it in every respect.

This dialog in letters is so vivid and so passionate that once again, as the author of this book, I feel drawn to enter into it and say: Bartel, you invoke "a man of democratic principles," your father as your important influence. However, you have not listened to your father in the most important matter of your life, when he "repeatedly and with a lot of emphasis insisted before the war that you should leave Germany." I agree with your father Theo and Uncle Jan: the "gangster-regime" (your words) occupying and terrorizing your Dutch people and other peoples of Europe was not the right place for a decent person like you. Some members of your family felt that it was "not done" by a good Dutchman like you to remain in Nazi Germany.[21]

Van der Waerden ends the letter with the major good news, promising an Amsterdam professorship to him very soon:

[21] Bartel's first Cousin Annemarie van der Waerden recalls the extended Van der Waerden family reaction to his decision to stay in Nazi Germany: "Definitely it was considered 'not done' that Bart stayed in Germany. Though he was excused probably by this committee – this must be the case considering the fact that he got a respectable job in Holland again – he stayed a disputed man. In the family some forgave him, some not. The ones that forgave him, that was also because he was such a sweet, innocent man." [WaD2].

> Revesz[22] informed me yesterday that the Amsterdam has recommended me for the appointment as a full professor to *B. en W.* Thus, things start moving now!

On December 22, 1945, Van der Waerden writes again. This four-page letter is full of technical negotiations. Van der Waerden discusses these details because in his mind his appointment at Amsterdam is a done deal.

In the end, Van der Corput is not completely satisfied with some positions of Van der Waerden the man. But Van der Corput has a great respect for Van der Waerden the mathematician, and he believes that he ought to help Van der Waerden get a position at Amsterdam. Van der Waerden would then spend his career there and thus would greatly benefit their Homeland.

It is worthwhile to note here that Van der Waerden is much more open and harsh in his criticism of the Dutch people in his November 20, 1945, letter to Richard Courant – who is far away in New York – than in his entire correspondence with Van der Corput:

> The Dutch are completely crazy. They have no concept in their heads except "cleansing" ("*Sauberung*"): they punish all those who had worked together with the Germans. There are managers, bosses who would not employ any workers who were forced to work in Germany.[23] There are more political prisoners in Holland than in all of France, even though the Dutch showed much more character in the war than the French did. So is my appointment to Utrecht, which ran into great difficulties, even though it was a done deal with the faculty for years. I am very happy that I currently have a pleasant job in industry and can await the return to normal circumstances.

And while Van der Waerden demands "a complete exoneration" from Van der Corput, he sounds much more conciliatory in his December 29, 1945, letter (in English) to Courant:

> I am much pleased that you have the intention to resume the old friendship with me and other old friends as far as possible, and that old Göttingen will keep a warm place in a corner of your heart. And just for that reason, I am convinced that you at least will understand a little bit what my other friends in America could not grasp, namely "why I as a Dutchman chose to stay with the Nazis."
>
> Look here, I considered myself in some sense as your representative in Germany. You had brought me into the *redaction* [editorial board] of the *Yellow Series* and *Math. Annalen*, I thought, in order to watch that these publications were not Nazified and that they might maintain their international character and *niveau* [standard] as far as possible. This I considered to be my task, and together with Hecke and Cara[théodory] I have done my best to fulfill it, which I could do only by staying in Germany. [It] is not that plain and easy to understand, apart from other sentimental and familiar [familial] links attaching me to Germany. I have made some mistakes perhaps, but I have never pacified the Nazis.

Indeed, in the mid-1933 and 1934, Courant envisaged Van der Waerden as his representative in the *Mathematische Annalen* and the *Yellow Series.* However, on August 20, 1935, Courant hinted to Van der Waerden to leave Germany:

---

[22] Geza Révèsz (1878–1955), the founding psychology professor at the University of Amsterdam.

[23] Van der Waerden refers here to *Arbeitseinsatz*, the Nazi forced labor program.

> I wish everybody could get out of this stuffy atmosphere. I must admit I cannot understand those who remain in Germany, unless they do it out of conviction or strong patriotism or from a willingness to fight. It seems to me more and more that remaining there as a civil servant is impossible without compromises.

With "other sentimental and familiar [Van der Waerden means *familial*] links" to Germany, Van der Waerden no doubt refers to his "German wife" and raising his children "pure German," and possibly to his sense of belonging to the German culture in general, and the German mathematics in particular. For the first time, Van der Waerden admits making "some mistakes."

The Dialog in Letters presented here is undoubtedly an important collection of documents for the history of the de-Nazification and for the reflection of the post-World War II search for moral standards. Furthermore, I hope it will prompt you, my reader, to define your positions on a number of fundamental moral issues, such as the relationship between the scholar and the state, in particular the place of a scientist in tyranny, duty to profession, patriotism vs. nationalism, etc. We will come back to these contemporary issues later in this book.

Chapter 42
In Search of Van der Waerden: The Unsettling Years, 1946–1951

42.1 The *Het Parool* Affair

When in May 1940 the Germans conquered our country, Mr. Van der Waerden was still standing behind his lectern at Leipzig.

– Het Parool, *January 16, 1946*

I find it surprising that the early press records have been completely overlooked and never mentioned by earlier biographers of Van der Waerden. Did they view the news reports to be too much off the cuff and not carrying lasting truths? Yes, the shelf life of a newspaper is one day, but it captures – and preserves – the *zeitgeist*, the spirit of the day, better than anything else available to a historian. Moreover, in our *Drama of Van der Waerden*, a newspaper was also an important *player*. I will therefore use newspapers liberally and unapologetically.

After the war, both East and West Germanies were quite soft even on Nazi collaborators, which Van der Waerden certainly was not. In addition, Van der Waerden's loyalty to Germany and German mathematics was unquestionably great. Holland was another matter. Its standards of "good behavior" during the Nazi occupation of Holland were much higher, especially when judged by the editors of a publication like *Het Parool*, a newspaper that had been heroically published underground ever since July 1940[1] and paid for it by lives and freedom of many of its workers. After the war and the occupation, at the circulation of 50,000–100,000 in Amsterdam alone and local editions appearing in more than ten cities [Kei], *Het Parool* had an enormous moral authority.

In early January 1946, everything was in place for appointing Dr. Van der Waerden to a professorship at the University of Amsterdam. The City Council's meeting with his appointment on the agenda was about to begin the afternoon of January 16, 1946, when just hours earlier a "bomb" exploded on page 3 of *Het Parool* [Het1]:[2]

[1] It was started by Frans Johannes Goedhart under the title *Nieuwsbrief van Pieter 't Hoen* on July 25, 1940, and became *Het Parool* on February 10, 1941 [Kei].

[2] In search for greater expressiveness, the editors included in this Dutch article some passages in German. I am leaving them in German and add translation in brackets. I also include in parentheses some Dutch expressions that are particularly hard to adequately translate into English.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_42

Him??
No, not him!

The proposal to appoint Dr B C [sic] van der Waerden as professor in the faculty of mathematics and physics at the University of Amsterdam should surprise all those who know that Mr. Van der Waerden served the enemy throughout the entire war. His "collaboration" is not today's or yesterday's news. When the war broke out in September 1939, and the Netherlands, fearing invasion, mobilized, Mr. Van der Waerden was standing behind his lectern at Leipzig University. He had stood there for years. And he *continued* to stand there. He saw the storm coming as well, but he did not think about coming back to his Fatherland. When in May 1940 the Germans conquered our country Mr. Van der Waerden was still standing behind his lectern at . . . Leipzig. And he continued to stand there. For five years the Netherlands fought Germany and for all those five years Mr. Van der Waerden kept the light of science shining in . . . Leipzig. He raised Hitler followers. His total ability – a very great one – and all his talent – a very great one – were at the service of the enemy. Not because Mr. Van der Waerden had been gang-pressed (*geronseld*) to the forced *Arbeitseinsatz* [labor service], not because it was impossible for Mr. Van der Waerden to go into hiding; no, Mr. Van der Waerden served the enemy, because he liked it at Leipzig; he was completely voluntary a helper of the enemy, which – and this could not have remained unknown to Mr. Van der Waerden – made all of higher education plus all results of all scientific work serve the enemy's "*totale Krieg*" [total war].

When asked, Mr. Van der Waerden cannot answer what an average German answers when he hears of the boundless horrors done in the country: "*Ich habe es nicht gewusst*" [I did not know]. In the middle of the war years Mr. Van der Waerden came back to the forgotten land of his birth and he heard and saw how disgracefully his patrons (*broodheeren*) were acting here. Did he not care at all? (*Liet het hem Siberisch koud?*) A few weeks later Mr. Van der Waerden was standing behind his lectern at . . . Leipzig again. In the Netherlands, firing squads shot hundreds. In the concentration camps, erected as signs of *Kultur* (culture) by the Germans in Mr. Van der Waerden's second Fatherland, many of the best of us died; as did a few Dutch colleagues of Mr. Van der Waerden. Did that do anything to him? The story is becoming monotonous: Mr. Van der Waerden raised the German youth from behind his lectern at . . . Leipzig.

However, that is where the house of cards collapsed. Germany, including Leipzig, surrendered. The Third Reich, which Mr. Van der Waerden had hoped would last, if not a thousand years, then at least for the duration of his life, became one great ruin. And at that very moment Mr. Van der Waerden remembered that there existed something like the Netherlands and that he had a personal connection to it. He looked at his passport: yes, it was a Dutch passport. He packed his bags. He traveled to "the Fatherland." Now Leipzig was not that nice anymore. All those ruins and all those occupying forces – yuk (*bah*). After five years of diligent service to the mortal enemy of his people, Mr. Van der Waerden was now prepared for the other camp.

There are more like him. But what is worse, the University of Amsterdam seems willing to give this Mr. Van der Waerden another lectern immediately. Mathematics has no Fatherland, you say? Yes, sir (*tot uw dienst*), but in the Netherlands in the year 1946 it should be desired of a professor of mathematics to have one, and to remember it more timely than on the day on which his lectern in the land of the enemy became too hot under his shoes.

This passionate article, circulated throughout the whole country, with "*Mr. Van der Waerden was standing behind his lectern at Leipzig"* repeating over and over like a refrain in a song, must have made the Amsterdam City Council concerned, if not embarrassed. While people who served in the German labor service (*Arbeitseinsatz*) among the faculty, staff, and students were to be removed from the university, the City Council was planning to approve the appointment of a professor who voluntarily served Germany the entire Nazi period, including the five years of the German occupation of Holland. The approval of Van der Waerden's appointment was postponed. The following day, on January 17, 1946, *Het Parool* reports the outcome [Het2]:

Prof. VAN DER WAERDEN NOT YET APPOINTED
Appointment halted

After the Amsterdam City Council convened yesterday afternoon in the Committee General (*Comité Generaal*), Mayor de Boer announced that the nomination to appoint Professor Dr. B. L. van der Waerden, Professor of Mathematics at Leipzig, as Extra-Ordinarius (*Buitengewoon Hoogleeraar*) at the University of Amsterdam has been put on hold.

Because of the publication in *Het Parool* about Professor Van der Waerden, the Council suggested that there should not be a rush action. Further information was demanded.

On behalf of *B. en W.*, City Alderman (*Wethouder*) Mr. De Roos responded that Professor Van der Waerden had good papers. Leipzig was a mathematical center. Beforehand many authorities were asked for information; among others also the Commission of Learned People (*Gestudeerden*) in Germany. The *College van Herstel* (College for Restoration) of the university and also the faculty supported the appointment. For now, however, the appointment has been halted; *B. en W.* will consult later with the *College van Herstel*.

Van der Waerden is outraged not only by the City Council's refusal to approve his appointment but also by such heavy and public accusations by the newspaper that was read and respected practically by everyone in the postwar Netherlands. On January 22, 1946, he briefs his friend Freudenthal on the state of events:

Amice,

Thank you for your kind letter. It did us a lot of good to have at least one loyal friend in the midst of this enemy world.

I have sent the enclosed rebuttal to *Het Parool* and to *Propria Cures*. Already before that I supplied Clay with the necessary data for the Alderman's[3] defense of [Van der Waerden]. I have the impression from the report of the council meeting in *Het Parool* that the Alderman is fighting for me like a lion.

The attitude of the students gives me great joy. As soon as I am there, I will win them for me completely. I am convinced of that.

I am not sure why Van der Waerden has gotten "a great joy" from the students' attitude. As we will soon see, students have presented a vocal opposition to his appointment. Please also notice Van der Waerden's line "I supplied Clay with the necessary data for the Alderman's defense": we will soon learn the contents of this data from a *Het Parool*'s article.

[3] Here, Van der Waerden refers to one particular Alderman (there were six): Mr. Albertus de Roos (1900–1978), the Alderman (1945–1962) for Education and Arts.

With this letter to Freudenthal, Van der Waerden encloses two documents – the "Defense" and "From a letter to Prof. J. G. van der Corput" that we have discussed in great detail in the previous sections, as well as the following handwritten letter to the editor, which he sent to both papers, *Het Parool*[4] and *Propria Cures*, even though the latter paper did not run any commentary on Dr. Van der Waerden's impending appointment:

Correction [*Rechtzetting*]

In the '*Het Parool*' dated Jan 16, my person was sharply attacked. I do not wish to go into this at great length. The question of whether or not I acted wrongly is being carefully researched by the concerned services.[5] But I have to correct two untruths. It is said that I hoped that the Third Reich would last for as long as I would. This is slander. I was known in Germany and outside as a strong opponent of the Nazi regime; I can prove this with witnesses.

It furthermore says that I returned because my lectern became too hot under my feet. This is also not true. I returned because the *Faculteit* of Mathematics and Physics of the State University of Utrecht asked me to take up a professorship in mathematics.
B. L. van der Waerden

There existed words – words about patriotism, love of the Fatherland, contributions of the Van der Waerden family to Holland, desire to return home, raise a new generation of scholars in the Netherlands – which could have touched the readers' and editors' hearts and made a strong case for Van der Waerden's acceptance. Van der Waerden's dry and proud prose about returning because of a job offer could not have possibly made things better for him.

The self-assessment as a "strong opponent of the Nazi regime" in Van der Waerden's letter to *Het Parool* also did not help, for it was certainly viewed as an exaggeration by the editors of an underground newspaper, who for five years daily risked their lives. Understandably, Van der Waerden's letter backfired.

Both *Het Parool* ("Prof. Van der Waerden defends himself," [Wae9]) and *Propria Cures*[6] ("Correction," [Wae10]) publish the complete text of Van der Waerden's "Correction" on February 1, 1946. *Het Parool* adds the following editorial response [Het4]:

We are pleased to give Mr. Van der Waerden the opportunity to defend himself. Has he made his case stronger with this? No, not quite. Unless there are Dutchmen who truly believe that the Germans from 1940 to 1945 allowed "strong" (!) opponents to occupy professorships. Which acts show this strong anti-Nazism of Mr. Van der Waerden? And the timing of his return to the Fatherland in 1945 is then one of those rare coincidences that one should believe as such... or not. Mr. Van der Waerden – and this is the heart of the matter – from the first until the last day of the war served science in the land of the enemy and this was compensated by the enemy's money. He who has voluntarily served the enemy from May '40 to May '45 is a bad Dutchman. Those who unleash him

---

[4] Van der Waerden's letter to *Het Parool* was dated January 21, 1945, as seen from *Het Parool's* January 23, 1945, acknowledgement sent to Van der Waerden and signed by Secretary Hoofdredactie: see ETH, Hs 652: 11631.

[5] Van der Waerden likely refers to the de-Nazification boards, *College van Herstel* of Amsterdam and Utrecht.

[6] University of Amsterdam students' weekly.

afterwards on the Dutch youth do not understand the demands of this time. And if the appointment of Van der Waerden is approved, then one should immediately stop objecting to workers and students who volunteered for *De Arbeitseinsatz* [the German Labor Service],[7] etc., for the *De Arbeitseinsatz* of Van der Waerden was more complete than that of any other Dutchman. "Rewarding" ("*Belooning"*) that with a professorship would mean that all the others who worked for the enemy voluntarily deserve a feather and a bonus.
– *Red* (Editors) *Het Parool*

Earlier, on January 25, 1946, *Het Parool* has already reported the postponement of the approval of Van der Waerden's appointment [Het3]:

Prof. Dr. B. L. van der Waerden

The nomination of *B. en W.* to appoint Prof. Dr. B. L. van der Waerden, which was put on hold at the previous session of the city council, because of the article in "*Het Parool*," does not appear on the agenda for January 30th. It was put there initially, but it has been scrapped off by *B. en W.*, from which it can be deduced that further consultation has not yet ended.

On February 13, 1946, *Het Parool* publishes its last commentary on the Van der Waerden affair [Het5]. From it we can understand what data Van der Waerden supplied to Professor Clay for Alderman Albertus de Roos' defense of Van der Waerden:

Concerning Van der Waerden

The city council has circulated a little piece of advertising for the benefit of Prof. Van der Waerden, of which the main points are that he protested against the firing of the Jews in 1934 (even though he himself continued teaching classes) and that during the war, with the exception of a family visit in November 1942, he was not allowed to leave Leipzig, while, the little piece says, at that moment "going into hiding was out of the question," so that it could not be expected of Van der Waerden to "go under," even less so because he would have had to leave [his] wife and children in Germany.

[7] Under the *Arbeitseinsatz* program, the Dutch (and other) peoples were sent to work in Germany (or "Greater" Germany). Those who went were punished after the war. In a 2004 email to me, Dr. Knegtmans comments as follows [Kne8]: "As far as I know, only very few people actually volunteered for the *Arbeitseinsatz*. Most (several hundreds of thousands) did so under pressure and among them were three thousand students of all Dutch universities and a few staff members. After the war, however, there was some criticism of these men. Could they not have evaded conscription, some asked publicly. I think they could not, because their names and addresses were known and most needed the income for their families. This was of course not the case with the students, but in fact most students fled from the *Arbeitseinsatz* in Germany back to Holland, while others did not return to Germany from their holidays. I think that none of the students, staff members, or professors of the University of Amsterdam was punished for voluntarily joining the *Arbeitseinsatz*. Probably no one did join voluntarily. But some of the Nazis among the students and staff joined the German army (or the Dutch Volunteer Corps) or para-military German organizations. The staff members among them were removed from the university, the students simply did not return to the universities."

> This writing makes us slightly nauseous. November 1942! Pieter ‘t Hoen[8] has been in prison for eleven months, Wiardi Beckman[9] is in prison, Koos Vorrink[10] is in hiding, indeed all *Parool* people are in hiding; the *O.D.*[11] trial is over [resulting in] 70 people shot. The entire *O.D.* leadership is in hiding. All *Vrij-Nederland* people and those of *De Geus*, and *Je Maintiendrai*, and *Trouw*, and *De Waarheid* are in hiding.[12] In hiding, leaving behind wives and children! No, the little piece of advertising says “going into hiding was out of the question.” And then the explosion comes: “. . . and there was also no clear resistance yet”!!! See above, reader! November 1942. Hundreds have been shot for the resistance. Thousands are in camps. Other thousands have gone under. The illegal press flourishes (*Parool* 15,000 copies!). “No, there was no clear resistance yet,” the writer of the little piece of advertising says.
>
> There was such a clear resistance that Van der Waerden was advised by his immediate environs not to return [to Germany]. He went anyway. For three more years he taught in the enemy’s country for the enemy’s money. Who could stomach to suspend an art student from the university for a few years while at the same time make Van der Waerden a professor?

Clearly, Van der Waerden’s statement conveyed to the Dutch people via Alderman de Roos that in November 1942 “there was also no clear resistance yet,” was untruth. Moreover, it must have been received in postwar Holland as the worst kind of slander of the Fatherland, which prompted such a powerful rebuttal from *Het Parool* editors.

Now that Van der Waerden has also initiated a discussion on the pages of the students’ weekly *Propria Cures*, he receives a published reply from P. Peters, apparently a student, in the next, February 8, 1946, issue of this weekly [Pete]:

---

[8] Pieter ’t Hoen was the pseudonym of the Amsterdam journalist Frans Johannes Goedhart (1904–1990), the founder of *Het Parool*, who was arrested in January 1942. Madelon deKeizer [Kei] reports that “Goedhart was one of the twenty-three suspects to be brought to trial before the German magistrate in the first *Parool* trial in December 1942. Seventeen death sentences were pronounced and thirteen *Parool* workers were executed by firing squad in February 1943. Goedhart managed to obtain a reprieve. He escaped in September 1943 and resumed his position on the editorial board.”

[9] Herman Bernard Wiardi Beckman, (1904–Dachau, March 15, 1945), a member of the Editorial Board of *Het Parool*, one of the intellectuals of the *SDAP* (*De Sociaal-Democratische Arbeiders Partij*), arrested in January 1942, he ended his life in the Nazi concentration camp Dachau.

[10] Jacobus Jan (Koos) Vorrink (1891–1955), a member of the Editorial Board of *Het Parool*, chairman of *SDAP* (*De Sociaal-Democratische Arbeiders Partij*) and later of *PvdA* (*De Partij van de Arbeid*, labor party), was arrested on April 1, 1943, and later sent to the Nazi Concentration Camp Sachsenhausen, from which he was liberated by the Soviet Army in 1945.

[11] “*O.D.*” stands for *Orde Dienst*, a national resistance organization.

[12] *Vrij-Nederland*, *De Geus*, *Je Maintiendrai*, *Trouw*, and *De Waarheid* were Dutch underground publications of the occupation period. Recall, Van der Corput served on the Board of *Vrij-Nederland.*

To Mr. Editor

During the last weeks there has been repeated mention in the press of the appointment of Prof. B. L. van der Waerden to a professor in group theory of algebra at our University. Still cloaked in the clouds of dust blown up by the return of other professors one should be surprised by the fact that no attention has been devoted by *P. C.* (*Propria Cures*) to the discussion of Prof. Van der Waerden.

Prof. Van der Waerden, as is well known, taught during the entire war at Leipzig University.

In *"Het Parool"* he recently declared having been anti-Nazi. Be it as it may, it is not entirely clear how to square this with his collaborative attitude, most tellingly illustrated by the fact that after the defeat of the Netherlands, he grew used to what he had been doing before that time, every single day he gave *Heil Hitler* salute (*Heil Hitlergroet*) in public at the start of his lectures to the enemy. Given the circumstances, it is hard to accept that he continued to fulfill his function in Germany under duress; even more so because, as was said, he was offered a professorship in the Netherlands. Subsequently, in his defense he does not discuss the voluntariness of his collaboration.

How tedious the subject of purification might have become, let there be no double standard.

Would it be therefore more tactful if the [City] Council, which is still contemplating his appointment, avoids the provocation here, and that Prof. Van der Waerden remains content with his present job [with *B.P.M.*] for now?

P. Peters

I do not know how reliable P. Peters' allegation was of Van der Waerden's daily use of the *Heil Hitler* salute at the start of his lectures. Van der Waerden did not send his rebuttal to *Propria Cures* as he did to *Het Parool* to refute its accusations he thought were false. He did not respond to the same allegation of using the *Heil Hitler* salute passed on to him by Van der Corput. I do know for a fact that Van der Waerden did use the *Heil Hitler* salute at the close of his official letters. Perhapshe did not think it was a serious enough accusation to merit a response? Van der Waerden's famous friend Werner Heisenberg did not think much of using the *Heil Hitler* salute. We have already read his nonchalant view.

P. Peters is correct in observing that "it is hard to accept that he [Van der Waerden] continued to fulfill his function in Germany under duress." And this is not an opinion of just one student: Peter J. Knegtmans in his monograph [Kne2] reports about the protest of the major students' organization *Algemene Studenten Vereniging Amsterdam* (*ASVA*):

> The *ASVA*[13] protested heavily against the coming of the mathematician Professor Van der Waerden to the University of Amsterdam because he had taught throughout the entire war at a German university.

Moreover, Knegtmans writes in an email to me [Kne3] that on February 5, 1946, *ASVA* wrote a letter to *B. en W*, the Executive Committee of the City of Amsterdam. According to Dr. Knegtmans notes (kindly translated by him for me from Dutch), the letter said:

[13] According to Dr. Knegtmans [Kne3], *ASVA*, a new general student union that had emerged from the circles in the Amsterdam student resistance. During the first postwar years, it was very keen on matters involving the behavior of old and new professors during the war.

> Word has reached the *ASVA* that *Burgemeester & Wethouders* have proposed Prof. Dr. B. L. van der Waerden as professor at the University of Amsterdam. This proposal has surprised the *ASVA*, considering the fact that during the war Prof. Van der Waerden has been professor at a German university.
>
> The *ASVA* is under the impression that the *College van Herstel* also had had some doubts before it eventually advised *Burgemeester & Wethouders* to go ahead with this proposal. However, the facts that have surfaced about Van der Waerden's behaviour during the war are so serious that his appointment would be unacceptable for the students, as long as the results of the investigations by the *College van Herstel* remain unknown.
>
> Therefore, the *ASVA* requests to reveal the grounds on which *Burgemeester & Wethouders* think Van der Waerden is qualified for a position of professor at a Dutch university.[14]

On February 15, 1946, Netherlands' Minister of Education, Culture and Science Van der Leeuw telephones Mayor of Amsterdam Feike de Boer and asks for information about Van der Waerden. The very same day, Mayor de Boer sends Minister Van de Leeuw a two-page glowing report, prepared by Van der Corput based on Van der Waerden's draft and signed by the Mayor. Mayor De Boer also sends the same report to Netherlands' Prime Minister Schermerhorn:[15]

> 15 February 1946.
>
> <u>Report on the actions of Prof. Van der Waerden.</u>
>
> <u>Confidential.</u>
>
> With respect to the request by telephone by your Excellency for information related to a possible position for Professor Dr. B. L. v/d Waerden as Professor at the University of Amsterdam, we have the honor to give you the following abstract of the results of our investigation by Professor Van der Corput and the Chair of the Mathematics and Physics Faculty of the University, who investigated the actions of Professor Van der Waerden from the beginning of the Nazi Government of Germany. The statements of Professor Van der Waerden are also partially included in this abstract.
>
> In 1934 Professor Van der Waerden spoke against dismissal of Jewish people openly during a faculty meeting in Leipzig. For this he received a reprimand from the Government of Saxony, and he was told that as a foreigner he should not meddle in the politics of Germany.
>
> In 1935 or 1936 Van der Waerden wrote a very nicely formulated eulogy for the Jewish professor Emmy Noether who left for the USA.
>
> The German Government considered Van der Waerden as not trustworthy and in 1939 did not allow him to visit Volta Congress, and give lectures for prisoners of war, and to participate in the Mathematical Congress in Rome. An invitation to the faculty at Munich University had not occurred because the party leaders found Van der Waerden "untrustworthy."

[14] Archives of the *ASVA* in the International Institute for Social History in Amsterdam.

[15] *Het Nationaal Archief*, Den Haag, finding aid number 2.14.17, record number 73 – dossier B.L. van der Waerden (Archive of the Ministry of Education).

When in 1940 the occupation of the Netherlands began, Van der Waerden was first interned and after that Germany banned him from leaving until the end of the war. Only after the death of his mother in November 1942 he was given permission to visit Netherlands for 6 days.

In 1941, Van der Waerden was still graduating "non-Aryan" [students]. Until 1942 he was accepting articles of Jewish people and people of mixed race to the *Mathematische Annalen* in his position as an editor.

Before the war Van der Waerden did not leave Germany because he believed that in his position of a scientist he would be able to defend culture against the culture destroying National Socialism; and only much later he came to an opinion that he was not able to do it.

During the short stay in the Netherlands [November 1942], he did not use this stay to go underground because not many people at that time did that, and also because his wife and children were in Germany and would not know what would happen to him.

In 1943 he did not accept the invitation by the faculty at Utrecht University because he did not want to accept a position from the government of the time.

Mrs. Van der Waerden is from Austria, and right from the beginning was very much against the Nazi regime.

Professor [Samuel] Goudsmit, who is chair of the American Bureau in Paris, had a task of investigating political activities of professors in Germany, has told Professor Clay and Professor Michels that his investigation did not show anything against Professor Van der Waerden. And a telegram was received by Clay from Goudsmit that said "Preliminary information favorable."

We would like to know based on the above what the Government position is with respect to granting professorship to Mr. Van der Waerden.
Mayor and Aldermen of Amsterdam
De Boer (stamped)
Secretary
Van Lier (stamped)

This two-page document was accompanied by a cover letter, which is of great interest to us, due to several consequent handwritten comments written on it. Let us start with the typed text:

To his Excellency the Minister of Education, Culture and Science.
15 February 1946
Result of the inquiry into the behavior of Prof. van der Waerden.
Confidential.
We have the honor to hereby forward Your Excellency a copy of a letter with information on Professor Van der Waerden, which our City Council sent to the Prime Minister in response to today's request for information.
Mayor and the Aldermen of Amsterdam
(signed) De Boer [Feike de Boer, the Mayor]
Secretary
(signed)

Five days later, on February 20, 1946, Mayor De Boer adds a note handwritten in pencil in the lower right corner of this letter (I am grateful to Dr. Peter Knegtmans for its translation). Mayor De Boer is concerned but still optimistic about approving Van der Waerden's professorship:

> Considering the report, it seems to me that objection against the appointment in Amsterdam cannot be maintained, albeit that the sentiment regarding v.d.W. [Van der Waerden] will at first not be favorable. If needed, a further reinforcing report by [Samuel] Goudsmit can be requested. (signed) De Boer 20.2
>
> In the upper right corner, I see a short handwritten note in ink added on February 26, 1946:
>
> Register as received (signed) De Boer 26.2

Why would one register on February 26, 1946, a letter written eleven days earlier? The answer comes from another document from the *Nationaal Archief*. This document, also dated February 26, 1946, arranged in landscape (horizontally), consists of two critically important letters written by the Netherlands' Prime Minister, Professor Willem Schermerhorn (1894–1977), who was also the Minister of War:[16]

> **[Left Side]**
> To the Minister of Education, Culture and Science
> February 26, 1946
> **SECRET** (stamp)
> Very Confidential
>
> With this I am sending you a copy of my letter that I have sent to the Mayor and the Aldermen of the City of Amsterdam with respect to a possible position for Mr. Van der Waerden as professor at the University of Amsterdam.
>
> Prime Minister
> (signature) W. Schermerhorn
>
> ---
>
> **[Right Side]**
> To the Mayor and Aldermen of the City of Amsterdam
> February 26, 1946
> Very Confidential
>
> In answer to your letter of 15 February [1946] regarding Case 0 nr. 13/I with respect to a possible position of Mr. Van der Waerden as professor at the University of Amsterdam, I can let you know that this kind of position will not be signed off.
>
> Prime Minister
> (signature) W. Schermerhorn

I am convinced that Van der Waerden's professorship was on the January 16, 1946, agenda of the City Council only because Mayor de Boer received an approval from the Minister of Education, Culture and Sciences Van der Leeuw. The fact that the Prime MinisterSchermerhorn overruled Minister Van der Leeuw's decision shows how powerful

[16] *Het Nationaal Archief*, Den Haag, finding aid number 2.14.17, record number 73 – dossier B.L. van der Waerden (Archive of the Ministry of Education).

the newspaper *Het Parool* was right after it emerged from the underground to the above-ground in the liberated Netherlands. Amsterdam's professorship for Van der Waerden has thus been closed for the foreseeable future.

Accordingly, on April 17, 1946, the *Burgemeester & Wethouders* advise the *College van Herstel* of Amsterdam, which served a dual duty of a de-Nazification Committee and the Board of Curators, of the withdrawal of Van der Waerden's nomination [Kne3]:

> *Burgemeester & Wethouders* inform the *College van Herstel* that they felt obliged to withdraw the nomination to appoint Dr. B. L. van der Waerden as extra-ordinary professor in group theory and algebra that they submitted to the city council on 4-January-1946, as it turned out that the government would withhold its assent in the event of an appointment of Dr. Van der Waerden.

This document demonstrates that Royal Assent was required for a professorial appointment at *any* Dutch university, including the municipal University of Amsterdam. Professor Van der Waerden, as well as, apparently, his colleagues Van der Corput and Clay, never understood this point, for even in 1993 Van der Waerden tells his interviewer Dold-Samplonius [Dol1] that "Amsterdam is a city university, and there the queen was unable to interfere." In fact, The University Historian of the University of Amsterdam Dr. Knegtmans, who knows best, advises me as follows [Kne5]:

> If Clay and Van der Corput really thought that an appointment as professor at the University of Amsterdam by the city council did not need approval by the queen, they were mistaken. It did so by law of 1876 and this procedure was not changed until sometime around 1980. However, approval by the queen did and does in fact mean approval of the minister (of Education, in this case). The queen was and is not supposed to have an opinion of her own. This [is] the minister's responsibility. It is the minister who advises the queen what to do: to give or not to give her approval. In Van der Waerden's case this meant that the then Minister of Education, Professor Gerardus van der Leeuw, Professor of Theology [as well as Religions and Egyptology] at the Groningen University, who was minister in the first postwar year, withheld his approval of Van der Waerden's appointment as professor in Utrecht as well as in Amsterdam. Van der Waerden was probably not appointed in Utrecht at all, because it was Van der Leeuw who had to appoint him. He was probably only proposed as professor by the *College van Herstel* in Utrecht.

Years later, I received documents of *College van Herstel en Zuivering*, the de-Nazification Committee of Utrecht, which specifically dealt with Van der Waerden's case among other matters. Thus, both de-Nazification Boards, those of Amsterdam and Utrecht, have investigated Professor Van der Waerden's behavior during the Nazi era. In the end, we see that the press and students held the feet of the academics and the governments to such a hot fire that the latter, convinced or not of the validity of the arguments, were so scared to err in the public eye on the serious issues raised by the press and students, that they gave up on trying to place Dr. Van der Waerden in any Dutch university.

On March 13, 1946, this was formalized in a letter from Dr. Gerardus J. van der Leeuw, Minister of Education, Culture, and Sciences to *College van Herstelen Zuivering* of Utrecht University:

> I am notifying you that the Council of Ministers has decided that persons, who during the occupation years have continuously worked in Germany out of their free will, cannot now be considered for government appointments.
>
> The reason for the decision was the discussion of a possible appointment of Dr. B. L. van der Waerden to professor in Amsterdam.
>
> It will be clear to you that the appointment of Dr. Van der Waerden either in Amsterdam or in Utrecht cannot take place.
>
> The Minister of Education, Culture and Sciences
> Signed for the Minister by Secretary-General H. J. Reinink

Astonishingly, Van der Waerden's *individual case* prompted the Government of the Netherlands to pass a new law, banning *all* "persons, who during the occupation years have continuously worked in Germany out of their free will" from *all* government jobs!

I have been unable to find Van der Corput's reaction to the *Het Parool* Affair, but I have found the next best thing: the opinion of the second major supporter of Van der Waerden at Amsterdam, Professor Jacob Clay. On March 19, 1946, just six days after the Minister's decision, Clay writes to Van der Waerden as follows:

> Dear v d Waerden,
>
> To my great regret our plan has not materialized at the last moment. The City government had already been convinced that the appointment was appropriate when the decision from the Minister came that nobody who has worked in Germany during the war, without any exceptions, for the time being would receive an appointment in public service. The response that I had prepared was not looked at, and in retrospect I am sorry that I have allowed the Alderman to keep me from responding to *Het Parool*. When so much time has passed, it seems better not to bring these things up again. I now hope very strongly that we will receive a better collaboration for the Mathematical Centre and that in time this matter will still work out OK, and I do not doubt that this is going to happen in time.

Nicolaas de Bruijn and my dear friend and coauthor Paul Erdős allow us an additional glimpse of Holland, year 1948. De Bruijn recalls [email to me of February 3, 2004]:

> You wanted to know more about my early contacts with Paul Erdős ... We met in person on several occasions, for the first time in 1948. I first saw him at his arrival in the harbor of Rotterdam, and took him to Delft, where he stayed a few days at our house.

Paul Erdős conveys a relevant detail of this 1948 visit:

> Once at the dinner table, when the conversation turned to Van der Waerden, Nicolaas' wife, Elizabeth "Bep" de Groot said,
>
> > If Van der Waerden were not such a fine mathematician, things would have been much worse for him [in the postwar Netherlands].

42.2 Job History 1945–1947

Upon his return to Holland in late June 1945, Dr. Van der Waerden needed a job as soon as possible. Hans Freudenthal introduced Van der Waerden to *Bataafsche Petroleum Maatschappij* (*B.P.M.*), today known as *Royal Dutch Shell*, and on October 1, 1945, Van der Waerden got his first post-World War II job as an analyst for *B.P.M.*

On July 30, 1946, Van der Waerden sends a letter to Leipzig Professor of Physical Chemistry Karl Friedrich Bonhoeffer, whose younger brother, the famous theologian Dietrich Bonhoeffer, was hung days before the end of World War II for his part in a conspiracy to assassinate Hitler. Van der Waerden expresses his condolences, criticism of the Nazi regime, criticism of the post-Nazi regime, difficulty of retrieving his bicycle from Leipzig, and difficulty convincing the world, a year after the horrific German brutalities, that there are still some decent Germans:

> Dear Mr. Bonhoeffer,
>
> We were very pleased to hear from you. But we were devastated by the fate of your four brothers and brothers-in-law murdered by the Nazis. That is terrible! The entire hopeless time that speaks from your letter and some other letters from Germany is very painful for us. How differently we imagined it when this hated gang would have gone away. And we would all work joyfully for the reconstruction of Germany of scholarship and of a better world. Now one year later I cannot even get permission to go to Leipzig in order to retrieve my bicycle and some papers and everyday things and to settle my relationship with the university. Everywhere there are walls of division, mistrust, and hate, and not much constructive work. Indeed, as you write, it is very difficult to make it clear to people everywhere that there are still decent Germans. Every individual half-way reasonable person admits it, but the general population does not want to see it.

From this letter, we also learn that Van der Waerden is happy with his industrial job and is offered an academic job in Graz, Austria (which did not work out). In 1946, a group of mathematicians led by Van der Corput establishes the *Mathematisch Centrum* (Mathematics Center), *MC* for short, in Amsterdam. As *MC's* first director, Van der Corput hires Dr. Van der Waerden to a part-time (one-day a week) position as the applied mathematics director of the *MC*.

At this point, Zurich enters the stage in our narrative. The lifelong ETH Professor Beno Eckmann (March 31, 1917, Bern–November 25, 2008, Zurich) kindly recollects for me [Eck1]:

> In 1944 the chair of applied mathematics became vacant. Lars Ahlfors was appointed in 1945, but he left after 3 semesters.

Olli Lehto writes [Leh1]: "Ahlfors did not stay long in Zurich; later he confessed that he did not have a good time there." Ahlfors explains (ibid.):

> I cannot honestly say that I was happy in Zurich. The post-war era was not a good time for a stranger to take root in Switzerland . . . My wife and I did not feel welcome outside the circle of our immediate colleagues.[17]

Consequently, Ahlfors gladly accepts an offer to return to Harvard University – where he worked 1935–1938 – and remains there for decades (1946–1977, plus afterward as an active Professor Emeritus). The University of Zurich upgrades Ahlfors' position (who was an extraordinary professor) to a full *ordinarius* and starts the new search.

Dr. Heinzpeter Stucki, *Universitätsarchivar* at Zurich, has found for me only one document directly related to this search, which, however, proved to be of a great significance: the 6-page July 15, 1946, report by Dekan Hans Steiner to Executive Authority (*Regierungsrat*) Dr. R. Briner of the Education Directorate (*Erziehungsdirection*) of the Zurich Canton. Steiner chooses two foreign mathematicians and recommends grabbing them as soon as possible, never minding the controversies surrounding these candidates:

> Prominent mathematicians are available today for a short time, and the two world-famous mathematicians in question are: <u>Rolf</u> <u>Nevanlinna</u>[18] (Finland) and Prof. <u>Van der Waerden</u> (Holland).

Steiner assesses the candidacy of Professor Nevanlinna first. After praising his mathematical achievements, Dekan addresses the personality of the candidate:

> He was born on October 22, 1895, in Joenuu (Finland) and for many years was Rektor of the University of Helsinki. He had to leave this position as a consequence of the political circumstances after the end of the war. Consequently, as he has briefly communicated, he is ready for an appointment at Zurich.

This is a rather short assessment: *born-rektored-forced to resign*. Looking at the 15-page summary [Ster] of the 317-page biography of Rolf Nevanlinna, written by his student (Ph.D., 1949) and *advocate* Olli Lehto, one is compelled to quote at least some information, which should have been relevant to the *neutral* Switzerland just one year after World War II:

> In 1933 Hitler became the German *Reichskanzler*. Up to the year 1943 Nevanlinna was of the opinion that Hitler [!] in German history could be compared to Friedrich the Great and Bismarck . . . He and other members of his family regarded the cause of Nazi Germany as their own cause. Germany was Nevanlinna's motherland (his mother was German) . . . This contributed to . . . his Nazi-friendly convictions in particular, which he expressed in a series of speeches and publications. Nevanlinna, however, has never been a member of a National Socialist party and did not hold anti-Semitic positions.

---

[17] Earlier, on September 18, 1938, Einstein expressed his distaste for the Swiss government: "I haven't forgotten that the Swiss authorities didn't stand by me in any way when Hitler stole all of my savings, even those designated for my children." (Letter to Heinrich Zangger [Ein1, p. 128]. In recent years, the cultivated for decades belief in Swiss neutrality during the war has been questioned.

[18] Rolf Herman Nevanlinna (1895–1980), a professor of mathematics (1926–1946) and *Rektor* (1941–1944) at Helsinki University; professor of applied mathematics at the University of Zurich (1946–1963, Honorary Professor starting in 1949).

When in Finland as well as in Germany the thought arose to establish a Finnish Volunteers Battalion, Nevanlinna welcomed this idea and agreed to the deployment of volunteers unreservedly. On the demand of [*Reichsführer SS*] Himmler there was developed the *SS* Battalion, and in the summer of 1942 Nevanlinna became the Chairman of the *SS* Volunteers Committee of this [*Waffen-SS*] Battalion!

Elsewhere [Leh2] Olli Lehto addresses the Nazi leadership role of his teacher Rolf Nevanlinna again:

In 1942, at the request of the Foreign Minister, Nevanlinna made himself available as chairman of the *SS* Volunteer Committee, which handled the recruitment of Finnish *SS* troops. After the war, Nevanlinna came in for especial condemnation [!] for his involvement in these activities.

My young readers may benefit from a very brief information about *SS*. The *Schutzstaffel* (*Protection Squadron*), abbreviated *SS* was a major paramilitary organization under Hitler and the Nazi Party. Under Himmler's leadership (1929–45), it grew to one of the most powerful organizations in the Third Reich. *SS* was responsible for many crimes against humanity during World War II. *SS*, along with the Nazi Party, was banned in Germany as a criminal organization after 1945. According to the Nuremberg Trials and many war trials conducted since then, *SS* was responsible for the vast majority of Nazi war crimes. *SS* was the primary organization that carried out the Holocaust.

I cringe while reading the Internet home pages of the International Mathematics Union (IMU), the highest organization of my profession:[19]

The Rolf Nevanlinna Prize in mathematical aspects of information science was established by the Executive Committee of the International Mathematical Union in April 1981. It was decided that the prize should consist of a gold medal and a cash prize similar to the ones associated with the Fields Medal and that one prize should be given at each International Congress of Mathematicians. The prize was named the Rolf Nevanlinna Prize in honor of Rolf Nevanlinna (1895–1980), who had been Rector of the University of Helsinki and President of the IMU and who in the 1950s had taken the initiative to the computer organization at Finnish universities.

I am compelled to ask the IMU executives: How can you ignore or minimize Nevanlinna's willing and eager service as the Chairman of the Finnish *SS* Troops Committee, his speeches in support of Nazi Germany, and on March 25, 1941, still claiming that Hitler saved European culture? Professor Nevanlinna was an excellent analyst with *no relationship to "Information Theory,"* once IMU President (1959–1962), and *the Finns offered to pay for the prize*, but mustn't we take into account the public deeds and moral bearings of the person whose profile we etch on our medals, let me repeat, *etch on our medals*? Or for the IMU executives, mathematics is above all moral concerns, *Mathematik über alles*? Didn't you understand that by "naming the prize in honor Rolf Nevanlinna" you are dishonoring Mathematics?

The talk, however, is cheap, and so I took upon myself to educate the Executive Committee of IMU on their own prize. What came out of it, I will convey in Section 42.10.

[19] http://www.mathunion.org/general/prizes/nevanlinna/details/

Let us roll back to the Swiss search, year 1946. Professor Nevanlinna is the first choice. *Dekan* Steiner then moves on to the second choice, Dr. Van der Waerden. Steiner admits that since he [Van der Waerden] became politically strongly disputed in Holland, the real state of affairs had to be clarified. Steiner then quotes a clarification supplied by the Dutch mathematician Jan A. Schouten,[20] who at that time lives in seclusion in Epe, Holland:

> Herr van der Waerden [...] remained during the war in Germany, to which he was, being exempt from [the Dutch] military service, fully entitled, and he always behaved there as an enemy of Nazism and in particular did much good for the Jews. The State Commission for Coordination of Higher Education [The Van der Corrput Commission], which has been established here after the war, and of which I have the honor to be a member, would have liked to have Herr van der Waerden in Amsterdam or Utrecht. The 'Purging Commission' that was installed after the liberation, with the task to test the heart and kidneys[21] of all Dutchmen, had declared him 'clean,' and the Minister of Education was ready to appoint him. Then a Jewish brother-in-law of Herr v. d. Waerden, who had already for years made enemies of him and particularly his (German) wife, unleashed a terribly dirty (*hundsgemeine*) agitation in the press. The Minister, who is no strong personality and who already had grave unpleasantness with other similar agitations, has thereupon given in to intimidation. You cannot at all imagine what sick conditions prevail here, dirty malicious agitation with self-interest and political purposes, often born from desire of revenge are the order of the day...
>
> Our main purpose was to keep Herr v. d. Waerden for Holland for the time being, and as soon as the wave of hatred and suspicion has subsided, he will get the *Ordinarius* Professor position, which he deserves as a great mathematician.

These harsh words of Schouten, directed at his recently liberated Motherland, were intended to make Van der Waerden appear as a victim of extremism. It must be said that Dr. Schouten peddled gossip to the Swiss: Van der Waerden had no sisters [!] and thus could not have had any brother-in-law, Jewish or otherwise. Regardless, so many Jews so recently had been killed, including circa 80% of the Dutch Jews, that it was in poor taste to blame a Jew for Van der Waerden's employment difficulties. But to claim that one ordinary person, Jewish or not, was able to "unleash a terribly dirty agitation in the press" meant to take Zurich Faculty for fools. Unbelievably, *Dekan* Steiner takes Schouten's shameless fabrication for truth and concludes Van der Waerden's political evaluation with.

> No reason is thus present to refrain from a possible appointment of Herr v. d. Waerden in Zurich.

Thus, two top choices, two world-class mathematicians, two individuals, whose political and moral choices have been questioned during the immediate post-World War II period, end up at the top of the Swiss wish list. Nevanlinna is chosen for the position, approved by the

---

[20] Jan Arnoldus Schouten (1883–1971), from a well-known wealthy family of shipbuilders, a professor of mathematics and mechanics at Delft Technical University (1914–1943), extraordinary professor (without teaching) of mathematics at the University of Amsterdam (1948–1953). Schouten was President of the 1954 International Congress of Mathematicians in Amsterdam.

[21] A biblical expression.

Government of the Canton Zurich and still in 1946 begins his Zurich professorship. However, on December 23, 1946, a member of *Züricher Kantonsrat* (Zurich Cantonal Council, a legislative body), Alfred White submits the following interpellation to *Regierungsrat* (Executive Authority):

> According to newspaper reports and letters from Finnish journalists, the newly-elected professor of mathematics at the University of Zurich Rolf Nevanlinna has operated as a recruiter for the sworn to Hitler Finnish *Waffen-SS*.

The Canton Government takes its time. Eventually, on March 14, 1947, the Directorate of Education (*Erziehungsdirection*) sends an inquiry to *Dekan* Steiner of Zurich University. Now they desire to receive the defense of Nevanlinna – and themselves – "as soon as possible." *Dekan* Steiner quotes Professor Fueter who assumes "full responsibility" (whatever this means) for Professor Nevanlinna's character, alleges that Nevanlinna's acceptance of Zurich job is a proof of the latter's interest in scholarship [sic], and minimizes Nevanlinna's sympathies toward and support of Nazi Germany. Fueter does not seem to understand the difference between patriotism and Aryan-kind of nationalism:

> "With Prof. Rolf Nevanlinna, both Prof. Speiser (former *Ordinarius* in Zurich, now Basel) and I have been friends for many years. We know him well and can assume full responsibility regarding his importance and his character . . .
>
> He has dedicated his entire life to scholarship; his acceptance of the Zurich offer confirms this once again, because he believes himself better able at this point to pursue his scholarly work here . . .
>
> The current *Rektor* [in Helsinki] was appointed to this position under the current government, which is strongly influenced by communists[22]. . .
>
> That Prof. Nevanlinna is in addition a great <u>Finnish patriot</u> will not be held against him in Switzerland. As such of course he tried to support his people with all [sic] available means in their struggle for existence. Obviously that was his duty. From the communist side in Finland, that is being held against him today. Any sympathies for National Socialist Germany played no role and were non-existent.[23]
>
> As for the precise accusations in the interpellation, we have no exact information about these things. It is certain that the selected Finnish soldiers were brought to Germany for further training (certainly not <u>before</u> the war, but mostly <u>between</u> the wars). Among them there were students. These soldiers were later integrated into the army and are supposed to have proven themselves as good soldiers. It seems doubtful [that we should] make use of the fact that they swore an oath to Hitler. There would need to be a proof of that. These soldiers were thus <u>non-German</u> *SS*, but were possibly only trained by such [German *SS*]. According to statements by Prof. Nevanlinna, he had simply nothing to do with this whole thing except that he was obligated as <u>*Rektor*</u> of the

[22] Repeatedly blaming "communists" could hardly fly. Wikipedia informs: "*Parliamentary elections"* were held in Finland on 17 and 18 March 1945. The broad-based center-left government of Prime Minister Juho Kusti Paasikivi (National Coalition/Independent) remained in office after the elections.

[23] This plainly contradicts Olli Lehto's writings that we have read earlier in this chapter.

university to place his name under a call to provide food for these people under orders to leave . . .

That today, after a lost war, political suspicions and pretensions are the order of the day is not surprising. It is clear that we in Switzerland should put an end to it. Above all we should steer clear of this foreign loose talk."

Thus, Rektor Nevanlinna "only" lent his name to the recruitment of "non-German *SS*" troops, presided over *SS* recruitment committee, gave speeches in support Nazi Germany, and praised Hitler as Friedrich the Great of his time. "After a lost [sic] war" as Steiner–Fueter put it, there were great mathematicians to be picked up by Switzerland, who were not wanted by the United States, Great Britain, etc., due to their questioned conduct. And so, we see in this letter facts bent to fit the desired goal of recruiting top mathematicians.

On May 14, 1947, based on the Steiner–Fueter letter, the Canton Government issues a self-serving, self-clearing response to the Alfred Weiss Interpellation (*Protokoll des Regierungsrates* 1947; 1631 *Interpellation. Am 23. Dezember 1946 reichte Kantonsrat A. Weiss-Zürich*). Thus, nearly a year after his hiring, the Nevanlinna Case is finally closed in Zurich.

I must have awakened Professor Beno Eckman's thoughts about the times past. In his email to me [Eck3], he volunteers a view of Zurich postwar hiring from his present standpoint:

If I may make a remark as I see it today [in 2004]: Politically Nevanlinna and vdW [Van der Waerden] were not easy cases for Switzerland one year after the war. But Universities tried to forget the past and look into the future. The decision for Nevanlinna must have been mathematical: he was absolutely world famous and at that time many mathematicians still considered analysis to be the most important part of mathematics – this has changed soon; algebra and topology became more and more important.

This affair shows that the famed Swiss neutrality was a pragmatic rather than a moral choice, façade rather than substance. Four years later, the new Dekan Boesch will write about this search as follows:

It is explicit from the Faculty proposal for filling a new position of Professor of Applied Mathematics dated July 15, 1946, that Prof. Van der Waerden was thoroughly considered.

Indeed, Prof. Van der Waerden was thoroughly considered, and the interest in hiring him was high. In four years, this 1946 consideration would bear fruit. Meanwhile, Van der Waerden continues his full-time work at *Bataafsche Petroleum Maatschappij* and part-time work at the *Mathematisch Centrum* in Amsterdam.

42.3 "America! America! God Shed His Grace on Thee"[24]

After the war, Van der Waerden desired a university professorship – he had held one ever since the tender age of 25 years. As we know from his letters to Lefschetz, Veblen, Neugebauer, and Courant, his first choice was an academic job in the United States. In

[24] From *America the Beautiful*, a song by Katharine Lee Bates, written on Pikes Peak, Colorado.

early 1947, Dr. Van der Waerden receives a letter from Baltimore, Maryland, that offers him both: a university professorship and an opportunity to live in America. Frank Murnaghan,[25] Johns Hopkins University's chair of mathematics, offers Van der Waerden the position of a Visiting Professor. In his May 5, 1947, letter, Van der Waerden informs Johns Hopkins' President Isaiah Bowman of his acceptance "with much pleasure."[26] Coincidentally, on the same day, May 5, 1947, the Board of Trustees of Johns Hopkins University approves the appointment. From their minutes, we learn that the appointment was effective from July 1, 1947, to June 30, 1948 (ibid). On May 13, 1947, Provost Stewart Macaulay specifies Professor Van der Waerden's salary at $6,500 for the year (ibid). The Van der Waerdens – Bartel, Camilla, Helga, Ilse, and Hans – board the ship called *Veendam*, which arrives in the Port of New York on September 29 or 30, 1947 (ibid).

At Johns Hopkins University, Van der Waerden is well respected and is offered a permanent professorship. This offer is made suddenly and is the result of an unspecified "emergency," as it is called in a number of documents,[27] which happened at Johns Hopkins University in the early February 1948. Naturally, I have tried to find out what the emergency was and came up with a conjecture. J. J. O'Connor and E. F. Robertson write as follows in The MacTutor History of Mathematics archive:[28]

> He [Murnaghan] held this post until 1948 when he retired after a disagreement with the President of Johns Hopkins University [Bowman], and went to Sao Paulo, Brazil.

The sudden departure of the chair of mathematics (chair did depart) is a serious loss for Johns Hopkins University. It creates a senior-level vacancy and most likely is the "emergency" that prompts President Bowman, a party to the disagreement, to rush and remedy the loss by making Professor Van der Waerden an offer of a permanent position. Let us take part in the emergency proceedings.

On February 6, 1948, President Bowman swiftly forms a special committee and writes to its members the following letter:

> An emergency has arisen in the Department of Mathematics that calls for early action on an appointment recommended by both Dr. Murnaghan and Dr. Wintner.[29] The candidate is Dr. Van der Waerden . . . You have received telephone notice of an Academic Council meeting at 8:30 a.m. on Monday, February 9, in Room 315 Gilman Hall. You will want to study the enclosed material on Professor van der Waerden before the meeting.

This is a short notice indeed. The next day (!), on February 7, 1948, the special committee, chaired by the chemist Alsoph H. Corwin, unanimously approves the mathematics department's recommendation without the usual in academia external letters of reference. The

[25] Francis Dominic Murnaghan (1893–1976), mathematics chair at Johns Hopkins University (1928–1948).

[26] Johns Hopkins University (JHU), The Milton S. Eisenhower Library, Record Group 01.001 Board of Trustees, Series 2, Minutes, May 5, 1947.

[27] JHU, Record Group 01.001 Board of Trustees, Series 2, Minutes, 2/9/1948.

[28] http://www-history.mcs.st-andrews.ac.uk/Mathematicians/Murnaghan.html

[29] Aurel Friedrich Wintner (Budapest, 1903–Baltimore, 1958), one of the leading mathematics professors at Johns Hopkins University (1930–1958).

following morning the Academic Council, also chaired by Professor Corwin, at its special 20-min meeting (8:30 A.M. to 8:50 A.M.) "Voted to suspend its hold-over rule and unanimously recommend to the president the appointment of Dr. Van der Waerden." The same day (!) the Board of Trustees approves the appointment of Professor Van der Waerden to a Full Professorship that pays "$8,000 first year; $9,000 second year; and $10,000 third year."[30]

Surprisingly, Van der Waerden turns this offer down and chooses to return to Holland. Instead of himself he recommends for the position Wei-Liang Chow, his former Leipzig doctoral student (Ph.D., 1936) and coauthor of several of his algebraic geometry papers. Chow will indeed be hired the following year and will serve as a professor at Johns Hopkins University for nearly three decades (1949–1977), including over ten years as the chair.

In 1945, Van der Waerden wanted badly to come to America. He has gotten his wish in 1947. Why then in 1948 does he decide to reject a prestigious, well-paying professorship at Johns Hopkins and leave America? He chooses to return to Amsterdam, where, rightly or wrongly, he has not been treated particularly warmly during 1945–1947. Has his treatment in the United States been worse? I tried – and failed – to find answers in the Archives of Johns Hopkins University. My investigative thread seems to have run into a dead end.

Time passed. One day in my University of Colorado office I glanced at my many books on the shelves and picked one to read at home. It happened to be *Heisenberg's War: The Secret History of the German Bomb* by the Pulitzer Prize winner Thomas Powers [Pow]. It was a great read; moreover, Van der Waerden makes a cameo appearance on the pages of the book. So far, there are no surprises, for we already know that Van der Waerden was a close friend of Heisenberg at Leipzig. However, in this book, Van der Waerden appears as Heisenberg's American pen pal in 1947–1948! The letters are quoted from the 1987 Princeton History Ph. D. thesis of Mark Walker, defended under the supervision of my dear late friend and the founder of Princeton's History of Science Program Charles Coulston Gillispie (1918–2015). I was intrigued, and so I googled and then telephoned Thomas Powers at his Vermont country home. Powers leads me to Walker; Walker sends me copies of the Heisenberg–Van der Waerden correspondence. The surprising answers are hidden in the Werner Heisenberg Archive in Munich, in the unpublished December 22, 1947, letter from Van der Waerden, who is in Baltimore, to his friend Heisenberg at Göttingen. I read in excitement and disbelief:[31]

> Dear Herr Heisenberg,
>
> On the 9th of October I sent you a care package, write to me please if it has arrived and how you are doing with groceries. I would be very glad to send you more next year. I am still in your debt: in the past when I was arrested, you helped me to something much greater, and that is freedom.

[30] JHU, Record Group 01.001 Board of Trustees, Series 2, Minutes, February 9, 1948.

[31] Van der Waerden, letter to Heisenberg, December 22, 1947, Private Papers of Werner Heisenberg, *Max Planck Gesellschaft*, Berlin-Dahlem. I am grateful to Prof. Mark Walker for sharing with me the 1947–1948 correspondence between Van der Waerden and Heisenberg, and Van der Waerden and Goudsmit. I also thank Dr. Helmut Rechenberg, Heisenberg's last Ph.D. student and former Director of the Werner Heisenberg Archive, for the permission to reproduce these materials.

> I need your advice: you are a reasonable man and at the beginning of this war, you predicted who in the end would be the victor. I think I will receive an offer to be a professor in Baltimore, and then I must decide either in favor of Baltimore Johns Hopkins or Holland. In Holland, I would do for the most part applied mathematics and I would train applied mathematicians at the newly founded Math Centrum and at my oil company. I like this work very well and my work at Johns Hopkins I like too, so this [aspect] is equal. The people here are unbelievably nice and helpful: you know that. Nevertheless, I would rather stay in Europe: I love Old Europe and so does my wife.

Thus, Van der Waerden likes his job at Johns Hopkins and considers American people to be "unbelievably nice and helpful." Yet, Bartel and Camilla prefer good "Old Europe." Fair enough, one can relate to that. However, his surprising main concern about living in Baltimore pops up in the next paragraph:

> Now my question: how do you judge the prospects for war, and how do you judge the question whether one could better safeguard one's family in America or Holland if the insanity would break out? The people here and in Europe are telling us that it is crazy, that it is insanity, and that if you have a possibility to stay in America, it is insanity to go back to Holland. Personally, I do not believe there will be a war, but if it nonetheless should come, then a big American city does not seem to me to be the most secure place in the world, but in the past I have been very mistaken in similar cases and do not want to have a responsibility on my shoulders for leading my wife and children to ruin. You understand more about nuclear physics than I do; what do you think about this?
>
> Here I have spoken with different people and gotten a definite impression that America would never start a war on its own, which has set me to rest.

Van der Waerden is afraid that in a large American city – Baltimore – his wife and children could be in a real danger of a Russian atomic attack! This may sound irrational to us looking from today at the year 1947. However, I recall similar fears experienced by Van der Waerden's successor at the University of Amsterdam N.G. de Bruijn, who wrote to me about it in his June 1, 2004, email [Bru8]:

> In 1952 I got a professorship in Amsterdam and . . . I preferred not to live in town but in a village 20 kilometers to the east of it. Nobody would believe now that one reason I had at that time was that in a Russian atomic attack my family would be pretty safe at that distance. A few years later atomic bombs would be hundred times as strong as the Hiroshima type, so the whole argument became utterly silly.

Van der Waerden concludes his Dec 22, 1947, letter to Heisenberg with the hope that Germany will be rebuilt, and they will once again work there together:

> They [Americans] even see in all seriousness a desire to support the reconstruction of Germany, which I am very happy about. Courant thinks that because of the Marshall plan, in some years Germany would once again reach the heights. Maybe we will get together again!

In the March 18, 1948, letter, Van der Waerden informs Heisenberg of his employment choice:

> In principle, I have accepted the job offer from [the University of] Amsterdam.

42.4 Werner Heisenberg's Unpublished Work "On Active and Passive Opposition in the Third Reich"

At about the same time as the letters exchange with Van der Waerden begins, Werner Heisenberg authors but does not publish a document that makes moral principles of this enigmatic man clearer than almost anything else. We have already discussed Werner Heisenberg's morality theory of Kill-One-Save-Ten, which he included in his 1971 book [Hei2]. In fact, that was not the first time he had written about it. Mark Walker was the first to discover it in Heisenberg's Munich archive and discuss [Wal1, pp. 335–338] the November 12, 1947, Heisenberg's unpublished 4-page paper *Die aktive und die passive Opposition im Dritten Reich*, with the subtitle "Written in the context of newspaper reports on the war crime trials in Nuremberg" [Hei1].[32] The paper is attached to the November 11, 1947, cover letter addressed to *Fräulein* Dr. H. [Hildegard] Brücher, a science editor of *Neuen Zeitung* in Munich. To the best of my knowledge, this cover letter has not appeared in print. I wish to present it here in its entirety:

> Dear *Fräulein* Dr. Brücher,
>
> Since you are taking the trouble in such a friendly manner to produce a fair report on the physicists, and since you so readily gave me information on the telephone regarding colleague Dölger, I would like once again to convey to you a wish that this time concerns a political problem.
>
> As you know, a war crimes trial is taking place at this time in Nuremberg against members of the Foreign Office. One of the main defendants is former Secretary of State Baron von Weizsäcker. Since I know Herr von Weizsäcker personally and believe I know his exact political views and know with what intensity he worked over many years to preserve the peace, I am completely convinced that the Nuremberg trial will end with his acquittal after even von Papen and Schacht have been acquitted. (I would like to mention here that back in 1937, when I had been rudely abused by the *SS* newspaper "*Das Schwarze Korps*," I received all the possible support from Herr von Weizsäcker.) For this reason I regret when the press is given one-sided information by the prosecution, and when reports about atrocities committed by the defendants, who have not been verified by any court, are already being published, before the defense has had a chance to say a word. I would be very grateful if in your newspaper you could bring about some moderation. Perhaps it would be more pleasant for the paper not to have published all the charges of the prosecution and then afterwards have to report the news of acquittal. Of course, I cannot foresee the result of the trial with certainty more than anyone else, but for that exact reason I would find it more correct if the newspaper reports were as neutral as possible. If you share this view, I would be very grateful for your support.
> Best regards, also to our common Munich acquaintances,
> Your,
> [signed] *H*

[32] Private Papers of Werner Heisenberg, *Max Planck Gesellschaft*, Berlin-Dahlem. I thank Prof. Walker for sharing with me this document, and Dr. Helmut Rechenberg and the Werner Heisenberg Archive he used to direct, for the permission to reproduce it here.

In the attached to this letter four-page essay, Heisenberg defends the Third Reich Secretary of State Ernst Baron von Weizsäcker, who was facing a Nuremberg Trial. As you recall, the physicist Carl Friedrich von Weizsäcker was Heisenberg's closest friend and a fellow researcher in *Uranverein* ("Uranium Club," a project trying to create atomic bomb and atomic reactor in Nazi Germany). While on the surface Heisenberg refers to "active opposition" of Ernst von Weizsäcker, he seems to count himself among the active oppositionists to the Third Reich too. In his commentary Professor Walker uses a few quotes from this essay. This is an established practice of scholars in history. However, I wish to share with you this entire document so that you can digest it thoroughly and gain your own insight into the fundamental moral positions of mysterious Heisenberg. I will share my view as well. Let us listen to Werner Heisenberg, one of the great minds of the XX century. To begin with, he defines his terms of active and passive opposition.

> If the overwhelming majority of the German people had turned away from the National Socialism immediately in 1933 and had refused every compliance, then a good deal of misfortune would have been prevented. In fact, this reaction did not take place. Rather, the system that in the most clever form knew how to blame its opponents for all of the misfortunes of past years, the system did not find it difficult to win the masses who for the most part lacked judgment. After this happened and after the power lay in Hitler's hands, there was a relatively thin stratum of people, to whom their sure instinct spoke, informing them that the new system was basically bad.
>
> This relatively thin stratum of people only had an opportunity of passive or active opposition. In other words, these people could either say that Hitler's system is basically bad and will lead to a huge catastrophe for Germany and Europe, but I see no way to change anything from inside Germany. So, I am going to exile or in any case I withdraw from all responsibility in Germany and wait until by means of war the system is overcome from outside (overcome by means of war and by means of unheard of war-related sacrifices of goods and blood). I would like to designate this way as the attitude of **passive opposition**.[33] The most extreme part of this group later decided to take part in the war on the side of the allies. Many were simply satisfied to enjoy safety from prosecution in a foreign country.
>
> Another group of people viewed things in the following way. A war, even when its subject is to overcome National Socialism, is such a terrible catastrophe and would cost so many millions of people their life, that I myself must do absolutely everything that is in my power to hinder this catastrophe, or if it has already taken place, to shorten it and to restrict it and to help the people who are suffering as a result of it. Many people who thought this way but did not know the stability of a modern dictatorship, tried in the early years the way of **open immediate resistance** and ended up in a concentration camp.[34] For others, who recognized the hopelessness of a direct attack on the dictatorship, to help suffering people, many of the people who thought this way but did not know the stability of a modern dictatorship, tried in the early years the way of open immediate resistance, and ended up in a concentration camp. For others who recognized

[33] Throughout this letter, the emphasis in bold is added by me for better clarity.

[34] In the next sentence Heisenberg repeats himself, but I am not here to copy-edit his text, and thus I am keeping his repetition.

> the hopelessness of this way, there remained another way, the attainment of a certain degree of influence, i.e., the attitude that had to appear on the outside like collaboration. It is important to be clear that this was in fact the only way to really change anything. This attitude that alone had contained the prospect of replacing National Socialism with something better but without enormous sacrifices, I would like to designate as the attitude of **active opposition**.
>
> On the outside the position of these people was much more difficult than that of the others. Remember, the active opposition had to repeatedly make concessions to the system on unimportant points in order to possess the influence to improve things on important points. In a certain sense he had to play a double game.

Dr. Heisenberg, you must have needed all your brilliant ingenuity to present collaboration *with* the Nazis as "active resistance" *against* the Nazis. Those who were forced out of Nazi Germany, you label as being in "passive [read: worthless] opposition," passive even those who took "part in the war on the side of the allies"! You even insinuate that they *chose* the exile. By 1947 you surely knew, if you did not know *much earlier*, that the Third Reich threw Jews and socialists out of their jobs denied them basic human rights, condoned pogroms, and let them leave without almost any property. Many of these exiles would have chosen to stay in Germany and fight the regime, but why would they risk their lives and freedom for the German masses who viewed these eventual exiles not as fellow-Germans but as alien-Jews or enemies-socialists? And you call this forced emigration a *choice*? Do you believe that Germany was any less theirs than yours? Do you believe the refugees from Germany *chose* to give up *their* country, their language, culture, friends, relatives and go to foreign lands that owed them nothing at all, and a professional job least of all? As once a refugee myself, I understand how unfair your view really is. And later there was no choice, for Germany closed the emigration and opened death and labor camps. Even the lucky survivors were scarred for a lifetime. Ralph Phillips recalls a faculty fired from your Leipzig University, who was lucky to survive and be accepted as a professor of mathematics at Princeton:

> I remember [Salomon] Bochner as a kind and friendly man, still [1939–1940] troubled by scars inflicted by Nazi anti-Semitism.

Those, who actively fought the regime, in your opinion "did not understand the stability of a modern dictatorship, tried the path of open immediate resistance during the first years and ended up in a concentration camp [read: worthless]." The President of West Germany Richard von Weizsäcker, a brother of your closest friend Carl Friedrich von Weizsäcker, disagrees with you. In his moving May 8, 1985, speech in the *Bundestag* during the Ceremony Commemorating the 40th Anniversary of the End of the War in Europe and of National Socialist tyranny, he says:

> As Germans, we pay homage to the victims of the German resistance among the public, the military, the churches, the workers and trade unions, and the Communists. We commemorate those who did not actively resist but preferred to die instead of violating their consciences.

Dr. Heisenberg, you praise "active opposition" as the behavior of the highest morality and ascribe to it collaboration with the Nazi regime, participation in the Nazi capital crimes in order to gain the Nazi trust, and then use it for "replacing National Socialism with something better but without enormous sacrifices." How unbelievably hypocritical it is to term a collaboration with the criminal regime as "active resistance" and put it on a pedestal of high morality! Yours must have been "active opposition" when you collaborated with the Nazi regime in creating an atomic bomb and atomic reactor, in order to achieve, may I ask, exactly what? Create the bomb and thus win trust of and influence on the Nazi government? However, let us return to your essay. You continue:

> One can understand the unavoidable difficult moral problem that was put before the member of the active opposition by means of the following constructed case, to which the reality may well have come close sometimes.
>
> Let us assume that a man wishing to save human life comes into a position where he can really decide about the life and death of other people. And further let us assume – and this is thoroughly conceivable in a really evil system such as National Socialism – that he can prevent the execution of ten innocent people only by means of signing a death sentence for another innocent person. He knows that the ten others will be executed through the action of someone else who will be put in his place if he does not sign the death sentence. The fate of the one is in any case sealed, no matter whether he signs or not, nothing is changed. So how should he act? Personally, I believe upon a conscientious reflection that in such a case signing a death sentence is demanded of **us** [sic], which entails of course **our** readiness to bear the consequences of that personally. Measuring this by the ultimate moral standards, it seems to me that a person who acts and thinks in this manner stands higher than the one who simply says, I do not want anything to do with all of this. Similar problems have occurred in the Third Reich if not always with this intensity.

All right, Dr. Heisenberg, you illustrate your idea of a high moral position by a hypothetical example. You find it acceptable – moreover, highly moral – to prove loyalty to the Nazi criminal regime by signing a death sentence to an innocent person, for this may allow to save other lives. It seems as if you are a theoretical arithmetician, for you justify collaboration with the Nazi regime and complicity in a murder of an innocent person by a simple arithmetic calculation $10 - 1 = 9$. Human life, in my opinion, carries infinite value, and if you were to understand that your arithmetic would have given an *undetermined* result: $10 \times \infty - 1 \times \infty$.

How could such a brilliant intellectual as you not understand that murdering one innocent man constitutes a capital crime and the ultimate collaboration with the Nazis in committing it? How could such a devout Christian man like you play God, even hypothetically, and decide which innocent man is to die and which to live? Indeed, Dr. Heisenberg, you seem to compete with God for employment! How can a conscientious man like you ignore the teaching of the Babylonian Talmud (Sanhedrin 37a):

> For this reason was man created alone, to teach thee that whosoever destroys a single soul . . . scripture imputes to him as if he had destroyed an entire world; and whosoever preserves a single soul . . ., scripture ascribes to him as if he had preserved a whole world.

Re-reading recently *Time within Time: The Diaries* by the great Russian film director Andrei Tarkovsky, I discover that he completely agrees with me. Pondering on the life and fate of Shakespeare's Hamlet, Tarkovsky addresses this very issue, as if he has heard your argument, Dr. Heisenberg, and replies to you [Tar]:

> Can a man judge another, can one man shed another's blood? I do not consider that he can, that he has the right … One drop of blood shed is equal to an ocean. I do not consider that a man has the right to kill another for the sake of the welfare of ten [!] people. If I am told – "Kill that man, and lots of people will be better off!" – I do not consider that I have the right to do so, and I would do better to kill myself, as one of our writers did at a particular moment of his life, after being obliged to sign death warrants. In the end he killed himself. Why … ? No one knows, but it seems to me that it was his inevitable end. The only pity is that he didn't come to that same decision at the moment when he had to sign the first death warrant.

In 1937 you, Dr. Heisenberg, sought – and in 1938 received – protection personally from the *SS Reichsführer* Himmler. Having attracted the high personal attention and patronage of Himmler, you could have hardly allowed yourself as much as a whisper of an opposition during the Nazi rein. But here, after the war, you insinuate in the following paragraph that you – and Ernst Baron von Weizsäcker – were heroes of "active resistance":

> In Germany there was a small stratum of people in high positions who from the beginning belonged to the active opposition and who for a certain amount of time really thought that they could turn the steering wheel of Hitler's policy of war. One of the best known of them is former Secretary of State Ernest von Weizsäcker who already in 1938 used his entire influence to prevent war, but also after the collapse of his political effort in the year 1939, it was self-evident for the small circle of people who "belonged to it" that one could turn to v.Weizsäcker with any good cause and he would listen and that he would help if there was a possibility of success. In many cases he actually became involved and successfully saved and helped people. For this reason it seems to me that it is based on a deep misunderstanding that now v.Weizsäcker as one of the accused for war crimes, stands in front of the Nuremberg Court, while there are so few people on earth who undertook as much as he did to prevent the war.
>
> To make the difficulty of the problem that I am describing clear, it may be permitted to recall a real issue of current politics. Everyone knows that there is a certain danger that the conflicts that have arisen between the East and the West will not be cleared up peacefully and that they could lead to armed confrontation. Everyone knows too that this would mean a terrible catastrophe for humanity. Are people in Russia now doing anything to prevent this catastrophe? Some of those who have openly acknowledged themselves as opponents of the Soviet system and who are now in Russian concentration camps, are completely disengaged, no matter how great our respect for their attitude and our concern for their suffering may be. They don't have the slightest influence on the policies of Russia.
>
> The only ones who can help are people who officially are regarded as Communists and make some concessions to the party line, but in their hearts possess moral standings of the Christian world and secretly do everything to hinder armed confrontation and to make possible a moderation of Soviet policy. We don't know if there are such people on the Russian side. In fact, it is part of the essence of what they are trying to do is not to let anyone know anything certain about them, and they apparently are playing a double game. Nothing would be more damaging to their intention than to have it openly acknowledged that they possess such moral principles.

In Germany we now know in retrospect that there were such people. If there are such people in Russia and if they are successful in their efforts, paradoxically one day they would be regarded as the real Communists as the representatives of policy of international cooperation that was always demanded of Communism. In reality they helped the good to victory and successfully protected the world from a huge catastrophe. But when they fail in their political initiative, should they then be put in front of a court as war criminals because they could maintain their influence on Soviet policy only by means of concessions? I have written these thoughts because the way that the problem of war crimes is being diverted in Nuremberg from the moral plane onto the political plane, fills me with a great deal of worry. One should not discourage the people who perhaps are now conducting in Russia the same desperate battle that in the past von Weizsäcker, von Hassell, Beck and others conducted in Germany.
Göttingen, 12 November 1947
W. Heisenberg

At the end of your essay, Dr. Heisenberg, you applaud smart Stalin's hatchet men. Collaborators and accomplices of the criminal regime are heroes, "active resisters;" dissidents, although merit respect, are stupid and worthless for they do not understand the stability of the regime; and the emigrants and refugees (like I) are worthless passive resisters. These views are not new; I have heard that before from the loyalists of the Soviet totalitarian regime as I was departing the Land of Soviet Promise.

I have got to quote here a passionate letter that the codiscoverer of nuclear fission, unfairly non-Nobeled Lise Meitner, wrote in late June 1945 to her coauthor Nobeled Otto Hahn. She addresses here Hahn, Heisenberg, and other scientists who worked for the Third Reich, and without even reading this Heisenberg's "Opposition" manuscript (as Heisenberg would write it two years later) she powerfully rebuts Heisenberg's pretense of any resistance, even a passive one [LS, p. 310]:

> You all worked for Nazi Germany and you did not even try passive [!] resistance. Granted, to absolve your consciences you helped some oppressed person here and there, but millions of innocent people were murdered and there was no protest. I must write this to you, as so much depends upon your understanding of what you have permitted to take place. Here in neutral Sweden, long before the end of the war, there was discussion of what should be done with German scholars when the war was over. What then must the English and the Americans be thinking? I and many others are of the opinion that one path for you would be to deliver an open statement that you are aware that through your passivity you share responsibility for what has happened, and that you have the need to work for whatever can be done to make amends. But many think it is too late for that. These people say that you first betrayed your friends, then your men and your children in that you let them give their lives in a criminal war, and finally you betrayed Germany itself, because even when the war was completely hopeless, you never once spoke out against the meaningless destruction of Germany. That sounds pitiless, but nevertheless I believe that the reason that I write this to you is true friendship. You cannot really expect that the rest of the world feels sympathy for Germany. In the last few days one has heard of the unbelievably gruesome things in the concentration camps; it overwhelms everything one previously feared. When I heard on English radio a very detailed report by the English and Americans about Belsen and Buchenwald, I began to

cry out loud and lay awake all night. And if you had seen all those people who were brought here from the camps. One should take a man like Heisenberg and millions like him, and force them to look at these camps and the martyred people.

Indeed, how could these brilliant scholars, by their silence and their work support the Nazi brutes gloating with cynicism, erecting "*Arbeit macht frei*" above the gates of Auschwitz and Dachau, Gross-Rosen and Sachsenhausen, Fort Breendonk and Theresienstadt? "Works makes one free"? Free in the slave labor of the Nazi concentration camps? Did these great minds approve of Buchenwald's "*Jedem das Seine*"? Everyone gets what one deserves? Do the innocents deserve torture and death, Professor Heisenberg?

I wish to note here, that, to my regret, the high moral authority of the Nazi years' Germany, Nobel Laureate and Einstein's friend Max von Laue, added his insult of exclusion and mistrust to the Nazi injury of Samuel Goudsmit when in 1948 he wrote a response to Goudsmit's book *Alsos* and its December 1947 review [Mor1] by Professor Philip Morrison of Cornell University [Lau]:

We do know that Goudsmit lost not only father and mother, but many near relatives as well, in Auschwitz and other concentration camps. We realize fully what unutterable pain the mere word Auschwitz must always evoke in him. But for that very reason one can recognize neither him, nor his reviewer Morrison, as capable of an unbiased judgment of the particular circumstances of the present case.

Earlier Heisenberg expressed the same opinion as von Laue that victims of Nazism, such as Goudsmit, have no right to be arbiters of the Nazi regime:[35]

Goudsmit's position can be explained only by the fact that he lost his two parents in Auschwitz and naturally is embittered toward Germany. It is at least understandable and pardonable that he finds it difficult in his bitterness to make a distinction between the different people of our country.

In my opinion, Morrison is absolutely correct in his powerful rebuttal of von Laue [Mor2]:[36]

I am of the opinion that it is not Professor Goudsmit who cannot be unbiased, not he, who most surely should feel an unutterable pain when the word Auschwitz is mentioned, but many a famous German physicist in Göttingen today [i.e., Heisenberg], many a man of insight and responsibility, who could live for a decade in the Third Reich, and never once risk his position of comfort and authority in real opposition to the men who could build that infamous place of death.

As to Heisenberg's concept of moral superiority of the German physicists over the Allied scientists, it is best refuted by Philip Morrison in his December 1947 review [Mor1] of Goudsmit's *Alsos*:

The documents cited in *Alsos* prove amply that, no different from their Allied counterparts, the German scientists worked for the military as best their circumstances allowed. But the difference, which it will never be possible to forgive, it that they worked for the

[35] Quoted from [Wal1], p. 340.

[36] Rebuttal, which was not published in Germany [Wal1, p. 360].

cause of Himmler and Auschwitz, for the burners of books, and the takers of hostages. The community of science will be long delayed in welcoming the armorers of the Nazis, even if their work was not successful.

Regretfully, Morrison's prediction is not materialized. Very soon, in 1950 – and again in 1954 – Werner Heisenberg is invited for a VIP lecture tours of the United States. On May 14, 1958, he is made a Foreign Honorary Member of the American Academy of Arts and Sciences. Heisenberg is offered a number of jobs in the United States, as are many Third Reich scientists and engineers. America is acquiring ammunition for the Cold War and paying for it a high moral price.

The clearest example of American hypocrisy is secretly bringing in the leading German rocket scientist Wernher von Braun (1912–1977), a member of the Nazi party and an officer in the *SS*, and his associates. They are brought in not for trial – but for building American rockets. Von Braun titles his autobiography "I Aim for the Stars," but he should have added "But Sometimes I Hit London," as is suggested by the American mathematician, pianist, and songwriter Tom Lehrer, who wrote a satirical song "Wernher von Braun":

Gather round while I sing you of Wernher von Braun,
A man whose allegiance is ruled by expedience.
Call him a Nazi, he won't even frown,
"Ha, Nazi schmazi," says Wernher von Braun.

Don't say that he is hypocritical,
Say rather that he's apolitical.
"Once the rockets are up, who cares where they come down,
That's not my department," says Wernher von Braun.

Some have harsh words for this man of renown.
But some think our attitude should be one of gratitude,
Like the widows and cripples in old London town,
Who owe their large pensions to Wernher von Braun.

You too may be a big hero
Once you've learned to count backwards to zero
"In German oder English I know how to count down
Und I'm learning Chinese," says Wernher von Braun.

Now that you have read the lyrics, enjoy Tom Lehrer performing his song; I found it for you: https://www.youtube.com/watch?v=TjDEsGZLbio

During the prewar visits of the United States, Heisenberg stayed at Goudsmit's home; they were old friends and shared many common friends in the world of leading physicists. Yet, the friendship between Heisenberg and Goudsmit was never quite renewed after the war. Yes, Heisenberg was upset over Goudsmit's criticism, especially unfair criticism of his war time physics efforts. But after giving it much thought, I see elsewhere the major reason for Heisenberg's displeasure. Goudsmit unearthed Heisenberg's pleas for help to Heinrich Himmler. The two of the most notorious Nazi murderers, Himmler and Heydrich, granted their cover to Heisenberg, and this had to be extremely embarrassing for Heisenberg when these letters appeared in Goudsmit's book *Alsos.*

Elisabeth Heisenberg in her memoirs [HeiE, p. 112] states that “Goudsmit later regretted having written the book, and apologized to Heisenberg for it; nevertheless, the book is one of the reasons for Heisenberg’s character falling into such ill repute.” Goudsmit had regrets but not due to writing the book. Werner and Elizabeth Heisenberg’s son, Physics Professor Emeritus at the University of New Hampshire, Jochen Heisenberg writes to me on February 4, 2011:

> Dear Alexander Soifer,
>
> During the time my mother wrote her book I was already living and teaching here in the US. Thus, I do not know the details of that apology. However, at a meeting of the APS [American Physical Society] in Washington D.C. that I attended, Goudsmit had asked to meet me. At that meeting he apologized to me for the difficulties he had caused to my father, his family, and also to me. This, however, had been a different incident, and in this conversation the book *Alsos* was not mentioned in a particular way.

As the scientific head of Alsos Missions, Goudsmit was instrumental in identifying the ten German scientists, who were held in the Farm Hall (a manor) near Cambridge, England, for exactly six months, and then released to live anywhere in Germany, except the Russian and French zones of occupation. Goudsmit must have felt responsibility for denying Heisenberg’s and other families their bread providers and causing them separation and hardship, and for that he apologized.

The ten distinguished scientists, including Werner Heisenberg, Carl Friedrich von Weizsäcker, Max von Laue, and Otto Hahn, were kept in captivity, in fine conditions, without being charged with any crime. The captives could have demanded to be charged or else released, but they probably realized that they just might get what they would ask for and be charged and tried at Nuremberg trials for their contribution to the German war efforts. And so, they did not object (except Heisenberg appropriately demanding that his wife and six children be taken care of). The British wanted to prevent these leading German scientists and their atomic bomb and reactor research from falling into the Russian hands – the Cold War has begun – or even the French hands. And this is how, in my opinion, this strange compromise of detention came about.

Heisenberg’s 1948 *New York Times* interview (in English) reveals his surprising to me interpretationof patriotism:[37]

> German sciences sank to a low ebb. I think I am safe in saying that, because of their sense of decency most leading scientists [in Nazi Germany] disliked the totalitarian system. Yet as patriots who loved their country they could not refuse to work for the Government when called upon.

These words explain the rationale of Heisenberg life’s choices. When his government – even the criminal Nazi government! – calls upon him, Heisenberg and “most leading scientists,” out of “their sense of decency” – *decency!* – “could not refuse to work for the Government”! He subscribes to the widely shared but false notion of patriotism, according to

[37] Kaempffert, W., “Nazis Spurned Idea of an Atomic Bomb,” *New York Times*, Dec. 28, 1948, p. 10.

which in times of war a true patriot has to rally behind his government, even if this government engages in ostensibly criminal activities.

As a civilized writer from a civilized country, I am expected to spare you discomfort, my reader. Yet, I have got to give you – and Werner Heisenberg – the taste of what blind love of the country and "patriotic" obedience can produce, an opportunity to touch the evil, to quote however briefly from the 1946 voluminous ca. 500-page document, *The Black Book: The Nazi Crime against the Jewish People* [BB].

The following quote is short but extremely disturbing, and so I will let you decide whether to read or skip it. It describes some of the countless German atrocities in the Majdanek Concentration Camp, which became known as *Vernichtungslager* (extermination camp), where people were murdered on an industrial scale [BB, p. 384]:[38]

> Heinz Stalbe, of the German *Kampfpolizei*, stated at a plenary session of the Commission[39] that he himself saw the director of the crematorium, *Oberscharfürer* Mussfeld, tie a Polish woman hand and foot and throw her alive into the furnace. Witnesses Jelinski and Oleh, who worked in the camp, also tell of the burning of living people in the crematorium furnaces.
>
> "They took a baby from its mother's breast and killed it before her eyes by smashing it against the barrack wall," said witness Atrokhov.
>
> "I myself," said witness Edward Baran, "saw babies taken from their mothers and killed before their eyes: they would take a baby by one foot and step on the other, and so tear the baby apart."

Dr. Heisenberg, in the waning hours of 1948, when the German crimes against humanity have been thoroughly established at Nuremberg trials and other courts and documented in many books and reports, you are telling the New York Times that as a decent [sic] and loving patriot you "could not refuse to work for the Government." Could you refuse your share of responsibility for what your government has done on behalf of all Germans, on your behalf, Dr. Heisenberg?

42.5 Professorship at Amsterdam

By 1948, the de-Nazification of the Netherlands was over, and the institutions *College van Herstel* were gone. In addition, the American acceptance improved Van der Waerden's standing in Europe. However, L. E. J. Brouwer eloquently objects to Van der Waerden's appointment in his April 15, 1948, letter to the Minister of Education, Culture and Science Jos Gielen; moreover, Dirk van Dalen observes that "the feelings expressed in this passage perfectly reflected the general opinion of the Dutch, and in particular the students, in the matter":

[38] The quoted material was included in *The Black Book* [BB] from the statement of the Polish-Soviet Extraordinary Commission for the Investigation of Crimes Committed by the Germans in the Extermination Camp of Majdanek in the Town of Lublin.

[39] See the previous footnote for the description of the Commission.

> From a researcher like Professor Van der Waerden, who is only theoretically, but not experimentally active, the scientific influence is almost independent of personal presence. Thus, as soon as a materially and scientifically favorable position has been secured, the question of his presence here in the country loses all scientific and national importance, and it becomes almost exclusively a matter of national prestige. From a viewpoint of national prestige, the motivation of his appointment here in the country seems, however, extremely weak to the undersigned. For if it is claimed that by the presence of Professor Van der Waerden in Amsterdam the strength of our nation is enhanced, the reply is forced upon us that in that case the national strength of the German empire has been enhanced during the whole period of the Hitler regime by the presence of Professor Van der Waerden in Leipzig. And if it is argued that if Professor Van der Waerden is not offered a suitable position in the Netherlands, this will be done by America, the reply is forced upon us that if at the moment there are positions open to Professor Van der Waerden in America, this should not have been less the case between 1933 and 1940, when many prominent and right-minded German scholars and artists were welcomed with open arms in America, and that therefore one has to assume that Professor Van der Waerden had not felt the desire to turn his back on the Hitler regime.

The Dutch Professor Herman Johan Arie Duparc (1918–2002) wrote down for me the following recollections of the year 1948 during our September 1996 meetings in his apartment in Delft [Dup]:

> Van der Corput and others feared again difficulties. He said to me: "Tomorrow vd Waerden gives his first lecture; interesting; let us go there." So, we went there. There were no difficulties . . .
>
> Then Van der Corput and vd Waerden had a common room in Amsterdam University. When vd Corput went to the US in September 1950, I had to take over his work and met vd Waerden regularly there.

According to Duparc, in 1948 Van der Waerden was appointed a *bijzonder* (special) professor of applied mathematics at the University of Amsterdam. This part-time ("one day a week," according to Duparc) position was paid by the Foundation, and thus did not require an approval by the Queen. This was a far cry from a tenured full professorship at Johns Hopkins University that Van der Waerden turned down, but this was a start. Plus, this time Van der Corput hired Van der Waerden as a full-time director of applied mathematics at Amsterdam's *Mathematisch Centrum*, where Van der Waerden worked part-time in 1946–1947.

How good a professor was Bartel L. van der Waerden at the University of Amsterdam? This is a hard question for us to answer in 2023, except by good luck or providence. Dirk van Dalen, my good luck, was Van der Waerden's student at Amsterdam during the fall 1950 semester, and so he could answer this question for you and me. In his January 14, 2011, email, Dirk recollects:

> My own memories of Van der Waerden are rather limited. I took his analysis course as a freshman, and the next year he was in Zurich. He was a gifted teacher, if you heard his lecture the material became quite clear. His style was, if I may compare it, like that of the [analysis] book of Courant. I guess that this was the general continental style. One thing was rather unusual: when a new edition of his *Modern Algebra* came out, he offered

> students a copy for a reduced price. Later he told me that he had made this a condition with the publisher. So that is when I got my copy.

Yes, we can all relate to the gift of clarity in Van der Waerden's expositions from reading his many books. We also learn here that Van der Waerden cared about his students and even arranged for a student discount with Springer.

Then there came a prestigious membership in the Royal Dutch Academy of Arts and Sciences, which had to be – and was – approved by the Queen. This, however, was not the same Queen Wilhelmina, who in 1946 rejected civil service appointments for Van der Waerden and others who voluntarily worked for the German occupiers. Her daughter Queen Juliana, who took over in 1948, presided over less principled and less emotional times. In 1949, Dr. van der Waerden became a member of the Royal Dutch Academy of Sciences.

On January 31, 1950, Minister Rutten informed Van der Corput that he would have no objections to a "real" university professorship. Nearly four months later, on May 16, 1950, the Mayor and Aldermen of the City of Amsterdam appointed Van der Waerden to a professorship, contingent on the Royal assent, which came on June 19, 1950. And so, five years after the war's end, Van der Waerden is finally appointed to a full professorship at the University of Amsterdam, effective October 1, 1950. It appears that the relationship between Holland and her prodigal son Bartel has been restored and would likely grow closer with time. Van der Waerden has a fine job and talented and very supportive colleagues. Yet, he chooses to leave his Homeland and accept a chair at the University of Zurich.

Van der Waerden de facto includes his notice of resignation in his inaugural [sic] speech "Concerning the Space" [Wae12], given on Monday, December 4, 1950, at 4 o'clock in the afternoon at the University Auditorium:

> Eminent Clay and Van der Corput,
>
> With undaunted energy you both have organizationally prepared my appointment to a Professor regardless of all difficulties and you have finally reached your goal. I appreciate this very much and will remain grateful to you forever for it. Even though now I will soon be going to Zurich, I trust that someone else would take over my job on this faculty, which was organized by your ideas.

On March 21, 1951, Professor Van der Waerden formally asks for his resignation from the University of Amsterdam, which is granted effective May 1, 1951.

Van der Corput has been proven wrong: he did all he could to support Van der Waerden in academia and in government; he closed his eyes on his disagreements with some of Van der Waerden's moral positions and life's choices; and yet, in the end he has not won Van der Waerden for Holland. Nicolaas G. de Bruijn, who in 1952 became that "someone else [to] take over [Van der Waerden's] job on this faculty," writes to me about the understandable disappointment of Van der Waerden's mathematical colleagues in the Netherlands [Bru7]:

> I had regular contact with some mathematicians who knew him [Van der Waerden] better than I did, like Kloosterman, Koksma, Van Dantzig, Freudenthal, Van der Corput, who were disappointed by his leave after they had gone into so much trouble to help him with jobs in the Netherlands.

De Bruijn continues [Bru8]:

> Actually, I do not remember anything from my own experience. I only remember that people like Koksma, Van Dantzig, Schouten confidentially complained that Van der Waerden disappointed them after all the trouble they had taken. I suppose they had to fight unwilling authorities in order to let them forget the objections from the past. Step-by-step they got him a position with the Shell Company, a part-time professorship at the University of Amsterdam, the membership of the Royal Dutch Academy of Arts and Sciences (which had to be signed by the queen) and finally the full professorship. The people who all went through this trouble of course felt they lost their face with respect to all those authorities when Van der Waerden unexpectedly left them in the lurch . . .
>
> As a part-time professor Van der Waerden taught applied mathematics, maybe mainly from a pure mathematician's perspective. As a full professor he had not even started; around that time he decided to leave for Zurich. So there was hardly a Van der Waerden tradition of courses in Amsterdam.

Amsterdam appears to have been used by Professor Van der Waerden merely as a steppingstone in his career.

During the Dutch years of Bartel L. van der Waerden, 1945–1947 and 1948–1951, the brothers Bart, Coen, and Ben and their families were close. Bartel's family was given the gorgeous parents' house *Breidablik* to live in – until the time came to sell it. Ben as a brother and a lawyer stepped in to help Bart when the latter had difficulties with obtaining a visa from the United States Consulate due to Bart's attempt to conceal his years of living in Nazi Germany. We witness a noble, brotherly defense, and Ben's desire to explain away Bartel's incorrect information given to the American Consulate. We also see how bitter Bartel van der Waerden still is in 1949, four years after his return home, and three years after a very public debate of his life's choices on the pages of *Het Parool*. "My brother doesn't wish to hear anything more about this issue or to discuss it," writes Benno van der Waerden. Bartel would rather not go to Seattle than discuss his life in Nazi Germany.

Another Beno, Professor Beno Eckmann of ETH and Bartel's Zurich friend for nearly half a century, 1951–1996, writes to me that Bartel and Camilla van der Waerden always avoided any mention of their time in the Third Reich [Eck0]:

> We never really talked about his time in Leipzig, in any case not about politics. He and his wife seemed to avoid these themes.

Bartel's persistent decades-long silence about his Nazi years seems to convey his regrets or embarrassment louder than any words could. As to Seattle, Bartel and Camilla were granted American visas; we see their happy faces on the photo taken in Seattle in 1949 in [Dol1]. Their son, Hans van der Waerden, kindly contributes his view of his father's Dutch postwar years [WaH1]:

> It would have been impossible for a Nazi collaborator to get a professorship in Amsterdam at any time after 1945.
>
> Of course, this could not be done without doubts and hesitations. The resistance against my father's appointment was a very natural and logical one: my father could not expect Dutch authorities to act as if nothing at all happened! The mere fact that he had served, though indirectly, a government that suppressed his compatriots, could not but arouse a wave of suspicion. But the fact that – after not more than five years – he was again trusted [with] a responsible public position, shows that the suspicions obviously could not be verified in any detail.

42.6 Escape to Neutrality

Mathematics has no Fatherland, you say?
– Het Parool Editors, *1946*

On my request, Professor Beno Eckmann recollects the *Universität Zürich* 1950 succession [Eck1]:

> In 1950 Fueter retired. Shortly before I was offered that position (and to be "director"). Then the position was offered to vdW [Van der Waerden] who accepted but his appointment was finalized only in 1951 (I vaguely remember that there were discussions among Zurich authorities whether it would be appropriate to appoint a man who had remained in Nazi Germany during the war).

In fact, Eckmann is the early first choice [Eck2]:

> I was asked either in 1949 or early in 1950 whether I would accept (I really cannot remember when this happened – Rolf Nevanlinna talked to me personally, had I said yes, I would have received that position).

I am holding in my hands a voluminous file of Rudolf Fueter's succession. It opens with *Dekan* of Philosophical *Facultät* II Hans Boesch's May 5, 1950, letter calling the meeting of the Mathematics Commission for Monday, May 8, 1950, at 1400 hours in *Dekanat* room 13. Next there comes a mysterious page containing only names and numbers, the handwritten in pencil super-concise stenography of this meeting that would delight any professional or amateur paleographer – let me try my hand on it. The Commission considers young Swiss mathematicians, such as Nef, Häfeli, and others, but only three candidates are numbered, clearly in the order of ranking:

1. Van der Waerden (03), Ord. Leipzig, Hollander
2. Pólya (62) [should be 1887], Stanford University
3. Eckmann (17), ETH

References, which are to be asked to evaluate the above candidates, are also listed on this page:

> Fueter, Speiser, Hopf, Ahlfors, ~~Erhard Schmidt~~, Schouten

At the bottom of the page the final list appears again, without the stricken Erhard Schmidt of Germany. Schouten's name is separated by a line from the other four names, with an arrow going from Finsler to Schouten, for the latter is to be asked by Finsler only about the current political opinion about Van der Waerden in Holland.

The following day Dekan Boesch sends identical letters to Professors Van der Waerden and Pólya, inquiring whether they would like to be considered for professor and director of the mathematics institute in succession to the retiring Professor Fueter. What the file is missing, is telling as well: it does not contain a similar letter to Professor Eckmann of ETH: he has already turned down this position, for he has been quite happy at the ETH, where he will later found *Forschungsinstitut für Mathematik*.

On the same day Dekan Boesch also sends letters to the four official references. The long-term Van der Waerden pen pal on matters of algebraic geometry (at least since 1936), Professor Paul Finsler, and not the *Dekan*, writes to the fifth, personal reference, Jan A. Schouten of the Netherlands.

Shortly, letters of reference pour in. ETH Professor Heinz Hopf recommends considering only the top three candidates, Pólya, Van der Waerden, Eckmann. He showers all three with high praise. Lars V. Ahlfors, Chairman of Mathematics at Harvard University, expresses an opinion similar to Hopf's. Professor R. Fueter, whose seat is the object of this search and whose influence as the past *Rektor* is very strong, shockingly, has nothing positive to say about George Pólya. Reading his letter, I wonder why on earth they invited Pólya to apply:

> Prof. Dr. Pólya, during his first years in a Zurich position [at the ETH] attempted to work together with us, but then in many situations worked against Speiser and myself and fought with our students. In this situation I would also like to point out some of Prof. Speiser's views regarding this.

Fueter much prefers Van der Waerden or else one of his own former doctoral students. Fueter knows that Eckmann has already refused the position, and so he writes nothing about Eckmann in his letter. Andreas Speiser praises Van der Waerden and the young Swiss candidates and shockingly and unfairly puts down Pólya as a *mathematician*:

> Of the foreigners Pólya does not even come into view. He has dealt with an enormous amount of small problems but has never seriously worked in a serious area and would rapidly sink the level of mathematics at the University. Opposite to this, Van der Waerden is an apt (*trefflicher*) mathematician, whom one would have to recommend.

Evaluating Pólya unfairly is not the only deplorable aspect of Speiser's letter. Following praise for the (Jewish) mathematician Richard Brauer, Speiser uses – in the year 1950! – the *Nazi Deutsch* to describe Brauer as "not Aryan (*nicht arisch*)". Truly, old habits die hard!

Summing up, Professor Van der Waerden is the unanimous choice of the four references. Only one question remains: has Van der Waerden been sufficiently "purified"? It is to be answered by Professor Schouten. The latter sends his handwritten reply to Professor Finsler on May 12, 1950. It deals exclusively with Van der Waerden the person, and not at all with his mathematical work. The following is its complete text:

> Dear Herr Colleague!
>
> I have received your friendly letter of May 9. A few weeks ago Herr Van der Waerden has been named *Ordinarius* in Amsterdam. Political reservations do not apply here [in the Netherlands] against him. I should actually say that they do not apply anymore, because certain circles had earlier tried completely without justification to raise their voice against him. But that has all now passed and he is also now a Member of the Royal Amsterdam Academy.
>
> Even though I hope that you will not snap this man away from us, I must absolutely tell you my opinion that he is completely politically harmless (*unbedenklich*).
>
> With friendly greeting to the entire Zurich circle,
>
> Yours most respectfully
>
> J. Schouten

Thus, Professor Van der Waerden is cleared for the Swiss employment again. The Mathematics Commission consists of Professors Paul Karrer, Paul Niggli, Paul Finsler, Rolf Nevanlinna, and Walter Heitler. They meet on June 3, 1950, and end up with exactly the same slate and order of the three candidates they started with. On June 9, 1950, Dekan Boesch reports the faculty findings to the Education Directorate (*Erziehungedirection*) of the Canton of Zurich in a 5-page letter. He reserves the highest compliments for "Herren Van der Waerden, Pólya and Eckmann [who] would be the candidates for this Mathematics Professor position, whereby Herr Van der Waerden would be in first place, (Herr Pólya in second place)." Professor Van der Waerdengets a clean bill of political health from Dekan Boesch (ibid):

> Certain problems found in Herr Van der Waerden's working at Leipzig University during the war which were focused on by Holland are no longer applicable according to the communication that Prof. Schouten has forwarded. On the contrary, it is explicit from the [Zurich University] Faculty proposal for filling a new position of Professor of Applied Mathematics dated July 15, 1946, that Prof. Van der Waerden was thoroughly considered.

As we know, Eckmann turns down the offer before the search began; Pólya is rejected by Fueter and Speiser, who certainly knew in advance that they did not wish Pólya back in Zurich. From day one of the search, Van der Waerden has been listed as number one candidate. Thus, the elaborate smokescreen of a search seems to have been invented to satisfy the rules but has had only one goal from the beginning – to hire Van der Waerden. He is offered the job on September 20 and accepts it with "heartfelt gratitude" on September 24, 1950.

Van der Waerden could realize his Swiss dream right away, without spending another year at the University Amsterdam. Apparently, *he* does not agree to an early Zurich start. I can venture a conjecture: Van der Waerden desires a vindication for the *Het Parool*ean humiliation, and the Amsterdam full professorship with its Inaugural Lecture ceremonies in December 1950 provides such an opportunity. Van der Waerden wants to leave his Homeland but leave it as a winner, by willingly giving up Holland's highest academic credentials he has finally earned.

For a decade I have been absorbed with the following question: why did Van der Waerden leave Holland for good in 1951? Was the University of Zurich (which, in my opinion, was no match to its famed neighbor ETH) a better place than the University of Amsterdam? This was not at all obvious to me, and so I asked N.G. de Bruijn, who replied as follows [Bru8]:

> We were looking at the U.S. and Switzerland as a kind of paradise. Whether in the long run Zurich would be much better than Amsterdam may be open to discussion. In 1950 Amsterdam had lost the glory of Brouwer's days of the 1920's . . .
>
> By the way, I really do not know the order of the events. The offer from Zurich may have come at a time when the procedures for getting him the full professorship at Amsterdam had hardly started. He may have kept the Zurich offer secret for a time, in order to keep both possibilities open. If it had happened to me, I would have felt a moral pressure against letting Amsterdam down.

Yes, Nicolaas de Bruijn would not have let Amsterdam down. Why didn't Van der Waerden feel "a moral pressure against letting Amsterdam down"? H. J. A. Duparc recalls and writes it down for me [Dup]:

> Van der Waerden's wife, Rellich, was German and had many difficulties in normal life in Holland because of her speaking the German language (Holland was occupied 5 years by the Germans).

N. G. de Bruijn [Bru9] adds:

> Justified or not justified, those anti-German feelings were very strong indeed. I can understand that Camilla was treated as an outcast, and that she therefore disliked living in Holland.

Hans van der Waerden, the son of Bartel and Camilla, gives us a most thoughtful, psychological, and convincing explanation [WaH1]:

> Why did my parents leave Holland for Switzerland? The reason my mother told me was that she could not stand the rainy, windy Dutch weather. I don't think that was all. I imagine, my mother did not feel at home for language reasons as well: she had to learn to speak Dutch, and by her accent everyone could instantly recognize her German (or Austrian) origin, which after 1945 was compromising and made her feel uneasy.
>
> Furthermore, Switzerland at that time had a reputation as almost a paradise: sound landscape, sound towns, a sound politics (so it seemed to be), sound economy… This, I suppose, was extremely tempting. I imagine – please take this as my imagination, not more – that my parents longed to live in a new society, where they were no longer confronted with this perpetually underlying question: "Have you been, or have you not been a Nazi collaborator?"
>
> This most probably applies to my mother, but also, in some deeper sense, to my father, who was extremely vulnerable to accusations of this kind. For him, <u>living life in honor and moral integrity was the most important thing on earth</u>, more important than material comfort, relations, or even scientific research. He was a dogmatic about that. That is why suspicions of the kind mentioned above – that he could ever have "collaborated" or at least "contributed" to such a horrible thing as the Third Reich – not only saddened or infuriated him: they shook him to the roots of his personality.
>
> This very vulnerability, besides, probably made him react in a somewhat naïve or helpless way to the feelings of his countrymen after 1945. He could not allow himself to admit that perhaps there were good reasons for negative feelings against him because of his behavior, because of moral.

Bartel van der Waerden's niece, Dorith van der Waerden, conveys family memories and impressions, so intangible and yet so helpful for our psychological insight [WaD1]:

> After the war Bart wanted to come back to Holland when this was possible again. He moved into the house of his parents in Laren. During these few years that they lived here we stayed there during the holidays and were in good relations with them as far as I remember.
>
> But he and his family must have had a very hard time. In the first place he was suspected because of coming from Germany. I never heard about it that he was scrutinized by a committee, but this is very likely because that is what happened to everybody about whom there were doubts about their behaviour during the war. The ones that had actually helped the Germans went to prison and camps and so on, and suffered for many years because they and their families were not accepted.
>
> I believe that they found that absolutely nothing was wrong. I know my father [Benno] was convinced of that, and he did everything he could to help him [Bartel] and clear him of accusations. I have one letter proving that. And I happen to know my father, he was so very honest, he would never have helped Bart if he had any doubts.

> But still Bart had a German wife, and children that came here and went to school but spoke with a German accent. In the years after the German occupation this was not accepted in Holland, and they all suffered a great deal. And probably the period of doubts about him took a long time too.
>
> So, the whole family must have been very hurt. Finally, they could escape from Nazi Germany and now people here thought about them as Nazis. Though his brothers supported him, there were other family members, cousins, who did not forgive him for staying there.
>
> In this marriage of Bart and Camilla she was the practical one, she took care of everything, so he could do his work. She protected him also against the outside world, I think.
>
> So, I believe that they were so unhappy in Holland that they looked for another country again, not Germany of course, and found Switzerland where the spoken language was also German. This of course was easy for Camilla and the children, and there no one would think they were not okay for speaking German.
>
> I believe that for Camilla the period in Holland must have been so painful that she never wanted to have anything to do with Holland and even with Bart's family. She wanted it to be over and forget about what to her seemed utterly unfair toward her husband, herself, and her children.

While the role of Camilla in the decision to leave Holland must have been very significant, such an important step was ultimately Bartel's to make – he was the one who almost simultaneously accepted two job offers, from Amsterdam and from Zurich.

At all times he desired to be at the best place for doing mathematics, which according to him has now moved to Switzerland and the United States. Which one should he claim? He aspired to belong to the German culture; it was important – perhaps too important – to him. The decision to move to Switzerland was the last key decision of Van der Waerden's life and career. He chose to leave the Motherland of Suffering for the Land of Neutrality, the Land of the German Language but not Germany.

42.7 Zurück nach Zürich

Van der Waerden aspired to be part of the German culture, live in a land of German language, and his desire is granted. He arrives in Zurich with his wife Camilla and children Ilse and Hans.

In search for information, I approach Van der Waerden's close personal and professional friend of his forty-five Zurich years, ETH Professor Beno Eckmann, who on December 7, 2004, generously shares information with us [Eck0]:

> Yes, I knew vdW [Van der Waerden] very well, until his death. But I met him only after he came to Zurich. I and my wife saw him and his wife at various occasions, mathematical and private... His interests moved later from Algebraic Geometry to Probability and then to History.

During the May 4, 1993, interview [Dol1], Camilla van der Waerden told the interviewer and her husband: "I have always preferred that he were more involved in mathematics. He didn't do it. I have always said he spends too much time on history and truly too little on mathematics." Camilla is correct: mathematically the Third Reich years were more productive for Van der Waerden.

Van der Waerden writes with an impressive breadth and fine detail series of historical books, *Science Awakening II: The Birth of Astronomy*, 1974; *Geometry and Algebra in Ancient Civilizations*, 1983, [Wae22]; and *A History of Algebra*, 1985, [Wae23]. While in Zurich, Van der Waerden was in touch with great physicists while editing the important 1967 source book of quantum mechanics. In 1973, van der Waerden retires from his chair at Zurich at the mandatory retirement age of 70 years.

On my request, Bartel's son, Hans van der Waerden contributes a valuable perspective in his June 20, 2004, letter from Switzerland [WaH1]:

> There can be no doubt about my father's unshakable anti-fascist convictions – I think, in this we agree. I remember him, till the end of his life, becoming furious, when anybody dared to compare irresponsible political activities of whatever kind with Nazi crimes. I further remember that – sometime about 1980 – he declined an invitation to Leipzig University, saying: "I have lived long enough under dictatorship, I need not see any more of it." From his father, who at his time was a socialist rejecting bolshevism, he had inherited a strong conviction that one-party-government is the worst kind of government at all – he used to quote an article of his father on this issue, and my mother recalled that when she and her mother had been listening to Hitler speeches on the radio, and the two of them got into doubts, asking "Couldn't there be perhaps some truth in it, anyway?" – my father vividly explained to them that Hitler was wrong in every respect.

Bartel Leendert van der Waerden, ca. 1980, *Courtesy of Leipzig University*

In 2010, Hans van der Waerden adds [WaH2]:

I remember him as a perfectly honest man with a high, sometimes almost fanatical sense of duty and moral integrity; and as an extremely modest man, never jealous of other people's achievements or liable to exaggerate his own.

Indeed, Bartel van der Waerden strives to be a highly moral individual, a fitting member of his great family of Dutch public servants. Later, after 1935, we witness his compromises with the Nazi authorities, instances of insensitivity, declared desire to save the German culture and little effort to contribute to the culture of his Motherland that has been served with such a high distinction by the rest of the Van der Waerden family.

He clearly sees Nazi Germany for what it is. In the early years of Nazism, he criticizes the regime. But the regime easily finds Van der Waerden's soft spot: his clinging to a German professorship. Once warned not to interfere in the German "internal" affairs or lose his professorship, Van der Waerden no longer speaks out publicly in Germany.

Van der Waerden writes to Van der Corput, "Germany attacked the Netherlands and shamefully abused it, and the whole German people are also responsible for that." Exactly right, Bartel. However, you too lived in that Nazi Germany the entire 12+ years of the Third Reich, and retaining Dutch citizenship is a lame excuse. You ought to accept your small share of responsibility for what your Germany did *on your behalf*, with your silent approval, to the German people and the peoples of your beloved Europe.

I faced a similar Hamletian question in the Soviet tyranny: to leave or not to leave, to go to barricades or to the airport? It was unbearable to see in August 1968 how *on my behalf* my country drove tanks through the heart of Czechoslovakia. I started to openly criticize the regime for not living by its constitution. However, sometimes strangers on the street told me, "this is not your country, you are not Russian, go to your Israel." And so I decided not to pay with my life and freedom for the freedom of those who did not consider me an equal citizen. I terminated my Soviet citizenship, scientific degree, career, and left as a refugee protesting tyranny, left without any job (let alone Princeton or Utrecht), money, connections, and language. How did you feel, Bartel, when *your Germany* drove through the lives of tens of millions of peoples of Europe, good "old Europe" you said you so much loved? Is German professorship worth the price of responsibility for Europe's suffering?

I was asked on a number of occasions, what could have Bartel done alone? This reminds me a play I saw in the spring 1969 in the Moscow State University theater.[40] A man comes on stage and thinks aloud: "There were times when writing was a dangerous profession. Fyodor Dostoyevsky was sent to four years of hard labor in Siberia; Alexander Herzenwas forced to live life in exile. And now? But *what can I do alone?*" The second man appears on stage, and the two walk around as if not seeing or hearing each other, each exclaiming, "*What can I do alone?*" The third, fourth, etc., people appear onstage. Soon we witness some thirty men and women walking randomly and randomly complaining "*What can I do alone!*" The whole scene is full of random motion of people exclaiming "*What can I do alone!*" Slowly, unnoticeably they form rows and columns, marching and chanting together, "*What can I do alone! What can I do alone!*" Half of the audience sat in grave silence, while the other half

[40] I was an undergraduate student when Mathematics Professor Abram Khaimovich Livshitz invited me to see his performance at the Moscow State University's (MGU) student theater *Nash Dom* (*Our Home*). This theater-studio was founded by Mark Rozovsky in 1958 when he was still a student of journalism. The theater was shut down by the Soviet totalitarian authorities on December 23, 1969, a few months after my visit. By the spring 1969 all previous plays were banned, leaving measly scenes collected under the title "Take Old Staff and Show." The scene I describe was originally written by Novosibirsk student theater's authors Evgeny Vishnevsky and Vadim Sukhoverkhov.

loudly applauded. Millions of good Germans were righteously exclaiming, just like the actors in this production, "*What can I do alone!*"

Van der Waerden chooses to stay, because he believes that even during the Nazi era *Germany is the best place for doing mathematics*. "Why would I go to Holland where oppression became so intolerable and where every fruitful scientific research was impossible?" he writes to Van der Corput without realizing that the intolerable oppression of his Homeland was inflicted by the very country he served!

The great anthropologist and my dear friend James W. Fernandez, upon reading the early version of this book, summarized my findings concisely during our "Fang Summit" in early August 2007: "Frailty of Brilliance!"

In the Story of Van der Waerden, I confirmed one lesson of my own life: Silence in the face of a tyranny makes one a slave, an accomplice, and an executioner. I have thought about the following simple formula for a very long time. It has evolved, and it has inspired, to my satisfaction, an ongoing debate:

> One's response to living under tyranny without willingly supporting it can only be to leave, to engage in resistance, or to compromise.[41]

---

[41] Bartel's son, Hans van der Waerden contributes his view [WaH1]: Let us turn to the underlying general question, whether it was right or wrong for my father to stay in Germany after 1933, and even more so after 1940. I am glad to hear you pronounce your personal opinion on the subject (a moderate and carefully deliberated opinion indeed). Allow me to add some of my personal reflection too.

Judging the behavior, decision, "Life choices" of other people can only be done by applying general principles, which must be true not only in one place, but in every place on earth at any time. How, which could have been the general principle stating as a moral imperative for my father to leave Germany after 1933? Could it be this: "When the government of a country is turned into cruel and criminal tyranny, all intellectuals serving that government are obligated to emigrate, otherwise they become guilty of 'contributing' (as you put it) to the dictatorship"? Is this really a general principle, applied all over the world and at any epoch? I only heard it being pronounced for Germany, and only after 1945 in retrospection, and even that not to everybody, and not applied to everybody. I never heard the principle being applied to the USSR under Stalin (whose dictatorship was as horrible as Hitler's, if comparing the devil to satana is possible at all). Under Stalin, some intellectuals emigrated (as a personal choice) or were forced to do so. But never has anybody been blamed for <u>not</u> emigrating and so "contributing" to the Staling tyranny.

Allow me to add yet another, even more general consideration. In a wider sense, every intellectual in a public position "contributes" to the government he is working for. If this government – even without mutating into open tyranny – commits criminal actions on a larger scale, the intellectual gets involved and makes himself responsible, unless he "acts bravely" by openly protesting (or emigrating, if protesting seems too dangerous). This applies, for instance, to the actual US government.

In 2010, Hans van der Waerden returns to this topic [WaH2]: As a crude approximation, your three-cross-road theory may be of some use; it is inadequate when it comes to really understand day-to-day life in a totalitarian system. Because there is a fourth way, chosen by many who wished to preserve both life and soul. It says: "Stay in the country, avoid great gestures of opposition, but quietly and persistently show by small signs that you disagree, and so give hope and comfort to others." Under a perfectly organized surveillance system as Stalin established in the USSR, this sideway too apparently was barred; in Hitler's Germany, however, thousands of anti-fascists have followed it, thus surviving and uniting in an invisible network of free thinking and breathing.

Van der Waerden chose the compromise between his high moral aspirations and his desire to do mathematics in Nazi Germany. The struggle between these two conflicting goals produced the drama – perhaps, the tragedy – of the life of Bartel Leendert van der Waerden, one of the great mathematicians of the XX century, the century marked by merciless tyrannies and brutal wars.

42.8 Today I: The Scholar and the State

Our history will be what we make it. If we go on as we are, then history will take its revenge, and retribution will not limp in catching up with us.

– Edward R. Murrow

Unless the direction of science is guided by a consciously ethical motivation, especially compassion, its effects may fail to bring benefit. They may indeed cause great harm.

– Dalai Lama

I write my books and essays with the assumption that mathematicians are human beings and as such ought not to be entitled to Ivory Towers, but to be involved with the world. As you have seen, Van der Waerden's four chapters have not only been about his personal life and the life of his friend Werner Heisenberg, but moreover about moral obligations of a scholar in the world. Similarities between the past and the present are so striking that I have got to address here the most important time, the time we live in. If not here, where? If not I, who?

And so, this and the following two sections are my wake-up call for solving what I consider to be *the hardest open problem of mathematics: establishing and maintaining high moral grounds of the profession* – victims of which were, for example, Grigory Perelman and a person you have met on the pages of this book, Dmitry Raiskii.

Most people desire to be with the winners. It is, perhaps, a self-preservation instinct of *Homo sapiens*. As soon as Hitler's ascent to power became assured, most German professors, lawyers, doctors, and writers – the intellectual elite – jumped on Hitler's bandwagon. I was reminded about it when in March 2014 Russian President Putin invaded Eastern Ukraine's region of Donbass and annexed Crimea, both parts of sovereign Ukraine. I hoped cultural icons, pride and joy of Russia, would oppose the war, at least in word. Most of them, instead, supported criminal actions of their president, just as the German elites supported Hitler.

On March 12, 2014, the Russian Ministry of Culture, on its official Internet site published a group letter "*Russian cultural figures – in support of the position of the President [Putin] on Ukraine and Crimea*," signed within two days by 511 [!] prominent creative people, including celebrated and beloved movie stars, film directors, and presidents of major museums and theaters, including Oleg Tabakov, Alexej Batalov, Pavel Lungin, Alexej Uchitel, Vladimir Khotinenko, Karen Shakhnazarov, Valery Gergiev, and Yuri Bashmet. The letter proclaimed (I am translating from Russian):[42]

[42] https://www.bbc.com/russian/russia/2014/03/140312_russian_artists_letter

> In the days when the fate of Crimea and our compatriots is decided, Russian cultural figures cannot be indifferent observers with a cold heart. [. . .] That is why we firmly reiterate support for the position of President of the Russian Federation on Ukraine and Crimea.

Minister of Culture V.P. Medinsky could not hide his delight in his interview entitled "There will be no war [!], don't fantasize":[43]

> The workers of culture – public opinion leaders, enjoyed considerable moral weight and influence. [. . .] The more intense the political moment is, the more tangible is the need [of their support of the President]. "A poet in Russia is more than a poet."

I am compelled to respond to Medinsky: Evgeny Evtushenko, whom you quote without credit, meant that a poet in Russia is a prophet – not a conformist!

Kiril Serebrennikov is a fine theater and film director. On his Facebook timeline, he defended Oleg Tabakov, director of the prestigious Moscow MXAT Theater, who publicly and continuously supported Putin and cosigned the letter of 511. Serebrennikov argued that it was all right for Tabakov to support Putin and thus get government money to fund the theater and feed the troupe. A heated discussion erupted, which prompted my response:

> Dear Viktor Balabanov, you write: "Theater directors and the like leaders of Centers for the Arts, worry about preserving the culture, their nest, and fear that the Usurper [Putin] will deprive them of this opportunity." And what of it? Is it seemly to support people's tyranny in order to carry culture to those same people? Is culture worth tyranny? I dedicated my life to culture and education, but I do not support corruption with good intentions. It may pave the road to hell, as is well-known. Praise those who did not sell out: Yuri Shevchuk, Andrey Zvyagintsev, Boris Akunin, etc., and not the artists with a price tag sewn to them.

On March 14, 2014, a group of ca. 150 celebrated members of Russian intelligentsia responded to Putin's war and the letter of 511 Putin's supporters:[44]

> "*Intelligentsia of the Russian Federation:*
> *Do not bend, do not succumb to a lie*"
>
> Not for the first time in Russian history, people who disagree with the aggressive imperial policy of the state are declared defeatists and enemies of the people. It is not the first time that loyalty is valued above citizenship. Events in Crimea are developing rapidly and are fraught, if not with bloodshed, then with disgrace for Russia and troubles for the peoples of the two countries. Hopes to stop what is happening with the arguments of the mind are becoming less and less. But it is all the more shameful to be silent and passively stand aside. We, who do not call ourselves "workers" of culture or science, but simply Russian intellectuals, each working in his own field, declare:
>
> We are against the invasion of the territory of another state.
>
> We are against war with Ukraine and enmity with the world community.
>
> We are in solidarity with everyone who does not bend and does not succumb to lies.

---

[43] www.gazeta.ru/culture/2014/03/14/a_5949581.shtml

[44] https://blogs.pravda.com.ua/authors/haran/532456382c4fc/

I was happy to see among the 150 signatures of song-writers-performers Yuri Shevchuk and Boris Grebenshchikov, Nobel Laureate Lyudmila Ulitskaya, writer Boris Akunin, and my Moscow friend and great film director Andrey Zvyagintsev.

You may think that in the Free World of the United States and Canada, the opposition to the Russian annexation of parts of Ukraine will be met with a unanimous support, especially in intellectual circles that include mathematics professors, right? Wrong. In 2017, when the dust of invasion settled, *Congressus Numerantium*, a US–Canada mathematical journal, insisted that Putin did not invade Crimea in 2014! Permit me to put a mirror in front of my colleagues.

In April 2017, I submitted to the journal *Congressus Numerantium* my talk "Pontryagin-Kuratowski-Zykov-Harary, Kantorovich, Shafarevich, et al" given in March 2017 to a receptive audience at the Southeastern International Conference on Combinatorics, Graph Theory, and Computing, Florida Atlantic University, Boca Raton. In it, I conveyed ethical indiscretions in Russian mathematics, such as anti-Semitism, plagiarism, public scandals, and backroom stabbing. On September 18, 2017, I received an email from Professor David Allston, Managing Editor of the *Congressus Numerantium* containing a referee's report:

> Comments: This is a mixture of personal reminiscence and "fake news." It could nevertheless be published for its interest once section 11 has been removed. I would like to see the paper again before acceptance.
>
> Editorial Decision: not accepted yet.
>
> Would you please send me a revised pdf file and I will send it to the referee.

It was the first time anyone – and that includes Donald Trump – accused me of reporting "fake news." I asked for clarification. On September 19, 2017, I received it:

> Professor Soifer
>
> I received the following from the referee:
>
> Section 11. It is stated as fact that Russia invaded Crimea. However, there is no evidence that there is any truth to this. The journal should not publish dubious or incorrect statements as facts. Mathematics does not work like that.
>
> David Allston

You will find more examples of present-day censorship in our own Western World in *Geombinatorics* [Soi54].

A timid response of the United States and Europe to the Russian 2008 military "excursion" into Georgia (which continues still today) and the 2014 annexation of parts of Ukraine must have assured Putin that the West is weak, the West is scared of Russian nuclear arsenal, and Putin will get away with anything. And so, in the wee hours of February 24, 2022, Putin waged an all-out criminal war on Ukraine.

In early March 2022, just a week after the start of the war, I received an email from a talented Russian mathematician, inquiring whether I will now ban him from publishing in *Geombinatorics*. "Write to me that you support Putin's war on Ukraine, and I will ban you," was my reply. Instead, he sent me another fine paper for the April 2022 issue of the journal. However, his question showed me the need for *Geombinatorics* to have Editorial Policy, addressing authors' positions on the criminal war, which I published in April 2022 issue, entitled "Taking Sides: *Geombinatorics*' Response to the War and Boycotts"

[Soi53]. Welcome to read my 8-page statement in *Geombinatorics*. Here, I am including what in bureaucratic jargon is called "executive summary":

> Rector [President in American terminology] of Moscow State University Viktor Sadovnichiy, who happens to be a *Professor of Mathematics*, is perhaps the most politically influential mathematician of Russia. He is also President of the Russian Union of Rectors, which on March 4, 2022, published a letter signed by 184 [**!**] university rectors supporting Putin's criminal war on Ukraine.
>
> On the other hand, On March 7, 2022, the Ukrainian Film Academy started a petition calling on Film Institutions (including film festivals) and Film Professionals to ban *all* Russian films (as of March 20, 2022, the number of signatures reaches 10,192).
>
> The most famous Ukrainian film director Sergei Loznitsa disagreed with a blanket boycott: "Among Russian filmmakers, there are people who have condemned the war, who oppose the regime and openly expressed their condemnation. And in a way they're victims of this whole conflict like the rest of us. I was hoping that society these days is more intelligent, more sophisticated, than to apply this collective guilt to an entire community."
>
> I too am against an indiscriminate boycott of all Russian mathematicians based solely on their citizenship. I agree with Sergei Loznitsa, and will decide the Hamletian question, to ban or not to ban, based not on the author's passport, but on the author's words and deeds. Starting immediately, *Geombinatorics* will not knowingly publish authors who in the past did or presently support the Russian wars on Ukraine or Georgia, or unprovoked wars on other countries that Russia could wage in the future.

The great sage Dalai Lama warns that without "consciously ethical motivation, especially compassion" science "may indeed cause great harm." Exactly right. We have seen throughout history, time and again, how evil usage of science and technology can be if it is not built on a foundation of high morality. Atrocities of Nazi Germany alone provide countless examples of science, technology, and even arts and literature used for ill deeds. I value education, however, I must admit that

> Fine education does not guarantee high culture,
> And high culture does not guarantee humanity.[45]

We ought to be principled, for there is no appreciation of the good without recognition of the evil. And the principled scholars cannot afford to be silent. We ought to never be silent accomplices of injustice, as the Holocaust survivor and Peace Nobel Laureate Elie Wiesel so eloquently argues:[46]

> I swore never to be silent whenever and wherever
> human beings endure suffering and humiliation.
> We must always take sides.
> Neutrality helps the oppressor, never the victim.
> Silence encourages the tormentor, never the tormented.

[45] A. Soifer, *Charge to the Winners*, The 30th Colorado Mathematical Olympiad, May 3, 2013.
[46] Elie Wiesel, The Nobel Peace Prize Acceptance Speech, December 10, 1986.

42.9 Today II: "The Silent Agreement"[47]

False opinions are like false money, struck first of all by guilty men and thereafter circulated by honest people who perpetuate the crime without knowing what they are doing.
– Joseph-Marie, Comte de Maistre

If faith can move mountains, disbelief can deny their existence. And faith is impotent against such impotence.
– Arnold Schoenberg, June 1924, [Scho]

I was born in 1960 into a country in which virtually everyone of the older generation was declared free of any serious guilt, except the few obvious villains.
– Moritz Epple[48]

The Fool's Gold of Silence

Yes, I know, I've heard this wisdom in Russia, "Silence is Golden," and in America, "A closed mouth catches no flies." However, my heroes Edward R. Murrow, Elie Wiesel, and Albert Camus rejected this fool's gold. Camus eloquently conveys his vision of creator's duties in the society in his Uppsala University lecture on December 14, 1957, just 4 days after his acceptance of the Nobel Prize [Cam2]:

> The writer can no longer hope to stand aside and pursue reflections and images dear to him. Until now, for better or worse, abstention was always possible in history. The person who did not approve could often remain silent or speak of something else. Today everything has changed, and even silence takes on a daunting significance. From the moment when abstention itself is considered as a choice, punished or praised as such, the artist is conscripted whether he wants it or not. "Conscripted" strikes me as more accurate here than "committed" . . .
>
> To tell the truth, this is not easy, and I understand that artists may regret [losing] their former comfort . . .To create today is to create dangerously. Every publication is an act, and this act exposes you to the passions of a century which forgives nothing.

All true, Cher Monsieur, and if not us, who? If not now, when?

[47] The title is a quote from [Epp].

[48] [Epp].

G.M. Ziegler, An Important Operative

The first edition of *The Mathematical Coloring Book* [Soi44] became a standard text, beloved by numerous colleagues around the world. The Ramsey theory leader Ronald L. Graham told me he bought three copies in order to have it handy in each of his three offices. Legendary Helge Tverberg sent his high praise from Norway. Peter Mihók writes in *Mathematical Reviews* of the American Mathematical Society [MR 2458293]:

> The beautiful and unique *Mathematical coloring book* of Alexander Soifer is another case of "good mathematics," containing a lot of similar examples (it is not by chance that Szemerédi's Theorem story is included as well) and presenting mathematics as both a science and an art. It is easy to find a lot of information about this book, including the three excellent forewords of B. Grünbaum, P. D. Johnson and C. Rousseau on the Internet (see www.springerlink.com/content/m615m7/front-matter.pdf). Let us mention here, in the words of the author, that "this book includes not just mathematics, but also the process of investigation and psychology of mathematical invention, . . . it presents mathematics as a human endeavor, . . . it explores the birth of ideas, . . . and moral dilemmas of the times between and during the two World Wars."

Springer was so pleased with the book's resonance that they signed me to a contract for this much expanded edition. There was, however, one man in Germany who almost 6 (!) years after my book's appearance, on September 18, 2014, published in English an untruthful personal attack on me under a disguise of a "book review" [Zieg]. I have never before responded to reviews of my books. This time, I had to reply because this operative from mathematics held a high position of the President of the *Deutsche Mathematiker-Vereinigung* (DMV, German Mathematical Society) three times, 2006–2008, is a member of the Executive Committee of the International Mathematics Union, has been President of Freie Universität Berlin 2018–2022 and again 2022–2026, and held many other important appointments. His high positions may influence some of our lower-information colleagues to take his word without verifying it. His name is Günter M. Ziegler. Ziegler wrote a review of my book in English and blocked my response by publishing it in *Jahresbericht der Deutschen Mathematiker-Vereinigung*, whose Editor Hans-Christoph Grunau refused to publish my response or even my Letter to the Editor. Grunau copied his rejection to Ziegler and declared that all I can do is to ask (read: beg) Ziegler to correct his mistakes. Thus, my reply had to be published elsewhere; it appeared in *Geombinatorics* in print and online [Soi52].

You would expect *math* professor Ziegler to point out *math* mistakes, right? Wrong. "After struggling with the book for 2 1/2 years" (his words) he found none in the book of 640 pages. What then were his "struggles" with? Ziegler alleges: "This book has many faults, starting with *the title* [sic], *the dedication* [sic] and the many *prefaces* [sic], but more seriously with the *selection of the material*." Ziegler criticizes the title! He criticizes the dedication of the book to my late father – how dare he!! Ziegler can't stand three detailed forewords written by much more competent mathematicians than he is. He expresses his dislike of "the material" – as if someone forced Ziegler to read my book.

Having grown up in a totalitarian state, I am not surprised by Ziegler's untruths, but neither am I going to be silent. Having sacrificed everything for freedom on the other side of the Atlantic, I will never give up my hard-earned freedom of speech. Let us take a closer look at Ziegler's so-called "review."

Four-Color Problem

Ziegler alleges that the Four-Color Problem (4CP) is not good enough a problem for the inclusion in my book. Yet he himself admits how important and influential it has been for a century and a half (p. 263): "Was this a good problem? Certainly it was important, as it has driven the development of graph theory to a large extent. Nevertheless, it has made little connections to other parts of Mathematics." 4CT has "connections" to topology, graph theory, and other branches of discrete mathematics, and this is plenty, as much as most mathematical results have. Moreover, this is *my book* and thus my choice of what to include.

Ziegler then complains that Heinrich Heesch did not get a grant to work on 4CC, and thus Heesch "couldn't complete his proof, and the fame for solving the problem instead [sic] went to Hermann [sic] Haken and his team." How can Ziegler be certain that with a grant Heesch would have found a proof? Does he think money guarantee proofs in mathematics? Isn't Ziegler's "instead" hints to a stolen credit? Didn't Wolfgang Haken and Kenneth Appel deserve the highest honors for conquering the problem that had withstood all assaults for 124 years? They did give generous credit to Heesch. It must be embarrassing for president of a German university Ziegler and his German publisher Hans-Christoph Grunau not to know the name of a German American celebrity Wolfgang Haken and list instead his cousin physicist Hermann Haken.

Ziegler mentions in passing, "the [Appel–Haken] proof was reworked [sic] later by Robertson et al." While we do not have in mathematics a definition of a "different proof," Ziegler's remark is a scandalous understatement. The 1997 Robertson–Sanders–Seymour–Thomas proof was dramatically better than that of Appel–Haken. Instead of 486 secondary discharging rules (those unfamiliar with the terminology can think of these rules as "ideas"), the new proof used just 20. When 13 years later, in 1989, Appel–Haken's proof finally appeared in print, it filled 741 oversized book pages, whereas the new proof comprised a very readable journal article of 43 pages. Moreover, the new proof was *verifiable*, for the authors achieved a clean separation of what they did by hand (better said, by mind) and what their computer did and placed their software on an anonymous ftp for anyone to verify.

In the premier 1981 issue of *Combinatorica*, Paul Erdős gives the highest praise to 4CC: "The most famous conjecture of graph theory or perhaps of the whole mathematics, the four-colour conjecture, became recently the theorem of Appel and Haken."

On August 14, 1991, Erdős wrote to me "I would be much happier with a computer-free proof of the four-color problem, but I am willing to accept Appel–Haken proof – beggars cannot be choosers." Ziegler appears to be a choosy beggar, who dismisses as no good this major, influential, celebrated, classic problem of mathematics.

Chromatic Number of the Plane

Ziegler believes that finding the chromatic number of the plane (CNP) is also a bad problem (p. 265): "The chromatic number of the plane: Is this a good problem? Again this is a question of taste. In my view the fact that there is so little progress on the original problem in so many years, and progress only on variations, and that the answer might depend on set theory all indicate that it is not a productive, helpful problem."

Ziegler’s logic is absurd. The indication that the problem is hard and consequently takes a long time to be conquered, he uses as a “proof” of “not a productive, helpful problem”! If “the fact that there is so little progress on the original problem in so many years” were to mean the problem is bad, then *all* the great classic problems of mathematics are bad, from Fermat’s Last Theorem, which required ca. 360 years, to the Goldbach Conjecture, Riemann Hypothesis, Poincare Conjecture, etc., etc., etc.

Yes, it is a matter of taste: Ziegler’s taste puts him, perhaps, in a minority of one, when he suggests that CNP problem is bad. The greatest problem creator of all time Paul Erdős liked the CNP problem so much that he included it in his numerous problem papers and talks, and so did the leader of Ramsey Theory Ronald L. Graham. The CNP problem was selected for the inclusion in the well-known problem books “Unsolved Problems in Geometry” by Croft–Falconer–Guy, Springer, 1991, and “Old and New Unsolved Problems in Plane Geometry and Number Theory” by Klee–Wagon, Mathematical Association of America, 1991. I was invited to write Chapter 8 for the book “Topics in Chromatic Graph Theory,” Cambridge University Press, 2015, edited by Lowell W. Beineke and Robin J. Wilson. The Nobel Prize (1994) and Abel Prize (2015) laureate John F. Nash, Jr. liked CNP problem so much that he invited me to write a chapter on it [Soi50] for the 2016 Springer book edited by Nash and Michael Th. Rassias on most famous “Open Problems of Mathematics,” where other chapters are dedicated to such celebrated classic unsolved problems as the Riemann Hypothesis, the Goldbach Conjecture, the P versus NP Problem, the Hadwiger Conjecture, etc. In March 2021, Alfred Rényi Institute of Mathematics of the Hungarian Academy of Sciences invited me to give a “Public Lecture” on CNP and related problems for a worldwide audience [Soi56].

Ziegler complains that “the answer [to CNP] might depend on set theory” to dismiss CNP as a bad problem. The opposite is true. Shelah–Soifer papers reawakened the mathematical world that ever since the 1930s mostly fell asleep on **ZFC** as *the* foundation of set theory and showed that by using other axioms for sets, we could build many exciting buildings of mathematics, free from counterintuitive paradoxes caused by **ZFC**. The prominent French mathematician Jean-Paul Delahaye published a deep, long review of Shelah–Soifer work; he views Shelah–Soifer papers as highly important mathematically and philosophically [Del]:

> When Gödel and Cohen proved independence of **AC** from the rest of the axioms **ZF** of set theory, they created a parallel, so to speak, between **AC** and the parallels postulate. As so, when Shelah–Soifer came out, it showed that various buildings of mathematics can be constructed.

Sorry to disappoint you, Dr. Ziegler: your complaint about the lack of CNP results in general case is no longer true. In 2018, Aubrey de Grey achieved a major breakthrough by constructing a 5-chromatic unit-distance graph, thus reducing the range of possibilities to 5, 6, or 7. He was followed by a stream of 5-chromatic constructions by Marijn Heule, Jaan Parts, Geoff Exoo, and Dan Ismailescu, and a number of other scholars. Read about it later in this book and in Polymath Project.

In 2 ½ Years of "Struggling," Ziegler Found No Mathematical and No Historical Mistakes

Poor Dr. Ziegler: after "struggling with the book for 2 ½ years (!) on the way to this review" (p. 262), he found *no* mathematical and *no* historical errors in the book. He had to criticize something, and so he objected to the title, the dedication, the choice of problems, etc.

Ignoramus in History, Ziegler Attempts to "Correct" My History and Fails on All Counts

Ziegler quotes me and then *alleges* to refute my statements, (pp. 266–267):

> Soifer reports that Bartel Leendert van der Waerden [. . .] proved this pioneering result while at Hamburg University and presented it the following year at the meeting of *D.M. V.*, *Deutsche Mathematiker-Vereinigung* (German Mathematical Society) in Berlin. The result became popular in Göttingen, as the 1928 Russian visitor of Göttingen A. Y. Khinchin noticed and later reported [Khi1], but its publication [Wae2] in an obscure Dutch journal hardly helped its popularity. [. . .]
>
> This report gets a number of facts wrong. For example, the DMV meeting 1928 was held in Hamburg, and Aleksandr Khinchin writes that the result was obtained in Göttingen. The "obscure Dutch journal" was *Nieuw Archief voor Wiskunde*.

Not Soifer – Van der Waerden himself wrote the Story of Creation of this proof in 1926 in Hamburg with the aid of Emil Artin and Otto Schreier. Did Ziegler read my book that he is reviewing?

Not Soifer, and not 1928 – Alfred Brauer wrote that Van der Waerden found his proof in 1926 and presented it "*the following year* at the meeting of *D.M.V.*" The following year here obviously meant 1927, thus Ziegler's statement that the 1928 meeting was held in Hamburg is totally irrelevant, as Ziegler tried to correct incorrectly.

Not Soifer – Khinchin incorrectly stated that the result was obtained in Göttingen – Ziegler should send his complaints to Khinchin's heirs.

Not Soifer – the Dutch Nicolaas G. de Bruijn called *Nieuw Archief voor Wiskunde* an "obscure Dutch journal," and he certainly knew that journal much better than Ziegler.

How Important Is the Authorship of a Conjecture?

Ziegler rhetorically asks (pp. 266–267), "Why this urge to prove Van der Waerden wrong about the origin of the conjecture, if he apparently heard it from Baudet?"

If Ziegler read my book attentively "for 2 ½ years," he must have learned that P.J.H. Baudet passed away in 1921, while Van der Waerden heard the conjecture in 1926. Therefore, Van der Waerden did *not* hear the conjecture from Baudet and moreover wrote to me about it.

Ziegler then (ibid) declares a rhetorical question, "Does it really make sense to talk about the 'authorship of the conjecture'?"

Yes, it most certainly does. What would a prover be proving if someone did not create a good conjecture? I generally view creation of a good conjecture as important as proving it and hence systematically give a joint credit for a theorem to the author of the conjecture and its prover. Shouldn't we give credit where credit is due?

Good conjectures inspire and direct research, and at times it is very hard to envision the future, i.e., to create a good conjecture. If it is always easy, as Ziegler insinuates, why wouldn't he, for example, conjecture for us criteria for a graph to be Hamiltonian?

Francis Bacon points out to Ziegler [Pet, p. 494]:

A prudent question is one-half of wisdom.

Ziegler Slanderously Accuses Van der Waerden of Anti-Semitism

Ziegler pleads impartiality toward Van der Waerden (p. 267): "I have no stakes in Van der Waerden, I have never met him, and I cannot (and dare not) judge him, neither his contributions to Mathematics, nor what he did or didn't do for example as a professor in Leipzig 1931–1945." Yet, *Ziegler then falsely accuses Van der Waerden of anti-Semitism* (!) without any substantiation when he claims (p. 267) that "Some of his [Van der Waerden's] actions seem to have harmed Jewish colleagues (but I don't know and can't judge whether any of this was intentional or even done knowingly)."

Where are the facts to back such a horrible slander? I spent over 20 years researching Van der Waerden's life with the assistance of thousands of documents, members of his family, and eyewitnesses and showed clearly in my two books [Soi44, 47] that Van der Waerden had *never* been an anti-Semite. Moreover, he was prevented from succeeding Constantine Carathéodory at Munich precisely because he was perceived as a philo-Semite. In 1935, Van der Waerden bravely published a eulogy for his beloved Jewish teacher Emmy Noether in *Mathematische Annalen.* In my two books I describe at a great length the May 1935 Faculty Meeting at Leipzig, where Van der Waerden, Werner Heisenberg, and three more scholars publicly (!) protested the firing of five Jewish professors from Leipzig University. When their protest did not succeed, they even contemplated a group resignation. Ziegler owes a profuse apology to the Van der Waerden family and all of us!

Ziegler Then Groundlessly and Redundantly Blames Soifer

Ziegler apparently attempts an old trick: he repeats false accusations many times, in hopes it will become "truth" to the less informed. It-will-not, Mr. Ziegler! Ziegler, who slandered Van der Waerden by the accusation of anti-Semitism (!), now baselessly and with a great redundancy accuses Soifer of "badly disliking" Van der Waerden:

> "The only plausible reason I can see for Soifer's passion and persistence in his investigations and his attempts to find fault with Van der Waerden is that he badly dislikes him." (p. 267)
>
> "He [Soifer] badly tries to find fault in his stay at Leipzig University during Nazi times, and so on." (p. 267)
>
> "It cannot be good if a historian has an ax to grind, if from the outset he wants to prove things about his subject of study, since this will color his judgement." (p. 267)
>
> "The impression remains of a personal war." (pp. 267–268)

"This passion and scornfulness against Van der Waerden." (p. 268)
"Soifer's persistent personal campaign against Van der Waerden." (p. 268)

A historian, Mr. Ziegler, is not in the business of liking or disliking the subject of his research. As a historian, I paid the ultimate, highest respect to B.L. van der Waerden by telling the truth, grounded in facts, revealed by decades of archival research and eyewitness testimonies, including memoires of his son, his niece, and his nephew.

Concerns about Van der Waerden's presence in Nazi Germany for the entire duration of the Third Reich were raised contemporaneously by Otto Neugebauer, Richard Courant, Johannes G. van der Corput, editors of *Het Parool*, and others. I communicated their concerns in my two books [Soi44, Soi47]. I also included instances of brave and honorable conduct exhibited by Van der Waerden during those horrific Nazi times.

Ziegler Admits He Is *Not* a Historian and Has *Never* Been to Archives

Ziegler admits (p. 267): "At this point, I must say that I am not a historian, I have not read all materials and I have not been to the archives, so I can't (!) really judge this."

However, judging Van der Waerden and Soifer is what Ziegler is doing! We witness the case of an unapologetic judgmental ignoramus of history, elected to the presidency of DMV and Freie Universität Berlin. I urge his better educated subordinates to stress upon him that history deserves respect and rigor every bit as much as mathematics.

Since It Is *Not* Math and *Not* History, What Is the Goal of Ziegler's 6-Year-Late 9-Page-Long Review?

Ziegler found no grounds to correct mathematics or history presented in my book. What is then the goal of Ziegler's criticism of *The Mathematical Coloring Book* [Soi44] and my 2015 book *The Scholar and the State: In Search of Van der Waerden* [Soi47], which Ziegler mentions in his review?

Is it my opinion that too many potentially good Germans – including the majority of professors – joined the Nazis or remained silent and thus made Nazism in Germany possible?

Is it my questioning the International Mathematics Union (IMU), which ever since 1981 had been etching on its prestigious gold medals the profile of the Finnish *Waffen SS* Volunteer (Recruitment) Committee Chairman Rolf Nevanlinna, that same Nevanlinna who in his speeches and articles praised Adolf Hitler as the Savior of Europe?

Is it my concern with the 2002–2013 Director of the *Mathematisches Forschungsinstitut Oberwolfach* Prof. Dr. Dr. h.c. Gert-Martin Greuel starting the history of the Institute in 1946 (http://link.springer.com/chapter/10.1007%2F978-3-642-25710-0_26), thus concealing (!) its start in 1944 by the Nazi Wilhelm Süss, with the approval and funding by the high Nazi authority Hermann Göring?

Is it my 2014 book review where I objected to Roquette–Frei–Lemmermeyer fabricating a hero out of Nazi-collaborating anti-Semite and racist Hasse? Do the readers know that this review was published on June 21, 2014 https://zbmath.org/?q=an:06214484, censored and removed off the zbMATH website on July 9, 2014, and published again by the Editor-in-Chief Greuel on September 4, 2014 Zbl 1294.01004? (In the end, Greuel republished my review and "thanked" me by expulsion from the reviewers of zbMATH[49]). Ziegler published his review *days after*, on September 18, 2014.

What was Ziegler's goal in waging such an all-out attack, consisting of fabrications, untruths, and irrelevancies? Was it his hope that if to throw enough accusations, no matter how false, something will stick? This is an old "counterfeiters" trick, described centuries ago by Joseph-Marie, Comte de Maistre (1753–1821) in *Les soirées de Saint-Pétersbourg*, Ch. I:

> *False opinions are like false money, struck first of all by guilty men and thereafter circulated by honest people who perpetuate the crime without knowing what they are doing.*

Ziegler is not alone. He quotes the German historian Reinhard Siegmund-Schultze (henceforth S-S) while suspiciously S-S quotes Ziegler. S-S wrote in his emails to me numerous high compliments for my four articles about Van der Waerden[50] that were the foundation of my book [Soi47], but in his 2015 AMS *Notices* review insinuates that I could not possibly understand Germany and Germans because I did not live there. By this (il)logic, we may not research Ancient Greece and Rome, for none of us lived there. He essentially demands a monopoly for the German authors in researching German history. Monopoly does not breed trust; moreover, monopoly allows cover-up of the truth. The American Mathematical Society Publisher Sergei Gelfand called me when S-S review came out and said:

> While writing a book like this, you shouldn't be surprised; he [S-S] followed "party" orders.

Ziegler's malicious failed attempt to silence me is symptomatic of serious problems of Germany dealing with its past, even now, 90+ years after Hitler's assent to power. On April 23, 2014, I received an email [Epp] from the well-known German scholar Moritz Epple, Professor of History, specializing in the History of Mathematics at Goethe University Frankfurt. I did not communicate with him before Springer Birkhäuser invited Prof. Epple to be the official referee of my book *The Scholar and the State: In Search of Van der Waerden* [Soi47], and Epple asked Springer to share with me his name and email. Epple raises the veil off the "*Secret Life of the Postwar Germany*" and allows us to understand many actions (including Ziegler, Siegmund-Schultze, Greuel, Roquette, Frei, etc.) that before I could only guess about:

> I was born in 1960 into a country in which virtually everyone of the older generation was declared free of any serious guilt, except the few obvious villains whose involvement in atrocious and – for me as a young person – completely unfathomable crimes was so obvious that no one could get around it. But all the others, the van der Waerdens, own family members, older teachers and later even some professors: What about them?
>
> . . . well, to put a long story short: To NOT talk about the moral problems that their earlier lives involved seemed to be the silent agreement that kept (and to some extent still keeps) this society going.

[49] On March 16, 2021, the new Editor-in-Chief Klaus Hulek reversed the expulsion: "I confirm that your reviewer account is now active and I have informed the relevant editors of this fact."

[50] https://geombina.uccs.edu/siegmund-schultze

> Ever since I understood this (if I really understood – who can be sure) I felt the need to join those who addressed these issues with careful, but sharp judgement, and to break, rather than to prolong, the silent agreement of suspending judgement. Conflicts were the unavoidable consequence for all of us.
>
> After the first few chapters of your book I understood that your challenge to the reader was exactly this: To provoke her or his moral judgement, on the basis of a wealth of relevant information.
>
> The more I read, the more I enjoyed reading your text. You do make a strong case.
>
> I think the need to make such cases about life in Nazi Germany, and in the occupied countries, still is and remains great. Of course at some point in history other aspects of our complicated present and recent past may require similar attention – keeping the borders of the richer countries of our world shut for refugees comes to mind – but the Nazi past still haunts us in so many ways. And especially us who were born into families who in some non-zero degree were involved in, and responsible for, the reality or at least the possibility of the crimes of this period.

It is apparently not enough for the Zieglers to keep silent about the German past and its manifestations today – they want to silence the Soifers. Surely, at times my books deliver truths, inconvenient for some and treasured by others. I learned that there is no such thing as a free free speech, and thus I am now less surprised when an invoice for my exercise of free speech arrives. Ziegler's review vividly illustrates how important and timely my two books are *for Germany* in particular and the world in general. The treatment of the German past affects the integrity of all aspects of Germany today, including German scholarship. I deeply appreciate the brave honest paths paved by Herbert Mehrtens, Moritz Epple, and some other German scholars.

One of the main reasons I researched archival documents for over 20 years was to learn important lessons of history and apply them to today's world, where the 2014 Russian annexation of Ukrainian Crimea eerily reminds me the 1938 Nazi Germany's annexation of Czechoslovakian Sudetenland; and the 2022 Russian war on Ukraine resembles the 1939 Nazi and Soviet war on Poland. Sadly, in both periods, the majority of the intellectual elites, artists and scholars, supported their criminal leaders. Both times the world hoped to satisfy the insatiable appetites of the tyrants by throwing Poland to Hitler and Stalin then, while some willing to throw Ukraine to Putin now, in the years 2022–2023. Complicity and conformism will not pave a path to a brighter future. We ought to look the past straight in the eyes, learn from it, and strive not to repeat mistakes of the past. The German unwritten agreement to conceal its tragic past will not work. *Truth, like water, will find its way out.*

I am leaving you here with a few lines from a genius long 1970 poem "KADDISH, Dedicated to Janusz Korczak" by Alexander Galich, a dissident poet, songwriter, performer, screenwriter, and playwright. In 1974, he was forced to leave the Soviet Union. In 1977, Galich was found dead in his Paris apartment under suspicious circumstances. I am translating for you from the Russian original.

> "It's time," one day said noble prince,
> "To overpaint this dirt."
> The painter said: "It's time, my prince,
> Long overdue, my lord."

And dirt became all dirty-white,
And dirt became all dirty-green,
And dirt became all dirty-blue
Under the painter's brush.
It all because the dirt is dirt,
Whatever color you insert.

42.10 Today III: From Rolf Nevanlinna Prize to Abacus Medal: A Noteworthy History of an IMU Prize[51]

The past is never dead. It's not even past.
– William Faulkner
Requiem for a Nun, 1951

The world is a dangerous place to live; not because of the people who are evil, but because of the people who look on and don't do anything about it.
– Albert Einstein

Movement 1: Why I Objected to the Rolf Nevanlinna Prize's Name

Yes, Faulkner is correct, the past is always alive. It continues to live within the present. No matter how many times we pledge "never again," we are predisposed to repeat mistakes of the past as poor students of history. You met Rolf Nevanlinna in this book during the 1946 job search at Zurich University. Now I will tell you the present *Tale of Rolf Nevanlinna and the IMU Prize named in his honor*.

Sometimes moments occur when I feel the need to act, I just cannot let a problem be, especially when no one else does anything about a moral issue of high importance to me. This is what I felt when I discovered a great honor bestowed on Rolf Nevanlinna by the International Mathematics Union (IMU) by establishing in 1981 the prize and gold medal in his name. People often pour cold water on their impulse to act by repeating the old rhetorical question "What can I do alone?" I prefer an alternative principle: *I will do all I can and let the chips fall where they may.*

To give you a taste of Nevanlinna's flavor, let me quote his March 25, 1941, letter to Helmut Hasse, where he praises Hitler and Nazism. Enjoy the sing-along duet of the two active Nazi supporters:

> You know, dear Herr Hasse, your remarks about the hypocritical and stupid "moral indignation of Western politicians, who try to hide their hate against Germany under the mantle of nice phrases," correspond completely to what we feel here and say to

[51] This section has concurrently appeared in *Geombinatorics* in July 2023 [Soi57].

ourselves daily. You know those deeply rooted sympathies which connect us Finns with Germany, these bonds are today stronger than ever now that the easily understandable irritation caused by our difficult time a year ago has died down . . .

It is absolutely clear to us that only a strong and powerful Germany, the heart of Europe, is capable of forming the fate of European community in the way, which the interest of all European nations of culture demands. Personally, I am firmly convinced thereof and I believe to see a total justification of this conviction in European history, namely that Germany is today summoned not only to save European culture, which already happened in 1933, but to lead it to an undreamt-of blooming. The world-historic significance of the present hour is immense.

And this active Hitler supporter was chosen by International Mathematics Union (IMU) to be etched on its Gold Medal, awarded to the best theoretical computer scientists!

Movement 2: Rolf Nevanlinna and His Prize: A Brief Excursion in History

As you know, in 1981 IMU Executive Committee decided to create the Rolf Nevanlinna Prize for "Mathematical Aspects of Information Sciences," i.e., Mathematical Aspects of Computer Science, even though Nevanlinna made *no contributions* to theoretical computer science. Helsinki University, Finland, offered to pay for the prize in honor of a Finn (a gold medal with Nevanlinna's profile, and cash to match the Fields Medal, ca. $15,000). The IMU Executive Committee (EC) accepted the Finnish offer, allowed the Helsinki University seal to appear on this international medal, and has been awarding the Rolf Nevanlinna Prize once every four years at the International Congress of Mathematicians (ICM), most recently on August 1, 2018, in Rio de Janeiro.

The IMU Executive Committee was duty bound to consider the moral bearings of the person they chose for a high prize – and they failed. Have the Executive members knowingly chosen a willing Nazi collaborator for the IMU Prize, or is their ignorance is the defense of their integrity? Let us be charitable and presume Executive members' ignorance of history in 1981.

I wrote all this and more about Nevanlinna in my 2015 book [Soi47; pp. 189 and 286–288] and urged the IMU Executive Committee to change the name on the prize. But whoever reads 500-page books, and furthermore, remembers a few pages after reading such a substantial dense volume!

Movement 3: IMU President Shigefumi Mori

Meanwhile, I was elected President of the World Federation of National Mathematics Competitions (WFNMC) in 2012 and as such was asked in July 2016 to give my organization's report to the General Assembly of the International Commission on Mathematical Instruction (ICMI) during its Hamburg quadrennial Congress. Right before my report, I had a brief exchange with the IMU President Shigefumi Mori:

– Mr. President, may I have your address, I would like to mail you a letter?
– What about?
– About one of your prizes.
– Which one?

– Rolf Nevanlinna Prize.
– You know, I cannot do anything by myself, but I will present your letter to the Executive Committee.

I had a feeling that President Mori *knew* what I was going to write about, for otherwise how would he know – without asking me – that I will complain about the name of the prize and the face on the medal? You will see in this section that *Mori acknowledged his and IMU's Executive Committee prior knowledge and years of silence.*

I presented the WFNMC report and then told the roomful of the delegates about the Nazi collaboration of Rolf Nevanlinna. I ended with my personal impassioned call to change the name of the Rolf Nevanlinna Prize. A long silence fell on the room, followed by enthusiastic applause and private remarks "I did not know" from some delegates to me.

Mori acknowledged the receipt of the letter [Soi37] I sent on behalf of the Executive Committee of WFNMC and promised to put it on the agenda of the next IMU Executive Committee meeting, 8 months later, in April 2017.

I sent my second letter [Soi38], this time a personal one to Mori and his EC with two essential points. I offered to personally pay $15,000 to IMU every four years to eliminate IMU's dependence upon Finnish funding. For someone, who started his American life from scratch as a refugee, this was a substantial expense, which, as the saying goes, put my money where my mouth was. I also observed that while the 1981 EC that established the prize could have pleaded ignorance, now EC could not do so, for I informed them of the Nazi collaboration of Rolf Nevanlinna.

I stressed to Mori and his EC that there is a popular misconception that one who does nothing, does nothing wrong. "In fact, now that you know the truth," I continued my letter to Mori, "doing nothing would transform an innocent mistake of 1981 into an intentional stain on IMU and on all mathematicians." Remember Grigory Perelman's refusal to accept the Field's Medal and the Millennium Prize, and his exodus from mathematics? Now you understand why this great mind did not wish to be a "poster boy" for mathematics, where the majority condones immorality of the minority. *Keeping the Nevanlinna name on an IMU prize would stain mathematics forever*, I concluded in my letter.

Imagine, EC keeps dates and locations of its meetings in secret. Only in late April 2017, did I learn that EC meeting took place on April 1–2, 2017, in London and asked President Mori to share with me their decision. His April 24, 2017, reply was a riddle. On the one hand, he wrote,

> We did discuss the issue regarding the Nevanlinna Prize at our recent EC meeting, and we made a decision.

On the other hand, he was not going to disclose that decision to me:

> But, as I am sure you understand, we need to discuss this with the partners involved. Before we have reached an agreement with them, we will not go public. We ask for your understanding of this way to proceed.

I met this part without understanding. "What if you do not reach an agreement with partners?" I asked Mori, who went non-communicado for what felt like eternity.

Movement 4: Open Letter to the Nine Winners of Rolf Nevanlinna Prize

In view of a long Mori silence, on September 28, 2017, I sent "*An Open Letter to All Nine Rolf Nevanlinna's Prize Laureates*" Robert Tarjan, Leslie Valiant, Alexander Razborov, Avi Wigderson, Peter Shor, Madhu Sudan, Jon Kleinberg, Daniel Spielman, and Subhash Khot [Soi41]. I hoped they would join me in opposing Hitler supporter's profile on their medals and his name of their prizes:

> Ever since 2010, I wanted to write to you. I decided to first do everything I could on my own and see whether I would succeed without becoming your messenger of negative information. The time has come to share with you my concerns about *The Rolf Nevanlinna Prize*, one of your highest awards.
>
> In the 20 years of my writing "The Mathematical Coloring Book" (Springer, 2009) and "The Scholar and the State: In Search of Van der Waerden" (Birkhäuser, 2015), I looked into the life of Professor Nevanlinna.

I then quoted for the nine laureates a passage from "The Scholar and the State" [Soi47] about Nevanlinna's admiration of Hitler and his recruitment of Finnish SS troops demanded by Himmler. I ended my letter observing that the greatest moral authority in this matter rests with the laureates. I urged them to join in and be counted, for the integrity of our profession was on the line.

The first reply came from Peter Shor of MIT:

> Dear Alexander Soifer:
>
> Finland was in a rather terrible position during World War II. It was caught between the Soviet Union (headed by Stalin, who was responsible for the deaths of tens of millions of people, including the deliberate decision to let around 7 million Ukrainians starve to death in the Holodomor in 1933) and Nazi Germany. For Finland, the Soviet Union was by far the more serious and immediate threat.
>
> Does this excuse Nevanlinna's actions? I don't know. However, without evidence that Nevanlinna supported the anti-Semitic [sic] aspects of the German fascist government, I don't believe that I feel warranted in taking any action at this time.
>
> From Wikipedia:
>
> "Neither the unit [The Finnish Waffen-SS] nor any of its members were ever accused of any war crimes."[52]
>
> Peter Shor

On the same day, I responded to Professor Shor with copies to all laureates:

> Dear Peter Shor,
>
> I appreciate your prompt and thoughtful reply. Permit me to comment on your two arguments.

[52] Shortly after, Wikipedia replaced its statement by this one: A 2019 report by the National Archives of Finland concluded that "at least some of the cases show that Finnish volunteers did participate in carrying out atrocities against Jews and civilians." International sources say that the Finnish soldiers were likely involved in atrocities.

Yes, Stalin was a major criminal as was Hitler. Stalin caused Holodomor and also deaths of tens of millions in concentration camps of the GULAG, murder of ca. 21,000 prominent Polish people in Katyn and elsewhere, etc., etc. However, is being against one mass murderer justifies collaborating with another mass murderer, deserves a profile etched on our medals?

Anti-Semitism was one of the Third Reich's main policies, which resulted in ca. 6 million Jewish deaths. I will stipulate that Professor Nevanlinna was not anti-Semitic. However, he recruited for those who were guilty of Jewish deaths. Moreover, Nazi Germany is responsible not only for taking lives of 6 million Jews, but for tens of millions of Jewish, Polish, Russian, Dutch, Danish, Norwegian, Belgian, French, Czechoslovakians, African, etc., deaths.

Finally, we are not facing a binary choice between such compromised persons as, say, a Hitler's collaborator Nevanlinna and a Soviet anti-Semite Pontryagin. We do not need a stain on our profession of either kind. We can select for our medal a profile of a person we will all be proud of, such as Alan Turing, Claude Shannon, John von Neumann, Norbert Wiener, and a good number of other noble human beings and great scholars.

Best wishes,
Alexander Soifer

On September 30, 2017, Avi Wigderson of the Institute for Advance Study, Princeton, entered the discussion:

Dear Alexander,

I was approached about this issue about 10 years ago by someone I never knew, who told me more or less the story you tell, and asking how could I, as a Jew, accept this medal. I looked into the matter then, read some more, and basically got to essentially the same conclusion Peter did.

Let me summarize how I view it. I have no idea what I would have done in Nevanlinna's shoes as a Finn during Nazi rule of Finland. E.g., I cannot tell what pressures on his family he may have felt in that terrible time, which caused him to support in the way you carefully describe (like many Finns) Nazi Germany, indirectly helping their terrible deeds. I also cannot tell what else he did with his position during that time.

E.g., I recently learned that he is solely responsible to saving the life of Andre Weil from execution by the Nazis in Finland and arranging his return to France. I have no idea if he did other similar things. What mattered to me (and both my parents lost all their families to the Nazis) is that outside unbelievable unimaginable times of WWII and occupation of his country[53], what I know of Nevanlinna's life shows no sign of problematic opinions or behavior in my eyes. I cannot say that he is not overall "stained" by his actions during the war, but I can't say it is a part of his personality, and I can't say either that good people would behave otherwise then. Clearly, it would be nicer for the prize and for his memory in general if this part of his history was not there.

[53] This is incorrect: Finland was not occupied during WWII; in 1940, Soviet Union annexed part of the Finnish territory, while Finland remained a sovereign country.

Anyway, this is my personal view, and I completely understand and respect others, in particular your view and actions regarding it. Of course, there are many *professional* reasons why some mathematicians (as the ones you suggest) are better suited, to name a medal for work on the theory of computing.
Best,
Avi

About three hours later, Alexander Razborov of the University of Chicago, joined the conversation:

I would also throw in here imminent threats to the "professional family": students, colleagues etc. Most likely, the Weil story (of which I did not know before) is only a tip of an iceberg consisting of largely undocumented things.

Peter and Avi said most of what I thought of saying. The most important point in my view is that it is not fair to draw comparisons between "stained" people and stellar scientists whose nobility was simply never tempted on such an unprecedented level.

Let me perhaps also add my personal "credentials": I never knew my grandfathers, both of them died during the war. But I do not have any issues with this Finnish battalion fighting in [the] Soviet Union as long as they did not commit any war crimes[54]. Judging from what I saw on the Wikipedia, the Nazis used them mostly as a cannon feeder, and they never were a part of any occupation force. Perhaps, for a good reason:

I feel that the civil population would have been much better off under Finnish occupation (and sorry Alexander I do not mean abstract principles, just saying that more people would have survived).
Sasha

So, Alexander Razborov believed that Russia would have been better off under the Finnish occupation than under Stalin; not a very patriotic proposition. While this could be true, there had never been as much as a whisper that tiny Finland would invade the giant Soviet Union. There was a realistic chance of Adolf Hitler's rule in Russia, and this would have meant tens of millions more of dead Russians and a brutal slavery for the rest.

Summing up, the laureates believed that Rolf Nevanlinna was not bad enough to justify the renaming of the medal and the prize. It seems that if Nevanlinna were an anti-Semite, the laureates would be in favor of the prize's name change. Even though the Soviet authorities labeled me "Jewish" in my passport, and in spite of the horrors of the Holocaust, I for one do not consider anti-Semitism (contrasted to the "Final Solution") to be the main disqualification in choosing the prize's name. Evil wears many clothes.

In the end, the laureates were content having an active Nazi supporter on their medals and prizes.

Movement 5: Executive Committee and General Assembly of IMU

On August 10, 2018 (yes, over a year later), IMU President Mori reappeared:

Dear Professor Soifer,

[54] We know now that they did.

This is to let you know of the decision that IMU has finally [!] made at GA [General Assembly of IMU, July 30–31, 2018].

It is the Resolution 7 of the attached "RESOL2018.pdf," which you can also find under the item Resolutions of "18th GA in São Paulo, Brazil" in the URL https://www.mathunion.org/organization/general-assembly

Best regards,
Shigefumi Mori
President of the International Mathematical Union

Let me reproduce for you Resolution 7 approved by the IMU's General Assembly:

Resolutions of the IMU General Assembly 2018
Resolution 7

The General Assembly requests the 2019–2022 IMU Executive Committee, giving due consideration to all the issues involved, to determine and set up statutes for a prize continuing and with the same purpose and scope as the Nevanlinna Prize but with a new name and appropriate funding to be secured. The statutes of the new prize will be sent to the Adhering Organizations for approval by a postal ballot.

In 2018 it looked as if the Executive Committee recommended the General Assembly (GA) to replace Nevanlinna's name with something decent, and GA agreed. It proved to be not true. In fact, the sailing to this destination was very stormy. A year and a half later (!), **on January 10, 2020**, we learn troubling details "*As approved by the IMU Executive Committee Report of the 18th General Assembly of the International Mathematical Union (IMU) São Paulo, Brazil July 29, 2018*" [IMU1]. Let me quote only the relevant part of the Report, emphasis in bold is mine:

Report of the 18th General Assembly of the International Mathematical Union (IMU) São Paulo, Brazil July 29, 2018. (09:00–18:15 hrs) July 30, 2018. (09:30–17:00 hrs)[55]

- Adjustment of the Nevanlinna Prize, Presented by Shigefumi Mori, IMU President. The IMU President read the following statement. It was decided at the 88th meeting of the IMU Executive Committee (EC) in March 2018 in Montreal that **2018 would be the last year at which a Rolf Nevanlinna Prize be awarded by the IMU**. The unanimous vote of the EC [after all, I got through to every member of EC!] on the matter was made after consulting with and under the agreement of IMU's Adhering Organization in Finland which was and will be responsible for funding the prize until the ICM 2018.

  Background and rationale for the decision: The history of the prize is that in 1981 the Executive Committee of the IMU decided to establish a prize in information sciences and in 1982 accepted an offer from the University of Helsinki to finance it. Consequently [who would believe that Helsinki offered to fund the prize not knowing that it honored Finn Nevanlinna?], it was named after a former President of the IMU, Rolf Nevanlinna (1895–1980), who had been Rector of the University of

[55] chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/ https://www.mathunion.org/fileadmin/IMU/Organization/GA/GA_2018/18%20GA%20Report%20Final%20180729%20Sao%20Paulo.pdf

Helsinki and who had taken various initiatives to organize computer science in Finnish universities. Since its inception the prize laureates have all been individuals of outstanding ability who have made ground-breaking contributions to theoretical computer science and the IMU is extremely grateful to the Finnish community for its generous sponsorship of this initiative. However, **the IMU has been approached over a long period and by several individuals regarding the naming of this Prize. These individuals found it unacceptable that the IMU should award a prize that carries the name of Rolf Nevanlinna, due to allegations regarding his behavior during World War II**. [What an admission! "IMU has been approached over a long period and by several individuals" and IMU *did noting*, kept Hitler lover Nevanlinna on the prize and gold medal!] While the IMU is in no position to pass judgement in this matter, **the IMU EC decided that the matter could not be ignored as it could tarnish the reputation of the IMU and the Nevanlinna prize recipients.**

The Executive Committee asks the General Assembly to consider the question on the type of future involvement of the IMU in a prize in this area. Background for the question: After careful consideration, and in view of the need to emphasize the unity of mathematics that it stands for, the IMU EC felt that it is no longer appropriate for IMU to single out for recognition and encouragement of this particular and important area of mathematics only, in the same way as the Nevanlinna Prize had been set up to do. The EC felt that this prize has successfully served its purpose, to foster and encourage research in this direction, that has now become a core part of contemporary mathematics and that extraordinary work in this field can naturally be recognized by a Fields medal. The current rules for instance mean that a Nevanlinna Prize winner (say at the age of 35) becomes ineligible for a Fields Medal four years later, but there is no procedure in place to handle the case where the Fields and Nevanlinna committees choose the same name. Furthermore, the Nevanlinna Prize identifies two areas within Mathematical Aspects of Information Sciences, namely "mathematical computer science" and "computational mathematics." To date the Nevanlinna Prize has only been awarded in the first area. On the other hand, the Nevanlinna Prize is an established important prize for the theoretical computer science community, and it would be natural that some new prize in this area should be set up. IMU bears a responsibility for this important prize and to its past Nevanlinna prize laureates. In any case, setting up such a new prize is a delicate issue. Name, choice of funding and institutional partners, role of IMU and all the implications of these choices would need to be carefully thought through. The IMU is involved with some prizes that are not IMU prizes per se, but carry the IMU approval, e.g., the Ramanujan prize for young mathematicians from developing countries, which is awarded jointly by the Abdus Salam International Centre for Theoretical Physics (ICTP) in Trieste, Italy, the Department of Science and Technology of the Government of India (DST), and the IMU. The IMU EC therefore asks the Adhering Organizations of the IMU to decide between the following three alternatives:

Alternative 1: The IMU takes no further action. A possible successor of the Nevanlinna Prize would be without formal IMU involvement.

Alternative 2: The 2019–2022 IMU EC is asked to participate in the creation of a new prize in collaboration with another institutional partner. The scope of this prize and the IMU involvement would mean that this prize might be recognized as a natural successor of the Nevanlinna Prize. The statutes for this new prize would avoid the unwanted risk that work on Nevanlinna Prize themes might de facto be excluded from consideration for Fields Medals.

Alternative 3: The 2019–2022 IMU EC is asked, after it has given careful consideration to all the issues involved, to set up a prize with a new name and funding yet to be determined, with the same purpose and scope as the Nevanlinna Prize. The only difference would be the name and funding. The side effect that work on Nevanlinna Prize themes might de facto be excluded from consideration for Fields Medals, would persist. In case Alternative 2 or 3 is selected, the statutes of the new prize will be sent to the Adhering Organizations for approval by a postal ballot.
[It is clear that EC prefers the Alternative 1: to shut down the Rolf Nevanlinna prize and walk away.]

After discussion of the issue [that, according to Avi Wigderson's email to me, included a passionate address by László Babai and letters from the former Nevanlinna Prize winners in favor of Alternative 3] a vote was made in order to choose the favored Alternative to be put to vote afterwards.
VOTE (by show of hands) in order to select between Alternative 1, 2, or 3:

IN FAVOR of Alternative 1 = 10; IN FAVOR of Alternative 2 = 4; IN FAVOR of Alternative 3 = Majority [ca. 136].

The GA proceeded to vote on Alternative 3:

The General Assembly approved that the 2019–2022 IMU EC is asked, after it has given careful consideration to all the issues involved, to set up a prize with a new name and funding yet to be determined, with the same purpose and scope as the Nevanlinna Prize. The only difference would be the name and funding. The side effect that work on Nevanlinna Prize themes might de facto be excluded from consideration for Fields Medals, would persist.
VOTE (by show of hands): IN FAVOR = 140, OPPOSED = 2, ABSTENTIONS = 8.

Thus, we know now that "the IMU has been approached over a long period and by several individuals regarding the naming of this Prize." IMU and its EC did nothing, just ignored complainers like annoying flies. It took my commitment, with letters and emails, public speaking, and several articles published in *Geombinatorics* and posted on the Internet to *make* IMU EC realize that they *finally* must act.

Movement 6: IMU Abacus Medal[56]

So, the General Assembly of IMU finally voted to change the name of the Rolf Nevanlinna Prize and Rolf Nevanlinna Medal. However, the new name of the award puzzled many, me included. From now on, IMU will grant four Prizes for mathematical achievement: Fields Medal, Carl Friedrich Gauss Prize, Chern Medal Award, and ... IMU Abacus Medal. Imagine, "Abacus" is all that the 12 EC members were able to come up with!

Here is how EC explains its "abacuous" decision, reached during its March 1–3, 2019, Berlin meeting [IMU2]:

[56] See my first reaction to Abacus here [Soi42].

> **The naming of the prize.**
>
> There are many constraints to be considered, and the EC spent considerable time discussing various options. The name must not refer to commercial companies, any country, or any living persons. The name should be concise, comparable to other IMU awards. [Is "Abacus" really comparable to "Gauss," "Chern," Fields"?] An aspect to be considered is that if we name the award after a person, we should be convinced that the person has an unblemished reputation [sic]. We should avoid the name of any existing international prizes in related areas, both out of respect to other organizations and to escape misunderstandings. Also, generic names like IMU Prize in the Mathematics of Computation or IMU Prize on the Mathematical Aspects of Information Sciences were discussed but they were found to be too long.
>
> The name IMU Abacus Medal relates to the abacus, an ancient device that was used for numerical computations, and it underscores the importance of calculations already in early mathematics. The exact place and time of origin of the abacus is unknown, and it can be considered a truly global artifact associated with mathematics and computation.

It seems that my *public* revelations about the Nevanlinna past scared IMU Executives so much that they walked away from a tradition of having a person's name and profile on the medal, as they and others had almost always done, and came up with a very safe and very lame name "IMU Abacus Medal." Even being scared of using a person's name, any name, they could have chosen "Mathematical Computing Medal," which is not longer than "Rolf Nevanlinna Prize" or "Carl Friedrich Gauss Prize." Imagine, a scholar receiving Abacus for a major theoretical computer science achievement! Why not an even older one, a bone with markings, or a pile of pebbles? The dozen Executives must have loved ambiguity, for they proudly point out that "the exact place and time of origin of the abacus is unknown."

By the way, IMU's new name for its prize is not new: a simple Internet search shows that Upsilon Pi Epsilon, International Honor Society for the Computing and Information Disciplines had Abacus Award for a while [UPE].

IMU did the right thing in disassociating itself from a Nazi-supporter. However, IMU had an obligation to explain to the world the reasons for their rare and decisive correction of their old error. Yet, the fearful IMU failed to do it, as failed Heidelberg Laureate Forum. On May 14, 2019, The London Mathematical Society did not explain the reasons either, but at least included a hint in its announcement of IMU Abacus Medal, as if leaving a historical research to the homework of the readers:

> In 2018, the IMU Executive Committee took the decision to discontinue this prize because of historical issues arising from its name.

Not all hid behind the trees from responsibility. Wikipedia [Wiki] gets the Nevanlinna Prize story right. William Gasarch in Computational Convexity [Gas] presents very good arguments, using even mathematical constructions in the process.

Meanwhile, Rolf Nevanlinna is departing from the public domain, and this is good for the integrity of Mathematics. On July 3, 2022, Avi Wigderson wrote to me "I am glad the name was changed."

Ron Graham shared with me his opinion that while in my call for the name change, I was "making some good points, the chances of IMU changing anything are very slim." I thought so too. However, we ought to do all we can and let the chips fall where they may. Anything less would compromise our integrity and guarantee the victory of the status quo in this world that needs so much change. One person empowered by truth and glasnost, can affect a major change.

Chapter 43
How the Monochromatic AP Theorem Became Classic: Khinchin and Lukomskaya

> *What amazes us today is, of course, that no one in Hamburg (including Schreier and Artin) had known about Schur's work [1916]. In that connection we must realize that the kind of mathematics involved in the [Baudet–Schur] conjecture was not mainstream, and that combinatorics was not a recognized field of mathematics at all.*
>
> – Nicolaas G. de Bruijn (E-mail to A. Soifer, *January 5, 2004*.)

Is it possible to find a needle in a haystack? Yes, it is, with good fortune and great perseverance. We discussed the birth of this monochromatic AP theorem earlier in the book, so by now, it should be "old news" to you. In fact, it has taken 20 years and a Russian aid for this theorem to become classic. As you recall, its 1927 publication [Wae1] in a small-circulation Dutch journal hardly helped its popularity. Only two Japanese mathematicians [KM] and Issai Schur with his two students Alfred Brauer and Richard Rado realized its importance and improved upon Van der Waerden's result almost immediately; and later, in 1936, Paul Erdős and Paul Turán commenced density considerations related to Van der Waerden's result [ET].

In 1928, a Russian visitor to Göttingen and a fine analyst Alexander Yakovlevich Khinchin[1] (1894–1959) heard about Van der Waerden's proof and became very impressed by it. All right, so one Russian liked it. You may be wondering, what is a big deal? A very long time had passed, 19 years and a horrific war to be exact, but Khinchin remembered his Göttingen excitement and after World War II, in 1947, included Van der Waerden's proof in his little book *Three Pearls of Number Theory* as one of the pearls [Khi1]. The booklet was an instant success, and the second edition came out in Russian in 1948 [Khi2]. It included a new "much simpler and transparent" proof, in the opinion of Khinchin, found by the Russian mathematician M. A. Lukomskaya. Do you know who Lukomskaya was? No? You are not

[1] My wonderful high school mathematics teacher Tatiana Nikolaevna Fideli was an M.S. Khinchin's student at Moscow State University. It is a small world!

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_43

alone: I knew nothing about her and did not expect to ever find out when Google informed me that the biography of Van der Waerden in *The Mathematical Coloring Book* [Soi44] inspired a discussion on the Russian Scientific Forum http://dxdy.ru/topic19166.html. On January 14, 2009, someone named "*Geomath*" wrote (in Russian):

> In this translated into Russian [Soi39] biography of Van der Waerden, which is a part of the [English language] book [Soi44], its author, a mathematician-Jew, our former compatriot, researches in a most meticulous way and gives a moral assessment of the fact that Van der Waerden, while remaining a Dutch citizen, taught mathematics in Nazi Leipzig, even during the five years when Germany occupied the Netherlands.

The following day, *Geomath* continued:

> The "new and much simpler and transparent proof" of Van der Waerden's theorem was found by M. A. Lukomskaya and published in UMN in 1948 [Luk]. Who is Lukomskaya? What has happened with her? If she was young, then, with time, she had a good chance to develop into a famous mathematician . . . However, I was unable to find anything about her on the Internet. Perhaps, she changed her last name?
>
> By the way, the mentioned by me book *The Mathematical Coloring Book: Mathematics of Coloring and the Colorful Life of Its Creators* by Alexander Soifer (together with a biography of Van der Waerden in it) can be downloaded free, I have already done so.

I inquired from *Geomath* and the forum why he chose to characterize me as a "mathematician-Jew," but did not get an answer. A year later (!), on January 12, 2010, a surprising reply was posted by someone nicknamed Elena31. It was her first and only appearance in the Forum:

> Lukomskaya Mira Abramovna (my mother) was born on May 1, 1900, and passed away on October 30, 1976. She graduated from Leningrad [State] University, phys-math [faculty], and for many years worked as a docent (equivalent to an associate professor) at the Belarus State University. Her works were primarily on differential equations. I remember well how she solved the problem you mentioned.
> Sincerely, E. N. Lambina

Sherlock Holmes was reawakened in me. The same day, I sent a private e-mail to Elena:

> Dear Lena,
>
> Tell me please in detail about your mother and even in greater detail about when and *how* she worked on a solution of the problem.
>
> Among other things, I am a biographer of Van der Waerden and the author of "The Mathematical Coloring Book," mentioned in the discussion.
> All the best,
> Alexander

Two days later, on January 14, 2010, I learned from Elena about her mother, the author of the second proof of the Baudet–Schur–Van der Waerden Theorem (I am translating for you from her Russian original):

> Dear Alexander!

Thank you for your interest. I will try to answer your questions. My parents, Mira Abramovna Lukomskaya and Nikolay Venediktovich Lambin were both mathematicians. Mama was born in Bykhov of the Mogilev region [What a coincidence: Issai Schur was born in Mogilev!], in 1917 graduated from Mogilev Gymnasium for Women, and during 1920–1925 studied at the Petrograd University, renamed into Leningrad [State] University. There she met my papa, a native of Petersburg. Upon graduation from the university, they both worked at Pulkovo Observatory and in the Meteorological Institute. In 1930 they both moved to Minsk, where they worked at Belarus State University at phys-math [faculty] (renamed into math-fac, and then mech-math). During the war they worked in Kazan, at the Defense Institute of the USSR Academy of Sciences. (First war winter, mama and I lived in the village Kulaevo, 35 km from Kazan; Mama taught there almost all disciplines – from Minsk we were able to walk away on our own feet [Minsk was severely bombed on day one of the German invasion of Russia, June 22, 1941, and occupied four days later]. In the fall 1944, we returned to Minsk together with the university, where mama worked as a docent through mid-1960s, and papa until the early 1970s.

Now about the theorem. When mama was solving it, my brother and I were 16 years old each (the end of 1947) and therefore I can share with you only the following. In the first edition of Khinchin's book, mama read a proof of this theorem and right away said that one can prove it simpler . . . (Mama was interested in number theory, and in her youth even spent a week trying to prove Fermat's Last Theorem ☺), but her publications, except the one of your interest, belong to differential equations . . . Mama used to say that the essence of the theorem is this: "*Any chaos contains its own order.*"[2] She jokingly applied it to some chaos in our apartment (although, our apartment consisted then of only one room in the former kindergarten, where the returning to Minsk university employees were housed). Having proved the theorem fairly quickly (in ten days or so, as I recall), mama wrote to Khinchin at his MGU [Moscow State University] address; and got a reply where Khinchin approved her solution and offered some improvements. He asked for a permission to publish it in the new edition of his book (which is what he did), and also offered to publish it in [the journal] UMN (*Uspekhi Matematicheskih Nauk*), which is what mama did . . .

Respectfully, Elena Nikolaevna Lambina (I graduated from MGU in 1954, and several decades worked as a docent in the department of theoretical mechanics of Belarus Polytechnic Institute).

I asked Elena for copies of her mother's publication and her correspondence with Khinchin. On May 24, 2010, Elena kindly sent me the journal publication [Luk] of her mother's proof of the Baudet–Schur–Van der Waerden Theorem, and copies of the letters her mother exchanged with Khinchin. Now I can convey the rest of the story. As you already know, in late 1947–early 1948, Mira Lukomskaya sent her proof to Moscow State University Professor Alexander Khinchin, who replied on February 9, 1948 (I am translating from Russian):

[2] What a wonderful description of Ramsey Theory! Clearly independent, it reminds the famous description attributed to Theodore Motzkin: "Complete disorder is impossible."

Much Honored Mira Abramovna,

Your proof of Van der Waerden's Theorem, which was forwarded to me, is incredibly interesting. Based on the same idea as the original author's proof, it uses a much simpler and more transparent construction, whereby the proof is reduced to at most half the length and is much more accessible. I only think that your resorting to infinite fractions is unnecessary and only complicates the matter, and even raises some doubts (which probably are easily resolvable). I think that it is much more convenient to realize your construction directly on a finite segment and, so to speak, in reverse order (i.e., from large segments to small ones). I am mailing to you the corresponding presentation in its complete form on two pages. Of course, you will see right away that in spite of a different setup, it is not a new but exactly your construction. I am interested in knowing your opinion about my editing.

Are you going to publish your proof? At the moment, I have a favor to ask you. My "Three Pearls" will soon be published in the second edition, and I ask for your permission to allow me to include your proof in the first chapter, your proof instead of the old one (of course, with a clear indication of your authorship).

With sincere respect,
Khinchin

On June 30, 1948, Mira Abramovna replied to Khinchin:

Much respected Alexander Yakovlevich,

In accordance with your advice, I am sending you my work on the Theorem of Van der Waerden. I chose your method of presentation, as it is preferred over mine in its conciseness and clarity. If this work can be published in "Uspekhi [Matematicheskih Nauk"] ["Successes in the Mathematical Sciences"] or another journal, would you be so kind to forward it for publication. For this case, I am sending you two copies ...

I am reading this short 3 ¼-page article [Luk]. It actually contains a *generalization* of Van der Waerden's result, which I would call *The One-Dimensional Gallai Theorem,* obtained in 1947 and published in 1948 independently from Tibor Gallai! Let me translate the theorem and give it a well-deserved title:

The Lukomskaya Theorem 43.1 [Luk]. Given an infinite sequence of positive integers t_1, $t_2, \ldots, t_q, \ldots$ Then for any pair of positive integers k, l there is a positive integer $n(k, l)$ such that if any array of consecutive positive integers of length $n(k, l)$ is partitioned into k classes, there are in at least one class l numbers $c_1, c_2, \ldots, c_l$, satisfying the condition

$$(c_2 - c_1) : (c_3 - c_2) : \ldots : (c_l - c_{l-1}) = t_1 : t_2 : \ldots : t_{l-1}.$$

As you can readily see, Van der Waerden's result is but a particular case of Lukomskaya's Theorem for $t_1 = t_2 = \ldots = t_l = 1$.

Khinchin should have included Lukomskaya's Theorem in the new edition of his book, regretfully he did not. In the second 1948 edition [Khi2] of his *Three Pearls of Number Theory*, Khinchin chooses to include Lukomskaya's proof just for this particular Van der Waerden's case.

The success of this little booklet is hard to overestimate. In 1951, this second Russian edition of the book was translated into German [Khi3] and in 1952 into English [Khi4]. The

English edition becomes so popular that in 1956, the publisher issued the second printing. These translations prove instrumental in creating excitement about Ramseyan ideas all over the world. They even encourage the emergence of two more independent proofs of the Gallai Theorem, i.e., generalizations of Van der Waerden's result. The 1951 German translation inspires Ernst Witt, a former Emmy Noether student, to discover his proof ([Wit], submitted on September 21, 1951, and published in 1952). The 1952 English translation stimulates Adriano Garsia in finding his proof [Gar] in 1958. Khinchin writes [Khi2]:

> It is not out of the question that Van der Waerden's theorem allows an even simpler proof, and all efforts in this direction can only be applauded.

Witt [Wit] quotes this Khinchin's call to arms in his paper and happily reports:

> This was the occasion to strive for a new order of proof that then led directly to a more general grasp of the problem.

The great success of this booklet not only makes the Baudet–Schur–Van der Waerden Theorem famous – it heralds to the mathematical world the arrival of the new Ramseyan ideas.

Van der Waerden published his proof in an obscure little read Dutch journal. Now Khinchin's book and its popularity prompt Van der Waerden to reassess the value of his old theorem. He reconstructs the whole process of finding the proof in "How the proof of a Baudet's conjecture was found" by him with Emil Artin and Otto Schreier. Van der Waerden also speaks about the process of discovery at his inaugural lecture at Zurich University. He then publishes this fascinating story several times in two languages ([Wae13], [Wae14], [Wae16], and [Wae18]). I too love this story, and, with Professor Van der Waerden's and Academic Press London's permissions, included the complete story in this book (Chapter 35).

On November 10, 1953, Van der Waerden sends the proofs of the series of three articles tracing the processes of mathematical discoveries to a witness and coauthor of one discovery Emil Artin at Princeton University, accompanied by the following letter:

> Dear Herr Artin,
>
> My three articles "*Einfall und Ueberlegung in der Mathematik*" ["Sudden Insight and Reflection"] will be published in the Swiss journal "*Elemente der Mathematik*." The first one was my inaugural speech [at Zurich University]; I am sending you its proof. The second and third articles give two more examples, for which I give short descriptions, in particular, the second one is measuring a ball by Archimedes; the third is the Baudet Conjecture, for which the three of us found a solution. From the proof that I published in *Nieuw Archief*, no one can see how I came to it, and what role you and Schreier had in finding the solution. For the psychology of mathematical thought this case is particularly promising because all of our thoughts were immediately communicated to both of the others and thereby were held better in memory, as usual.
>
> I hope you will read it and tell me if everything is exactly in agreement with your memory.
>
> Heartfelt regards and to your wife as well,
> Your
> B.L. van der Waerden

Artin replies with an undated handwritten letter on the Princeton University stationary:

Dear Herr van der Waerden,

You have a better memory than I. I could have never reconstructed our conversation. Now that I have read your analysis, I remember it again. I have the impression that it is by and large correct. But you certainly know it best.

With many greetings,

Your

Artin

At this point in time, Van der Waerden realizes that in his young years he proved a beautiful theorem that has now become a classic, a "pearl of number theory." By now, he has also become a historian of mathematics. So, of all people, he is in the best position to research and present the history of mathematics of coloring, now known as Ramsey Theory. He has not done that. In fact, as he writes to me, he did not even know about Issai Schur's 1916 theorem, the first influential coloring result in history. And so, the job of researching and preserving the history of mathematics of coloring fell on me and materialized in the first 2009 edition of this book.

On May 27, 2009, during the international workshop "Ramsey Theory Yesterday, Today, and Tomorrow" that I organized on the request of Director Fred Roberts and DIMACS' Executive Committee, I got an additional confirmation of the influence of Khinchin's book. Leaders of Ramsey Theory Ronald L. Graham and Joel H. Spencer told me that this Khinchin's book introduced them both, for the first time, to the name of Van der Waerden, his theorem, and Ramseyan ideas!

Part VIII
Colored Polygons: Euclidean Ramsey Theory

There is a running discussion between Dieudonné and Branko Grünbaum. Dieudonné sort of says that geometry is dead and of course Branko Grünbaum disagrees with him. I think I am on the side of Branko Grünbaum and I hope that I will convince you that at least combinatorial geometry is not dead.

– Paul Erdős[1]

[1] [E83.03].

Chapter 44
Monochromatic Polygons in a 2-Colored Plane

We have already met briefly a 2-colored plane in Problem 2.1, which can be restated as follows:

Problem 44.1. For any positive d, any 2-colored plane contains a monochromatic segment of length d.

The next exercise is a homework. :)

Problem 44.2 For any positive d, any 2-colored plane contains a nonmonochromatic segment of length d if each of the two colors is present in the plane.

Let me remind you that in our discussions, a *triangle* stands simply for a 3-element set. When these three points are on a line, we will call the triangle *degenerate*. Accordingly, a set of n points in the plane will be called an *n-gon*. An *n-gon* with all n vertices in points of the same color is called *monochromatic*.

You may wonder why after discussing a multicolored plane should we now talk about a mere 2-colored plane? Would it not be more logical to put this chapter first in this book? Yes, it would. But this logical approach creates, as Cecil Rousseau put it ([Soi1], introduction), "books written in a relentless Theorem-Proof style." This logical approach ignores a higher logic of mathematical discovery.

For me, personally, a fascination with the chromatic number of the plane problem came first. Then I looked at a 2-colored plane. Why? If we can prove the existence of certain monochromatic configurations in any 2-colored plane, we will have tools to study a 3-colored plane. And some configurations present in any 3-colored plane may provide tools to attack a 4-colored plane. And it is a 4-colored plane where we 'only' need to find out whether a monochromatic segment of length 1 is necessarily present.[1] Then a 5-colored plane will hopefully appear in our field of vision.

With this rationale in mind, in 1989–1990 I proved some results, formulated conjectures, and thus rediscovered *Euclidean Ramsey Theory*. I published a problem essay [Soi2] about it

[1] Of course, others may have had different reasons for looking at 2-colored planes. Erdős at al. in their trilogy [EGMRSS] were pursuing expansion of Ramsey Theory to Euclidean Ramsey Theory.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_44

in the first issue of volume I of the newly founded research quarterly *Geombinatorics*, when on July 5, 1991, Ron Graham sent me a copy of the series of 3 papers by 6 authors, which broke the news to me: I was 15–17 years too late: Paul Erdős et al. were first to discover what they named *Euclidean Ramsey Theory*! Fortunately, some of my results remained new, and you will see them in this chapter (Problems 44.7, 44.14, 44.19), which is chiefly dedicated to Erdős et al. series of papers [EGMRSS]. Paul Erdős referred to the authors as "us", or "the six." The distinguished 6 authors deserve to be all listed here. They are Paul Erdős, Ronald L. Graham, P. Montgomery, Bruce L. Rothschild, Joel H. Spencer, and Ernst G. Strauss.

When three of these six coauthors wrote *Ramsey Theory* monograph (1st edition [GRS1] in 1980; 2nd edition [GRS2] in 1990), they did not include much of the trilogy [EGMRSS] results in their book. Perhaps, they viewed these results to be too 'elementary' for their dense monograph. On the other hand, they realized how difficult these 'elementary' problems can be, for Paul Erdős and Ron Graham included open problems of Euclidean Ramsey Theory in many of their (open) problem talks and papers. It seems that most of these results of "the six" and other results of Euclidean Ramsey have appeared in a book form for the first time in the first edition [Soi44] of this book.

Problem 44.3. (Erdős et al. [EGMRSS]). Two-color the plane to forbid a monochromatic equilateral triangle of side d.

Solution Divide the plane into parallel stripes, each $\frac{\sqrt{3}}{2}d$ wide ($\frac{\sqrt{3}}{2}d$ is the altitude of the equilateral triangle of side d), then color them alternatively black and white (Fig. 44.1). Include in each stripe region its left boundary line, and do not include its right boundary line, and we are done. ■

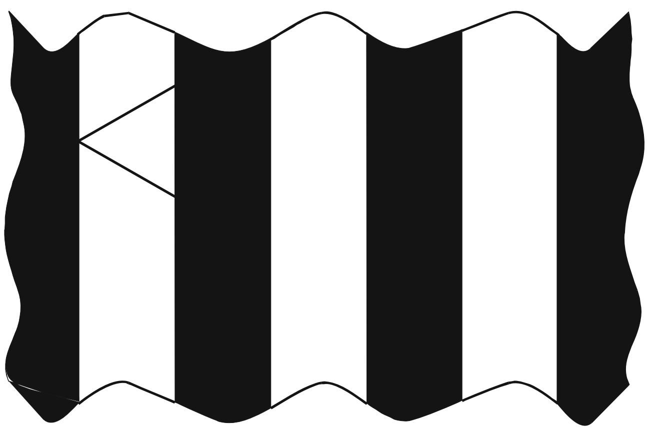


Fig. 44.1 Striped 2-coloring of the plane

Problem 44.4 (Erdős et al. [EGMRSS]). Find a 2-coloring of the plane, different from the one in the solution of Problem 44.3, that does not contain a monochromatic equilateral triangle of side d.

Solution Start with the coloring described in the solution of Problem 44.3 (Fig. 44.1). Draw a line making, say, a 60° angle with the boundary lines of the stripes (Fig. 44.2), and change the colors of the points of their intersections. It is easy to verify that, as before, the plane does not contain a monochromatic equilateral triangle of side 1. ■

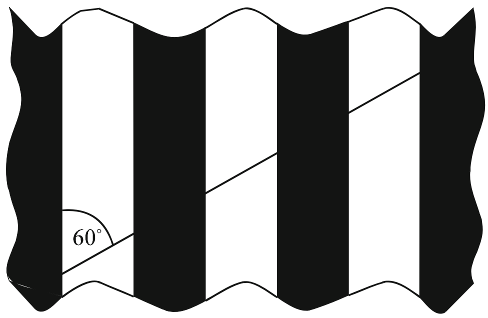


Fig. 44.2 Alternative striped 2-coloring of the plane

If you solved Problem 44.4 on your own, you have probably noticed that your and my solutions did not differ much from each other and from the solution for Problem 44.3. In fact, Paul Erdős et al. thought that the solutions cannot differ much.

Conjecture 44.5 ([EGMRSS], Conjecture 1 of Part III). The only 2-colorings of the plane for which there are no monochromatic equilateral triangles of side 1 are the colorings in alternate strips of width $\frac{\sqrt{3}}{2}$, as in the solution of Problem 44.3, except for some freedom in coloring the boundaries between the strips.

Decades had passed; Ronald L. Graham and Paul Erdős repeated problems and conjectures of the Euclidean Ramsey Theory, including 44.5, in their talks and papers (see, for example [E8303]), but no proof was found to these easy-looking, hard-to-settle triangular conjectures. However, in March 2006 a group of four young Czech mathematicians from Charles University located on Malostranské plaza (in 1996 I spent two months at this historic place as a guest of Jaroslav Nešetřil) Vít Jelínek, Jan Kyncl, Rudolf Stolar, and Tomás Valla [JKSV] disproved this 33-year-old conjecture!

Counterexample 44.6 ([JKSV, theorem 3.19]) Every zebra-like 2-coloring of the plane has a twin 2-coloring that forbids monochromatic unit equilateral triangles.

For definitions of "zebra-like" 2-coloring of the plane and of "twin" coloring, I refer you to the original work, which after over 3.5 years from the March-2006 submission was finally published in *Combinatorica* in November-2009. Fortunately, the authors made their paper available in arXiv in January-2007. Here I would like to show an example of a zebra-coloring, kindly provided to me by one of the authors, Jan Kyncl (Fig. 44.3).

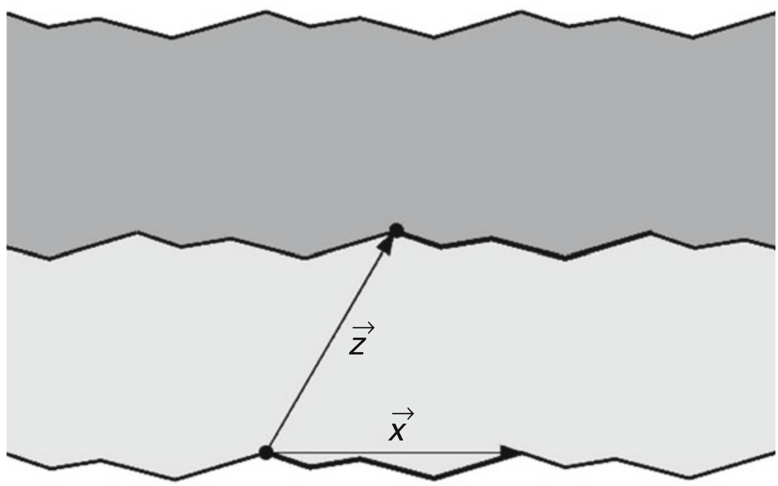


Fig. 44.3 Zebra 2-coloring of the plane

I congratulate the authors with introducing into the field a totally new rich class of 2-colorings of the plane, and with solving, in the negative, the old-standing conjecture by Erdős et al.

Any equilateral triangle can be excluded from appearing monochromatically by choosing an appropriate 2-coloring of the plane. Some triangles, however, exist monochromatically in any 2-colored plane. The first such example was found by Paul Erdős et al. [EGMRSS].

Problem 44.7. ([EGMRSS]). Any 2-colored plane contains a monochromatic triangle with a small side 1 and angles in the ratio 1:2:3.

My Solution [Soi9]. Pick a monochromatic segment *AB* of length 2 (Problem 44.1 guarantees its existence) and construct a regular hexagon *H* on *AB* as on the diameter (Fig. 44.4). If at least one more vertex of *H* is of the same color as *A* and *B*, we are done. If not, we are done too! ■

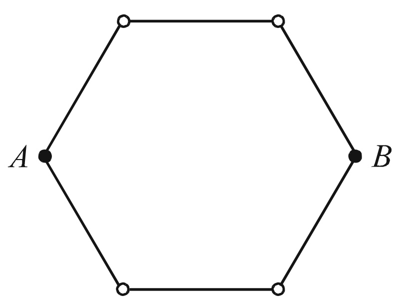


Fig. 44.4 2-coloring of the vertices of a regular hexagon

I offered Problems 44.7 and 44.8 to middle and high school students during the Colorado Mathematical Olympiad in 1990. The first was solved by several young Olympians. Nobody solved the second one:

Problem 44.8. ([Soi2]). Any 2-colored plane contains a monochromatic triangle with small side 1 and angles in the ratio 1:2:4.

Solution [Soi9] Assume that such a triangle does not exist in a 2-colored red and blue plane. Toss a regular 7-gon of unit side in the plane (Fig. 44.5). Since 7 is odd, two of its consecutive vertices will be of the same color. Say, A and B are blue. Then D and F must be red. Therefore, C and G are blue. We got a blue triangle CAG in contradiction to our assumption. ■

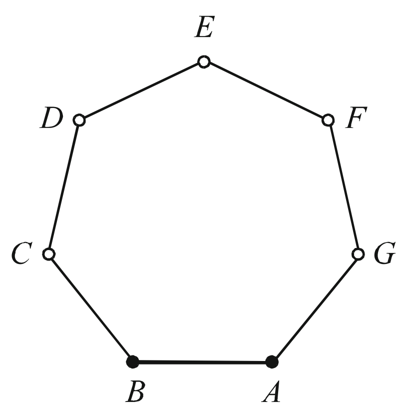


Fig. 44.5 2-coloring of the vertices of a regular heptagon

Problem 44.9 ([Soi2]). For any positive integer n, any 2-colored plane contains a monochromatic triangle with a small side 1 and angles in the ratio:

(a) $n : (n + 1) : (2n + 1)$;
(b) $1 : 2n : (2n + 1)$.

Proof Assume that a 2-colored plane P (red and blue) does not contain a monochromatic triangle with small side 1 and angles in the ratio $1 : 2n : (2n+1)$. Let the length of the main diagonals of a regular $(4n+2)$-gon of side 1 be d. Due to Problem 44.1, we can find in the plane P a monochromatic (say, red) segment S of length d. We construct on S as on a diameter, a regular $(4n+2)$-gon K. We then number the vertices of K starting with an endpoint of red diameter S (Fig. 44.6).

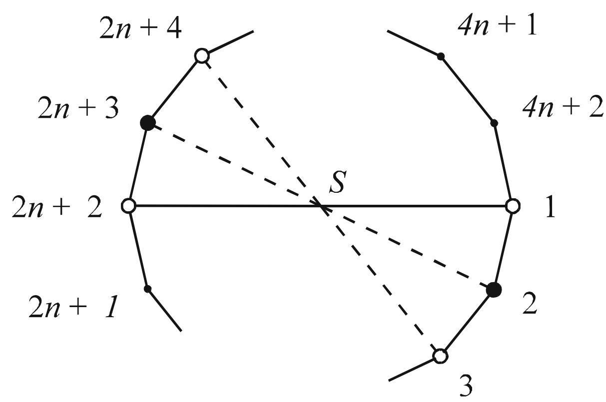


Fig. 44.6 The rotational argument

Now we start a rotational argument. The points 1 and $2n + 2$ are red, therefore the points 2 and $2n + 3$ are blue. Thus, the points 3 and $2n + 4$ are red, etc. Finally, the points $2n + 2$ and 1 are blue, which is a contradiction.

The existence of a monochromatic triangle with angles in the ratio $n : (n{+}1) : (2n{+}1)$ can be proved by a similar rotational argument. (Instead of adding 1 to the endpoint numbers of the diameter, we just add $n + 1$.) ■

Leslie Shader from the University of Wyoming proved the following result.

Problem 44.10 (L. Shader, [Sha]). For any right triangle *T*, any 2-colored plane contains a monochromatic triangle congruent to *T*.

As you can see, we have lots of examples of triangles that exist monochromatically in *any* 2-colored plane, and one example of a triangle (equilateral) that may not. Having realized this, I posed the following $25 problem to my university and high school students in 1989 (published in [Soi2]).

Open $25 Problem 44.11 [Soi2]) Find all triangles *T* such that any 2-colored plane contains a monochromatic triangle congruent to *T*.

Paul Erdős et al. tried to solve this very problem earlier. Moreover, they posed the following conjecture in 1973.

Conjecture 44.12 ([EGMRSS], Conjecture 3 of Part III). For any non-equilateral triangle *T*, any 2-colored plane contains a monochromatic triangle congruent to *T*.

This problem appears surprisingly difficult, and in 1979 Paul Erdős offered a prize for its first solution [E79.04]:

> Many special cases have been proved by us (i.e., the authors of [EGMRSS]) and others but the general case is still open, and I offer 100 dollars for the proof or disproof.

In 1985 Erdős increased the payoff [E85.01]:

> Is it true that every non-equilateral triangle is 2-Ramsey in the plane (i.e., Conjecture 44.12)? I offer $250 for a proof or disproof.

Let me formally attach Erdős' price tag to the above conjecture:

Paul Erdős' $250 Conjecture 44.13 Is it true that any 2-colored plane contains any non-equilateral triangle monochromatically?

Paul Erdős et al. also conjectured that any 2-coloring of the plane may not contain monochromatically at most an equilateral triangle of one size.

Conjecture 44.14 ([EGMRSS]). If a 2-colored plane P does not contain a monochromatic equilateral triangle of side d, then P contains a monochromatic equilateral triangle of side d' for *any* $d' \neq d$.

In 2003 Ron Graham [Gra5] offered \$100 for the proof:

Ronald L. Graham's \$100 Conjecture 44.15 Every 2-coloring of the plane contains a monochromatic copy of every triangle, except possibly for a single equilateral triangle.

My intuition regarding the above conjectures agrees with the authors of [EGMRSS], except I am not sure about degenerate triangles.

Open Problem 44.16 Is it true that any 2-colored plane contains a degenerate isosceles triangle of small side 1 (Fig. 44.7)?

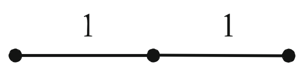


Fig. 44.7 A degenerate isosceles triangle

In order to solve the above open problems, you need tools. Here are two for you.

Let T be a triangle. Then T_m will stand for *the triangle whose sides are twice as long as the corresponding medians of* T (the medians of any triangle are themselves the sides of a triangle – prove this nice elementary fact on your own).

Tool 44.17 ([Soi2]). For any triangle T, any 2-colored plane contains a monochromatic triangle congruent to T or to T_m.

Proof Let the side lengths of T be a, b, and c, and P be a plane colored red and blue. If both colors are not present in P, we are done. Otherwise, by Problem 44.2, P contains a segment AE of length $2a$ with blue A and red E. The midpoint C of AE has the same color as A or E, let it be blue as A.

We pick B and D such that $ABCD$ is a parallelogram with side lengths b and c (Fig. 44.8). If at least one of the points B, D is blue, we get an all-blue triangle ABC or ADC. Otherwise, BED is an all-red triangle with side lengths twice as long as the corresponding medians of T (prove this nice geometric fact on your own). ■

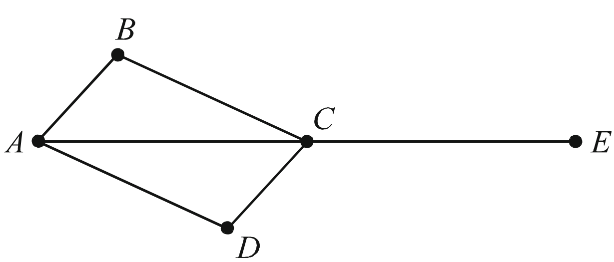


Fig. 44.8 The existence of a monochromatic triangle congruent to T or to Tm

Prove the following corollary of Tool 44.17.

Problem 44.18 Any 2-colored plane contains a monochromatic equilateral triangle of side 1 or $\sqrt{3}$.

Out of the many nice tools contained in [EGMRSS], I would like to share here with you my favorite. Erdős et al. prove it in a true Olympiad style, and so I am not changing a thing in it. However, I am adding an additional diagram showing that the statement is true for the case when the triangle K is *degenerate* as well.

Tool 44.19 ([EGMRSS] Theorem 1 of Part III). Let K be a triangle with sides a, b, and c, and let K_a, K_b, and K_c be equilateral triangles with sides a, b, and c respectively. Then a 2-colored plane contains a monochromatic triangle congruent to K if and only if it contains a monochromatic triangle congruent to at least one of the triangles K_a, K_b, K_c.

Proof Consider the configuration in Fig. 44.9. The six triangles HBC, ABD, CDE, EFH, DFG, AHG all have sides a, b, and c. The triangles ABH, DFE, BCD, FGH, HEC, ADG are equilateral with sides $a, a, b, b, c, c,$ respectively. We see that if one of the second six triangles is monochromatic, one of the first six must be monochromatic as well. The converse is true by a symmetric argument.

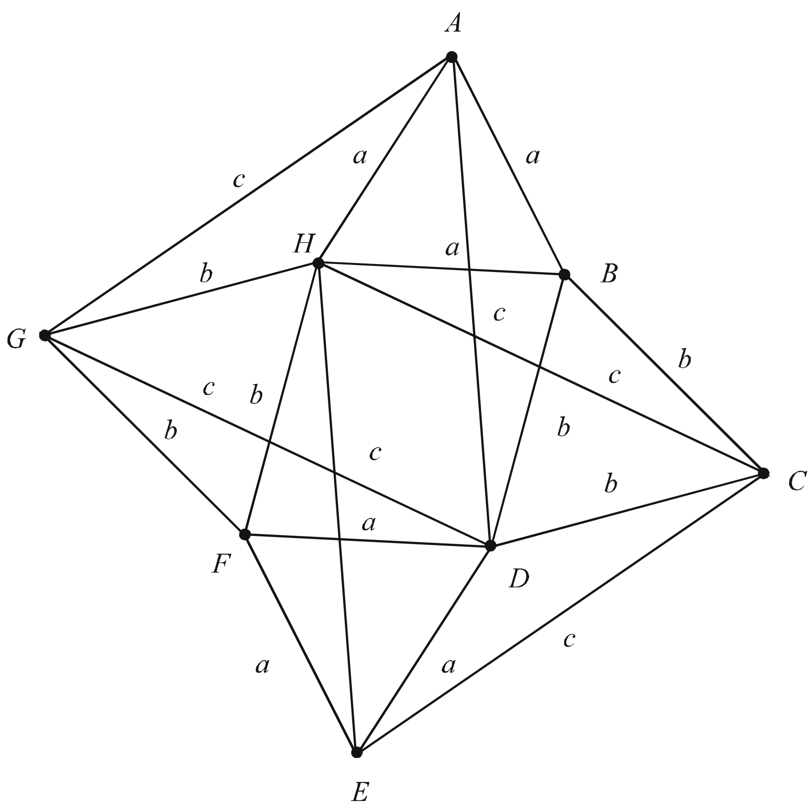


Fig. 44.9 Monochromatic triangles, non-degenerate case

If the triangle K is degenerate (this is one case the authors of [EGMRSS] did not explicitly address), look at the configuration in Fig. 44.10 that I added for you. No changes in the text of the proof are necessary while using Fig. 44.10 to prove a degenerate case. ■

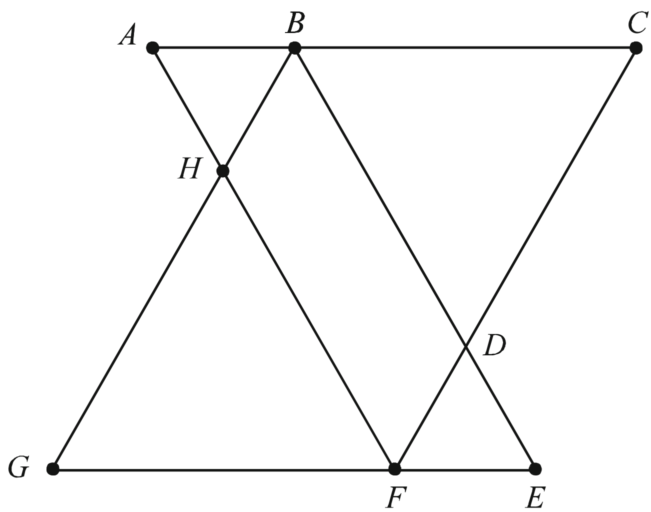


Fig. 44.10 Monochromatic triangles, degenerate case

The following problem is a good test of your skills: try it on your own before reading a solution.

Problem 44.20 (László Lovász, [Lov2]). Prove that any 2-colored plane contains a monochromatic triangle with side lengths $\sqrt{2}$, $\sqrt{6}$, and π.

Proof Given a 2-colored plane P. Due to Tool 44.18, P contains a monochromatic equilateral triangle of side $\sqrt{2}$ or $\sqrt{6}$ (just use as T an equilateral triangle of side $\sqrt{2}$; the sides of T_m will be equal to $\sqrt{6}$). In either case, due to Tool 44.19, the plane P contains a monochromatic triangle with sides $\sqrt{2}$, $\sqrt{6}$, π. ■

You may think that we are only concerned with triangles. We aren't. The following problem is a new form (and new solution) of a problem that the famous American problem solver and a long-term coach of the American team for the International Mathematics Olympiad, Cecil Rousseau, once created for the 1976 USA Mathematics Olympiad (USAMO).

Problem 44.21 (Cecil Rousseau; USAMO, 1976). Any 2-colored plane contains a $m \times n$ monochromatic rectangle such that $m = 1$ or 2, and n is a positive integer not greater than 6.

Proof Toss on a 2-colored plane (red and blue) a 2×6 square lattice (Fig. 44.11).

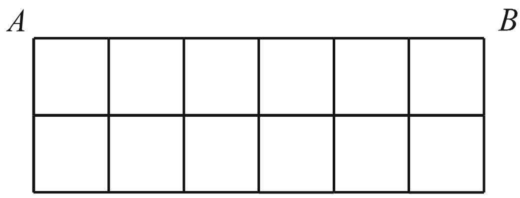


Fig. 44.11 The Rousseau argument 1

By the Pigeonhole Principle, out of 7 vertices (i.e., intersections of the lattice lines) in the top row AB, at least four must be of the same color, say, blue. We keep the corresponding 4 columns and throw away the rest (Fig. 44.12).

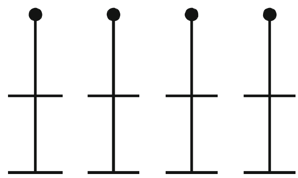


Fig. 44.12 The Rousseau argument 2

If the second *or* third row in Fig. 44.12 contains more than one blue vertex, we get a monochromatic blue rectangle, and the problem is solved.

If the second *and* third rows contain at most one blue vertex each, then we throw away the columns corresponding to these blue vertices. We are left with a monochromatic red rectangle located in the second and third rows. ■

I was able to strengthen this result in 1990 and offered it at the 1991 Colorado Mathematical Olympiad (CMO). Try to solve it on your own first.

Problem 44.22 (CMO 1990, [Soi9]). Prove the statement of Problem 44.21 with n not exceeding 5.

Proof Given a 2-colored plane. If one color is not present at all we are done. Otherwise, due to Problem 44.2, there are two points A and B of opposite colors distance 6 apart. Construct on AB a 2×6 square lattice like in Fig. 44.11 and repeat *word by word* the solution of Problem 44.21. ■

This train of thought naturally runs into the following open problems.

Open Problem 44.23 [Soi9]. Is the statement of Problem 44.21 true with n not exceeding 4?

Open Problem 44.24 Find the lowest upper bound for n, such that the statement of Problem 44.21 is true.

It is easy to prove the statement of open Problem 44.23 conditionally.

Problem 44.25. If a 2-colored plane P contains a monochromatic degenerate isosceles triangle of side 1 (Fig. 44.6), then P contains a $m \times n$ monochromatic rectangle such that $m = 1$ or 2, and n is a positive integer not exceeding 4.

Proof Let a 2-colored plane P (red and blue) contains a monochromatic, say blue, degenerate isosceles triangle T of side 1. We construct on T a 2×4 square lattice with T comprising the first column from the left (Fig. 44.13).

Fig. 44.13 A conditional reduction

If any of the other four columns contain at least two blue vertices, we get an all-blue rectangle. Otherwise, each of these four columns has at least two red vertices. But there are only $\binom{3}{2} = 3$ distinct ways to have two red vertices in a column. Therefore, at least two of the four columns have two red vertices in the same rows, i.e., we obtain an all-red rectangle.

The image of a Figure F under translation is naturally called a *translate* of F. Erdős et al. found a cute use of the Mosers Spindle.

Problem 44.26. ([EGMRSS] Theorem 3 of Part II). Given a 2-colored plane P (red and blue) and a triangle T in it. Then P contains a pair of red points distance d apart for every d, or a blue monochromatic translate of T. ■

Proof Let A, B, C be the vertices of T. Assume that for a positive d there is no pair of red points d apart. We drop the Mosers Spindle S (Fig. 2.2) of side d on the plane and denote by $S_1 = t_1(S)$ and $S_2 = t_2(S)$ the images of S under translations through $\overrightarrow{AB}$ and $\overrightarrow{AC}$, respectively.

Due to observation after Problem 2.2, any three vertices of the Mosers Spindle S contain a pair distance d apart. Therefore, each seven-point set S, S_1, and S_2 contains at most two red points. 2 + 2 + 2 aint equal to 7. :) Thus, there is a vertex, say A, of S such that all three vertices A, $t_1(A)$ and $t_2(S)$ are blue. They form a translate of T. ■

Problem 44.27 ([EGMRSS], Theorem 1' of Part II). Any 2-colored plane (red and blue) contains a red pair distance 1 apart, or 4 blue points on a line spaced by distance 1.

Proof Assume a 2-colored plane P does not have either a red pair of points distance 1 apart nor a four blue points on a line distance 1 apart. P must have a red point p. The circle C_1 of radius 1 and center at p must be entirely blue.

Add a concentric circle C_2 of radius $\sqrt{3}$ and an equilateral triangle a, b, c inscribed in C_2 (Fig. 44.14). Denote by d and e the points of intersection of C_2 and ab.

It is easy to confirm (please do) that $|ad| = |de| = |eb| = 1$.

Since both d and e are blue (they are on C_1), not both a and b are blue. This is similarly true for a and c and for b and c. Therefore, at most one of a, b, c is blue. Suppose a and b are red.

Now we rotate ab about p to its new position fg, such that $|af| = 1$. Then $|bg| = 1$. Therefore, f and g are both blue. So are h and i (they are on C_1). Thus, we get a blue quartet f, h, i, g distance 1 apart, in contradiction to our initial assumption. ■

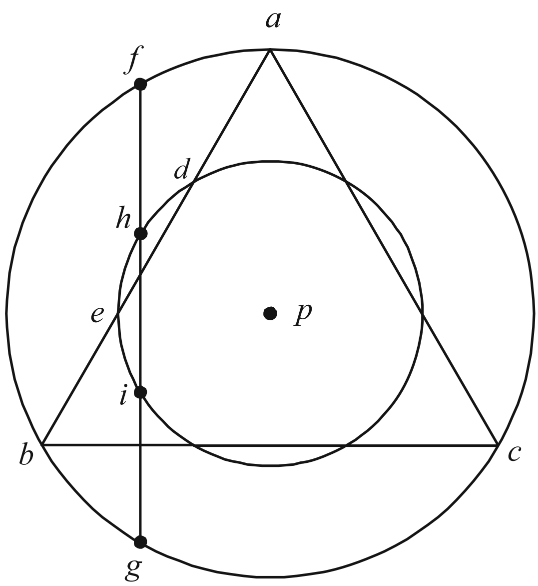


Fig. 44.14 A conditional result

Having proved 44.27, Erdős et al. [EGMRSS, part II, p. 535] formulate but could not decide the following question:

> Is it true that any 2-colored plane (red and blue) contains a red unit length segment or a blue unit square?

On March 25, 1977, the Hungarian mathematician Rozália Juhász submitted (and in 1979 published) an impressive paper [Juh], where in one stroke she proved a powerful generalization of Problem 44.27, and more than answered the above question by Erdős et al. in the positive:

Problem 44.28 ([Juh], Theorem 1). For any 4-gon Q, any 2-colored plane (red and blue) contains a pair of red points distance 1 apart, or a monochromatic blue 4-gon congruent to Q.

In the same paper, Juhász showed that the result of Problem 44.28 is not true for an n-*gon* where $n \geq 12$.

Counterexample 44.29 ([Juh], Theorem 2). There is a 12-gon K and a 2-colored plane P (red and blue) such that P does not have either a monochromatic unit-distant red segment or a blue monochromatic 12-gon congruent to K.

Construction First let us describe the 2-coloring of the plane that does the job. We start with a regular triangular lattice with distance 2 between nearest vertices (Fig. 44.15).

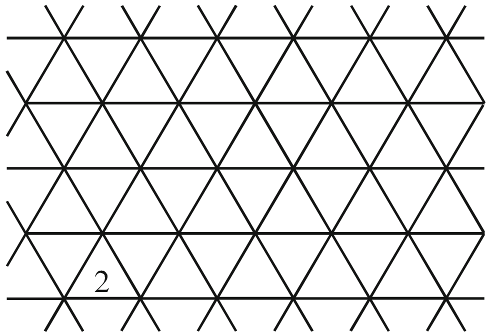


Fig. 44.15 The triangular lattice

We make every vertex of the lattice to be the center of a red circular disk of radius ½. With every disk we also color red half of its boundary under its horizontal diameter and the left vertex of that diameter. The rest of the plane we color blue (Fig. 44.16). You can easily verify that our 2-colored plane P has no red monochromatic segment of length 1. You can also show (do) that any closed disk (i.e., disk including its boundary circle) of radius $\frac{2}{\sqrt{3}} + \frac{1}{2}$ (a large circular disk shown in Fig. 44.16) in P must contain at least one of the red disks together with its boundary.

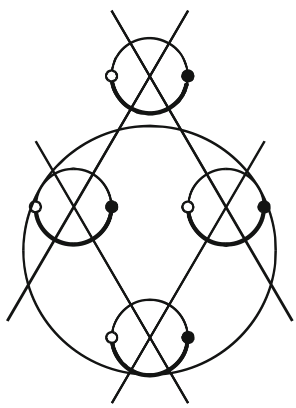


Fig. 44.16 The Juhász construction 1

Let us now define our 12-gon K. We draw a regular triangular lattice just like the one in Fig. 44.14, but with side $\frac{\sqrt{3}}{2}$, and a circle C of radius $\frac{2}{\sqrt{3}} + \frac{1}{2}$ with its center in the center of one of the triangles (Fig. 44.17). Inside C we have exactly 12 vertices of the lattice, they form our 12-gon K.

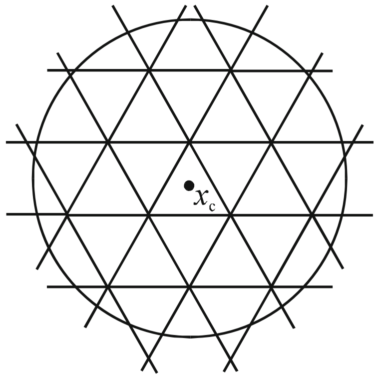


Fig. 44.17 The Juhász construction 2

All there is left to show is that the plane P does not contain a blue monochromatic congruent copy of K.

Place a congruent copy K_1 of K anywhere in the plane P, together with a surrounding circle C_1 congruent to C. As we noticed above, C_1 will contain completely at least one of the red disks C_2. Disk C_2 (with its red half-boundary) in turn will contain at least one of the vertices of K_1. Thus, at least one of the vertices of K_1 will be red! ■

About 15 years later, Rozália Juhász' 12-point counterexample was improved by two other Hungarian mathematicians, György Csizmadia and Géza Tóth. On January 15, 1991, they submitted, and in 1994 published [CT] an 8-point counterexample, thus "almost" closing the gap.

Counterexample 44.30 (Csizmadia and Tóth, [CT]). There is an 8-point set K in the plane (namely, a regular 7-gon and its center) and a 2-colored plane P (red and blue) such that P does not have either a monochromatic unit-distant red segment nor a blue monochromatic set congruent to K.

Problems 44.28 and 44.30 deliver the state of the art in this direction. Can we guarantee a monochromatic blue pentagon of at least one given shape and size in a 2-colored plane without a red monochromatic segment of unit length? Nobody knows! (So far pentagons have been slow to enter Euclidean Ramsey Theory.) A 3-number gap remains:

Open Problem 44.31 For which n in the interval $5 \leq n \leq 7$ is the following statement true: For any n-gon K, any 2-colored plane (red and blue) contains a pair of red points distance 1 apart, or a monochromatic blue n-gon congruent to K?

Chapter 45
3-Colored Plane, 2-Colored Space, and Ramsey Sets

In 1991 [Soi3] I named an n-gon K in a n-colored plane *representative* if all n colors are represented among its vertices. (Paul Erdős and Ron Graham preferred the term *rainbow*.)

Problem 45.1 [Soi3]. Any 3-colored plane contains a monochromatic or representative triangle T with small side 1 and angles in the ratio 1: 2: 4.

Proof. Assume that a 3-colored plane P (red, white, and blue) does not contain a monochromatic congruent copy of T. Toss a regular heptagon H of side 1 on the plane P.

H can have at most 3 vertices of the same color, because *any* 4 vertices of H contain a triangle congruent to T (prove it on your own). On the other hand, by the Pigeonhole Principle, H must contain at least 3 vertices of the same color. Hence, 3 it is: H contains, say, three red vertices.

There are only three ways to have 3 red vertices on H without red monochromatic copy of T (Fig. 45.1). Numbers of white and blue vertices must be 3–1 or 2–2 respectively. It is now easy to verify (do) that every completion of three colorings in Fig. 45.1, subject to the above constraints, contains a representative copy of T. ■

We probably cannot expect a guaranteed monochromatic copy of *any* triangle in a 3-colored plane. I would like to know which ones we can guarantee:

Open Problem 45.2 Find all triangles T, such that any 3-colored plane in which all three colors are present, contains a monochromatic or representative triangle congruent to T.

Ronald L. Graham believes that we can exclude any triangle by an appropriate 3-coloring. He formulated the following conjecture during our July 10, 1991, phone conversation (it appeared in 1991 in [Soi3]). Now [Gra7], [Gra8] Graham is offering $25 for it.

Graham's $25 Conjecture 45.3 (R. L. Graham). For any triangle T there exists a 3-colored plane that does not contain a monochromatic triangle congruent to T.

And now, as promised in the title of this section, let us peek at 2-colorings of the space E^3. Unlike the case in the plane E^2, we do get a unit monochromatic equilateral triangle in *any* 2-coloring of E^3.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_45

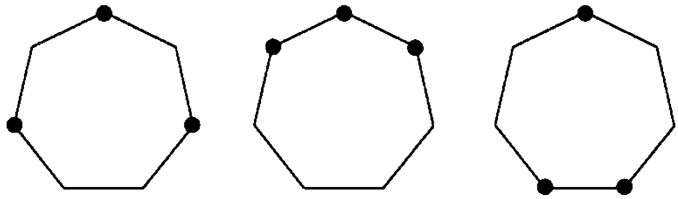


Fig. 45.1 Three possible positions of 3 red vertices

Problem 45.4 ([EGMRSS], Theorem 6 of Part I). Any 2-colored space E^3 contains a unit monochromatic equilateral triangle.

Proof. Let the space E^3 be 2-colored, red and blue. We pick two points A and B of the same color, say red, distance 1 apart (we can pick such A and B in *any* plane of E^3). If there is a third red point C at distance 1 from both A and B, we are done. Otherwise, we get a whole circle S_1 of blue points that lies in the plane perpendicular to AB through the midpoint O_1 of AB (Fig. 45.2).

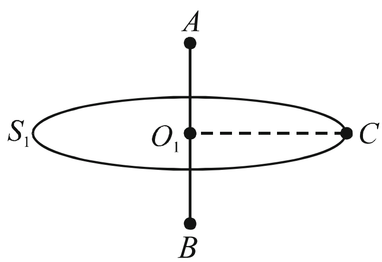


Fig. 45.2 Blue circle

The radius of this circle S_1 is $\frac{\sqrt{3}}{2}$. Now we pick a chord MN of S_1 of length 1. If there is a third blue point K at distance 1 from both M and N, we are done. Assume such a blue point K is not present, then there is a whole circle S_2 of red points in the plane perpendicular to the plane of S_1 (Fig. 45.3). The radius of S_2 is, of course, the same as the radius of S_1 (because we really used the same construction for both circles).

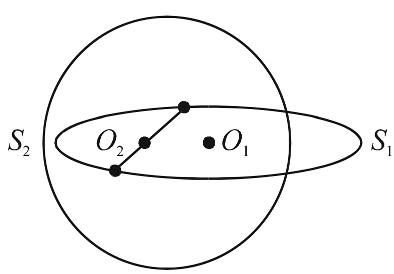


Fig. 45.3 Red circle S_2

We rotate the chord MN about O_1, the red circle S_2 will rotate about O_1 accordingly and will create a degenerate torus T (a torus without a hole in the middle due to self-intersection). Thus, we get a whole red torus T.

The largest horizontal circle (equator) S_3 on the torus T has diameter $d = \frac{\sqrt{2}+\sqrt{3}}{2}$ (verify that). We can inscribe in S_3 an equilateral triangle X of side $\frac{\sqrt{3}}{2}d = \frac{\sqrt{6}+3}{4} > 1$. Moving symmetrically the vertices of X along the surface of the torus T toward the middle of T (so that the plane determined by X remains horizontal), due to continuity, we will get an equilateral triangle X_1 of side 1 on T. Since X_1 is on T and the whole torus T is red, X_1 is the desired monochromatic triangle. ■

Paul Erdős et al. used a clever method similar to their solution of Problem 45.4, to prove the following stronger result. Try to prove it on your own.

Problem 45.5 ([EGMRSS]), Theorem 24 of part III). For any 2-colored space E^3, there is one color such that equilateral triangles of all sizes occur in that color.

This result, of course, makes one wonder whether a similar success can be guaranteed in the plane. This, however, is an open question:

Open Problem 45.6 [EGMRSS, part III, p. 579]. Is it true that for any 2-colored plane E^2 there is one color such that all triangles which occur monochromatically occur monochromatically in that color?

Now we can prove for the space what is still an open problem for the plane.

Problem 45.7 For any triangle T, any 2-colored space E^3 contains a monochromatic triangle T_1 congruent to T.

Proof Let T be a triangle with sides a, b, and c and the 3-space be 2-colored. By Problem 45.5, the space contains a monochromatic equilateral triangle K_a of side a. Since the plane P that contains K_a is 2-colored, due to Tool 44.19, we have in P a monochromatic triangle T_1 with sides a, b, and c, which is congruent to T. ■

For right triangles this result can be proved even for a 3-colored space, as Miklós Bóna and Géza Tóth showed in 1996 [BT]:

Problem 45.8 (M. Bóna and G. Tóth). For any right triangle T, any 3-colored space E^3 contains a monochromatic triangle T_1 congruent to T.[1]

In 2017, Andrii Arman and Sergei Tsaturian, both from Manitoba University, uploaded to arXiv the following result [AT]:

Problem 45.9 (A. Arman and S. Tsaturian). Any 2-colored, red and blue, space E^3 contains a red unit segment or six collinear blue points with unit distance between any two consecutive points.

In conclusion I would like to present here, without proofs, two main results and the main open problem of the Erdős at al. trilogy [EGMRSS] and related results by P. Frankl and V. Rödl, and I. Kříž.

[1] Compare this result to Shader's problem 44.10.

Generalizing the line E^1, the plane E^2 and the space E^3, we define the *n-dimensional space* R^n for any positive integer n as the set of all n-tuples $(x_1, x_2, \ldots, x_n)$, where $x_1, x_2, \ldots, x_n$ are real numbers. When the distance between two points $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$ of R^n is defined by the equality

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2} \qquad (*)$$

we get the *Euclidean n-dimensional space* E^n. In other words, E^n is the set R^n together with the distance d defined by (*).

Many notions are generalized from E^2 and E^3 straight forward to E^n. The *sphere* of radius r and center O in E^n is the set of all points of distance r from O. A 4-point set P in the plane is called a $d_1 \times d_2$ *rectangle* if it is congruent to the set $\{(0, 0), (d_1, 0), (0, d_2), (d_1, d_2)\}$. Similarly, a 2^n–point set in E^n is called a $d_1 \times d_2 \times \ldots \times d_n$ *rectangular parallelepiped* if it is congruent to the set

$$\{(x_1, x_2, \ldots, x_n) \mid x_1 = 0 \text{ or } d_1; x_2 = 0 \; or \; d_2; \ldots; x_n = 0 \text{ or } d_n\}.$$

A finite subset C of E^n is called *r-Ramsey for* E^n if for any *r-colorings* of E^n there is a monochromatic subset C_1 congruent to C. If for every r there is n such that C is r-Ramsey for E^n, then the set C is called *Ramsey*.

We are now ready for two main results by Paul Erdős at al.

Necessary Condition 45.10 ([EGMRSS], theorem 13 of part I). If a set C is Ramsey, then C must lie on an n-dimensional sphere for some integer n.

Sufficient Condition 45.11 ([EGMRSS], corollary 22 of part I). Any subset of a rectangular parallelepiped is Ramsey.

There is obviously a gap between the necessary and sufficient conditions for a finite set to be Ramsey. In fact, in 1986 Peter Frankl from France and Vojtěch Rödl from Czechoslovakia proved that the sufficient Condition 45.10 is not necessary by showing that even obtuse triangles (which cannot be embedded as subsets in a rectangular parallelepiped) are Ramsey:

Problem 45.12 (P. Frankl and V. Rödl, [FR1]). All nondegenerate triangles are Ramsey.

In their consequent paper, they generalized this result to n-dimensional Euclidean spaces.

Problem 45.13 (P. Frankl and V. Rödl [FR3]). Any nondegenerate *simplex* (i.e., n+1 points generating the whole n-dimensional Euclidean space) is Ramsey.

In 1991, Igor Kříž [Kri1], then from the University of Chicago (and later at the University of Michigan), published powerful results that imply the following:

Problem 45.14 (I. Kříž). Any regular polygon is Ramsey.

Thus, we finally got the first Ramsey pentagon: the regular one. Kříž‘s results also imply a similar statement in 3 dimensions:

Problem 45.15 (I. Kříž). Any regular polyhedron is Ramsey.

In 2007, this result was generalized by Kristal Cantwell [Can2] to all regular n-dimensional polytopes.

Problem 45.16 (K. Cantwell). All regular polytopes are Ramsey.

In his next, 1992 paper [Kri2], Igor improved Frankl–Rödl Result 45.11:

Problem 45.17 (I. Kříž). Any trapezoid is Ramsey.

As no criterion for a set to be Ramsey appeared, Paul Erdős attempted to speed up the process in 1985 [E85.01]:

> We (i.e., the authors of [EGMRSS]) do not know which (if any) of these alternatives characterize Ramsey sets, and I offer $500 for an answer to this question.

Paul Erdős' $500 Problem 45.18 Find a criterion for a set to be Ramsey.

Ever since 1993, if not before [Gra3], [Gra7], [Gra8], Ron Graham expressed his $1000 belief that the necessary Condition 45.9 is also sufficient:

Ronald L. Graham's $1000 Open Problem 45.19 Prove that all spherical sets are Ramsey.

He also offered a consolation prize for a partial result [Gra7], [Gra8]:

Ronald L. Graham's $100 Open Problem 45.20 Prove that any 4-point subset of a circle is Ramsey.

Chapter 46
The Gallai Theorem

46.1 Tibor Gallai and His Theorem

The Gallai Theorem is one of my favorite results in all of mathematics. Surprisingly, it is not widely known even among mathematicians. Its creator was Tibor Gallai, born Tibor Grünwald, a member of the Hungarian Academy of Sciences, who passed away on January 2, 1992, at the age of 79. His lifelong close friend and co-author Paul Erdős was staying me in Colorado Springs[1] when Professor Vera T. Sós called from Budapest to give Paul the sad news of Gallai's passing. Right then I asked Paul to write about Gallai for *Geombinatorics*. Here is Paul's *Obituary of My Friend and Coauthor Tibor Gallai* [E92.14] in its entirety, including the sketch for the Sylvester–Gallai result that Paul drew on the margin of his manuscript.

> I met Tibor Gallai in 1929 when we were both in high school. We knew of each other's existence since we both worked at the *Kozépiskolai Matematikai Lapok*, a journal for high school students which appeared every month and published problems and their solutions by students. This periodical had an immense influence on Hungarian mathematics; many children before the age of 15 realized that they wanted to be mathematicians, and many of the well-known mathematicians as young people worked in this journal. Gallai and I worked together on mathematics since 1930 and had many joint papers (for details, see my forthcoming obituary of Gallai in *Combinatorica* and also the article of Lovász and myself in *Combinatorica*, Vol. 2, 1982 written for Gallai's 70th birthday).
>
> Here I just want to state some of the elementary results of Gallai which can be easily understood by beginners. In 1933 I conjectured that if $x_1, x_2, \ldots, x_n$ are n points in the plane, not all on a line, then there is always a line which goes through precisely two of our points. I thought that I will prove this in a few minutes but, in fact, I could not prove it. I told my conjecture to Gallai who found a very nice proof of it which goes as follows:

[1] We were working on our join project, a book of Paul's open problems: *Problems of pgom Erdős*, which I hope to finish in the not too distant future.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_46

Project one of the points to infinity and join it to all the other points. If my conjecture would be wrong, we would get a set of parallel lines each of which contains at least 2 finite points. Consider now the oblique lines, each of them contains at least three points. Take the line which has the smallest angle (Fig. 46.1).

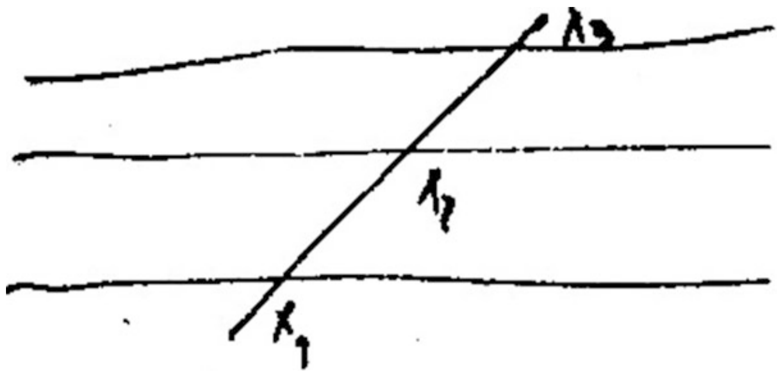


Fig. 46.1 Paul Erdős's drawing

On the line in the middle there must be another point besides x_2, say y and the one of the lines x_1y or yx_3, clearly gives a smaller angle. This contradiction proves the conjecture. A few years later, L. M. Kelly found that my conjecture is not really mine. It was conjectured in 1893 by Sylvester, but as far as we know, Gallai was the first who proved it. The simplest proof is due to L. M. Kelly. By the way, I observed that Gallai's theorem implies that n points not all on a line determine at least n distinct lines. There is a very nice related conjecture of Gabriel Dirac. Let $x_1, x_2, \ldots, x_n$ be n points not all on a line. Join every two of them. Then at least one of the points has $\frac{n}{2} - c$ lines incident to it. Beck, Szemerédi, and Trotter proved that there is a point with at least $c_1 n$ lines incident to it; their c_1 is positive, but it is very small.

Gallai was very modest – I would almost say abnormally so. Many of his beautiful results he published only with great delay. Often he did not publish them at all, and they were later discovered by others. He felt sorry for this only once. Dilworth in 1950 in the *Annals of Mathematics* proved the following classical theorem: Let ϑ be a partially ordered set. Assume that the maximal number of non-comparable elements is d, then ϑ is the union of d chains. In fact, Gallai and Milgram had a complete proof of this beautiful theorem in 1942. Milgram was a topologist who did not realize the importance of this result. Gallai wrote their joint paper in German. Milgram wanted to have it published in English and promised to rewrite it but delayed it until it was too late. I promised Gallai never to mention this in his lifetime since the theorem should clearly be known as Dilworth Theorem.

Hilbert, in his beautiful obituary of Minkowski in *Math Annalen* 1909 wrote "I can only be grateful that I had a friend and co-worker for such a long time." This is what I have to say about Gallai and "May his theorems live forever."

Paul Erdős added [E92.15]:

A few years before his death he [Gallai] finally accepted the degree of Doctor of the Academy and two years ago, much against his will, he was even granted the membership in the Academy.

Tibor Gallai, 1935–1936. Courtesy of Alice Bogdán

Indeed, Gallai discovered a number of fabulous results, some of which were named after other mathematicians: he preferred not to publish even his greatest results. Why? I learned the answer during Paul Erdős' 80th Birthday Conference that took place in the beautiful town Keszthely, Hungary, on Lake Balaton. On July 20, 1993, my (then) wife Maya, our baby Isabelle (in a stroller) and I took a stroll in the center of Keszthely when we were unexpectedly invited to join a couple for dinner at an outdoor section of a restaurant. Imagine, the couple that invited us was George Szekeres and Esther Klein, the legendary couple from the legendary circle of young Jewish mathematicians, who assembled in the early 1930s in Budapest! I was able to ask them about the friend of their youth.

"Gallai was so terribly modest," explained George Szekeres. "He did not want to publish because it would show the world that he was clever, and he would be restless because of it."

"But he was very clever indeed," added Esther Klein-Szekeres. Esther continued: "Once I came to him and found him in bed. He said that he could not decide which foot to put down first."

"Gallai was Paul Erdős' best, closest friend," continued George. "I was very close with Turán. It was later that Paul Erdős and I became friends." At this moment, the waiter served the four of us complimentary vodka. Silently, everyone moved vodka toward me, and so I got to consume four shots "to health" of our unforgettable company.

I always thought, as probably everyone, that hypergraphs were invented by the great French graph theorist Claude Berge. Amazingly, Gallai was first here too: at the age of 18–19 (Gallai was born on July 15, 1912), he introduced hypergraphs. Paul Erdős mentioned it in passing in his 1991 talk at Visegrád (Hungary) Conference, published 3 years later [E94.22]:

> As far as I know, the subject of hypergraphs was first mentioned by T. Gallai in conversation with me in 1931; he remarked that hypergraphs should be studied as a generalization of graphs. The subject really came to life only with the work of Berge.

Paul Erdős told me that Tibor Gallai discovered the theorem of our prime interest in the late 1930's. He did not publish it either. It *first* appeared in the paper [Rad2] by Richard Rado with a credit to *"Dr. G. Grünwald,"* which was Gallai's name then; the initial "G" should have been "T" and must be Rado's typo. Rado submitted this paper on September 16, 1939; it is listed in bibliographies as a 1943 publication, but in fact came out only 6 years later, in 1945 – World War II affected all facets of life and made no exception for the great Gallai result. I hope you will enjoy it as much as I have and try your wit and creativity in proving this beautiful and extremely general, classic result. If you are unfamiliar with n-dimensional Euclidean space, look up the definition in the previous chapter or assume $n = 2$: plenty of fun is to be found in the plane.

The Gallai Theorem 46.1 ([Rad2]). Let m, n, k be arbitrary positive integers. If the lattice points Z^n (i.e., the points with integer coordinates) of the Euclidean space E^n are colored in k colors, and A is a m-element subset of Z^n, then there is a monochromatic subset A' in Z^n that is homothetic (i.e., similar and parallel) to A.

In fact, with not too much effort the Gallai Theorem can be strengthened as follows:

The Gallai Theorem, a Strong Version 46.2 ([GRS2]). Let m, n, k be arbitrary positive integers. If the Euclidean space E^n is colored in k colors and A is a m-element subset of E^n, then there is a monochromatic subset A' in E^n that is homothetic to A.

As we have already observed in this book, the Russian mathematician Mira A. Lukomskaja from Minsk proved in 1947 what I named the Lukomskaya Theorem 43.1. It was precisely a one-dimensional version of the Gallai Theorem and was also published in an article form in 1948 [Luk] and in the 1948 Khinchin's book [Khi2], *much earlier* than Witt's 1952 publication presenting a two-dimensional Gallai Theorem [Wit] and Garsia's 1958 manuscript, where the n-dimensional Gallai Theorem is proved. So, four mathematicians made contributions related to this theorem, in chronologic order: Gallai, Lukomskaya, Witt, and Garsia.

How should we attribute credit for this classic result? Graham, Rothschild, and Spencer call it "Gallai's Theorem" ([GRS2]. Hans Jürgen Prömel with his coauthors Vojtěch Rödl and Bernd Voigt call it "Gallai–Witt's Theorem" [PR], [PV]. Prömel continues to seemingly insist on the coauthorship of Witt in 2005 [Pr1] and in the 2013 book [Pr2]. Let us take a look together at this 2013 book. Section 4.2.2 is called "Gallai–Witt's Theorem." It opens as follows:

> A multidimensional version of van der Waerden's theorem was proved independently by Gallai (=Grünwald), sf. Rado (1943) and Witt (1952).

Does Prömel not see the 9-year gap between the dates? Moreover, since he cites Rado-1943, he knows that Rado's paper [Rad2] with Gallai inside was submitted on 16 September 1939, when Nazi Germany's waged war delayed its publication for several years. Under bombardments, London was not up to publishing mathematics. When Rado's paper finally appeared in a prominent and well-read *Proceedings of the London Mathematical Society*. Witt had many years after the war when he could read it. It begs a question: is the reason to insist on including Witt of non-mathematical kind? Witt wrote his paper well, but in my opinion way too late for claiming authorship.

It is not a deciding factor for me that Gallai did not publish his proof – he shared it with the most reliable informants Erdős and Rado in the 1930s, and Rado included Gallai's proof in his paper submitted it in 1939 with credit to Gallai. It is not a deciding factor for me that Garsia did not publish his proof – he provided me with an old blue line, faded from age, copy of his 1958 proof and a contemporaneous communication of his proof to Van der Waerden. Lukomskaya, Witt, and Garsia appear to have discovered their proofs independently from Gallai, and their proofs constitute contributions to the field. However, a very significant time that lapsed between the 1930s and the late 1940s and 1950s, and the public availability of Gallai's published proof ever since 1945 must convince every fair person to give credit for the discovery of this all-important classic theorem to one person, and one person alone, and call it accordingly *The Gallai Theorem*.

46.2 Ernst Witt

Ernst Witt was born on the island of Alsen in 1911. Alsen, together with the rest of North Schleswig became part of Germany in 1864. The island was returned to Denmark in 1920. Two-year-old Witt went to China with his missionary parents. At 9 he was sent back to Germany to live with his uncle. Witt studied at the universities of Freiburg and Göttingen. His doctoral work at Göttingen was supervised by Emmy Noether. That was the dawn of Hitler's rein. Witt's former Ph.D. student at Hamburg (and presently Professor Emerita at Göttingen University) Ina Kersten writes in his biography [Ker] that on May 1, 1933, Witt joined the Nazi party and the storm troopers *SA* – observe, he did it days after his teacher Noether was fired by the Nazi regime. In his defense after the war, as a proof of how little the Nazi and *SA* memberships meant to him, Witt claimed that his family did not know about it [Ker]. He was not as considered toward his Jewish mentor Emmy Noether, fired from Göttingen [Bert]:

> Storm trooper Ernst Witt, resplendent in the Brownshirt uniform of Hitler's paramilitary, knocked on a Jew's apartment door in 1934. A short, rotund woman opened the door. Emmy Noether smiled, welcomed the young Nazi into her home, and started her underground math class. The Brownshirt was one of her favorite pupils.

Indeed, Witt must have been Noether's second most favorite student after Van der Waerden. Since Emmy Noether was forced out as Jewish and liberal, Witt defended his doctorate under Gustav Herglotz in July 1933 and joined Helmut Hasse's seminar, when the latter entered Göttingen in 1934 as the director of Mathematics Institute and professor.

Kersten [Ker] informs that in 1934 Witt became Hasse's Assistant at Göttingen. In 1937, Emil Artin left Hamburg for the United States. In 1939, Witt was appointed to the downgraded to an associate professor Artin's chair at Hamburg and worked there until his dismissal by the British Military Authority in fall 1945. However, the Brits could not keep long grudges against the Nazis in Germany: in 1947, Witt was reinstated in his position, in 1957 promoted to an Ordinarius, and remained on the job until his retirement in 1979.

Kersten describes Witt's 1960–1961 visit of the Institute for Advanced Study Princeton, and his "astonishment" at the negative reaction when Witt disclosed his Nazi past:

> One day during a discussion about a member of the National Socialist party, he [Witt] felt obligated to declare that he had also been a member of that party. To behave otherwise would have seemed insincere to him. He found, to his utter astonishment, that his contacts with his colleagues were suddenly severed.

Couldn't Witt comprehend that people at the Institute, some of whom escaped from the Nazis and had family and colleagues murdered by the German criminal regime, were "sincerely" shocked to find a former Nazi and moreover a Storm Trooper in their midst? I am utterly astonished that Witt was "utterly astonished."

In 1978, Ernst Witt was honored with membership in the Göttingen Academy of Sciences. He died in Hamburg on July 3, 1991 – of natural causes.

46.3 Adriano Garsia

Let us now turn our attention to Professor Adriano Garsia. The story of his discovery that he told me in the February 28, 1995, e-mail [Gars2], is almost as intriguing as the story Van der Waerden told us in Chapter 35:

> I discovered the result in the fall of 1958. I was then a Moore Instructor at MIT. We used to have fun at the time tossing each other problems at the common room. A student had asked the following question:
>
> *If we color the points of the plane in two colors, can we always find a square with vertices all of the same color?*

This problem frustrated everybody. . . including me. . . until Paul Cohen[2] solved it. I didn't want to know the solution since I wanted to solve it myself. . . After a few days of unsuccessful attempts, I finally asked somebody who knew Paul's solution how he did it!

I learned that he had used Van der Waerden's theorem on arithmetic progressions. I did not know of Van der Waerden's result at the time, so I was at disadvantage on this one. So. I got hold of Khinchin's book *Three Pearls of Number Theory* and studied Van der Waerden's proof very carefully.

I noticed then that the theorem could be generalized to higher dimensions to show that we could find any finite set of lattice points (up to scaling) with all elements of the same color.

I wrote up the proof and sent it to Van der Waerden who liked it and offered to publish it in the *Mathematische Annalen*. However, a few weeks later, I got another letter from Van der Waerden who had been doing some search on the literature on the subject and discovered that precisely the same generalization had already [been] published by T. [R.] Rado' . . . Under those circumstances he felt that although my proof was much neater . . . he didn't think it was worth publishing.

In the meantime, I had asked myself "what about a regular pentagon? . . ." In fact, what if we are given any geometric figure consisting of a finite set of points, can we find a stretch and translate of the figure with all elements of the same color? . . . Now it showed that my proof could be used under this more general situation as well. In fact, contrarily to P. Cohen or Rado who derived their result by *applying Van der Waerden's theorem*, I had obtained mine by *extending Van der Waerden's mechanism of proof.*

Basically, I showed that a sufficiently high "power" of the figure had to contain a monochromatic stretch and translate of the figure. (Power here means that we construct a figure of the form $F + a_1$, $F + a_2$, . . . $F + a_n$; with $A+B$ representing the vector sum of every point of A with every point of B.)

Although the version I had sent to Van der Waerden did not specifically address itself to this more general situation, very little needed to be added to include this. Nevertheless, after Van der Waerden's second letter I gave up on the idea of publishing the result. I have still some duplicates of seminar notes in which the more general result is presented. In fact, the summer of 1959 I did give a lecture at Bell Labs on it. I believe G. Rodemich who is now at JPL [Jet Propulsion Laboratory], perhaps Henry Pollack was also at that lecture . . . I don't quite remember others. Jurgen Moser was at MIT at that time, and I remember discussing my result with him in great detail.

---

[2] In 1963, the American mathematician Paul Joseph Cohen (April 2, 1934 – March 23, 2007) invented a technique called *forcing* and used it to prove that neither the continuum hypothesis nor the axiom of choice can be proved from the standard Zermelo–Fraenkel system of axioms (***ZF***) for set theory. In the summer 1966, he won Fields Medal for this great achievement at the International Congress of Mathematicians in Moscow. Cohen will appear again in a later chapter of this book.

This is the story. I am presently visiting UQAM and University of Montreal and it is difficult from here to locate those notes. I will get back to San Diego at the end of March. Send me your address and I will mail you a copy.

The idea of the proof is noticing that the same pigeonhole argument of the original proof of Van der Waerden can be used in this more general situation. Inductively, we consider "colored" powers of the figure as "colors" assigned to say the center of the power. Then having proved the result for any number of colors and all figures with $n-1$ points, we construct in a sufficiently high power of that power a monochromatic configuration of centers that is similar to the given figure minus a point. However, monochromatic centers now mean that the corresponding powers centered at those points are all colored the same way! . . .

At this point we then use the Van der Waerden idea . . . which is well explained in Khinchin's book. Incidentally, Khinchin states that he is presenting a simpler proof but Van der Waerden himself assured me that his proof was identical . . . I never did see Van der Waerden's original proof.

That is the story as I can remember it . . .

Best wishes on your book,

Garsia

PS: I am surprised that you call this Gallai's theorem . . . I was under the impression that a formal language version of the result which could be easily translated into mine (by sending letters into vectors) was due to Graham and Rothschild and a 3rd author I can't remember [Joel Spencer – A.S.].

To complement this fabulous story, Adriano Garsia sent me the *original*, faded with age, 10 mimeographed blue-lettered pages of his notes, as he wrote on April 20, 1995, in another e-mail:

I finally found the notes from which the paper I sent to Van der Waerden was written. I don't seem to have any copy of that paper. The notes are a bit faded but still readable. I am mailing them today. Best of luck in deciphering them.
– Garsia

Adriano Maria Garsia was born in Tunisia on August 20, 1928. He received his secondary and college education in Rome, Italy, and Ph.D. from Stanford University in 1957. He was a professor at the California Institute of Technology (1964 –1966), and since 1966 has been a professor at the University of California San Diego. In 2012, he became a Fellow of the American Mathematical Society.

46.4 An Application of Gallai

A beautiful application of the Gallai Theorem was found by Alexej Kanel-Belov (listed as just Belov in this article) and S.V. Okhitin in 1992 [BO].

Theorem 46.3 ([BO]). Each cell of an [infinite] square grid contains an integer. For any given nonzero integer n, there is a square with sides parallel to the lines of the grid, such that the sum of all integers inside it is divisible by n.

Proof Affix x and y axes along the lines of the grid. Now we "color" each unit cell (x, y) of the grid in one of n colors by assigning to it the remainder $S(x, y)$ upon division by n of the sum of numbers located in all cells with coordinates (a, b) such that $0 \le a \le x$; $0 \le b \le y$.

The first quadrant of the grid is now colored in n colors. By the Gallai Theorem, there is a monochromatic square, whose vertices have coordinates, say, (x, y), $(x + k, y)$, $(x, y + k)$, $(x + k, y + k)$. But this is all we need to prove the result, for it is easy to notice that the sum of all numbers inside this square is

$$S(x, y) - S(x + k, y) - S(x, y + k) + S(x + k, y + k),$$

and this sum is congruent to zero modulo n. ■

The authors generalize this theorem on two counts at once:

Theorem 46.4 ([BO]). Each cell of an [infinite] k-dimensional square grid contains an integer. For any given non-zero integer n and a positive integer m there is a positive integer $L= L(k, m, n)$ such that the grid contains a k-dimensional cube of side Lm with all edges parallel to the lines of the grid, which is partitioned into m^k "little" cubes of side L, such that the sum of all integers inside each "little" cube is divisible by n.

Hint Instead of claiming a monochromatic square, as we did in the proof of 46.3, we can now use the Gallai Theorem to claim the existence a monochromatic subgrid homothetic to the k-dimensional square grid of side L (which consists, of course, of m^k cells of the same color). ■

Theorem 46.5 ([BO]). Each cell of an [infinite] k-dimensional square grid contains a real number. For any given positive integer m and a (small) positive ε, there is a positive integer $L = L(k, m, \varepsilon)$ such that the grid contains a k-dimensional cube of side Lm with all edges parallel to the lines of the grid, which is partitioned into m^k "little" cubes of side L, with the sum of all numbers inside each "little" cube differing from an integer by less than ε.

Hint This proof repeats the proof of the previous result with the more delicate interpretation of coloring. We partition a unit segment [0,1] into $N > 2^k m^k \varepsilon$ equal "little" segments – they are our "colors" – and determine the color of a cell of the grid with coordinates $(x_1, \ldots, x_k)$ by the "little" segment into which the fractional part falls of the sum of numbers in the grid's cells with coordinates $(a_1, \ldots, a_k)$, where $0 \le a_i \le x_i$ for $1 \le i \le k$. ■

Of course, the Gallai Theorem allows us to generalize Theorems 46.3, 46.4, and 46.5 further and use k-dimensional parallelepipeds of the given in advance ratio of sides. I leave this development to you.

In 2000, Mark Walters [Wal] published combinatorial proofs of polynomial versions of several important theorems. Here is one of them.

Define a *D-dimensional integral polynomial* $p(n)$ to be a polynomial in n with coefficients in Z^D, which is zero at zero. (As usual, Z stands for the set of integers, and N for its positive part.)

The Polynomial Gallai Theorem 46.6 ([Wal]). Let $p_1, p_2, \ldots, p_m$ be D-dimensional integral polynomials and N^D be finitely colored. Then there exists $a \in N^D$ and $d \in N$ such that the set of points $\{a\} \cup \{a+p_i(d)\colon 1 \le i \le m\}$ is monochromatic.

Of course, Theorem 46.6 implies the polynomial generalization of the Baudet–Schur–Van der Warden theorem.

The Polynomial Baudet–Schur–Van der Waerden Theorem 46.7 ([Wal]). Let $p_1, p_2, \ldots,$ p_m be integral polynomials and N be finitely colored. Then there exist positive integers a and d such that the set of points $\{a\} \cup \{a+p_i(d)\colon 1 \le i \le m\}$ is monochromatic.

46.5 Hales–Jewett's Tic-Tac-Toe

Surely, you played Tic-Tac-Toe in your tender years (Fig. 46.2). The goal is to mark a line of cells with your sign. In the "normal" Tic-Tac-Toe, the line can be horizontal, vertical, and diagonal (there are two diagonals). In fact, we can represent the cells by nodes, and replace X's and O's by two colors. The game then asks two players to color the nodes in turn. The winner is the one who creates a monochromatic line in his color. We will accept all the usual lines except one of the diagonals, going from the upper left to the lower right corner (Fig. 46.3).

In 1963, two young mathematicians, Alfred Washington Hales and Robert Israel Jewett, published the result that raised the game of Tic-Tac-Toe to the level of a mathematical result of Ramsey Theory, the result of great importance. Informally speaking, they proved that the n-dimensional, r-player generalization of Tic-Tac-Toe cannot end in a draw, no matter how large n is, and no matter how many people r play so long as the playing board has a

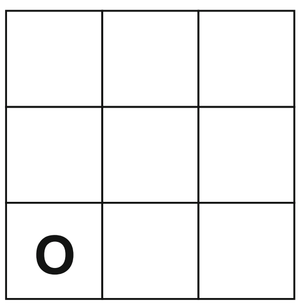


Fig. 46.2 Tic-Tac-Toe

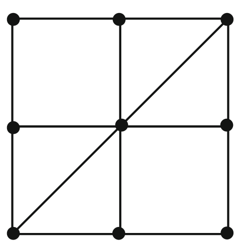


Fig. 46.3 Tic-Tac-Toe without one diagonal

sufficiently high dimension. (In fact, the first player has a winning strategy due to the strategy-stealing argument.) As is often the case in mathematics, this is an existence result: no algorithm is known for a winning strategy.

In order to present the theorem formally, we need to define an *n-dimensional* cube and a *combinatorial line*, or simply a *line* in it. Given a fixed finite set, often called *alphabet*, $A = \{a_1, a_2, \ldots, a_m\}$, the *n-dimensional cube* on the alphabet A is, expectedly, the set $A^n = \{(x_1, x_2, \ldots, x_n) : x_i \in A\}$. Given a set S of coordinates, $\varnothing \neq S \subseteq \{1, 2, \ldots, n\}$, a line L is a set of the form

$$L = \left\{(x_1, x_2, \ \ldots, x_n) : x_i = x_j \text{ for } i, j \in S; \text{and } x_l = a_l \in A \text{ for } l \notin S\right\}.$$

We are ready to formulate the Hales–Jewett Theorem.

The Hales–Jewett Theorem 46.8 [HJ]. For any finite set A and positive integer k, there exists an integer $N(A, k)$ such that for $n \geq$, $N(A, k)$ any k-coloring of A^n contains a monochromatic line.

A very clear "sketch of proof" can be found in [Gra1].

This result – as is often the case in mathematics, was obtained by young mathematicians: Alfred W. Hales was 23 and Robert I. Jewett 24. In the January 3, 2007, email to me, Alfred recalls how it all came about:

> Bob and I were undergraduates at Caltech[3] together – he was a year ahead of me. We had common interests in both math and volleyball. We also both worked in Sol Golomb's[4] coding theory group at the Jet Propulsion Laboratory (JPL, affiliated with Caltech) during summers, and we continued doing this when we were in graduate school – he at the University of Oregon and I at Caltech.
>
> The strong connection between error correcting codes and combinatorics led Sol to steer us in various combinatorial directions and this led (eventually) to our joint paper written at JPL in 1961.

In the December 17, 2007, email, Al adds:

> I did ask Sol [Golomb] about this – You recall that he was our supervisor in the Jet Propulsion Laboratory's coding theory group. He seems to remember that a problem in Martin Gardner's column suggested to him the possibility of generalizing van der Waerden's theorem in some way, with applications to games and to coding in mind. He thinks he discussed this with us, and we proceeded to formulate and prove the eventual result.

---

[3] California Institute of Technology.

[4] We have already met Solomon Golomb in Chapter 2 of the book.

Robert I. Jewett (left) and Alfred W. Hales in 1991, Courtesy of Al Hales

In their standard text on Ramsey Theory [GRS2], Graham, Rothschild, and Spencer give a very high praise to the Hales–Jewett Theorem:

> In its essence, van der Waerden's theorem should be regarded, not as a result dealing with integers, but rather as a theorem about finite sequences formed from finite sets. The Hales-Jewett theorem strips van der Waerden's theorem of its unessential elements and reveals the heart of Ramsey theory. It provides a focal point from which many results can be derived and acts as a cornerstone for much of the more advanced work. Without this result, *Ramsey theory* would more properly be called *Ramseyan theorems.*

In 1971, the Hales–Jewett Theorem earned the authors the George Polya Prize, which they shared with Ronald L. Graham, Klaus Leeb and Bruce L. Rothschild, the authors of the *Affine Ramsey Theorem*, which is a vast generalization of the Hales–Jewett Theorem.

I met Al Hales in the fall 1978 in his beautiful Pacific Palisades home, shortly after my arrival in the United States as a "parolee refugee." It was clearly Al's recommendation that prompted UCLA mathematics chair to offer me my first American professorial job at UCLA. Unfortunately, my romanticism took over, and I chose the mountains of Colorado over the fine Los Angeles campus, decorated by original bronzes of the great British sculptor Henry Moore.

There is a noteworthy connection between two celebrated results (see proof in [GRS2, pp. 40–41]):

Connection 46.9 The Hales–Jewett Theorem implies the Gallai Theorem.

Gian Carlo Rota conjectured what can be called the *Affine Ramsey Theorem Conjecture* (for this *Coloring Book*, I replaced "partitioning" by "coloring").

Rota's Conjecture 46.10 Let l, k, r be nonnegative integers and F a field of q elements. Then there is a number $N = N(q, r, l, k)$ depending only on q, r, l, k with the following property: If V is a vector space over F of dimension at least N, and all the k-dimensional subspaces of V are r-colored, then there is some l-dimensional subspace with all of its k-dimensional subspaces in the same color.

In 1971, this conjecture (and more) was proved by Ronald L. Graham and Bruce L. Rothschild [GR0]. Consistently with my convention of crediting authors of essential conjectures and its provers, we get the following theorem.

The Graham–Rothschild–Rota Theorem 46.11 Let l, k, r be nonnegative integers and F a field of q elements. Then there is a number $N = N(q, r, l, k)$ depending only on q, r, l, k with the following property: If V is a vector space over F of dimension at least N, and all the k-dimensional subspaces of V are r-colored, then there is some l-dimensional subspace with all of its k-dimensional subspaces in the same color.

Part IX
Colored Integers in Service of the Chromatic Number of the Plane: How O'Donnell Unified Ramsey Theory and No One Noticed

An interesting recent result of O'Donnell [Odo4,5], perhaps giving a small amount of evidence that $\chi(E^2) > 4$.

–Ronald L. Graham[1]

O for God's sake
they are connected
underneath

They look at each other
across the glittering sea
some keep a low profile

Some are cliffs
The bathers think
islands are separate like them.

–Muriel Rukeyser[2]

[1] [Gra6]

[2] *Islands*; in the book *The Gates*, 1976

Chapter 47
O'Donnell Earns His Doctorate

I agree with Ron Graham, Paul O'Donnell proved sensational results, showing that there are unit distance 4-chromatic graphs of girth 9, girth 12, and even of an arbitrarily high girth. These results did give some evidence that, perhaps, the search for a 5-chromatic unit-distance graph may celebrate its victory one day – this is the result O'Donnell was ultimately after but had not succeeded, no one had until 2018 (later about it).

The epigraph shows, of course, that Ron Graham appreciated the result, as did Paul Erdős, when I introduced Paul to Paul. What no one noticed, however, is how great Paul O'Donnell's *proofs* were. Just imagine, you create a huge 4-chromatic graph without cycles of size up to, say, 100. Now you need to embed it in the plane, so that every edge is a unit segment and coincidences are avoided. Wouldn't you feel that this is extremely hard, and messy, and you would likely waste much time, and possibly end up with nothing? Paul showed bravery and imagination when he plunged into unit distance embeddings, which we studied in Chapter 14.

He has also set world records of embedding smallest known unit distance graphs without small cycles, jointly with his friend and one-time roommate Rob Hochberg – we have seen those in Chapter 15. Decades later, Geoffrey Exoo and Dan Ismailescu broke one important record, as you have seen in the brand-new Chapter 16 of this expanded edition, dedicated to their results.

I appreciate O'Donnell's constructions presented in this chapter. Paul uses the power of classic results of integer coloring, such as the Baudet–Schur–Van der Waerden Theorem, great results related to the attempts to find a proof of Fermat's Last Theorem from Number Theory and Ergodic Ramsey Theory, such as the Mordell–Faltings Theorem and the Bergelson–Leibman Theorem. He applies this powerhouse of integer colorings and number theory sophistication to the problem of coloring the plane, the chromatic number of the plane problem. And by doing so, O'Donnell is *unifying Ramsey Theory* as nothing else could. This is what brought to my mind the second epigraph of this chapter, a beautiful poem by the American poet Muriel Rureyser *Islands*, islands that look separated by waters but in fact "they are connected underneath."

I followed Paul's research ever since February 1992 at Florida Atlantic University, where my talk inspired a number of young colleagues to join in researching the chromatic number of the plane. In particular, it inspired Paul to write his doctoral thesis on these problems. He visited me in Colorado Springs, and that visit, apparently, had exciting consequences.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_47

On May 25, 1999, Paul O'Donnell defended his doctorate at Rutgers University. It appears that my furniture had something to do with Paul O'Donnell's remarkable dissertation, for in the dissertation's Acknowledgements he writes:

Thanks to Alex. It all came to me as I drifted off to sleep on your couch.

I was a member of Paul O'Donnell's Ph.D. Defense Committee at Rutgers University, together with János Komlos, Michael Saks, and Endre Szemerédi. As soon as Paul finished his presentation and left the room, members of the Committee asked me with a touch of tension, "What do you think?" They considered me to be an informal supervisor of Paul's dissertation (the formal one was János Komlos) – perhaps, I was. I replied, "Paul exceeded all possible expectation of a doctoral dissertation." Even though I knew this dissertation well and followed it through many revisions, it took me time and writing the first edition of this book to fully appreciate how great Paul's methods were. Enjoy!

From the left: Paul O'Donnell, Alexander Soifer, Endre Szemerédi, János Komlos, and Michael Saks, May 25, 1999

Chapter 48
Applications of the Baudet–Schur–Van der Waerden

At the end of Chapter 14, I left you with the embedding in the plane of the 352,735-vertex Blanche Descartes graph by Paul O'Donnell. One may ask, would attaching longer k-cycles ($k > 7$) to the foundation vertices increase the graph's girth while keeping the chromatic number at 4? The answer is no – not if k-cycles were attached to *all* k-element subsets of the foundation set – because some k-cycles would have two or more vertex intersection that could cut down the girth of the graph. We would get a chance to succeed at this construction if we were to dramatically limit the number of attached k-cycles, by, say, allowing at most a single point intersection for the k-subsets of the foundation, to which k-cycles are allowed to be attached. This is exactly what O'Donnell implemented.

We met hypergraphs at the end of Chapter 27; let us meet a special type of them here. A *k-uniform hypergraph* H is a family of k-element subsets of an n-element set S. The *vertices* of H are the elements of S. The *edges* (or *hyperedges*) of H are the k-element subsets. *A cycle* of length $k > 2$ in H is a sequence of distinct vertices and edges of H,

$$v_1, E_1, v_2, E_2, \ldots, v_k, E_k,$$

such that $v_{i+1} \in E_i \cap E_{i+1}$ for $1 \leq i \leq k$ (where the addition in the indices is done modulo k). The *girth* of a hypergraph is the length of its shortest cycle. The *chromatic number* of a hypergraph is the minimum number of colors needed to color the vertices so that no edge contains vertices, which are all colored the same color.

Let n be a positive integer, H a graph on k vertices ($k \leq n$), and $S \subseteq \binom{|n|}{k}$ a k–uniform hypergraph.[1] Then $G_{n,H,S}$ would denote the Blanche Descartes graph[2] built on the foundation vertex set $F = \{u_1^*, u_2^*, \ldots, u_n^*\}$ by attaching[3] copies of H to those subsets of F that are in S.

[1]Here the symbol $\binom{|n|}{k}$ stands for the set of all k-element subsets of the $|n|$-element set.

[2]Defined in Construction 12.10; see also examples of use 12.8 and 12.9.

[3]Defined in Section 14.1.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_48

In this notation, the 112-vertex graph constructed in Problem 12.8 can be recorded as $G_{7,3-cycle,\binom{|7|}{3}}$. The 6448-vertex graph, first embedded by Wormald (Section 12.3), can be recorded as $G_{13,5-cycle,\binom{|13|}{5}}$. The girth of 6,352,735-vertex Blanche Descartes graph embedded in the plane by Paul O'Donnell (see the end of Chapter 14) is encoded as $G_{19,7-cycle,\binom{|19|}{7}}$.

O'Donnell came up with a brilliant idea of attaching cycles only to certain arithmetic progressions (AP's) of the foundation set and to restrict AP's in the following two ways:

- The set D of allowable constant differences is chosen so that arithmetic progressions with distinct constant differences overlap by at most one element (overlaps by two or more vertices may create small cycles).
- Given D, the set S is constructed so that arithmetic progressions with the same constant difference do not overlap.

The distance between any two points in a k-term AP is ad, where $a < k$ and d is the constant difference. To prevent, two AP's from intersecting in two points, it suffices to ensure that $ad_1 \neq bd_2$ for all a,b less than k and distinct constant differences d_1 , d_2 from D. Formally, let D_j denote the set of allowable constant differences less than or equal to j. We define D_j recursively:

$$D_j = \left\{ \begin{array}{l} D_{j-1} \cup \{j\}, \text{if for all } d \in D_{j-1} \text{ and positive integers} \\ a, b \in [k-1], ad \neq bj; \\ D_{j-1} \text{ otherwise.} \end{array} \right\}$$

Then the allowable set of constant differences is $D = \bigcup_{j=1}^{\infty} D_j$.

How dense is D? If too many numbers are in D, then the graph may have short cycles. If too few numbers are in D, then the graph may not be 4-chromatic. So, we need to perform a balancing act. The following tool gives an idea of the density of D.

Tool 48.1 For all d, at least one of $\{d, 2d, 3d, \ldots, k!d\}$ is in D.

Proof If $k!d \in D$, then we are done. If not, then there exist positive integers a, $b \in [k-1]$ and $d_1 \in D$ with $d_1 < k!d$ such that $ad_1 = bk!d$. Solving for d_1, we get $d_1 = \frac{bk!d}{a}$. Since $a < k$, a divides $k!$, and thus d_1 is a multiple of d, as desired. ■

Once we get D, we can construct the set S of APs. Let us formally define S:

$$S = S(n, k, D) = \{\{a, a+d, \ldots, a+(k-1)d\} : d \in D, \\ a \equiv 1, 2, \ldots, d (\text{mod } kd), \ a + (k-1)d \leq n\}.$$

For example, if $D = \{1, 3, 4, 5, \ldots\}$ then $S(17,3, D)$ is:

{1,2,3}	{1,4,7}	{1,5,9}	{1, 6,11}
{4,5,6}	{2,5,8}	{2, 6,10}	{2, 7,12}
{7,8, 9}	{3,6,9}	{3,7,11}	{3, 8,13}
{10,11,12}	{10,13,16}	{4,8,12}	{4,9,14}
{13,14,15}	{11,14,17}		{5,10,15}

Now we need to check the chromatic number and the girth of the graph $G_{n,k\text{-}cycle,S}$ for appropriate k and n and verify that $G_{n,k\text{-}cycle,S}$ is a unit-distance graph.

It is a delight to see how Paul O'Donnell uses the Baudet–Schur–Van der Waerden Theorem to show that for some n, $G_{n,k\text{-}cycle,S}$ is 4-chromatic.

Theorem 48.2 There exists n such that $\chi(G_{n,k\text{-}cycle,S}) = 4$.

Proof By the Baudet–Schur–Van der Waerden Theorem, there exists n such that any 3-coloring of the integers from 1 to n contains a monochromatic AP of length $(2k-1)k!$.

Let d be the constant difference of this AP. By Tool 48.1, there exists $d' \in D$, such that d' is a multiple of d such that $d' \leq k!d$. Hence, there is a $(2k-1)$-term monochromatic AP of the foundation vertices with $d' \in D$

$$u_a, u_{a+d'}, \ldots, u_{a+(2k-2)d'}.$$

One of the first k of these indices is congruent to some element in $\{1, 2, ..., k\} (\text{mod } kd)$. The vertex with this index and the k–1 vertices after it (in the AP with the constant difference d') form a set in S. This set has a k-cycle attached. But if all of these foundation vertices are of the same color, there are only 2 colors remaining to color the odd cycle. This is not enough. Thus, at least 4 colors are necessary to color $G_{n,k\text{-}cycle,S}$. ■

Theorem 48.3 For odd $k \geq 9$, $girth(G_{n,k\text{-}cycle,S}) \geq 9$.

Proof A cycle containing no foundation vertices is a k-cycle. All other cycles consist of the foundation vertices separated by at least 2 vertices of an attached cycle. It is, therefore, impossible to have a cycle with only one foundation vertex.

A cycle has only two foundation vertices if the APs of the two attached cycles intersect in two places. However, our choices for D and S prevent this.

A cycle with at least 3 foundation vertices has at least 9 vertices. Therefore, the girth of our graph is at least $\min\{9, k\}$. ■

Observation Just like the Blanche Descartes construction, this method generalizes to the arbitrary chromatic number. By attaching girth 9, $(l-1)$-chromatic graphs to appropriate APs of the foundation vertices, we obtain girth 9, l-chromatic graphs. However, we need to embed our graphs in the plane as unit-distance graphs, and the 4-chromatic graphs seem to be the only reasonable candidates for it.

Theorem 48.4 There exists a girth 9, 4-chromatic unit-distance graph.

Proof As we have established above, for appropriate choices of k and n, the graph $G_{n,k\text{-}cycle,S}$ is 4-chromatic of girth at least 9. Given odd $k \geq 9$, let n_0 be the smallest such n. We show that $G_{n,k\text{-}cycle,S}$ is a unit-distance graph using an embedding procedure similar to that used for the Blanche Descartes graphs in Chapter 14.

By the choice of n_0, there is a 3-coloring of the foundation vertices labeled from 1 to $n_0 - 1$ such that no monochromatic set has an odd cycle attached. We place all the foundation vertices with color i in the δ-ball around C_i for $1 \leq i \leq 3$. We place vertex n_0 in the δ-ball around C_4. Since the vertices with a k-cycle attached are always in at least 2 δ-balls, the embedding tools of Chapter 14 allow the attachments of all cycles and removal of any coincidences. (Technically, if the girth is more than 9, we add a 9-cycle to get a girth 9 graph.) ■

Chapter 49
Applications of the Bergelson–Leibman and the Mordell–Faltings Theorems

To achieve a girth 12 unit-distance graph, Paul O'Donnell alters the set D of allowable constant differences. This changes which sets are in S (i.e., which sets of the foundation vertices get odd cycles attached). It is no longer enough for the sets in S to have intersection of size at most one, as we required in Chapter 48. In addition, O'Donnell requires now that no three sets in S intersect pairwise. How does one achieve this?

Unexpectedly, O'Donnell uses sophisticated results from Number Theory and Ergodic Ramsey Theory. He attaches k-cycles only to specified APs whose constant difference is an mth power, for he wants to make use of BLT's Corollary 38.10 of the Bergelson–Leibman Theorem!

As was done in Chapter 48, we will again use the Blanche Descartes $G_{n,k\text{-cycle},S}$ construction. We will then establish that the constructed graph is indeed 4-chromatic, girth at least 12 unit-distance graph.

However, before we dive into "*O'Donnellia*," we need to take a tour of number theory related to . . . Fermat's Last Theorem. As is customary in this book, we will include at least a brief history of this field in our excursion.

In 1922, Louis Joel Mordell (Philadelphia, 1888–Cambridge, 1972) conjectured [Mor] and in 1983 the 29-year-old German mathematician Gerd Faltings proved (and in 1986 was awarded the Fields Medal primarily for his proof) this very important result (in a more contemporary formulation than Mordell could have had). This result, among other consequences, was, of course, a major step in the ascent on Fermat's Last Theorem. In consistently following my view that creating a good conjecture is important (every theorem is preceded by a conjecture, and sometimes the conjecture is brought up by someone other than the one who proves it), I will call it the Mordell–Faltings Theorem. We need here precisely the consequence of this theorem that is relevant to Fermat's Last Theorem when we construct the set of allowable constant differences. It deals with (integer) solutions of Diophantine equations of the form

$$ax^m + by^m + cz^m = 0. \qquad (*)$$

Before we state the theorem, we need to introduce some preliminaries. A solution (x_0, y_0, z_0) of (*) is called *primitive* if $gcd\{x_0, y_0, z_0\} = 1$; and *trivial* if $x_0, y_0, z_0 \in \{-1, 0, 1\}$. Notice that if $\{x_0, y_0, z_0\}$ is a solution, then any integer multiple of this triple is also a solution. Thus,

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_49

if an equation (*) has one solution it has infinitely many. However, for an appropriate choice of m, it has only a finite number of primitive solutions. For a better choice of m all primitive solutions are also trivial. For the final choice of m, all equations (*) with $a, b, c \in \{-k, ..., k\}$ not all zero, have no nontrivial primitive solutions. This allows us to construct the set of allowable constant differences and the set of arithmetic progressions to which odd cycles are attached.

The Mordell–Faltings Theorem 49.1 A nonsingular projective curve of genus at least two over a number field has at most finitely many points with coordinates in the number field.

I refer you to contemporary number theory texts for definitions of terms used in 49.1. What we need here is the following corollary, obviously relevant to the assault on Fermat's Last Theorem:

Mordell–Faltings' Corollary 49.2 Given $a, b, c \in Z$ not all zero; for $m \geq 4$, the equation $ax^m + by^m + cz^m = 0$ has at most finitely many primitive solutions.

Tool 49.3 Given $a, b, c \in Z$, not all zero; there exists m such that the equation $ax^m + by^m + cz^m = 0$ has no nontrivial primitive solutions.

Proof The Mordell–Faltings Corollary 49.2 states that for $m \geq 4$, $ax^m + by^m + cz^m = 0$ has finitely many primitive solutions. Given a, b, c, let w be the integer of the largest absolute value in any primitive solution of $ax^4 + by^4 + cz^4 = 0$. Choose $l = l(a, b, c)$ such that $2^l > w$. We need the following claim to complete the proof:

Claim 49.4 The equation $ax^{4l} + by^{4l} + cz^{4l} = 0$ has no primitive solutions except possibly trivial ones, in which $x, y, z \in \{-1, 0, 1\}$.

Proof of 49.4 Assume $ax_0^{4l} + by_0^{4l} + cz_0^{4l} = 0$ with $gcd\{x_0, y_0, z_0\} = 1$. Then the equality $a\left(x_0^l\right)^4 + b\left(y_0^l\right)^4 + c\left(z_0^l\right)^4 = 0$ shows that x_0^l, y_0^l, z_0^l is a primitive solution of $ax^4 + by^4 + cz^4 = 0$. By the definition of w,

$$\max\left(\left|x_0^l\right|, \left|y_0^l\right|, \left|z_0^l\right|\right) \leq |w| < 2^l,$$

therefore, $x_0, y_0, z_0 \in \{-1, 0, 1\}$.

All there is left to complete the proof of the tool 49.3, is to choose $m(a, b, c) = 4\,l$, which in view of 49.4 satisfies the statement of tool 49.3. ■

Corollary 49.5 Given a positive integer k, there exists a positive integer m' such that none of the equations $ax^{m'} + by^{m'} + cz^{m'} = 0$ with $a, b, c \in \{-k, ..., k\}$ not all zero has nontrivial primitive solutions.

Proof Given a, b, c, by tool 49.3, there exists $m = m(a, b, c)$ such that $ax^m + by^m + cz^m = 0$ has no nontrivial primitive solutions. The same holds for any exponent which is a multiple of m. Hence,

$$m' = \prod_{\{\{a,b,c\}: a,b,c \in \{-k,\ldots,k\}, \text{not all } 0\}} m(a, b, c.)$$

suffices. ■

Everything is now ready for our construction. Given $m' = m'(k)$, we define $D = \{x^{m'} : x \in N\}$.This is the set of allowable constant differences needed to construct the set S of arithmetic progressions. Each arithmetic progression in S corresponds to a set of foundation vertices with an attached cycle.

Theorem 49.6 For odd $k \geq 13$, girth($G_{n,k\text{-cycle},S}$) ≥ 12.

Proof A few cases need to be addressed, depending upon the number of the foundation vertices in a k-cycle.

A cycle containing no foundation vertices is a k-cycle. All other cycles consist of foundation vertices separated by at least 2 vertices of an attached cycle. So, a cycle with at least 4 foundation vertices has at least 12 vertices.

A cycle has 3 foundation vertices if the APs of the three attached cycles intersect pairwise. Let a_i be the starting point and d_i be the constant difference, $1 \leq i \leq 3$, for the three APs. The pairwise intersections of the APs imply the existence of constants $c_1, c_2, ..., c_6$ between 0 and $k - 1$ such that

$$a_1 + c_1 d_1 = a_2 + c_2 d_2$$

$$a_2 + c_3 d_2 = a_3 + c_4 d_3$$

$$a_3 + c_5 d_3 = a_1 + c_6 d_1$$

Thus,

$$a_1 + a_2 + a_3 + c_1 d_1 + c_3 d_2 + c_5 d_3 = a_1 + a_2 + a_3 + c_6 d_1 + c_2 d_2 + c_4 d_3$$

or,

$$(c_1 - c_6)d_1 + (c_3 - c_2)d_2 + (c_5 - c_4)d_3 = 0.$$

Since the constant differences are all m'-th powers and the three foundation vertices are distinct, this is an equation of the form $ax^{m'} + by^{m'} + cz^{m'} = 0$ with integral coefficients $a, b, c \in \{-k, ..., k\}$ not all zero. By corollary 49.5, it has only trivial primitive solutions. Thus, any solution has all the d_i equal, yet in the construction of S, APs with the same constant difference do not intersect.

A cycle has only 2 foundation vertices if the arithmetic progressions of the two attached cycles intersect in two places. Let a_i be the starting point and d_i be the constant difference, $1 \leq i \leq 2$, for the two arithmetic progressions. The intersection of the APs implies the existence of constants c_1, c_2, c_3, c_4 between 0 and $k - 1$ such that

$$a_1 + c_1 d_1 = a_2 + c_2 d_2$$

$$a_1 + c_3 d_1 = a_2 + c_4 d_2$$

By adding up the respective sides of these equalities, we get

$$(c_1 - c_3)d_1 + (c_4 - c_2)d_2 = 0.$$

Since the constant differences are all m'-th powers and the two foundation vertices are distinct, this is an equation of the form $ax^{m'} + by^{m'} = 0$, with nonzero integer coefficients between $-k$ and k. As in the previous case, there are no nontrivial primitive solutions. The d_i must be equal, yet in the construction of S, arithmetic progressions with the same constant differences do not intersect.

A cycle with only one foundation vertex is not possible. Therefore, the girth is at least min $\{12, k\}$. ■

Observation Just like Blanche Descartes's girth 9 construction of Chapter 48, this method generalizes to arbitrary chromatic number. By attaching girth 12, $(l - 1)$-chromatic graphs to appropriate arithmetic progressions of the foundation vertices, we get girth 12, l-chromatic graphs. Again, the only reasonable candidates for embedding in the plane as unit-distance graphs seem to be the 4-colorable graphs.

Theorem 49.7 There exists n such that $\chi(G_{n,k\text{-cycle},S}) = 4$.

Proof By the BLT' Corollary 38.10 of the Bergelson–Leibman Theorem 38.9 (Chapter 38), there exists n such that any 3-coloring of the integers from 1 to n contains a $(2k - 1)$-term monochromatic arithmetic progression of the foundation vertices

$$u_a, u_{a+d}, \ldots, u_{a+(2k-2)d}$$

where d is a mth power. One of the first k of these indices is congruent to some element in $\{1, 2, ..., k\}$ (mod kd). The vertex with this index and the $k - 1$ vertices that follow it, form a set in S. This set has a k-cycle attached. But if all of these foundation vertices are of the same color, there are only 2 colors remaining to color the attached odd cycle. This is not enough. Thus, at least 4 colors are necessary to color $G_{n,k\text{-cycle},S}$. ■

We are ready for the embedding.

Theorem 49.8 There exists a girth 12, 4-chromatic unit-distance graph.

Proof From the preceding theorems, we know that for appropriate choices of k and n, the graph $G_{n,k\text{-cycle},S}$ is a 4-chromatic graph of girth at least 12. Given odd $k \geq 13$, let n' be the smallest such n. We will show that $G_{n',k\text{-cycle},S}$ is a unit-distance graph using an embedding procedure similar to that used in the previous chapter. By the choice of n', there is a 3-coloring of the foundation vertices labeled from 1 to $n'-1$ such that no monochromatic set has an odd cycle attached. We place all the foundation vertices of color i in the δ-ball around C_i, for $1 \leq i \leq 3$. We place vertex n' in the δ-ball around C_4. Since the vertices with a k-cycle attached are always in at least 2 δ-balls, the embedding tools of Chapter 14 allow the attachments of all cycles and removal of any coincidences. (Technically, if the girth is more than 12, we add a 12-cycle to get a girth 12 graph.) ■

Chapter 50
Solution of an Erdős Problem: The O'Donnell Theorem

In a surprising twist, the complete solution of Paul Erdős' old July 1975 problem about unit distance 4-chromatic graphs of arbitrary girth comes out to be simpler than all partial solutions, we have discussed in the previous two chapters. In another surprise, Paul O'Donnell uses in his solution the 1966 result obtained jointly by Paul Erdős and Andras Hajnal, the result that has been known all alone, but no one noticed its connection to the problem at hand. You may wish to revisit definitions of uniform hypergraphs in the beginning of Chapter 48.

Theorem 50.1 (Erdős–Hajnal 1966, [EH1]). For all integers $k \geq 2$, $g \geq 2$, and $l \geq 2$, there exist a k-uniform, girth g, l-chromatic hypergraphs.

This theorem gives the desired generalization of the girth 9 and girth 12 constructions. Instead of attaching cycles to arithmetic progressions, we attach cycles to the edges (hyperedges) of a hypergraph. Given k and g, let H be a k-uniform, girth g, 4-chromatic hypergraph. Let $n = |V(H)|$. Then $G_{n,k\text{-cycle},H}$ is the desired graph (reread its definition in Chapter 48 if need be).

Theorem 50.2 (O'Donnell). $\chi(G_{n,k\text{-cycle},H}) = 4$.

Proof Since H is 4-chromatic, any 3-coloring of the foundation vertices contains a monochromatic hyperedge. In other words, any 3-coloring of the foundation vertices has a monochromatic set with an odd cycle attached. That cycle cannot be colored with the remaining two colors, so $\chi(G_{n,k\text{-cycle},H}) \geq 4$. With four colors, one color can be used for the foundation vertices leaving three for the odd cycles. Thus, $\chi(G_{n,k\text{-cycle},H}) = 4$. ■

Theorem 50.3 (O'Donnell). $\text{girth}(G_{n,k\text{-cycle},H}) = k$.

Proof The approach is to show that $\text{girth}(G_{n,k\text{-cycle},H}) \geq \min\{k, 3g\}$ and choose $g \geq k/3$.

The only cycles containing no foundation vertices are the attached k-cycles. All other cycles consist of foundation vertices separated by at least two vertices of attached cycles. Since any two consecutive foundation vertices are in the shadow of an attached cycle G (i.e., appear in a hyperedge of H), the consecutive foundation vertices form a cycle (i.e., hypercycle) in H. So if the girth of H is g, the length of the cycle in $G_{n,k\text{-cycle},H}$ is at least $3g$. Thus, all cycles of $G_{n,k\text{-cycle},H}$ are either k-cycles or l-cycles for $l \geq 3g$. ■

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_50

The O'Donnell Theorem 50.4 [Odo3]; [Odo4, Odo5]. For any integer $k \geq 3$, there exists a girth k, 4-chromatic unit distance graph.

Proof Assume k is odd. Let H be a k-uniform, 4-chromatic hypergraph with girth $\geq k/3$ having the fewest vertices. Let $n' = |V(H)|$, then as we know from the previous theorems, $G_{n',k\text{-cycle},H}$ is a girth k, 4-chromatic graph. As in the previous two chapters, we use the embedding tool chest of Chapter 14. By the choice of n', there is a 3-coloring of the foundation vertices labeled from 1 to $n' - 1$ such that no hyperedge is monochromatic, in other words, no monochromatic set has an odd cycle attached. We place all the foundation vertices with color i in the δ-ball around C_i for $1 \leq i \leq 3$, and place vertex n' in the δ-ball around C_4. Since the vertices with a k-cycle attached are always in at least 2 δ-balls, the embedding tools allow the attachments of all cycles and removal of any coincidences. For an even k, a k-cycle is added to the 4-chromatic unit-distance graph of girth $> k$. ■

Would you like to *see* the embedded O'Donnell graph? Paul offers an illustration (Fig. 50.1).

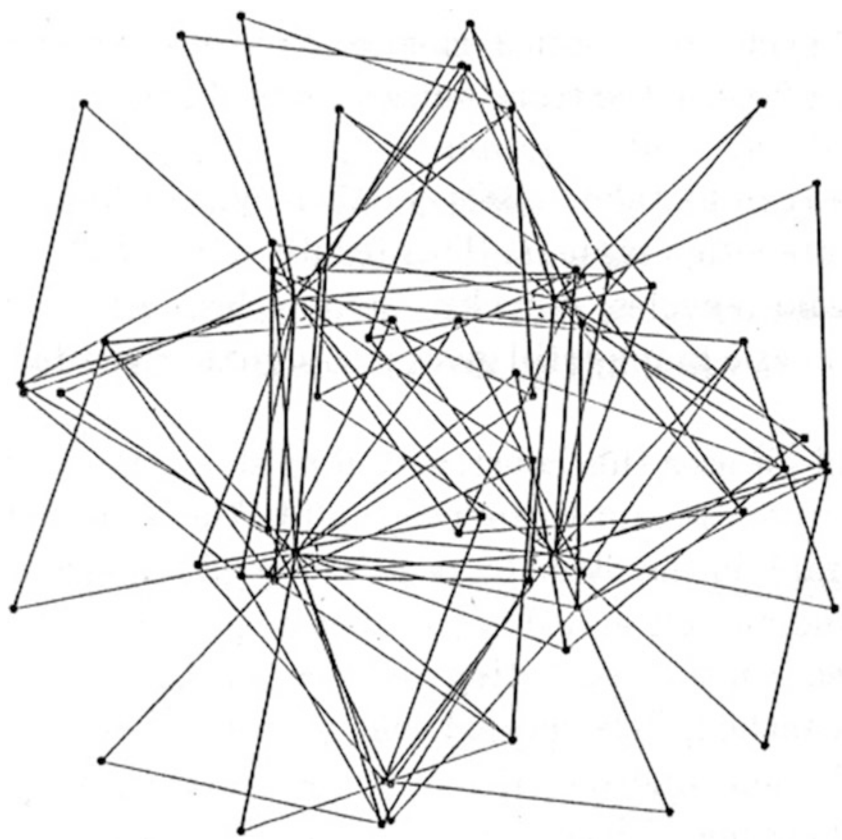


Fig. 50.1 A girth k 4-chromatic unit-distance graph in the plane. (Note: what looks like a vertex is many vertices; what looks like an edge is many almost parallel edges.)

50.1 Paul O'Donnell

My old request for a "self-portrait" Paul O'Donnell honored in the March 31, 2007, e-mail:

> I was born in New York City on April 18, 1968. I was adopted in October 1968 and grew up in Jackson, New Jersey. I received my undergraduate degree in mathematics and computer science from Drew University in 1989 and my Ph.D. in applied mathematics from Rutgers University in 1999. My doctoral thesis was on arbitrary girth 4-chromatic unit-distance graphs in the plane, from a problem posed by Paul Erdős.
>
> My interest in unit-distance graphs sprang originally from a Putnam exam problem about them, and from undergraduate courses taught by Linda Lesniak. This interest was reawakened after attending Alexander Soifer's 1992 presentation on the interesting history of this problem at a conference at Florida Atlantic University in Boca Raton. This marked the start of our friendship.
>
> The main idea for the arbitrary girth unit-distance graph work came a few years later while dozing off to sleep on Alexander Soifer's couch in Colorado Springs after watching the Derek Jarman movie "Wittgenstein" with him.
>
> I have taught at Rutgers University and Drew University. Currently, I am working in the Research & Development Equity Department of Bloomberg L.P. In my free time I play ultimate frisbee and am a theatre/movie buff (credits include work on the L.A. and off-Broadway productions of the musical Reefer Madness, and an appearance as an extra in the movie Army of Darkness). My wife Carmelita and I also teach ballroom/ Latin dance and are the proud parents of daughter Kimberly.

* * *

In 1998, Paul stayed with me for several days. I let him teach my class and found him a wonderful engaging lecturer. We watched together the film "Wittgenstein" by the unique British film director Derek Jarman. You have witnessed his excellence and will as a researcher. To my regret, Michael Bloomberg lured Paul away from Academia.

Paul O'Donnell

Part X
Ask What Your Computer Can Do for You

The fact that cinéma-vérité directors walk around with a Coutant camera under their arm or that they can film their researches into other people's lives makes no difference at all. They still have to be guided by an idea, an attitude, without which their camera would remain inert, just as the most powerful computer in the world remains inert, in spite of its superhuman memory, so long as it is not supplied with a programme.

– Michelangelo Antonioni, 1965 [Ant]

In the 1970s, many humans believed that they made fewer mistakes than computers. Many mathematicians and philosophers questioned Appel–Haken proof of the Four-Color Conjecture. Twenty years later, the new proof of 4CC was not questioned, partly because it was a better proof, partly because humans realized that computers simply do what humans tell them to do. As Paul Kainen amply noticed, computers are simply "computational amplifiers."

You would agree that computers do not solve mathematical problems – people armed with computers do. This new chapter of *The Mathematical Coloring Book* presents major breakthroughs achieved – as it happened – with the aid of a computer.

Computer-free proofs are preferred only because we, humans, are a curious bunch and a human-verifiable proof may satisfy our curiosity and shed light on why things are the way they are. Some computer-aided proofs can do it to.

Chapter 51
Aubrey D.N.J. de Grey's Breakthrough

In 68 years of the problem's life, many fine mathematicians obtained many beautiful results in special circumstances. However, no progress occurred in the general case until Aubrey de Grey (Ph.D. in Biology, University of Cambridge) succeeded in a joint effort of his imagination and a clever computer program he wrote for this purpose. On January 16, 2018, de Grey sent me the first version, followed by a corrected one [G1] on April 7, 2018, with the following introductory note:

> Dear Prof. Soifer,
>
> I append a copy of an email I sent you in January, which you may have overlooked. It's good that you did, because the graph that I told you about is in fact 4-colorable after all, and my failure to discover this was due to a bug in my code. However, I'm pleased to report that after fixing the bug I was rapidly able, using the same basic approach, to find somewhat larger unit-distance graphs that do indeed have no 4-colourings. My confidence that my code is not still lying to me arises largely from the fact that the 4-colouring of the earlier graph was found by Dr. Robert Hochberg, to whom I wrote at the same time as you; he became interested enough to spend time writing code that could test quite large graphs, and he has not found a 4-colouring of the (progressively smaller, but still four-digit) graphs I have been discovering since January even though his code can 4-colour all his previous attempts at 5-chromaticity in seconds. We appreciate that this is not a proof... but it makes us feel good enough about the graphs that I have now written up the discovery as a paper. I have just submitted it to the arXiv, and it is scheduled to go live on Monday. I still very much hope that you will be inclined to consider it for publication in *Geombinatorics*; I am in absolute awe of your 2008 book, and I hope that this might serve as some sort of mark of my gratitude.

I was not surprised by the result that the chromatic number of the plane was at least 5, for I conjectured two decades earlier that the chromatic number of the plane was 7. I was surprised by Aubrey de Grey's construction. Practically, all previous pursuits of a 5-chromatic unit-distance graph (see O'Donnell [O'D], for example) were based on an idea that the hardest part of creating a 5-chromatic unit-distance graph was its embedding in the plane. And so, all attention was concentrated on unit-distance graphs without 3-cycles, which could be easier to bend and embed in the plane. No triangles, of course, means no Mosers Spindles. De Grey goes the opposite way: he floods his construction with a high density of interlocking Mosers

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_51

Spindles. His goal is to force a certain coloring of a small number of vertices and then create contradictions to those forced colorations.

He constructs a unit-distance graph on 20,425 vertices. Can one check whether it is 4-colorable? Normally, one would not even try, especially when you do not have supercomputers, but ordinary "household" means: Mathematica 11 on a standard MacBook Air laptop. De Grey bravely goes for it, and with a clever use of specific properties of his graph succeeds in verifying that this giant graph is indeed not 4-colorable. Hence, we arrive at The de Grey Theorem.

The de Grey Theorem 51.1 $\chi(E^2) \geq 5$.

De Grey then dramatically reduces the size of his 5-chromatic unit-distance graph.

De Grey's Example 51.2 There is a 5-chromatic unit-distance graph on 1581 vertices.

With this size, the verification becomes within the reach of his computer program. What does Aubrey do next? I have known colleagues, who would keep their approach in secret, or worse, would publish a paper that is hardly comprehensible – in order to position themselves ahead of the competition. Aubrey de Grey is a true scholar. He does not wish to compete with others, but rather invites others to join in to conquer mathematics herself. He succeeds in commencing a Polymath Project where new blood is attracted to try their wits on the problem. And try they do.

Polymath Projects had a precursor. In 1991, I started ***Geombinatorics***, a research quarterly dedicated primarily to problem-posing essays in combinatorial geometry, jointly with the great geometer Branko Grünbaum, in the style of Paul Erdős' problem-posing talks and essays. I found existing mathematical journals to be like cemeteries for honorable burials of finished research. My goal has been to publish research in progress, so that people can join in the efforts. Clearly, *Geombinatorics* was a precursor to Polymath-type blogs. With your contributions and enthusiasm, *Geombinatorics* will continue to be our meeting place, the melting pot of ideas it has been for 33 years, 131 issues and counting.

Chapter 52
De Grey's Construction

I am presenting a slightly edited Aubrey de Grey's crisp and clear description of his construction [G1], translated from the British to the American English. :)

52.1 The Plan

Note that the 7-vertex, 12-edge unit-distance graph H consisting of the center and vertices of a regular hexagon of side-length 1 can be colored with at most four colors in four essentially distinct ways (up to rotation, reflection, and color transposition). The upper two of these colorings contain a monochromatic triple of vertices and the lower two do not (Fig. 52.1).

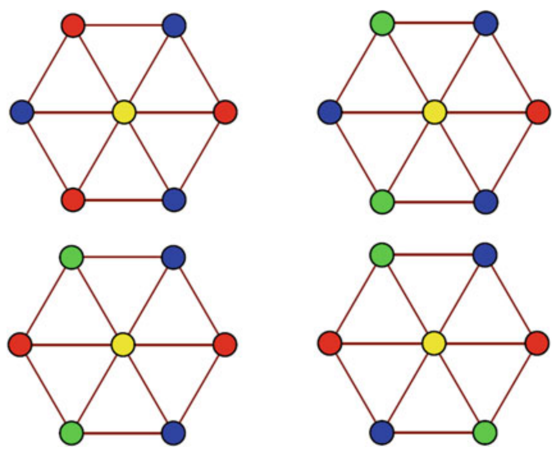


Fig. 52.1 The essentially distinct ways to color H with at most four colors

De Grey constructs a unit-distance graph L that contains 52 copies of H and shows that in all 4-colorings of L, at least one copy of H contains a monochromatic triple.

He then constructs a unit-distance graph M that contains a copy of H as a subgraph and shows that there is no 4-coloring of M in which that H contains a monochromatic triple. Thus, the unit-distance graph N created by arranging 52 copies of M so that their counterparts of

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_52

H form a copy of L is not 4-colorable. This completes the demonstration that the chromatic number of N and therefore of the plane is at least 5.

Finally, de Grey finds smaller non-4-colorable unit-distance graphs, first by identifying vertices in N whose deletion preserves the absence of a 4-coloring, and then by more elaborate methods. He is now ready for the implementation of the plan, let us follow him.

52.2 The 4-Colorings of *J* in Which No Copy of *H* Contains a Monochromatic Triple

Define the graph J, shown in Fig. 52.2, containing a copy of H in its center, six copies centered at distance 1 from its center, and six copies centered at distance $\sqrt{3}$ from its center.

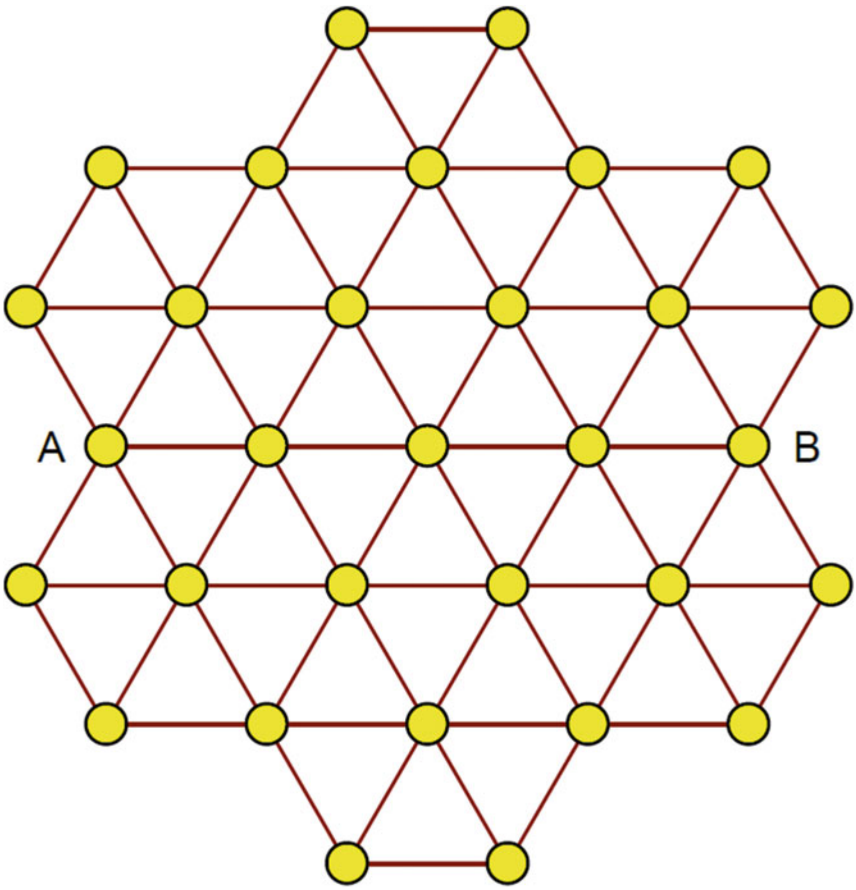


Fig. 52.2 The graph J, containing 31 vertices and 13 copies of H

Figure 52.3 shows six essentially distinct (up to rotation, reflection, and color transposition) 4-colorings of J in which no copy of H contains a monochromatic triple. The vertices colored black in Fig. 52.3 will be of no concern.

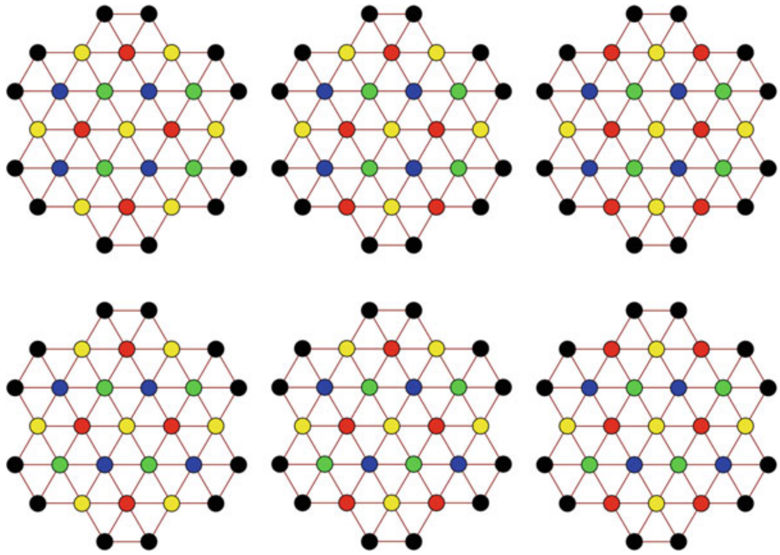


Fig. 52.3 The essentially distinct colorings of *J* in which no copy of *H* (including the ones containing some black vertices) contains a monochromatic triple. Black vertices can assume any color so long as no connected vertices are the same color

That these are the only such colorings can be checked by grouping the possibilities according to whether the central copy of *H* has two (the top row in Fig. 52.3) or no (bottom row) monochromatic pairs of vertices at distance $\sqrt{3}$ and, in the case where it has none, whether all the copies of *H* centered at distance 1 from the center also have none (bottom left coloring) or some have two (last two colorings). The bottom center coloring is the reason we need the vertices colored black; though those vertices can be colored in many ways, it turns out that if they were deleted then there would be additional 4-colorings of the remaining graph in which none of the seven remaining copies of *H* contained a monochromatic triple, and some of those new colorings lack *the key property* shared by all those in Fig. 52.3.

The key property of the colorings in Fig. 52.3 that will be the focus henceforward is the feature that only three essentially distinct colorings of the vertices at distance 2 from the center have, which shall be hereafter called *the linking vertices*. There are three possibilities:

(a) The linking vertices are all the same color as the center (left-hand cases in Fig. 52.3), or,
(b) Four consecutive linking vertices (when enumerated going clockwise around the center) are the same color as the center and the other two are a second color (middle cases), or,
(c) Two opposite linking vertices are the same color as the center and all the other four are a second color (right-hand cases).

Hereafter a pair of linking vertices located in opposite directions from the center, such as those labelled *A* and *B* in Figs. 52.2 and 52.4, will be called *a linking diagonal.*

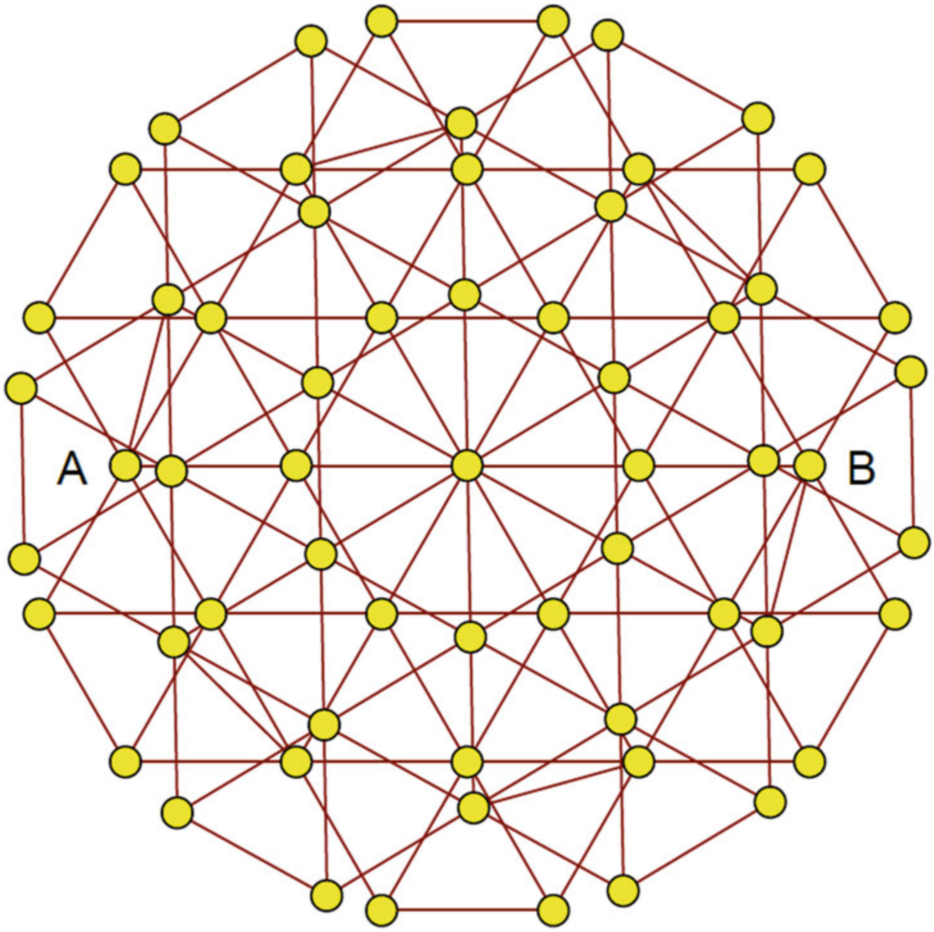


Fig. 52.4 The graph K, containing 61 vertices and 26 copies of H

52.3 61-Vertex Graph K Assembled from Two Copies of J

De Grey then constructs graph K as the union of J with a copy of J rotated clockwise around the origin by $2{\cdot}\arcsin(1/4)$. This rotation causes corresponding linking vertices to lie at distance 1 from each other; see Fig. 52.4. This construction is a generalization of Leo and Willie Moser's construction of the Mosers Spindle; it has become known as *spindling*. Later in his book, I will let Exoo and Ismailescu to formally define *spindling* and prove its property.

Observe that, in any 4-coloring of K in which none of the 26 copies of H contains a monochromatic triple, both copies of J must have their linking vertices colored according to option (c) above. This is of interest because in option (c) each of the three linking diagonals of J is monochromatic. Thus, in all 4-colorings of K where no copy of H contains a monochromatic triple, all six linking diagonals are monochromatic.

52.4 121-Vertex Graph *L* Assembled from Two Copies of *K*

Finally, de Grey constructs graph L as the union of K with a copy of K rotated around A by $2 \cdot \arcsin(1/8)$. This rotation causes the counterpart of B (denote it B') to lie at distance 1 from B – see Fig. 52.5, in which L has been translated and rotated to give it symmetry about the y-axis (spindling again).

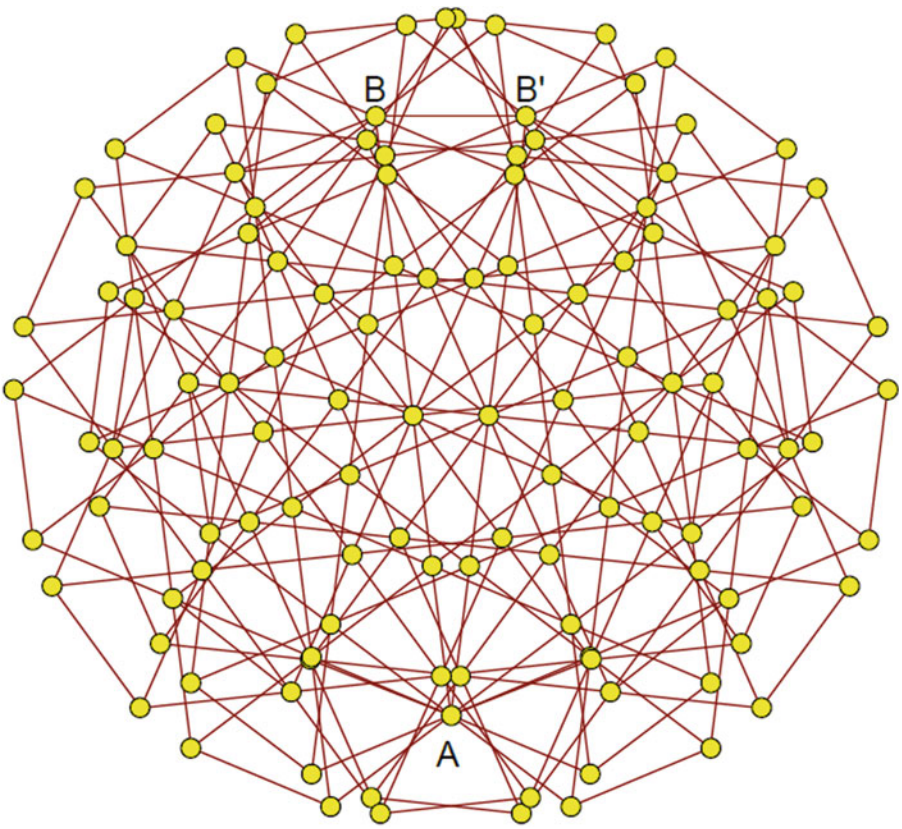


Fig. 52.5 The graph L, containing 121 vertices and 52 copies of H

The property of graph K observed above guarantees that in no 4-coloring of L do all of its 52 constituent copies of H lack a monochromatic triple. Either B or B' must be of a different color than A, so one of the copies of K must contain a non-monochromatic linking diagonal, thus it must contain a copy of H with a monochromatic triple.

52.5 In Search of Graph *M*

In seeking graphs that can serve as *M* in de Grey's construction, he focuses on graphs with a high density of Mosers Spindles – this is where de Grey departs from the previous attempts to construct a 5-chrinatic unit-distance graph. He explains his motivation for exploring such graphs by the fact that Mosers Spindle contains two pairs of vertices distance $\sqrt{3}$ apart, and these pairs cannot both be monochromatic. Intuitively, therefore, a graph containing a high density of interlocking spindles might be constrained to have its monochromatic $\sqrt{3}$-apart vertex pairs distributed rather uniformly (in some sense) in any 4-colouring. Since such graphs typically also contain regular hexagons of side-length 1, one might be optimistic that they could contain such hexagon that does not contain a monochromatic triple in any 4-colouring of the overall graph, since such a triple is always an equilateral triangle of edge $\sqrt{3}$ and thus constitutes a locally high density, i.e. a departure from the mentioned earlier uniformity, of monochromatic $\sqrt{3}$-apart vertex pairs.

52.6 Graphs with High Edge Density and Spindle Density

His search for graphs with high spindle density, de Grey begins by noticing two attractive features of the 9-vertex unit-distance graph *T* that is obtained by adding two particular vertices to the Mosers Spindle (see Fig. 52.6, left). These added vertices, *P* and *Q* are such that they:

1. Form an equilateral triangle with the tip *X* of the spindle
2. Lie on (the extension of) the line forming the base *Y Z* of the spindle

Thus, de Grey can construct a 15-vertex unit-distance graph *U* (Fig. 52.6, right) that contains three Mosers Spindles and possesses rotational and reflectional triangular symmetries. These symmetries suggest that graphs formed by combining translations and 60-degree rotations of *U* might have particularly high edge and spindle density.

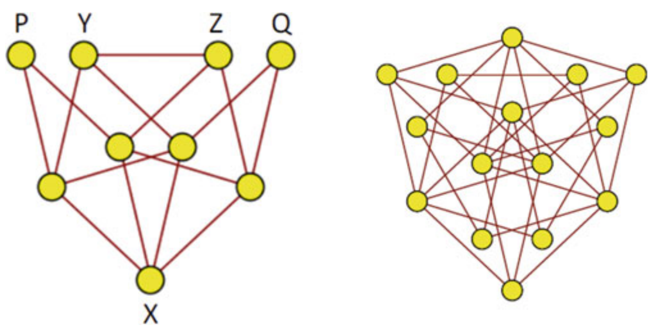


Fig. 52.6 The graphs *T* and *U*

52.7 Construction of a Graph That Serves as *M*

The expectation Aubrey de Grey just mentioned turned out to be true. For example, he found a 97-vertex graph containing 78 spindles (not shown). And so he writes a custom program (outlined in the next section) to test graphs of this form for possession of a 4-coloring in which the central H contains a monochromatic triple. However, he did not find a graph of this form that enforces sufficient uniformity of the distribution of $\sqrt{3}$-apart monochromatic vertex pairs to deliver the property required for M, even though he reports checking examples with over 1000 vertices.

A modification was needed. Graphs arising from the construction described thus far can have spindles in only six different orientations, with edges falling into just three equivalence classes where equivalence is defined as rotation by a multiple of 60 degrees (Fig. 52.7 left). De Grey adds new classes based on the relative orientations of spindles that share a lot of vertices, such as in Fig. 52.7 (middle). The maximum possible degree of a vertex in a graph constructed from these edges interpreted as vectors increases from 18 to 30; see Fig. 52.7 (right), which we denote as *graph V*. The angles of the edges relative to the vector (1,0) are $i \cdot \arcsin(\sqrt{3}/2) + j \cdot \arcsin(1/\sqrt{12})$, where $i \in \{0 \dots 5\}$, $j \in \{-2 \dots 2\}$.

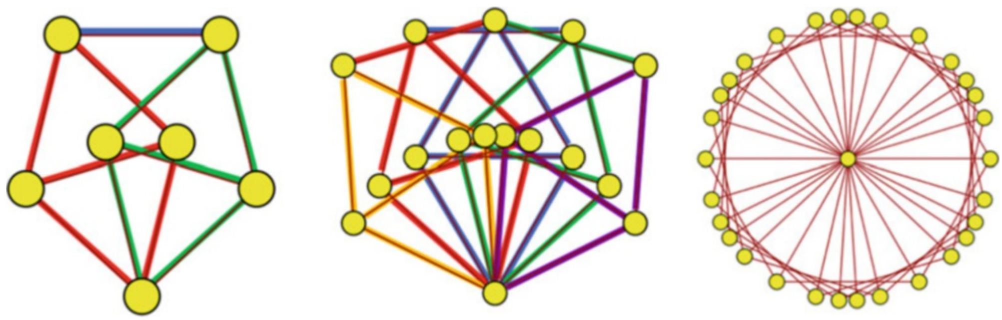


Fig. 52.7 The vector classes present in one (left), or three tightly linked (middle), Mosers Spindles; the graph V (right)

This turns out to suffice. Let W be the 301-vertex graph consisting of all points at distance $\leq \sqrt{3}$ from the origin that are the sum of two edges of V (interpreted as vectors). The 1345-vertex graph M shown in Fig. 52.8 is the union of W with its six translates, in which the origin is mapped to a vertex of H. Aubrey's program did not find any 4-colouring of this graph M in which the central H contains a monochromatic triple, so it indeed can serve as graph M. In other words, he can create a non-4-colorable unit-distance graph N as the union of 52 copies of M, translated and rotated so that each instance of H in L coincides with the central H of a copy of M. After merging coincident vertices arising from different copies of M, this graph N has 20,425 vertices. (De Grey does not show a picture of N, because it is visually impenetrable, and also because he shortly discovered smaller examples).

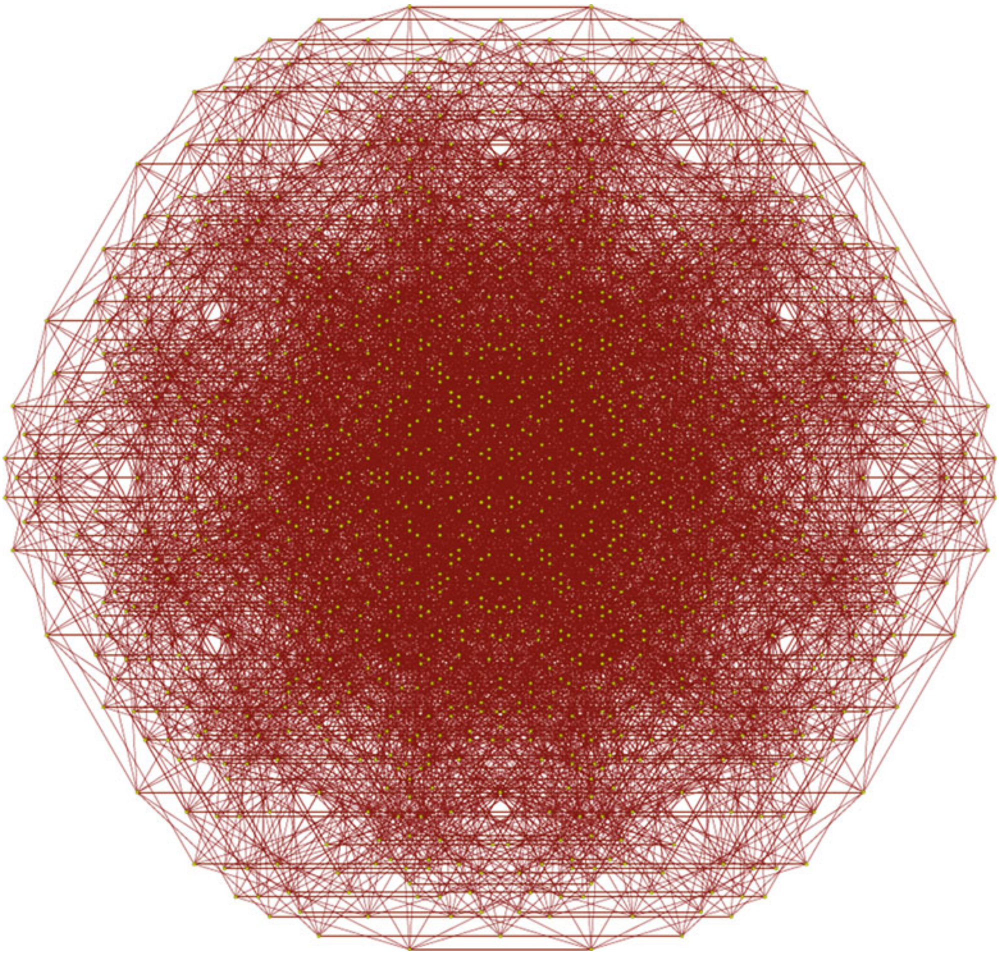


Fig. 52.8 The 1345-vertex Graph M

52.8 Testing 4-Colorability of Edge-Dense, Spindle-Dense Graphs

In general, it is computationally challenging to determine the chromatic number of a graph with over 1000 vertices by simplistic search methods, let alone one with 20,425 vertices. De Grey thus develops a custom program to test graphs for possession of the property required for our graph M, taking advantage of certain properties of our candidate graphs.

Because de Grey is only asking whether a specific number of colors is or is not sufficient, and also because of the high density of edges and spindles in his target graphs, the required test turns out to be computationally far cheaper than a general determination of chromatic number of comparable-sized graphs. It can be performed rapidly by a simple depth-first search optimized only slightly, as follows:

1. Since the question is whether there is any 4-coloring of M in which the central H contains a monochromatic triple, de Grey begins by specifying the colorings of the vertices of that central H, which he terms *the initializing vertices* (there are 7 of them). Since M has the

same symmetries as *H*, Aubrey only need to check the two essentially distinct triple-containing colorings of those vertices.
2. He orders the remaining vertices according to a hierarchy (most significant first) of parameters (all decreasing): how many Mosers Spindles they are part of, their degree, and how many unit triangles they are part of.
3. He colors the next not-yet-colored vertex (initially vertex 8, the first non-initializing vertex) with the first color that he has not already tried for it (initially color 1).
4. He checks each not-yet-colored neighbor (if any) of the just-colored vertex to see how many colors are still permissible for it. If any such vertex already has neighbors of all four colors, he will need to backtrack (see step 6 below). If any has neighbors of exactly three colors, he assigns the remaining (forced) color to it.
5. If he does not need to backtrack, but he did color some vertices in step 4, he repeats step 4 for each such vertex.
6. If he needs to backtrack, he uncolors everything that he just colored in the most recent iteration of steps 3 and (any resulting iterations of) 4.
7. If he just did an uncoloring and the just-uncolored vertex that was colored at step 3 has no colors that have not yet been tried, he repeats step 6 for the next-most-recent iteration of steps 3 and 4 unless he has already backtracked all the way to the vertices of *H*. Otherwise, if there are still some uncolored vertices Aubrey returns to step 3.
8. He terminates when he gets here, i.e., when either all vertices are colored or he has backtracked all the way down to the vertices of *H*.

Aubrey implemented this algorithm in Mathematica 11 on a standard MacBook Air; it terminated in only a few minutes for de Grey's candidate *M*, without finding a 4-coloring starting from either of the triple-containing colorings of its central *H*. Essentially, the speed increases because the fixing of only 20 or so vertex's colors at step 3 typically lets almost all remaining colors be forced at step 4.

52.9 Identification of Smaller Solutions

No one knows the order of the smallest non-4-colourable unit-distance graph. The smaller the size the better are the chances to use standard algorithms (SAT solvers for example) instead of inventing custom software. [We can dream of a hand-verification if the order of a graph shrinks substantially and symmetries provide further aid – A.S.]

The most direct way to find smaller graphs is to seek a succession of small simplifications of *N*. Many approaches to this are evident, some much more computationally tractable than others. Aubrey employed only a small range of strategies that take advantage of the stepwise method by which *N* was constructed: he identified vertices of subgraphs such as *M*, whose removal preserves the property required of them in the construction and thus also the chromatic number of *N*, and then he sought new vertices, one at a time, whose addition allowed the removal of more than one pre-existing vertex.

52.10 Status of Shrinking the Graph N

De Grey has shrunk N by a factor of nearly 13, to the 1581-vertex graph G that (more for artistic than expository reasons) is shown in Fig. 52.9. It can be constructed as follows:

1. Let S be the following set of points:

$$(0,0), (1/3,0), (1,0), (2,0), \left(\left(\sqrt{33}-3\right)/6, 0\right), \left(1/2, 1/\sqrt{12}\right), \left(1, 1/\sqrt{3}\right), \left(3/2, \sqrt{3}/2\right), \left(7/6, \sqrt{11}/6\right),$$
$$\left(1/6, \left(\sqrt{12}-\sqrt{11}\right)/6\right), \left(5/6, \left(\sqrt{12}-\sqrt{11}\right)/6\right), \left(2/3, \left(\sqrt{11}-\sqrt{3}\right)/6\right), \left(2/3, \left(3\sqrt{3}-\sqrt{11}\right)/6\right),$$
$$\left(\sqrt{33}/6, 1/\sqrt{12}\right), \left(\left(\sqrt{33}+3\right)/6, 1/\sqrt{3}\right), \left(\left(\sqrt{33}+1\right)/6, \left(3\sqrt{3}-\sqrt{11}\right)/6\right), \left(\left(\sqrt{33}-1\right)/6, \left(3\sqrt{3}-\sqrt{11}\right)/6\right),$$
$$\left(\left(\sqrt{33}+1\right)/6, \left(\sqrt{11}-\sqrt{3}\right)/6\right), \left(\left(\sqrt{33}-1\right)/6, \left(\sqrt{11}-\sqrt{3}\right)/6\right), \left(\left(\sqrt{33}-2\right)/6, \left(2\sqrt{3}-\sqrt{11}\right)/6\right),$$
$$\left(\left(\sqrt{33}-4\right)/6, \left(2\sqrt{3}-\sqrt{11}\right)/6\right), \left(\left(\sqrt{33}+13\right)/12, \left(\sqrt{11}-\sqrt{3}\right)/12\right), \left(\left(\sqrt{33}+11\right)/12, \left(\sqrt{3}+\sqrt{11}\right)/12\right),$$
$$\left(\left(\sqrt{33}+9\right)/12, \left(\sqrt{11}-\sqrt{3}\right)/4\right), \left(\left(\sqrt{33}+9\right)/12, \left(3\sqrt{3}+\sqrt{11}\right)/12\right), \left(\left(\sqrt{33}+7\right)/12, \left(\sqrt{3}+\sqrt{11}\right)/12\right),$$
$$\left(\left(\sqrt{33}+7\right)/12, \left(3\sqrt{3}-\sqrt{11}\right)/12\right), \left(\left(\sqrt{33}+5\right)/12, \left(5\sqrt{3}-\sqrt{11}\right)/12\right), \left(\left(\sqrt{33}+5\right)/12, \left(\sqrt{11}-\sqrt{3}\right)/12\right),$$
$$\left(\left(\sqrt{33}+3\right)/12, \left(3\sqrt{11}-5\sqrt{3}\right)/12\right), \left(\left(\sqrt{33}+3\right)/12, \left(\sqrt{3}+\sqrt{11}\right)/12\right), \left(\left(\sqrt{33}+3\right)/12, \left(3\sqrt{3}-\sqrt{11}\right)/12\right),$$
$$\left(\left(\sqrt{33}+1\right)/12, \left(\sqrt{11}-\sqrt{3}\right)/12\right), \left(\left(\sqrt{33}-1\right)/12, \left(3\sqrt{3}-\sqrt{11}\right)/12\right), \left(\left(\sqrt{33}-3\right)/12, \left(\sqrt{11}-\sqrt{3}\right)/12\right),$$
$$\left(\left(15-\sqrt{33}\right)/12, \left(\sqrt{11}-\sqrt{3}\right)/4\right), \left(\left(15-\sqrt{33}\right)/12, \left(7\sqrt{3}-3\sqrt{11}\right)/12\right), \left(\left(13-\sqrt{33}\right)/12, \left(3\sqrt{3}-\sqrt{11}\right)/12\right),$$
$$\left(\left(11-\sqrt{33}\right)/12, \left(\sqrt{11}-\sqrt{3}\right)/12\right)$$

2. Let S_a be the unit-distance graph whose vertices consist of all points obtained by rotating the points in S around the origin by multiples of 60 degrees and/or by negating their y-coordinates. S_a has 397 vertices.
3. Let S_b be S_a rotated counterclockwise about the origin by $2{\cdot}\arcsin(1/4)$.
4. Let Y be the union of S_a and S_b with the vertices $(1/3, 0)$ and $(-1/3, 0)$ deleted.
5. Rotate Y counterclockwise about $(-2,0)$ by $\pi/2 + \arcsin(1/8)$ to produce Y_a.
6. Rotate Y counterclockwise about $(-2,0)$ by $\pi/2 - \arcsin(1/8)$ to produce Y_b.
7. Let G be the union of Y_a and Y_b.

The graph G has turned out to be within the reach of standard SAT solvers, with which others have now confirmed its chromatic number to be 5 without the need to resort to using custom code or checking weaker properties of subgraphs.

De Grey believed it was highly likely that examples smaller than G existed. Indeed, a Polymath project [DHJP] has been created to seek such graphs, as well as to seek ones whose lack of a 4-coloring can be shown without a computer. Concise and explicit descriptions of certain 5-chromatic unit-distance graphs would seem to be a promising way to attack the question of whether 6-chromatic examples exist.

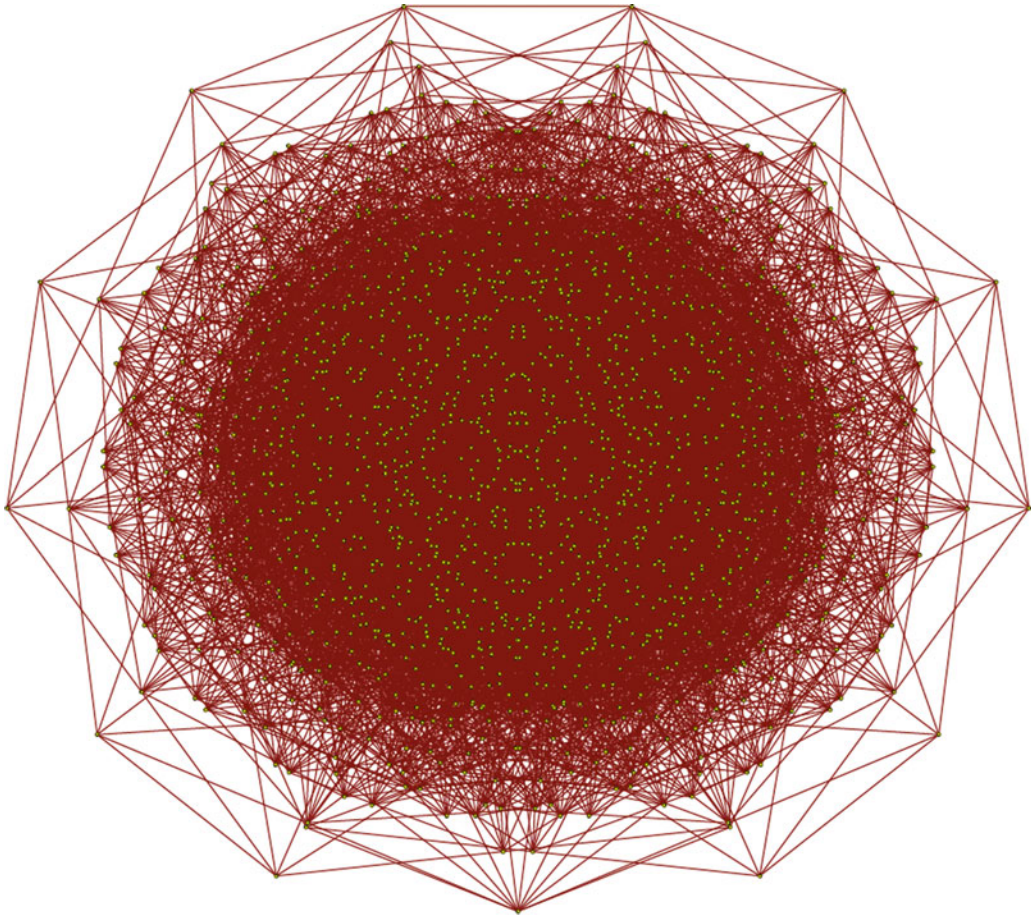


Fig. 52.9 The 1581-vertex 5-chromatic unit-distance *De Grey Graph G*

52.11 Reception of de Grey's Breakthrough

Aubrey de Grey's result answered in the negative May 4, 2002, $1000 problem of Ronald L. Graham, who asked whether it was possible to 4-color the plane to forbid a monochromatic unit distance [Soi44].

In 2018, the British gerontologist Aubrey de Grey achieved the first in 68 years breakthrough in the chromatic number of the plane problem *in general case*, the problem that essentially goes through this entire book. In July 2018, Aubrey asked Ron Graham whether he won a prize. Ron in turn asked me how much to pay Aubrey. I proposed a $1000 prize and asked them both for a photo of Aubrey receiving a check from Ron, to be published in the new expanded edition of *The Mathematical Coloring Book*. On September 9, 2018, Ron replied:

Hi Sasha,

I think your suggestion is good. As a matter of fact, I live in Berkeley some of the time, so it would not be difficult to have a photo of me presenting Aubrey with a check. This might take a month or two before I am in Berkeley again.

What do you think?

Ron

The problem creator and the problem solver met in San Diego, where Aubrey de Grey received the $1000 prize from the hands of Ronald L. Graham. On my request, they captured this landmark event, and thus you can, in a sense, participate in it. Both Ron and Aubrey sent me the photos of the check presentation.

Ronald L. Graham presents Aubrey D.N.J. de Grey the Prize: $1000, San Diego, September 22, 2018

"I will certainly be adhering to the convention of framing the check rather than cashing it," wrote Aubrey to Ron and me.

Aubrey de Grey was awarded the David Robbins Prize by the Mathematical Association of America, at its January 2020 meeting in Denver, Colorado, for

"The Chromatic Number of the Plane is at least 5"
"*Geombinatorics*, XXVIII(1), (2018), 18–31"

Having received the prize on January 16, 2020, Aubrey left the meeting before the banquet – he headed to Colorado Springs to spend a couple of days with me to discuss the composition of this new edition of *The Mathematical Coloring Book*. During our productive meeting, I seemed to recall that Aubrey used an electronic copy of the original *Mathematical Coloring Book.* And so, I presented Aubrey the prize of the inscribed for him hard copy of "The Mathematical Coloring Book." As you can see, it took place in the Italian pizzeria *Il Vicino.* The pizza was quite good too.

Aubrey de Grey and Alexander Soifer, *Il Vicino*, January 18, 2020

52.12 Aubrey D.N.J. de Grey

Aubrey David Nicholas Jasper de Grey

My request to write his life story, Aubrey de Grey fulfilled on August 22, 2018. The rest of this section is his.

Growing up in London, I was always reasonably good at mathematics but never exceptional. My journey to graph theory can be said to have begun in 1978, at the age of 15, when I was introduced to the board game Othello; My interest in it waned rather soon, but it was revived at Cambridge when I became friends with some of the top players in the UK, who were mathematicians mostly in the year above me. Two of them, Imre Leader and Graham Brightwell, became doctoral students of the distinguished Erdős protege Bela Bollobás, so it was inevitable that I would be exposed to graph theory and combinatorics – and it was to my and everyone's surprise that I turned out to be quite good at it, sometimes solving what were thought to be really quite hard problems more rapidly than my friends.

The result has been that I have dabbled with these fields, purely recreationally, through my whole adult life. Sometimes I have focused on well-known open problems, but definitely never with any hope of making real progress on them: my enjoyment derives merely from the process of gaining a rudimentary understanding of why they are so hard.

It remains a mystery to me why I did not encounter the celebrated question of the chromatic number of the plane decades ago. It was certainly known to my friends; I suppose somehow it just never came up in conversation. Even more strangely, I am also unable to work out precisely how I stumbled upon it in 2017 – or even precisely when, though it was some time between July and September. What I do know, however, is what led me to be captivated by the problem: as for so many before me, it was the discovery of the first edition of the book whose second edition you are now reading. All fields need great expositors if they are to attract the interest of a wider community, and as a scientist who successfully switched fields mid-career I am well-placed to see the benefit that a field gains from this accessibility.

For a few months I barely looked at the CNP question, because of pressure of my day job. However, I was able to devote much of the Christmas break to it, and was led uncannily rapidly to my successful approach by Alexander's descriptions of the failed attempts of the past, especially that of Paul O'Donnell. The idea of making a unit-distance graph non-rigid, and thus amenable to deformations to introduce new edges, seemed so attractive - and yet it apparently could not succeed. Why? Evidently the flexibility thus introduced was outweighed by the competing rise in the number of ways to four-colour such graphs. Well, I reasoned, in that case let's just try going in the opposite direction: accept rigidity, and indeed introduce as high a density as possible of rigid 4-chromatic graphs with shared vertices. From there to the solution took only a few weeks. I actually put the final pieces together (to construct the 20,425-vertex graph referred to in my eventual paper) in a pub a few miles from my home, to which I had retreated because of a power outage. Someone else at the bar saw some of the diagrams I was playing with and asked as to their relevance. . . I never caught his name, but he was the first person to know that this 68-year-old problem had finally seen some progress.

How soon will we reduce the range of CNP uncertainty further? I will be astonished if the next step forward is other than the creation of a 6-chromatic UDG, eliminating the possibility that CNP $= 5$ – but I suspect that it may be some years before this is achieved. Graphs not much larger than the 5-chromatic ones discovered by me and others are computationally intractable, and I will not be at all surprised if the smallest 6-chromatic UDG contains millions of vertices.

52.13 The Effect of the Breakthrough on Predictions of Many

"DIMACS" stands for the Center for Discrete Mathematics and Theoretical Computer Science. It was founded in 1989 by Princeton and Rutgers Universities, AT&T, and Bellcore, with other research institutions joining later. It is located on the Piscataway campus of Rutgers University. I spent 3 enjoyable years there as a Long-Term Visiting Scholar, concurrently with my research years at Princeton. My sponsors were Saharon Shelah and Fred Roberts. After I left both places in 2007, DIMACS' Executive Director Fred Roberts sent a mass email asking for events proposals for DIMACS. I replied concisely: "Ramsey Theory." Unexpectedly, Fred replied "DIMACS Executive Committee is interested in your proposal. Please, elaborate. We want a non-generic, original view of the field." "Ramsey Theory Yesterday, Today, and Tomorrow," replied I: "Day One 'Yesterday,' Day Two 'Today,' and Day

3 'Tomorrow.'" Fred and his executives called on me to organize and run this 3-day international workshop. I accepted, contingent to Ron Graham accepting my invitation to be one of the plenary speakers. Ron did, and I was in business.

On May 29, 2009, "Day Three: Tomorrow,"[1] I got an additional confirmation of the great influence of Khinchin's book. Leaders of Ramsey Theory and my plenary speakers Ronald L. Graham and Joel H. Spencer told me that this Khinchin's little book introduced them both, for the first time, to the name of Van der Waerden, his theorem, and Ramseyan ideas!

On May 28, 2009, "Day Two: Today," in the middle of my plenary talk on the chromatic number of the plane,[2] given 11:15–12:15, I asked the distinguished audience of leaders of Ramsey Theory and most talented graduate students, 30 participants in all, to determine the chromatic number of the plane by democratic means of a vote. The voters possessed a great intuition and a formidable wealth of knowledge. They included Ronald L. Graham, Joel H. Spencer, Jaroslav Nešetřil, Stanisław P. Radziszowski, Peter D. Johnson, Jr., András Gyárfás, William Gasarch, Vera Rosta, Jacob Fox, Dmytro Karabash, Colton Magnant, Bert Randerath, Marcia J. Groszek, Andrzej Dudek (joint work with Peter Frankl, and Vojtech Rödl), Vadim V. Lozin, Lynn Scow, Brian Hopkins, and others. One young Lady (Lynn Scow from the University of California Berkeley, if memory holds) voted for 6; I voted for 7; the rest of the workshop participants equally split between 4 (including Peter D. Johnson Jr. and Dmytro Karabash) and 5 (including Ronald L. Graham). I was therefore able to announce:

The democratic value of the chromatic number χ of the plane is 4.5.

Laugh or cry, but these lower values were a dominant expectation of the value of χ.

In 2018, Aubrey de Grey's Graph encouraged others to try their hand (or mind). A number of people created a number of 5-chromatic unit-distance graphs. This excitement propelled a hope and desire to quickly construct a 6-chromatic unit-distance graph. However, soon came a realization: not so fast. Many colleagues – perhaps, most – now expect the final answer to be 7, as I conjectured in 2002. Those who tried, however, do not expect the arrival of a 7-chromatic unit-distance graph during our lifetimes. So, either Aubrey de Grey will teach us how to prolong a meaningful life (Gerontology is his main field) or we will have to rely on future generations, who no doubt will have faster computers at their disposal and, perhaps, new bright ideas.

Marijn Heule, a Dutch computer scientist, working in the United States, enters the scene next.

[1]Center for Discrete Mathematics and Theoretical Computer Science, a joint project of several research institutions, including Princeton and Rutgers Universities. See [Soi49] for the texts of the plenary talks of this workshop.

[2]Published in [Soi49].

Chapter 53
Marienus Johannes Hendrikus "Marijn" Heule

53.1 The Records

Marijn Heule, a virtuoso computer scientist, who does not always rely on existing software but rather creates his own. We, mathematicians, are fortunate that he became excited about CNP problem, for he has produced a series of 7 world records for the smallest (in vertex count) known 5-chromatic unit-distance graphs. Let us document the succession of his seven 5-chromatic unit-distance graph records and then have Marijn present one of them in detail.

874 vertices and 4461 edges on April 14, 2018
826 vertices and 4273 edges on April 16, 2018
803 vertices and 4144 edges on April 30, 2018
633 vertices and 3166 edges on May 6, 2018
610 vertices and 3000 edges on May 14, 2018
553 vertices and 2722 edges on May 18, 2018

More world records were achieved independently by two researchers between July 2019 and March 2020 – we will address them later.

Rolling back in the spring-early summer 2018, I envisioned the Special Issue of *Geombinatorics* XXVIII(1), 2018, dedicated exclusively to the progress in the chromatic number of the plane problem. In it, my summary essay is followed by pioneering papers by Aubrey de Grey, which you have just seen, then the "553" record paper [Heu3] by Marijn Heule (Aubrey and Marijn were very pleased to appear in the same issue), and concluded with the paper by Geoffrey Exoo and Dan Ismailescu [EI2].

Marijn Heule presents his computer science approach that allows him to reduce the world record several times to a 553-vertex 5-chromatic unit-distance graph. It [Heu1] appears here in Marijn's words, slightly abridged. (The record was later reduced several times by Marijn Heule and Jaan Parts – more about it later.)

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_53

53.2 The Plan

Our method exploits two formal methods technologies: the ability of *satisfiability* (*SAT) solvers* to find a short refutation for unsatisfiable formulas (if they exist) and *proof checkers* that can minimize refutations and unsatisfiable formulas.

The refutations emitted by SAT solvers are hardly minimal. Depending on the application from which the formula originates, typically 10% to 99% of the refutation can be omitted. Several techniques have been developed to avoid checking irrelevant parts of a refutation [HHW]. These techniques minimize proofs in order to share and revalidate them. For example, the proofs of the Boolean Pythagorean Triples [HKM] and Schur Number Five [Heu1], [Heu2], problems are enormous, even after minimization: 200 terabytes and 2 petabytes, respectively.

Here we use clausal-proof-minimization techniques for a different purpose: shrinking graphs. Given a unit-distance graph with chromatic number 5, we first construct a propositional formula that encodes whether there exists a valid 4-coloring of this graph. This formula is unsatisfiable, and we can use a SAT solver to compute a refutation. From the minimized refutation, we extract a subgraph that also has chromatic number 5. We then apply this process repeatedly to make the graph ever smaller.

53.3 A Few Definitions

We will use three operations to construct larger and larger graphs: the Minkowski sum, rotation, and merge. Given two sets of points A and B, the *Minkowski sum* of A and B, denoted by $A \oplus B$, equals $\{a + b | a \in A, b \in B\}$.

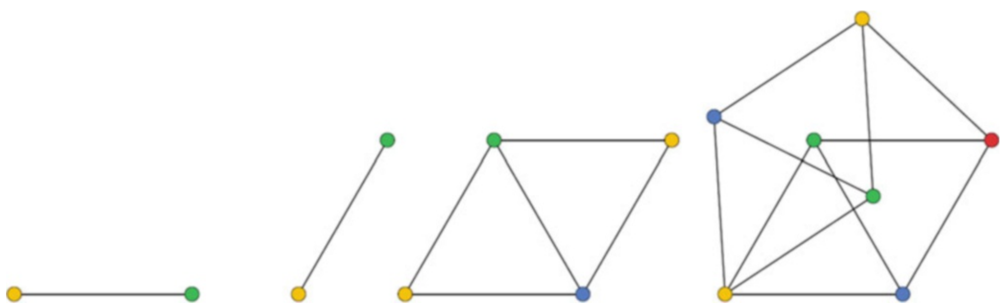


Fig. 53.1 From left to right: illustrations of A, B, $A \bigoplus B$, and the Moser Spindle. The graphs shown have chromatic number 2, 2, 3, and 4, respectively

Given a positive integer i, we denote by θ_i the rotation around point (0, 0) by the angle $\arccos\left(\frac{2i-1}{2i}\right)$ and by θ_i^k, which is the application of θ_i k times. Let p be a point at the distance $\sqrt{i}$ from (0, 0), then the points p and $\theta_i(p)$ are exactly distance 1 apart and thus would be connected with an edge in a unit-distance graph.

Consider again the set of points $A \oplus B$ above. The points $A \oplus B \cup \theta_3(A \oplus B)$ form the Mosers Spindle. Figure 53.1 shows visualizations of these sets with connected vertices colored differently.

53.4 Propositional Formulas

We will minimize graphs on the propositional level. We consider propositional formulas in *conjunctive normal form* (CNF), which are defined as follows. A *literal* is either a variable x (*a positive literal*) or the negation $\bar{x}$ of a variable x (*a negative literal*). The *complement* $\bar{l}$ of the literal l is defines as $\bar{l} = \bar{x}$ if $l = x$ and $\bar{l} = x$ if $l = \bar{x}$. For a literal l, $var(l)$ denotes the *variable* of l. A *clause* is a disjunction of literals, and a *formula* is a conjunction of clauses.

An *assignment* is a function from a set of variables to the truth values 1 (*true*) and 0 (*false*). A literal l is *satisfied* by an assignment α if l is positive and $\alpha(var(l)) = 1$ or if it is negative and $\alpha(var(l)) = 0$. A literal is *falsified* by an assignment if its complement is satisfied by the assignment. A *close is satisfied* by α if it contains a literal that is satisfied by α. Finally, a *formula is satisfied* by an assignment α if all its clauses are satisfied by α. A *formula is satisfiable* if there exists an assignment that satisfies it and otherwise it is *unsatisfiable.* Two formulas are *logically equivalent* if they are satisfied by the same assignments; they are *satisfiability equivalent* if they are either both satisfiable or both unsatisfiable.

53.5 Clausal Proofs

Here we introduce a formal notion of clause redundancy. A *clause C is redundant with respect to a formula F* if F and $F \wedge C$ are satisfiability equivalent. For instance, the clause $C = x \vee y$ is redundant with respect to the formula $F = (\bar{x} \vee \bar{y})$ since F and $F \wedge C$ are satisfiability equivalent (although they are not logically equivalent). This redundancy notion allows us to add redundant clauses to a formula without affecting its satisfiability.

Given a formula $F = \{C_1, \ldots, C_m\}$, a *clausal derivation* of a clause C_n from F is a sequence $C_{m+1}, \ldots, C_n$ of clauses. Such a sequence gives rise to formulas $F_m, F_{m+1}, \ldots, F_n$, where $F_i = \{C_1, \ldots, C_i\}$. We call F_i the *accumulated formula* corresponding to the ith proof step. A clausal derivation is *correct* if every clause C_i $(i > m)$ is redundant with respect to the formula F_{i-1} and if this redundancy can be checked in polynomial time with respect to the size of the proof. A clausal derivation is a *proof* of a formula F if it derives the unsatisfiable empty clause. Clearly, since every clause-addition step preserves satisfiability, and since the empty clause is always false, a proof of F certifies the unsatisfiability of F. The proofs computed in this paper show that the chromatic number of a given graph is at least 5. We will also refer to proofs as *refutations* as they refute the existence of a valid 4-coloring.

53.6 Clausal Proof Minimization

SAT solving techniques are not only useful to validate the chromatic number of a graph but they can also help reduce the size of the graph while preserving the chromatic number. The method works as follows. Given a graph G with chromatic number k, first generate the propositional formula F that encodes whether the graph can be colored with $k - 1$ colors. This formula is unsatisfiable. Most SAT solvers can emit a proof of unsatisfiability. There exist several checkers for such proofs, even checkers that are formally verified in the theorem provers ACL2, Coq, and Isabelle. We used the (unverified) checker DRAT-trim [HHW] that allows minimizing the clausal proof as well as extracting an unsatisfiable core, i.e., a subformula that is also unsatisfiable. From the unsatisfiable core, one can easily extract a subgraph $G^{'}$ of G such that $G^{'}$ also has chromatic number k.

53.7 Encoding

We can compute the chromatic number of a graph G as follows. Construct two formulas, one asking whether G can be colored with $k – 1$ colors and one whether G can be colored with k colors. Now, G has chromatic number k if and only if the former is unsatisfiable while the latter is satisfiable.

The construction of these two formulas can be achieved using the following encoding. Given a graph $G = (V, E)$ and a parameter k, the encoding uses $k|V|$ Boolean variables $x_{v,c}$ with $v \in V$ and $c \in \{1,..., k\}$. These variables have the following meaning: $x_{v,c}$ is true if and only if vertex v has color c. Now we can encode whether G can be colored with k colors:

$$G_k := \bigwedge_{v \in V} (x_{v,1} \vee \ldots \vee x_{v,k}) \wedge \bigwedge_{\{v,w\} \in E} \bigwedge_{c \in \{1,\ldots,k\}} (\overline{x}_{v,c} \vee \overline{x}_{w,c})$$

The first type of clauses ensures that each vertex has at least one color, while the second type of clauses forces that two connected vertices are colored differently. Additionally, we could include clauses to require that each vertex has at most one color. However, these clauses are redundant and would be eliminated by blocked clause elimination, a SAT preprocessing technique.

We added symmetry-breaking predicates during all experiments to speed-up solving and proof minimization. The color symmetries were broken by fixing the vertex at (0, 0) to the first color, the vertex at (1, 0) to the second color, and the vertex at $\left(1/2, \sqrt{3}/2\right)$ to the third color. These three points are at the distance 1 from each other and occurred in all our graphs. The speedup is roughly a factor of 24 $(4 \cdot 3 \cdot 2)$, when trying to find a 4-coloring.

We did not explore yet whether an encoding based on Alexander A. Zykov’s contraction [Zyk1] would allow shorter proofs of unsatisfiability. In essence, such an encoding would add variables and clauses that encode for a pair of vertices whether they have the same color. Solving graph coloring problems using such an extended encoding has been successful in the past.

53.8 Graph Trimming

Modern SAT solvers can emit clausal proofs. We used the SAT solver Glucose to produce the proofs. The most commonly supported format for clausal proofs is DRAT, which computes the redundancy of clauses using the resolution asymmetric tautology check. Some DRAT proof checkers can extract from a refutation an unsatisfiable core, i.e., a subformula that is still unsatisfiable. When the formula expresses a graph coloring property, the unsatisfiable core represents a subgraph with the same coloring property. The absence of the clause $(x_{v,1} \vee \cdots \vee x_{v,k})$ in the core shows that vertex v can be removed, while the absence of all clauses $\left(\bar{x}_{y,c} \vee \bar{x}_{w,c}\right)$ with $c \in \{1, \ldots, k\}$ shows that edge $\{v, w\}$ can be removed. When trying to find a small unit-distance graph with a given chromatic number, we are interested in reducing the number of vertices. Although the proof checker can be easily modified to ensure that no edges are removed, we achieved larger reductions by allowing edges to be deleted and then restoring edges between vertices that survived the shrinking.

53.9 Randomization

SAT solvers and clausal-proof-minimization tools are deterministic. To increase the probability of finding small unit-distance graphs with chromatic number 5, we want to randomize the process and minimize many clausal proofs.

The proofs produced by SAT solvers depend heavily on the ordering of the clauses in the input file. The initial heuristic ordering of the variables is based on their occurrence in the input file. The earlier a variable occurs in the input file, the higher its place in the ordering. Although more sophisticated initialization methods have been proposed, this method is effective in practice. The effectiveness is caused by the typical encoding of a problem into propositional logic where one starts with the more important variables. However, for our application there are no clear important variables.

Based on these observations, we applied the following lightweight randomization. First, we shuffle the input formula and apply graph trimming on the result. When the clausal-proof-minimization tool is no longer able to remove vertices from the graph, we shuffle the clauses of the current formula and produce a new clausal proof. Then we continue graph trimming using the new formula and proof. This process is repeated until randomization cannot further reduce the size of the graph.

53.10 Critical Graphs

A graph is vertex/edge *critical* with respect to a given property if removing any vertex/edge would break that property. Here we are interested in vertex critical graphs with respect to the chromatic number. Graph trimming as described above would remove most redundant vertices of the graph and the randomization method allows shrinking the graph even further. However, in most cases, the reduced graphs are not critical: There still exist some vertices that can be removed while preserving the chromatic number.

Both the SAT solver and the clausal-proof-minimization tool aim to find a relatively short argument (i.e., clausal proof) explaining why to imply the fewest number of involved vertices. In fact, an argument can frequently be shortened by using redundant (non-critical) vertices.

In the final step, we therefore make the graph critical by the following procedure. Randomly pick a vertex from the graph and determine the chromatic number of the graph without it. If the chromatic number is not changed, then the vertex is removed from the graph. This process is repeated until all remaining vertices have been determined to be critical.

Instead of using this naive method to make the graph critical, we could have used more sophisticated tools that compute a minimal unsatisfiable core from the propositional formula. However, these tools did not improve the performance or the size in an observable way.

53.11 Validation

Determining whether a set of points forms a unit-distance graph with chromatic number 5 requires two checks: (*i*) whether the corresponding graph has chromatic number 5 and (ii) whether the distance between two connected points is exactly 1. The techniques discussed in this paper can easily perform the first check. SAT solvers can compute valid 5-coloring for the critical graphs in a fraction of a second. The proofs showing that there exists no 4-coloring are actually quite small: between 14,000 and 19,000 clause addition steps. Proofs of that size can be checked in roughly a second even with formally verified checkers. We used the DRAT-trim tool [HHW] to validate them. Proofs of recently solved hard-combinatorial problems, such as the Pythagorean Triples and Schur Number Five, are much larger: roughly 1 trillion and 10 trillion clause addition steps, respectively [HKM], [Heu2].

For the second check, we used a tool based on Gröbner basis, available at http://fmv.jku.at/dist1sqrtgb/, to validate for every edge in the graph that the corresponding points are exactly 1 apart. The tool produces files that can be validated using Singular and pactrim. There is no need to check whether all edges are present as missing edges can only decrease the chromatic number. Checking only the correctness of the edges in the graph is cheap. The total validation time for our smallest critical graphs is about a second or two.

53.12 Results

In this section, we discuss the various techniques developed to obtain small unit-distance graphs with chromatic number 5. The techniques were originally designed for verification purposes and applying them to graph minimization is novel and unexpected. The main strategy is to start with a large graph and shrink it using clausal proof minimization. We minimized several large graphs with various heuristics and most reduced graphs consisted of 800 to 900 vertices. However, we were able to produce several graphs of 553 vertices using three techniques. The first technique (see next Section 53.12.1) enabled producing graphs with less than 700 vertices consistently. Second, we obtained graphs with just over

600 vertices by shrinking merged copies of graphs with less than 700 vertices (see Section 53.12.2 below). Finally, we added some points far away from the origin in order to eliminate more points close to the origin (Section 53.12.3). The graphs and corresponding proofs mentioned in this section are available at https://github.com/marijnheule/CNP-SAT.

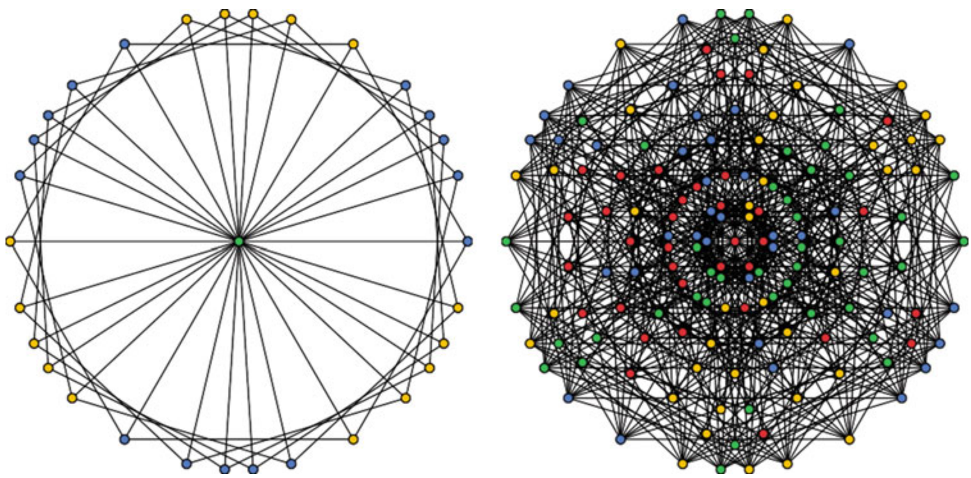


Fig. 53.2 Left, a 3-coloring of de Grey's graph V_{31}. Right, a 4-coloring of V_{151} being $V_{31} \bigoplus V_{31}$ without the vertices more than unit distance apart from the center

53.12.1 Finding a Small Symmetric Subgraph

The main building block of our graphs is de Grey's V_{31} [G1]: five 7-wheels with a common central vertex. The graph V_{31} has 31 vertices, 60 edges, is 3-colorable, and all points are in the field $Q\left(\sqrt{3}, \sqrt{11}\right)$. These points can be obtained by applying $\theta_1^j \theta_3^k$ on point (1,0) around (0,0) with $j \in \{0, 1, 2, 3, 4, 5\}$ and $k \in \left\{-1, -\frac{1}{2}, 0, \frac{1}{2}, 1\right\}$. A visualization of this graph is shown in Fig. 53.2 (left). During an early stage of the experimentation, we observe that graph $(V_{31} \oplus V_{31} \oplus V_{31}) \cup \theta_4(V_{31} \oplus V_{31} \oplus V_{31})$ has chromatic number 5. Furthermore, all points that are further away than 2 of the center can be removed without affecting the chromatic number.

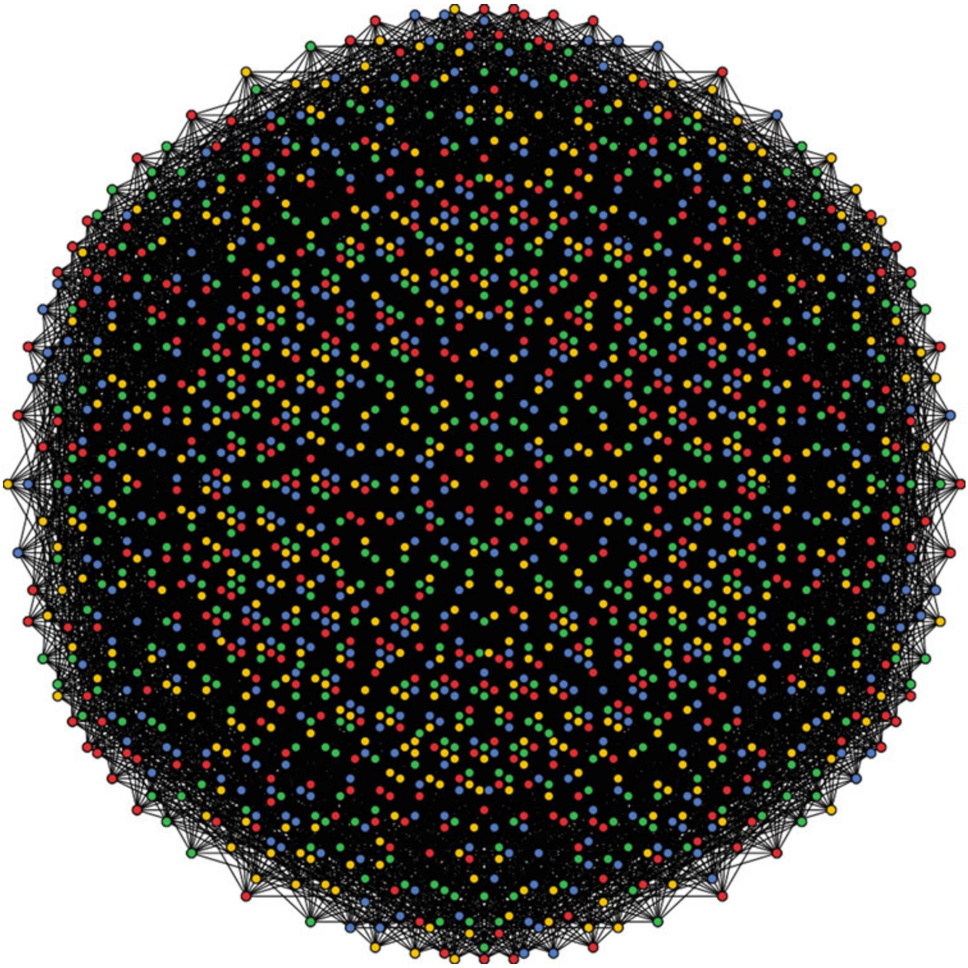


Fig. 53.3 A 4-coloring of V_{1939}, which is the Minkowski sum of V_{31} and V_{151}

Instead of removing the points at distance larger than 2 from the center, we constructed the following graph. Let V_{151} be the Minkowski sum of V_{31} and V_{31} without the points at distance larger than 1 from the center. This graph has 151 vertices and 510 edges and is shown in Fig. 53.2 (right). Now let V_{1939} be the Minkowski sum of V_{31} and V_{151}. This graph is shown in Fig. 53.3. The graph $V_{1939} \cup \theta_4(V_{1939})$ has chromatic number 5 as well.

We applied clausal proof minimization on the formula that encodes whether the graph $V_{1939} \cup \theta_4(V_{1939})$ is 4-colorable. Most random probes of clausal proof minimization produced a subgraph of $V_{1939} \cup \theta_4(V_{1939})$ with slightly more than 800 vertices. Occasionally, it produces graphs with fewer than 700 vertices, while never producing graphs in the range of 700 to 800 vertices.

Closer examination of the minimized graphs with fewer than 700 vertices revealed that only a small fraction of the points (always less than 200 vertices) are in the field $\mathbb{Q}\left[\sqrt{3}, \sqrt{11}\right]$.

These points originate from the subgraph V_{1939}, while the other points originate from the subgraph $\theta_4(V_{1939})$. Other patterns can be observed in the graphs with fewer than 700 vertices: there were at least 12 points in the field $\mathbb{Q}\left[\sqrt{3}, \sqrt{11}\right]$ at distance 2, while the graphs with more than 800 vertices had fewer than three such points. Hence, keeping the points at distance 2 appears crucial to find smaller graphs. Rotation θ_4 does not only add edges between points at distance 2 (by construction) but also between points at other distances. In fact, half the edges between points in V_{1939} and $\theta_4(V_{1939})$ are due to points that are closer to the center: i.e., at $\frac{\sqrt{33}+1}{2\sqrt{3}}$ and $\frac{\sqrt{33}-1}{2\sqrt{3}}$ from the origin. Figure 53.4 shows the newly introduced edges to θ_4.

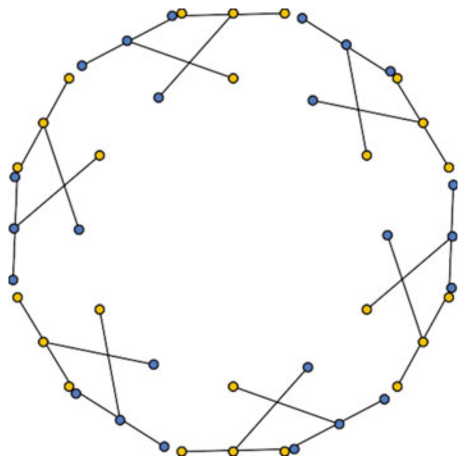


Fig. 53.4 A visualization of the edges between $S_{199} \cup \theta_4(S_{199})$ and the involved vertices. Vertices originating from S_{199} and $\theta_4(S_{199})$ are colored yellow and blue, respectively

Visualizing the points in the field $Q\left(\sqrt{3}, \sqrt{11}\right)$ reveals that they are highly symmetric: both reflection in the horizontal axis and a rotation of $\theta_1 = 60^\circ$ map the points onto themselves. Figure 53.5 shows this visualization. Shown is a 199-vertex graph with 888 edges at unit distance, which we call S_{199}. The minimized graphs did not fully produce S_{199}, but always yielded a subgraph that missed a handful (up to a dozen) vertices in various locations. There exist many 4-colorings of S_{199}, but we observed no clear pattern.

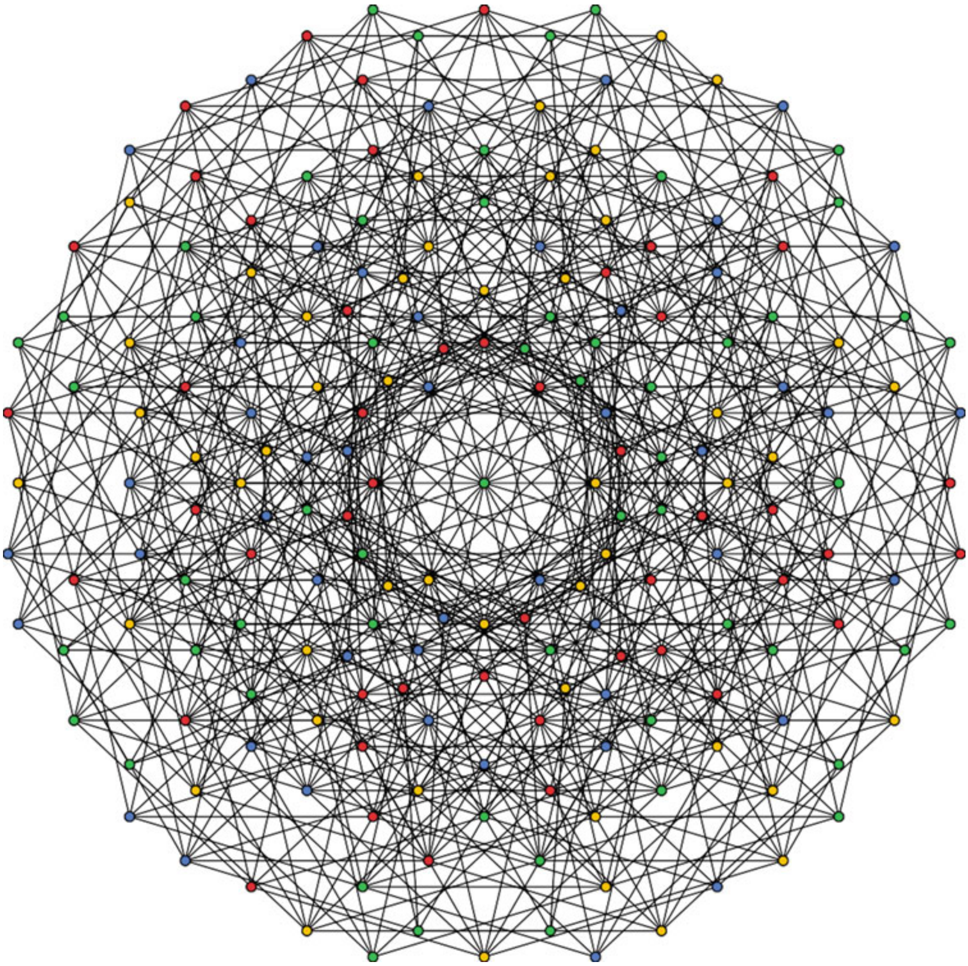


Fig. 53.5 A 4-coloring of the graph S_{199}, a symmetric subgraph of V_{1939}, that occurred in several early records

Interesting patterns emerge when merging S_{199} and $\theta_4(S_{199})$, as shown in Fig. 53.6. Notice that points that are close to each other frequently have the same color. More importantly, roughly half of the vertices that are close to distance 2 from the center have the same color as the central vertex. In later experiments, we minimized the graph $V_{1939} \cup \theta_4(S_{199})$, which allowed us to consistently produce unit-distance graphs with fewer than 700 vertices. We suspect that the above-mentioned patterns contribute to the lack of a 4-coloring of $V_{1939} \cup \theta_4(S_{199})$. Notice that V_{1939} has S_{199} as a subgraph.

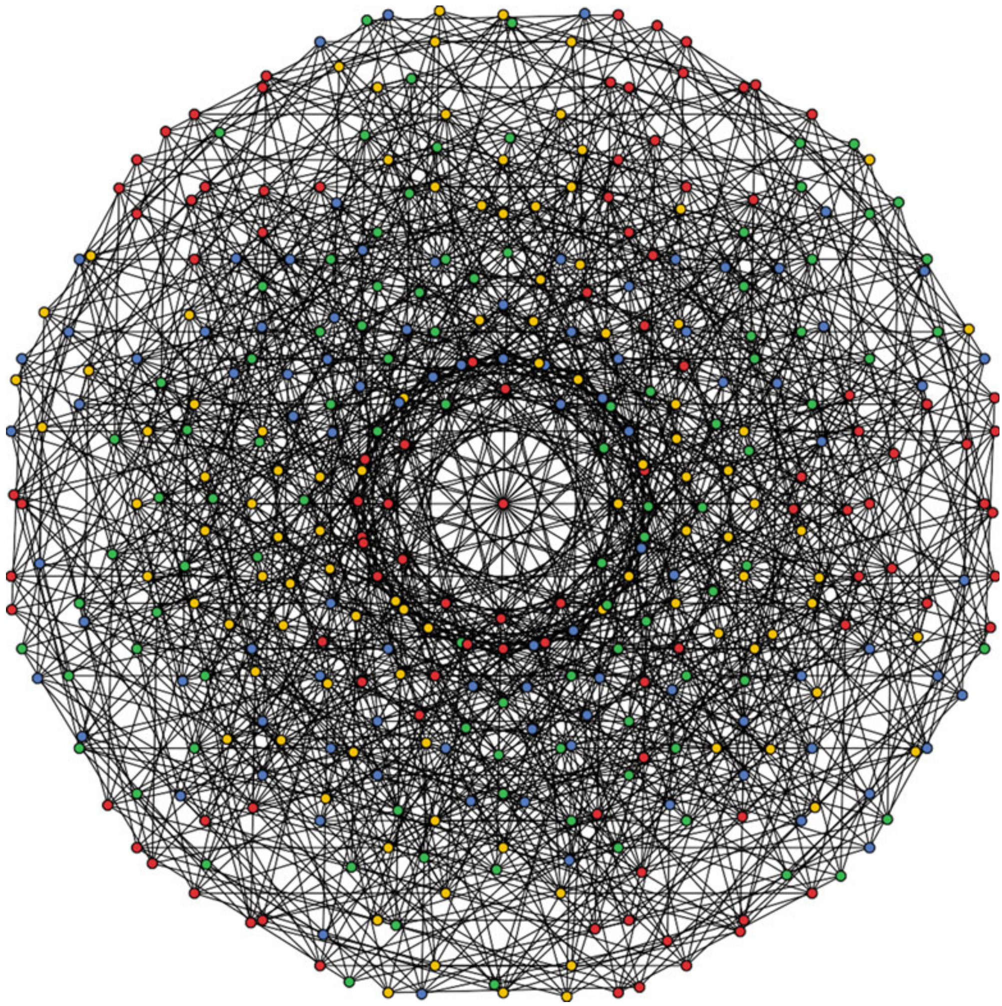


Fig. 53.6 A 4-coloring of the graph $S_{199} \cup \theta_4(S_{199})$

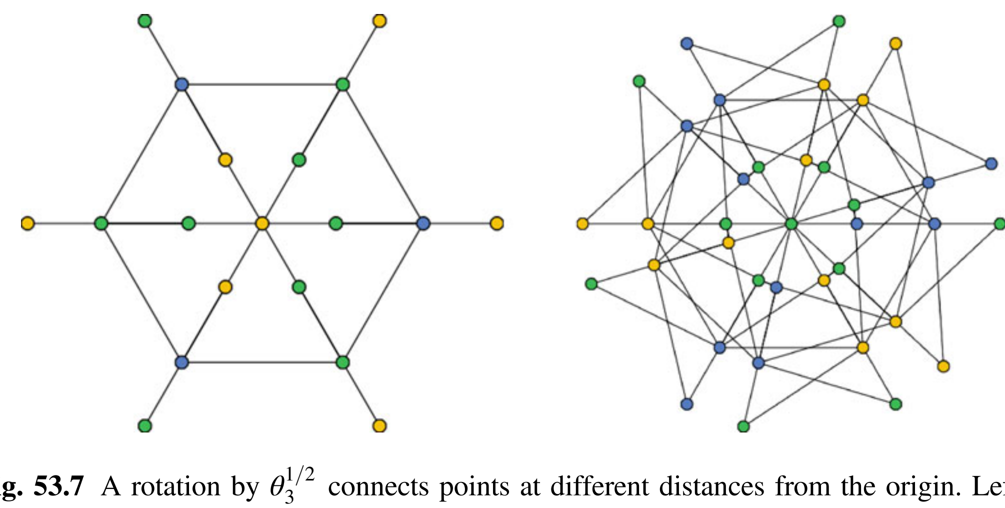


Fig. 53.7 A rotation by $\theta_3^{1/2}$ connects points at different distances from the origin. Left, a subgraph of $V_{31} \oplus V_{31}$ consisting of three 7-wheels with radii 1, $\frac{\sqrt{33}+3}{6}$, and $\frac{\sqrt{33}-3}{6}$. Right, two copies of that graph with one copy turned by this rotation

53.12.2 Merging Critical Graphs

In order to further produce smaller unit-distance graphs with chromatic number 5, we selected two critical graphs obtained earlier, merged them, and applied clausal proof minimization again. There are a significant number of options to merge two graphs and we experimented with a variety of these. The most effective merging strategy in our experiments turned out to be rotating the graphs along the central vertex in such a way that a vertex in one graph at unit distance from the center is merged with a vertex from the other graph at unit distance. Although two different critical graphs can be used for merging, we observed that it is also effective to merge two copies of the same critical graph.

The minimization procedure frequently produced a graph that was larger than both of the critical graphs that were merged. Only three kinds of rotations occasionally resulted in smaller graphs. The first two rotations are θ_1^k and θ_3^k for a small value of k. These rotations clearly increase the average vertex degree by merging vertices and introducing edges between points at distance $1\left(\theta_1^k\right)$ and $\sqrt{3}\left(\theta_3^k\right)$ from the origin. Most other rotations result in little interaction between the two critical graphs and thus hardly increase the average vertex degree. The most effective rotations introduce edges between the points at distance 1 and the distances $\frac{\sqrt{33}+3}{6}$ and $\frac{\sqrt{33}-3}{6}$ from the origin, thereby increasing the average degree significantly. Figure 53.7 illustrates this by showing three 7-wheel graphs with the radii 1, $\frac{\sqrt{33}+3}{6}$, and $\frac{\sqrt{33}-3}{6}$ (left) and two copies of this graph rotated in such a way that the points on these distances become connected (right). The graph on the left has average vertex degree $\frac{36}{19}$, while the graph of the right has average vertex degree $\frac{120}{37}$. A rotation by $\theta_3^{1/2}$ for example achieves this and maps point (1,0) onto point $\left(\frac{\sqrt{33}}{6}, \frac{\sqrt{3}}{6}\right)$. Both points are at unit distance from the origin and both are part of V_{31} and of most other graphs that we used in the experiments.

The combination of merging and minimization only introduced vertices in the field $Q\left(\sqrt{3}, \sqrt{11}\right)$. The smallest graphs contained roughly 50 vertices that do not occur in $(V_{31} \oplus V_{31} \oplus V_{31})$ and thus not in V_{1939}.

The smallest graph that we found using the techniques discussed so far contains 610 vertices and 3000 edges. This graph is shown in Fig. 53.8 using a 5-coloring in which only the central vertex has the fifth color. Recall that this graph is vertex critical. Hence, our graph possesses such a coloring in which any vertex can be the only one with the fifth color.

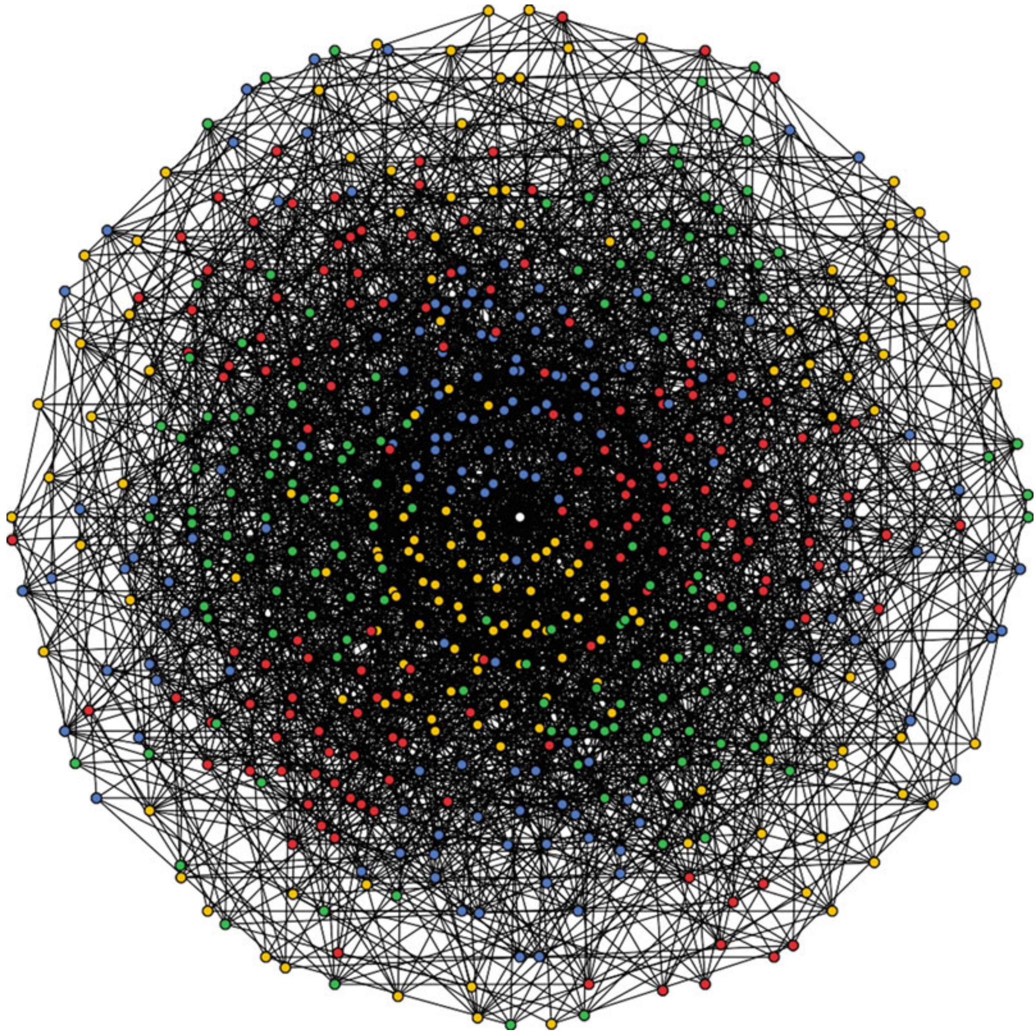


Fig. 53.8 A visualization of a 610-vertex unit-distance graph with chromatic number 5. Five colors are used for the vertices. Only the center uses the fifth color (white)

53.12.3 Minimizing the Small Part

The critical graphs found so far can be partitioned into two parts: a subgraph of $\theta_4(S_{199})$ and the subgraph induced by the remaining vertices. We refer to the former as the small part, as it consists typically of only 187 vertices, and to the latter as the large part. In all statements regarding the size of these graphs, we count the central vertex in both parts. All points in the smallest critical graphs are at a distance 2 or less from the center. Several approaches have been examined in order to find unit-distance graphs with fewer than 600 vertices. Only one approach was effective.

We focused on adding points that are further away than 2 from the origin in order to remove more inner vertices. Adding points from the field $Q(\sqrt{3}, \sqrt{11})$ may allow reducing the large part, but none of the experiments were successful. However, we were able to substantially reduce the small part using this strategy. The most effective approach was as follows. We first constructed the Minkowski sum of $\theta_4(S_{199})$ and $\theta_4(S_{199})$ and removed all points that were less or equal than 2 away from the origin. This graph consists of 2028 vertices. All points were added to the smallest critical graphs that were found in the earlier steps, followed by clausal proof minimization. This resulted in a dozen [of unit-distance] graphs with 553 vertices and (on average) 2720 edges. Figure 53.9 shows one of these graphs, which we refer to as G_{553}. Practically, all vertices that were removed during minimization originated from the small part. This part was reduced to 133 or 134 vertices. The 553-vertex graphs appear less symmetric compared to the earlier graphs. This is caused by the few vertices that are far from the origin.

The unit-distance graphs with 553 vertices are vertex critical, but not edge critical. The proofs of unsatisfiability show that many of the edge clauses can be removed without introducing a 4-coloring. Randomly removing edges until fixpoint eliminates about 270 edges (close to 10%) of these graphs.

Remarkably, all critical graphs have a handful of vertices with degree 4. If we removed such a vertex from the graph, all its four neighbors would have a different color in all valid 4-colorings. Graph S_{199} has even 12 vertices with degree 4. Reducing a 553-vertex graph to

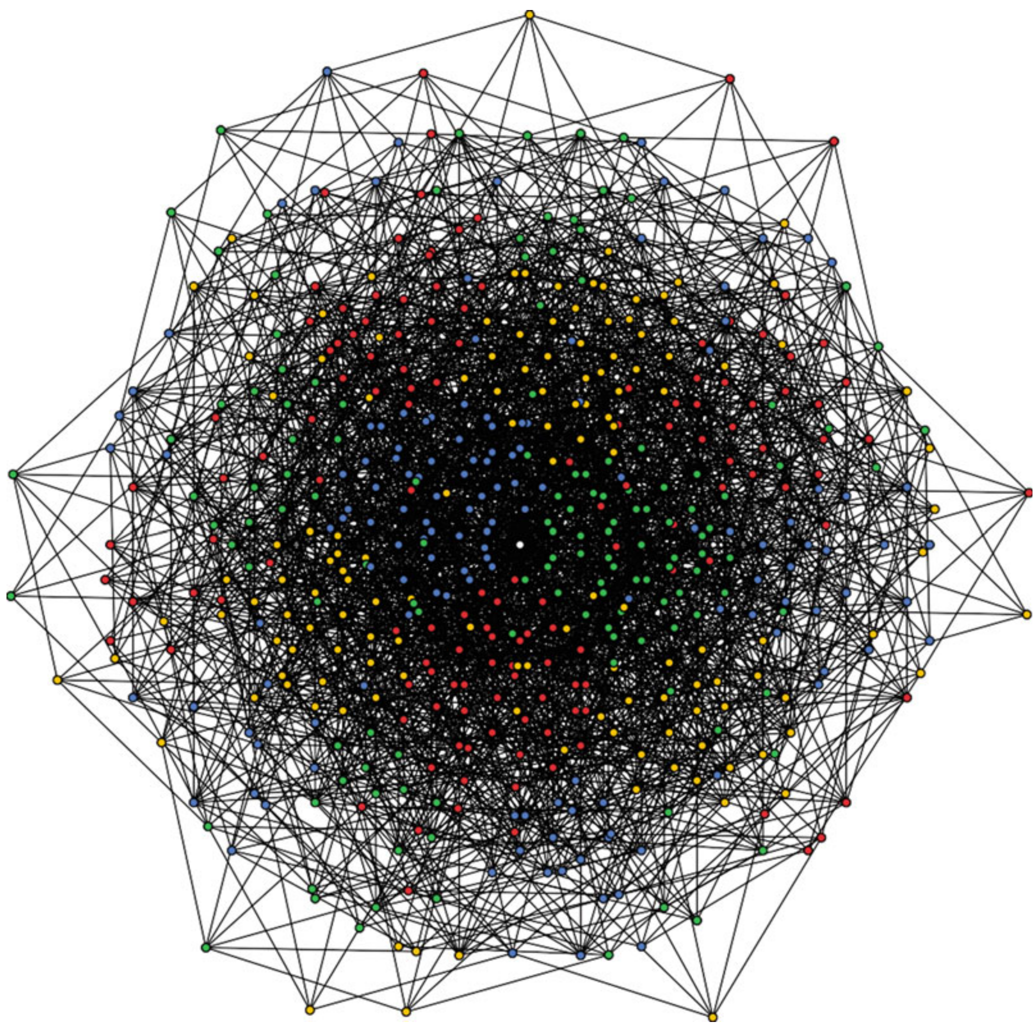


Fig. 53.9 A visualization of a 553-vertex unit-distance graph with chromatic number 5. Five colors are used for the vertices. Only the center uses the fifth color (white)

become edge critical will increase the number vertices with degree 4 to roughly 12. These vertices tend to be evenly distributed between the small and large parts of the graph.

53.12.4 Analysis

The sizes of the small and the large parts of the 553-vertex graphs suggest that they play a different role in eliminating 4-colorings. We analyzed the 4-colorings of both parts when restricted to the key vertices, i.e., the ones that connect the parts. These are the vertices at distance $\frac{\sqrt{33}-1}{2\sqrt{3}}$, $\frac{\sqrt{33}+1}{2\sqrt{3}}$ and 2 from the origin. Recall that Fig. 53.4 shows the interaction between these vertices. The 553-vertex graphs have all 24 vertices occurring in $S_{199} \cup \theta_4(S_{199})$ at distance 2 from the origin and most of the vertices at distance $\frac{\sqrt{33}-1}{2\sqrt{3}}$, $\frac{\sqrt{33}+1}{2\sqrt{3}}$ from the origin.

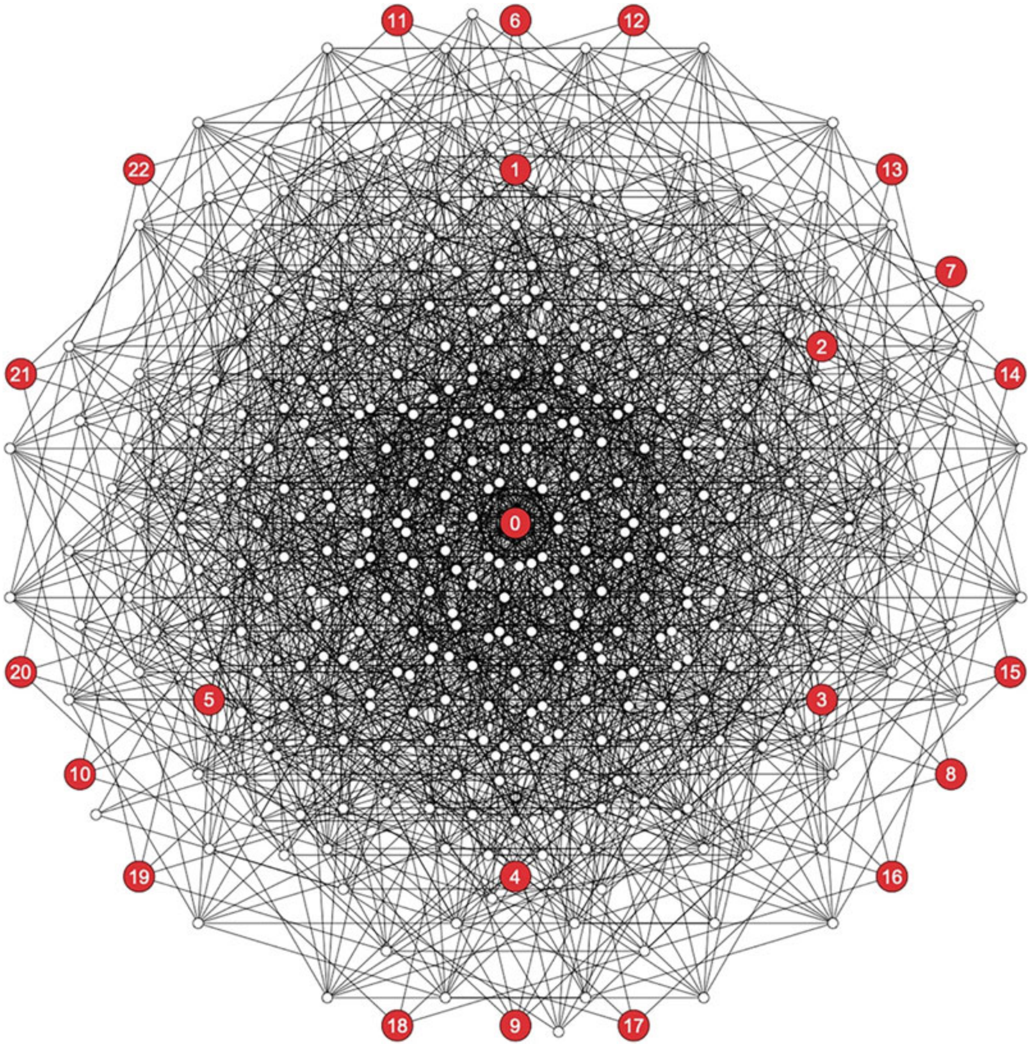


Fig. 53.10 The 420-vertex large part of G_{553} with the key vertices marked with numbers

Figure 53.10 shows the large part of G_{553} in which the central and key vertices are numbered. The other vertices in the large part significantly restrict the number of different 4-colorings of these key vertices. In fact, there are only twenty 4-colorings of these vertices such that they either have the same color as the central vertex or a different color.

53.13 Conclusions

We demonstrated that clausal proof minimization can be an effective technique to reduce the size of graphs with a given property. We used this method to shrink graphs while preserving the chromatic number. This resulted in a dozen of unit-distance graphs with chromatic number 5 consisting of 553 vertices – a reduction of over 1000 vertices compared to the smallest previously known unit-distance graph with chromatic number 5.

A main goal of this research is to obtain a human-understandable unit-distance graph with chromatic number 5. Although that goal has not been reached yet, the experiments produced some interesting results. For example, either all vertices at distance $\frac{\sqrt{33}-1}{2\sqrt{3}}$ or all vertices at distance $\frac{\sqrt{33}+1}{2\sqrt{3}}$ from the central vertex are forced to the same color as the central vertex by the large part of the minimized graphs. Also, our research produced a symmetric graph of 199 vertices that was vital for the reduction. We will study this graph in more detail to determine which properties make it so useful. Moreover, we found two rotations that connected points at multiple distances, thus increasing the average vertex degree of unit-distance graphs. Finding more such rotations may allow us to shrink the graphs even further.

Applying clausal-proof techniques to provide mathematical insights is an interesting twist in the discussion about the usefulness of mechanized mathematics. It has been argued that computers are just “ticking off possibilities” [Lam]. In this case, however, they reveal important patterns. The techniques described in this paper may actually produce the most clean and compact proof that the chromatic number of the plane is at least 5.

Finally, all graphs used in our experiments could be easily colored with 5 colors, even the ones with many thousands of vertices. However, we observed that this does not hold for the graph $(S_{199} \oplus S_{199}) \cup \theta_4(S_{199} \oplus S_{199})$. This graph is 5-colorable, even when requiring two colors for the central vertex, but computing such a coloring is expensive. Consequently, such colorings may be rare and thus may contain certain patterns. This could point to the existence of unit-distance graphs with chromatic number 6 with thousands of vertices.

53.14 Marijn Heule’s Summing-Up

On September 10, 2020, Marijn Heule shared with me a summary of his views and records related to CNP:

> My interest in the chromatic number of the plane started by an email from Scott Aaronson on April 10, 2018. We worked both at UT Austin at the time. Scott asked whether I could validate the new lower bound by Aubrey de Grey. He also mentioned

that I was well-placed to improve on the new result due to my expertise with satisfiability (SAT) technology.

It turned out that SAT is indeed effective for this problem. We can determine the chromatic number of a graph using two propositional formulas. For example, if we want to establish that a graph has chromatic number 5, then we ask whether there exists a 4-coloring of the graph (the answer should be no) and whether there exists a 5-coloring of the graph (the answer must be yes). If there is no 4-coloring, then the proof produced by the solver lists all vertices involved in the reasoning. All vertices that were not involved, can be removed while preserving the chromatic number. This method can be used to iteratively produce smaller and smaller graphs.

In the days and weeks that followed, I was able to reduce the smallest known unit-distance graphs with chromatic number 5. First down to 874 vertices (April 14), then down to 826 vertices (April 16), followed by 803 (April 30), 633 vertices (May 6), 610 vertices (May 14), and 553 vertices (May 18). Apart from refining the SAT approach, I studied the graphs to look for patterns. I used these patterns to construct large graphs as a starting point for the reduction. Over the summer of 2018, I ran the approach on the TACC cluster of UT Austin. That resulted in a unit-distance graph of 529 vertices and 2670 with chromatic number 5.

The first reduction of that graph took almost a year. Jaan Parts reported the construction of a unit-distance graph of 529 vertices and 2630 edges with chromatic number 5 on July 4, 2019. I found a unit-distance graph with 517 vertices by starting from a new large graph in late July [2019]. A week later, both Jaan [Parts] and I got down to 510 although with different methods. On March 7 [2020], Jaan was able to get it down to 509 vertices. Unit-distance graphs with chromatic number 5 with fewer than 500 vertices may exist but constructing them will likely require a different method.

I expect that the chromatic number of the plane is 7. However, that might be the answer that is the hardest to prove. Assuming that unit-distance graphs with chromatic number 7 exist, constructing one appears extremely hard. On top of that, even if we can construct a graph with chromatic number 7, there may be no technique that can prove the chromatic number.

Probably, the best result we can hope for in the foreseeable future is improving the lower bound to 6, which is already an enormous challenge. It is even not known whether there exists an odd-distance graph with chromatic number 6. In an odd-distance graph, a pair of vertices is connected if they are an odd distance apart from each other. The smallest known odd distance graph with chromatic number 5 has 21 vertices, so an order of magnitude smaller compared to the smallest known unit-distance graph with chromatic number 5. Constructing odd-distance graphs with chromatic number 6 is currently an important hurdle we need to overcome to make further progress on this intriguing problem.

Chapter 54
Can We Reach Chromatic 5 Without Mosers Spindles?

Is there a unit-distance 5-chromatic graph (in the plane) without Mosers Spindle? I was sure there was. This is a natural question that naturally came to my mind while posing relevant open problems. After all, Aubrey de Grey used a dense in spindles construction because it worked, not because it would not work otherwise. Without Mosers Spindles, we achieve much gain in flexibility of embedding, and thus, possibly, getting a smaller graph. And so, I posed the following problem:

Open Problem 54.1 (Soifer 2018). Find a 5-chromatic unit-distance graph in the plane of the smallest order that has no Mosers Spindle subgraph.

April 2020 issue of *Geombinatorics* saw the appearance of the new author, D. H. J. Polymath, who summarized his/her/their results as follows:

> We present a survey of the state of knowledge concerning the chromatic number of regions of R^2 bounded by a circle or by two parallel lines. Several of the designs we describe are hitherto unpublished and improve on previously known bounds. The authorship denotes that this work is a product of the Polymath 16 project.

Once Aubrey de Grey constructed the first 5-chromatic unit-distance graph, he proposed to create a blog, "Polymath 16" where enthusiasts could share their ideas about the problem of finding the chromatic number of the plane and many related problems.

I refer you to the original essay [Poly] in *Geombinatorics* for a systematic presentation of many new results and improvements of prior achievements, all related to circular disks and infinite strips in the plane. My interest here is in the authors' finding a strip thin enough to forbid a Mosers Spindle and wide enough to contain a 5-chromatic unit-distance graph (Fig. 54.1).

Let us look at the author(s) construction of a 5-chromatic unit-distance graph in such a strip.

– We start by seeking a 4-chromatic graph with two specific vertices that are colored the same in any 4-coloring. Given such a graph that fits into a strip of a given height, we can construct a 5-chromatic graph within the same height plus ε. We have identified such a graph in a strip of height $\frac{\sqrt{11}+7\sqrt{3}}{12} \approx 1.286748$ constructed as follows: we define 54 unit

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_54

vectors emanating from the origin, whose rotations from the (1, 0) are $\frac{m\pi}{3} + n\arccos\left(\frac{5}{6}\right)$; $m \in [0, \ldots, 5]$, $n \in [0, \ldots, 4]$.

– We keep all points that are the sum of at most 4 such vectors and construct the graph G with those points as vertices and in which an edge joins each pair of points at unit distance. We find that the vertices at (0, 0) and (8/3, 0) have the same color in any 4-coloring of G.
– We discard all vertices whose y-coordinate is outside the range $\left[-\frac{\sqrt{11}+7\sqrt{3}}{12}, \frac{\sqrt{3}}{2}\right]$, thus leaving 61,216 vertices.

We observe that the vertices at (0, 0) and (8/3 , 0) still have the same color in any 4-coloring of this graph. Clearly, even a high-resolution picture of a graph with that many vertices and edges will fill the entire area and thus will not be a good illustration of the construction.

Thus, we get a 5-chromatic unit-distance graph without a Mosers Spindle. ■

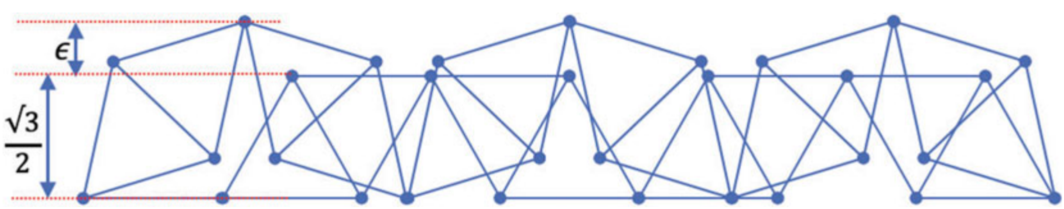


Fig. 54.1 A 4-chromatic unit-distance graph inside a strip slightly taller than $\frac{\sqrt{3}}{2}$

During 2021–2022, a group of Russian mathematicians, V.A. Voronov, A.M. Neopryatnaya, and E.A. Dergachev, constructed a series of 5-chromatic unit-distance graphs without a Mosers Spindle on 64,513 vertices [VND]. Their graphs "live" in $Q\left[\sqrt{2}, \sqrt{3}, \sqrt{5}\right] \times Q\left[\sqrt{2}, \sqrt{3}, \sqrt{5}\right]$, which guarantees the absence of Mosers Spindles, because the appearance of the latter requires $\sqrt{11}$.

In 2021, Marijn Heule enters the field and constructs the first relatively small 5-chromatic unit-distance graph without Mosers Spindles of order 1441 that appears in the October 2021 issue of *Geombinatorics* [Heu5].

In consultations with Vsevolod Voronov, he looks for a reasonably small set of points in $Q\left(\sqrt{2}, \sqrt{3}\right) \times Q\left(\sqrt{2}, \sqrt{3}\right)$ (observe, no Mosers Spindles there), such that a pair of points at distance 2 is monochromatic in all 4-colorings. Heule then applies SAT solving techniques that were described in [Heu3] to further trim the graph. He ends up with a graph on 721 vertices and 3948 edges. This graph, called T_{721}, is shown in Fig. 54.2. The *spindle argument* turns it into a 5-chromatic unit-distance graph T_{1441} with 1441 vertices.

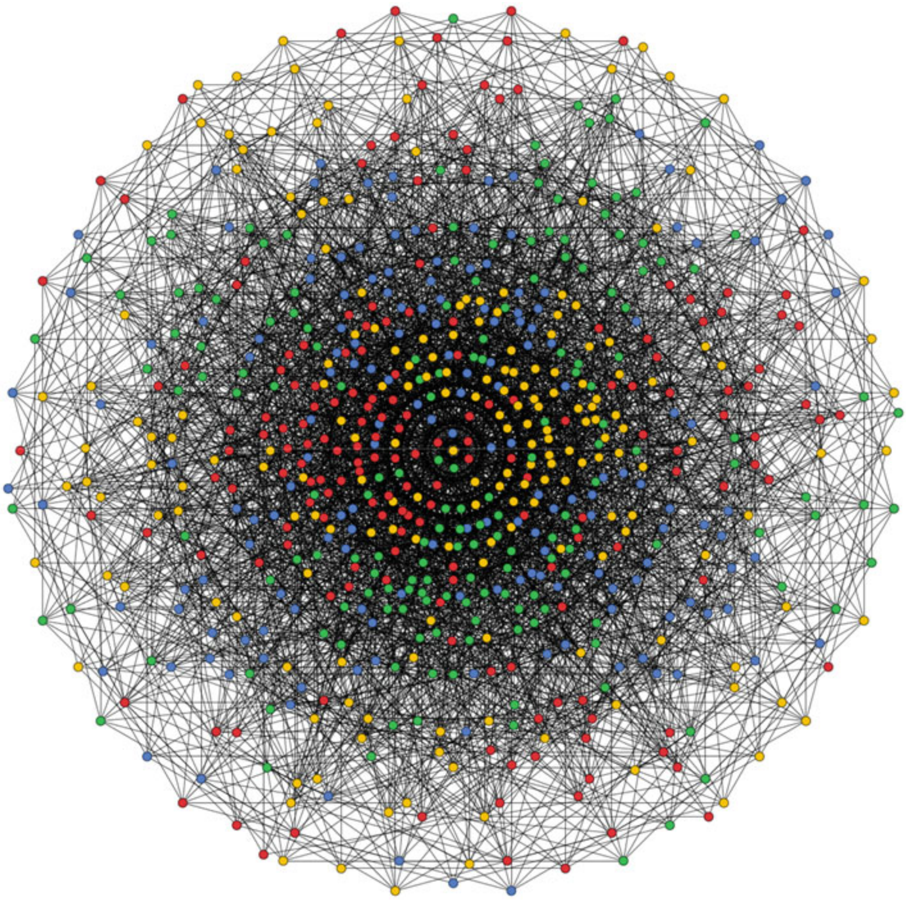


Fig. 54.2 A 4-coloring of UD graph *T*721

You may be wondering what is "*spindle argument*"? It is the idea used by Leo and Willie Moser in 1961 to create the Mosers Spindle, used in more sophisticated settings. Take a unit-distance graph *G* that includes vertices *a* and *b*. Rotate *G* about *a* until the distance between *b* and its rotation image *b'* becomes 1; call the rotated image *G'*. The final result is the union of *G* and *G'* with all the newly created unit edges, such as *bb'*.

On my request, on March 28, 2023, Marijn kindly sent me a visualization of his 1441-vertex graph T_{1441} for the inclusion in this book. It was not shown in the original essay [Heu5] You can see T_{1441} in Fig. 54.3.

Marijn observes:

> This graph is symmetric: a rotation by 60 degrees maps it onto itself. The symmetry is not a coincidence, but due to the trimming procedure: Every time a vertex-critical graph G was generated, the graph was extended by merging it with five copies of the graph rotated by 60, 120, 180, 240, and 300 degrees, respectively. The extended graph was used for the next iteration. The extension procedure improved the overall trimming effectiveness and produced a symmetric graph.

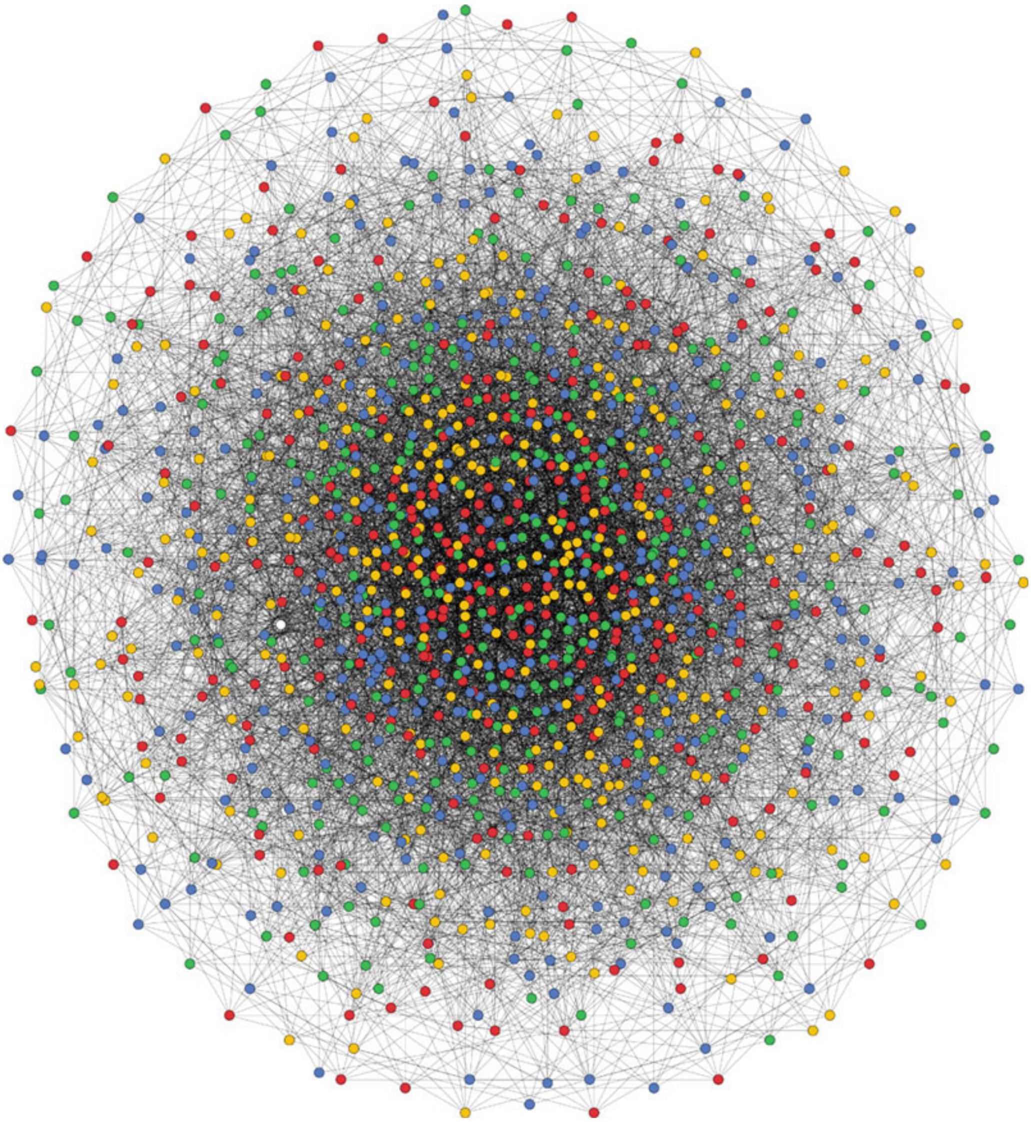


Fig. 54.3 5-chromatic unit-distance graph on 1441 vertices without Mosers Spindles

While the graph T_{1441} has no Mosers Spindles, it has plenty of unit equilateral triangles. Moreover, its building blocks are 4-chromatic unit-distance graphs on 10 vertices that are similar to the Mosers Spindle. Let us give it a name: The Heule Spindle. You can see it in Fig. 54.4. This kind of constructions were used before, the special in this graph is that it fits in $Q\left(\sqrt{2}, \sqrt{3}\right) \times Q\left(\sqrt{2}, \sqrt{3}\right)$.

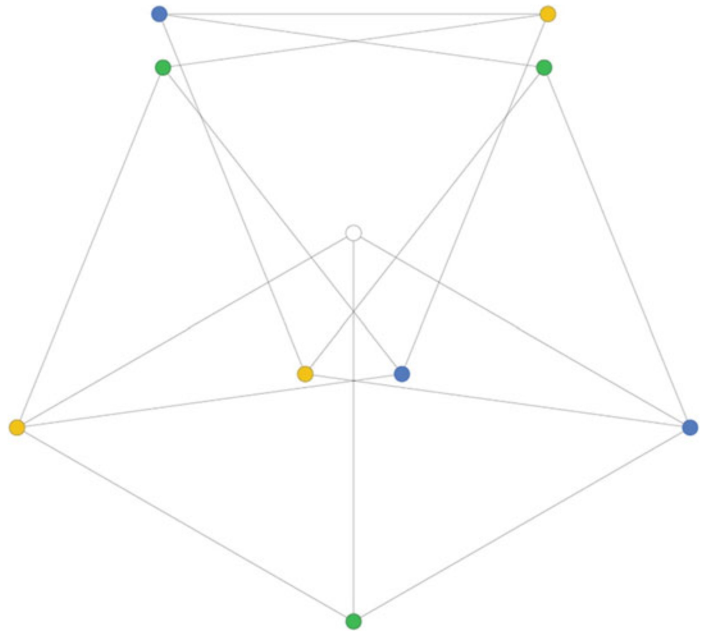


Fig. 54.4 The Heule Spindle

Chapter 55
Triangle-Free 5-Chromatic Unit Distance Graphs

Of course, I enjoyed constructions of Mosers-Spindle-free 5-chromatic graphs in the previous chapter. However, they included numerous unit triangles, and I asked for more in my 2019 problem paper [Soi46].

Construct 5-Chromatic Triangle-Free UDG 55.1 (2019, [Soi46]). Construct a triangle-free 5-chromatic unit-distance graph.

Smallest 5-Chromatic Triangle-Free UDG 55.2 (2019, [Soi46]). Find a 5-chromatic triangle-free unit-distance graph of the smallest order.

I do not really expect the smallest order graph to be found, but we ought to take steps in this direction.

Why do I pose these problems when the smallest order 5-chromatic UDG without a triangle-free condition is not larger than the graph with this condition? The triangle-free condition makes the graph easier to embed. Moreover, the Exoo–Ismailescu result of Chapter 16 allows for a relatively small building block: the smallest unit-distance 4-chromatic triangle-free graph has only 17 vertices, whereas without a unit-distance requirement, the Grötzsch graph is not much smaller at 11 vertices. I, therefore, believe that in triangle-free unit-distance 5-chromatic graphs, we may succeed in lowering the order of the 5-chromatic unit-distance graph sooner than in the general case.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_55

Chapter 56
Jaan Parts' Current World Record

56.1 The Record

In April 2020 issue of *Geombinatorics*, Dr. Jaan Parts, a microchip designer-engineer from Kazan, the capital of Tatarstan, Russia, presented an overview and classification of the state of hunting for 5-chromatic unit-distance graphs. I refer you to his 30-page *Geombinatorics* article [Par2] for important details. Parts' record holder (in terms of the smallest number of vertices) is a graph on 509 vertices with 2442 edges (visualized in Fig. 56.1), created using a large subgraph on 374 vertices with 1860 edges (visualized in Fig. 56.2), and a small subgraph on 136 vertices with 564 edges (visualized in Fig. 56.3).

© Alexander Soifer 2024
A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_56

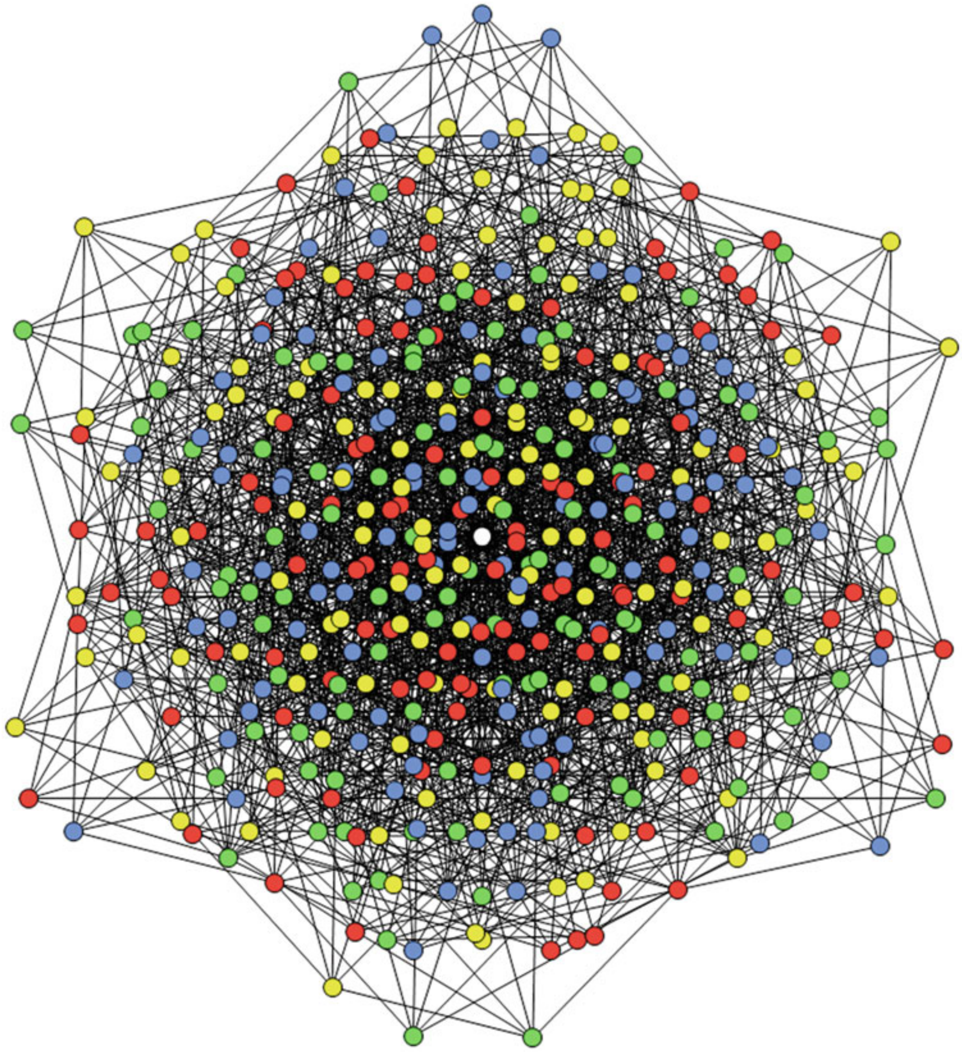


Fig. 56.1 5-chromatic unit-distance graph on 509 vertices with 2442 edges

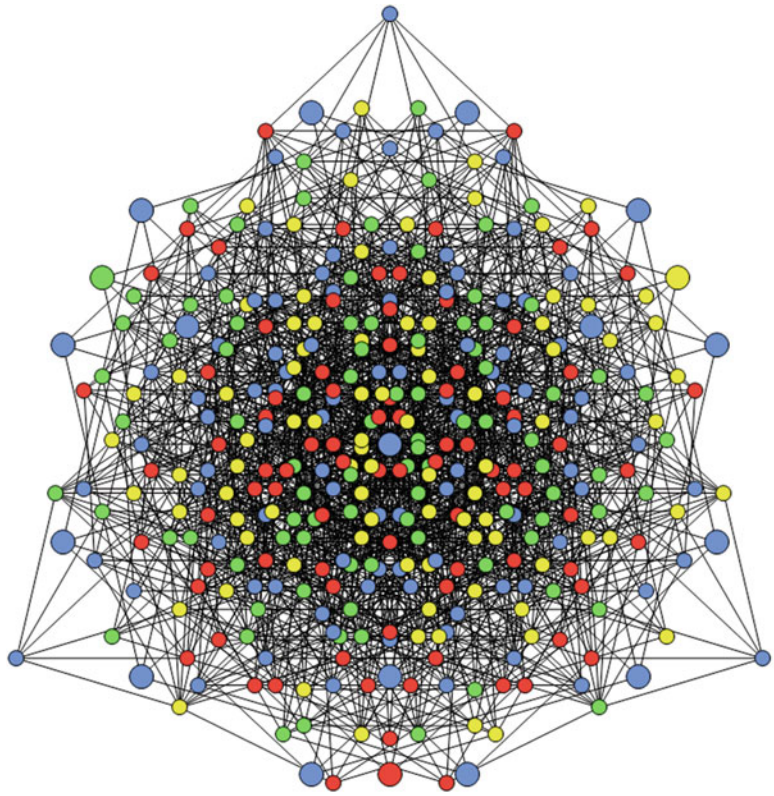


Fig. 56.2 Large subgraph L_{374} with 374 vertices and 1860 edges

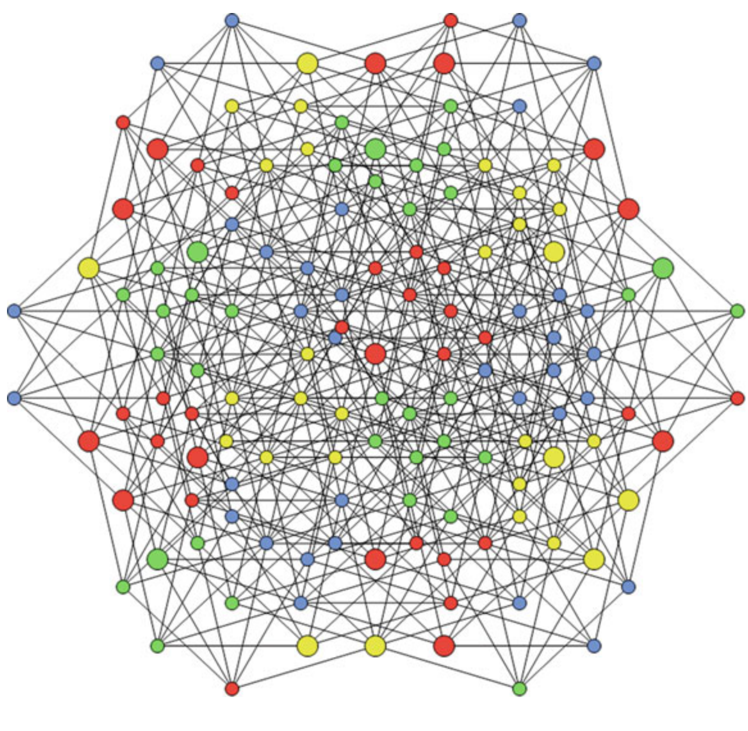


Fig. 56.3 Small subgraph S_{136} with 136 vertices and 564 edges. Vertices which can be used for connection to a large subgraph, are enlarged

Parts elaborates [Par3]:

> All calculations were performed in Mathematica 10 on a laptop with processor Intel Core i7-2670QM, 2.2GHz, 4 cores/8 threads, 6 GB RAM . . .
>
> We have introduced a new graph minimization method. This method can be useful in various minimization problems, where it is possible to effectively check whether a necessary property is preserved or not.
>
> Initially, our method was conceived as not requiring special programs and large computing power, while providing a good local minimum. In this sense, our approach can be seen as an alternative to the approach of Heule [Heu3]. With regard to achieving record results, both approaches give comparable results after proper tuning. There are certain difficulties in comparing their computational efficiency. Heule estimated the total costs to compute his 529-vertex graph as 100,000 CPU hours. Later, he rated 1,000 hours for a 510-vertex graph. We spent on the development of our approach roughly 1,000 laptop hours, including writing of programs and the study of different graphs. A 509-vertex graph can be found from scratch in about 100 laptop hours.
>
> In contrast to the approach of Heule with randomization and extraction of the (first available) unsatisfiable core, our reduction method looks through all the solutions and finds a global minimum for a given graph (but cannot work with as many vertices). The success of reduction is largely determined by the choice of the initial graph. Our approach reduces the impact of this choice. This allowed us to find a more efficient set of orbits and move a little further.
>
> As for the numerical results, they were obtained literally at the last minute. Not all options have been studied yet, which leaves the possibility of further progress.

The July 2018 Special Issue of *Geombinatorics* was dedicated to the chromatic number of the plane and related problems. At that time the world record for the smallest 5-chromatic unit-distance graph belonged to Marijn Heule and stood at 553. I wrote in that issue "I am sure the size of the smallest 5-chromatic unit-distance graph will go down, perhaps to the neighborhood of ca. 300 vertices." We have here a natural open problem, in which any result below 509 would be of interest, in number and even more so, in new ideas of achieving it.

Open Problem 56.1 (Soifer 2018). Find a 5-chromatic unit-distance graph of the smallest order.

56.2 Jaan Parts

Jaan Parts

My request for his life story, Jaan Parts honored in 2020 and updated on February 20, 2023.[1]

I smeared the map of daily life,
By splashing paint from a glass.
– V.V. Mayakovsky

When you have a blank sheet of paper, you think: well, what is there to write about? A work for a few minutes. Will fit in a couple of lines. But then the first words appear, like the first drops of rain, thoughts thicken into clouds, and before you have time to blink, puddles of colorful memories are already pouring in front of you – neither pass nor drive through.

I was born exactly in the middle of winter 1976. It happened in the city of Kazan, in the best country in the world. My daughter Alice (who lives in the same city) is now one-two-three-four years old (on the fingers of her right hand). If you ask her where we live, she will say, on planet Earth.

[1] The autobiography has been abridged in length while IMHO retaining its main ideas (A.S.)

Teaching is a special kind of activity, it requires not only deep knowledge but also passion for the subject, the ability to look at it from different angles, the ability to constantly find something new and unexpected in it. This is the only way to ignite the light of knowledge in a student. Otherwise, it will only be a smoke, which will quickly disappear. You cannot teach anything by force, maybe only to march. When I entered graduate school at the Faculty of Radio Electronics (precisely in order not to march) and had to give lectures as a duty, it was a painful task for me.

When America decided to destroy the education system in Russia (on the wave of success after the lucky collapse of the Soviet Union), it proposed: let us introduce our progressive system in your educational institutions. Of course, we immediately agreed: after the mountains of gum fell on us and rivers of Coca-Cola flowed, after so many years of thirst, we were looking forward to something new and progressive. Now, however, many would like to return to the old system. I did my teaching duty for some time, but then I seized the moment and ran away.

From aviation (which our country at that time began to joyfully destroy), I did not get much – I studied radio electronics. However, one of our teachers used to say: only beautiful planes fly. In my opinion, this statement can be safely extended to everything that is done by a person.

In the first years of study at the university, we enthusiastically assembled ZX Spectrum computers from a heap of small logic microcircuits. It was a good school, first of digital electronics, and then of programming. From that time, I have retained my passion for optimization. Those computers had only a few tens of kilobytes of memory, the clock speed was just over a dozen megahertz, but for us these were inexhaustible resources. We created complex effects, while we used several tens of bytes of program code, and everything worked. Our computers are now millions of times more powerful and have a million times more memory. And they barely crawl. Maybe because beauty disappeared somewhere in our hurry? Recently, I read in Wikipedia that Aubrey de Grey also worked in the very company that produced the good old ZX.

Now I continue to develop electronics and write programs for microcontrollers. From time to time, I find myself wanting to do science again. One time about three years ago, I watched a series of popular lectures on mathematics. The lecturer proudly said that mathematicians are not concerned with the question of whether it is possible to derive any practical benefit from mathematics. Wow, I also want to work on pure mathematics.

I took up chromatic numbers by accident, inspired by the same lectures. The task was formulated quite simply so that one could understand what is required. Of course, not even a week had passed since I successfully coped with it and proved that 7 colors are needed. Though, by this moment, I was already careful enough to realize that if something worked out so quickly, then it must be a consequence of dark ignorance. And so it turned out. Knowledgeable people advised Soifer's book on coloring and added that one should not chase 7 colors at once, even 5 colors would be a great achievement.

Aubrey's achievement soon happened, and the Polymath project, which immediately impressed me with a meaningful and friendly communication. As usual, I did not understand much, but I got caught up in a small task to find the maximum fraction of the plane, which can be guaranteed to have a proper coloring in five colors. I presented my results in the project, not hoping that they would interest someone. Suddenly, they arose a keen interest of Aubrey himself, the great magician and wizard. Sometime later, I got the courage and asked him to teach me how to calculate chromatic numbers. He readily provided me with a working tool.

Gradually I got involved in the work on the project, leaving behind heaps of blunders, among which sometimes good ideas occurred. By my rough estimate, only one idea in ten worked. All the rest was either known for a long time or contained gross miscalculations or did not lead anywhere. But, like any novice who is not burdened with knowledge, I had no shortage of ideas. At some point, I became so insolent that I tried to compete with another great sorcerer, Marijn Heule, the ruler of all computers in the world. The funny thing is, not without some success.

I already thought that it could not be any cooler, when at one point I received a letter from another celestial. This time, Alexander Soifer himself suggested me to write an article and publish it in the journal. In response spawned a large monster that could not fit in any gate. I had a natural assumption that after editorial changes, no more than 10% of the article will remain, and the rest will go straight to the trash can. Imagine my surprise when I saw that the article was published in full! Since little changed in my mathematical abilities lately, I have only one explanation for why this article was published. Apparently, Alexander was conquered by my literary talents.

Some time passed, and I dropped kilobytes of eloquence on Alexander. He does not even suspect that I have accelerated and am preparing another epic work for his misfortune. I had the idea of going to the other extreme, but then I discovered that the shortest article in the world had already been written, and it was done by none other than Alexander himself.

By now, I have the following evidence of my inordinate audacity: two great people translate my articles, catch tons of mistakes in them, and publish them in a journal. I feel that I have found myself in a dangerous position of the old woman from Pushkin's fairy tale about the fisherman and the fish. The golden fish regularly builds palaces for me, and my appetites grow. It looks like it is not far from the broken trough. But it is already hard to stop. And now Alexander wants to insert me into his book.

P.S.: More than two years have passed since then. A lot has changed. In co-authorship with Alexander and Aubrey, we wrote an article so short that it was even refused to be accepted in arXiv. Some of my modest achievements in the field of chromatic numbers have already become obsolete. The planet has changed. Only our inability to find a construction that brings the chromatic number of the plane to at least 6 does not change. What are my predictions? Apparently, it will be quite large.

Aubrey lit the way for us. But this is a way in a dark labyrinth. Therefore, after a couple of turns, we have to go by touch again. Have you tried juggling balls? As long as you have three balls, this is quite easy because most of the time only one ball is in the air, the other two are in your hands. About the same happens with $\chi \geq 4$. Three colors are in your hands immediately, and it remains to toss one more. With four balls, you have to keep track of two balls at once, which is much more difficult. The first to learn how to toss two colors at once was Aubrey, showing $\chi \geq 5$. But now try to manage five balls!

As experience (of our research) shows, if we limit ourselves to 4 colors, we can add vertices sequentially to a small graph in such a way that almost every new vertex will take on a strictly defined color, and only sometimes it will be necessary to choose one of two colors. It takes only a few dozen such forks and less than 1000 vertices, and we arrive at Aubrey's construction proving $\chi \geq 5$. We now know that a minimal 5-chromatic graph contains from

22 to 509 vertices. Now if we repeat the same approach, limiting ourselves to 5 colors, then the forks threaten to form at each of the first 500 steps and not only. But even 2^{100} coloring options are already too many.

However, Aubrey de Grey predicts that the $\chi \geq 6$ conjecture will be proved during his lifetime. But Aubrey is a biogerontologist and claims that a person can very well extend his life to 1000 years, so such conjecture does not sound very promising. So, I'm making a more optimistic conjecture: proof of $\chi \geq 6$ will come before I can juggle 5 balls. Although, to be honest, I do not force the training too much.

Part XI
What About Chromatic 6?

Isn't 6 a perfect number?

In this chapter we will look at some of the approaches toward constructing a 6-chromatic unit-distance graph. Perhaps, the third edition of this book will present a success in the direction of the perfect number 6. The 'magnificent seven' construction appears to be out of reach of today's people, even those who are armed with computers and supercomputers.

Chapter 57
A Stroke of Brilliance: Matthew Huddleston's Proof

In February 2008, I sent the manuscript of *The Mathematical Coloring Book* to Springer. At about the same time, the Problems Section of the *American Mathematical Monthly* published the solution [OTH] of our interest here. I read it and was amazed at the Olympiad-like beauty of the short Matthew Huddleston's solution of the part C of the problem proposed by Jim Owings. (It seems, no one else solved part C of the problem, including the proposer.) I have got to share it with you, for beauty is a rare commodity on our planet. Note that it appeared BG, i.e., Before de Grey, a whole 10 years before.

It is convenient for me to introduce new natural definitions. Let us call a graph G *two-distance graph* and denote it by $G\{1, d\}$, $1 < d$, if its vertices are distance 1 or distance d apart. You understand, of course, that we do not need to use analogously defined $G\{c, d\}$ because a scaling by c brings the smaller distance to 1. In *coloring a two-distance graph*, we forbid monochromatic pairs of distance 1 and distance d. *Chromatic number of a two-distance graph* G is naturally the minimum number of colors assigned to the vertices of G that forbids monochromatic unit and monochromatic d.

In 2006, Jim Owins asked, among other, the following question, which I am reformulating to serve our purposes.

Problem 57.1 (Proposed by Jim Owins in 2006 [OTH]). Is there a $d > 1$ and a two-distance graph $G = G\{1, d\}$, such that $\chi(G) \geq 6$?

Solution by Matthew Huddleston (February 2008, [OTH], Washington State University, Pullman). *American Mathematical Monthly* does not report receiving any other solutions.

Huddleston sets out to prove the existence of the two-distance graph with the vertex set E^2 with the chromatic number at least 6:

$$\chi\left(E^2\left\{1, \frac{1+\sqrt{5}}{2}\right\}\right) \geq 6.$$

Note: all pairwise vertex distances in a regular unit pentagon are 1 and $\frac{1+\sqrt{5}}{2}$.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_57

Proof Assume the plane is colored in 5 colors without creating monochromatic pairs of distance 1 and $\frac{1+\sqrt{5}}{2}$. Let S be the set of five points forming the vertices of a regular pentagon of side 1. Let Q be the set of ordered quintuples chosen from S, so that Q has 5^5 elements. Color Q by assigning to each of its 5-tuples the color of the sum of its five entries in the original coloring of the plane.

For any 4-tuple of points in S, note that adding the sum of all its entries to each of the five points of S produces a regular pentagon of side 1, so these five points have different colors. Therefore, each of the five colors is assigned to 5^4 of the 5^5 elements of Q. On the other hand, permuting an ordered 5-tuple in Q cannot change its color.. The number of permutations of a given quintuple is a multinomial coefficient of the form

$$\binom{5}{a_1, a_2, a_3, a_4, a_5}$$

where a_j is the number of occurrences in the quintuple of the jth element of S. This multinomial coefficient is a multiple of 5 except for the cases in which one of the a_j is 5 and the rest are 0. In order for the sizes of all the color classes in Q to be multiples of 5, these 5 exceptional cases must all be assigned the same color. Equivalently, the points of $5S$ all have the same color. This shows that in any regular pentagon with side 5, all vertices have the same color, so that any two vertices with distance 5 have same color. An isosceles triangle with sides of length 5, 5, 1 thus has vertices of the same color, and that is a contradiction. ■

Chapter 58
Geoffrey Exoo and Dan Ismailescu, or 2 Men for 2 Forbidden Distances

We have a team of a *Geombinatorics Editor* Geoffrey Exoo and Dan Ismailescu. You may recall Chapter 16 of this book dedicated to their construction of unit-distance graphs of successive orders 21, 19, and 17 of girth 4 and proving that 17 is best possible, thus settling my open problem 15.4 posed in the first edition of this book [Soi44]. You also saw Geoffrey's lower bounds for Ramsey Numbers, the Schur number 5, and many more results.

Geoffrey tells me – and I absolutely trust him – that he and Dan had their 5-chromatic unit-distance graph early, but life's events interfered, delayed the final steps, and they barely missed to be first. Having acknowledged – as they ought – de Grey's priority, they published their graph elsewhere [EI3], even though I was willing to publish it in *Geombinatorics*. In *Geombinatorics*, they are building tools clearly aimed at constructing a 6-chromatic unit-distance graph. In the 2018 *Geombinatorics'* Special Issue XXVIII(1), they first construct (fairly) small two-distance graphs in the plane of the chromatic number at least 5 [EI2]. Summing up their results, they first prove the following theorem, which we have informally used earlier in this book.

58.1 The Spindling Method

We have already used *the spindling method* several times in this book. Let the authors of this chapter formalize it for us.

Theorem 58.1 [EI2]. Let G be a finite graph with vertex set $V = \{1, 2, \ldots, n\}$ and edge set E. Assume that the chromatic number of G, $\chi(G) = k$ and that in every k-coloring of G, vertices 1 and 2 are colored identically.

Let G' be a copy of G such that $1 = 1'$ and $2 \neq 2'$. Then the chromatic number of the graph H with edge set $E \cup E' \cup \{\{2, 2'\}\}$ is $\geq k + 1$.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_58

Proof Assume that H is k-colorable and let c be such a k-coloring. Then $c(1) = c(2)$ by assumption. At the same time, since G' is a copy of G, it follows that $c(1') = c(1) = c(2')$. But then vertices 2 and $2'$ have the same color, which violates the condition that the endpoints of the edge $\{2, 2'\}$ must be colored differently. ■

In a word, we rotate the graph G about its vertex 1 until the distance between vertex 2 and its image $2'$ under the rotation reaches the distance we desire. And we often desire the unit distance.:)

58.2 Two-Distance Graphs of Chromatic Number At Least 5

Result 58.2 [EI2]. $\chi(E^2\{1,d\}) \geq 5$ for the following values of d:

$$\frac{\sqrt{5}+1}{2}, \sqrt{3}, \frac{\sqrt{6}+\sqrt{2}}{2}, \frac{1}{2}\sqrt{3^{1/4}\cdot 2\sqrt{2}+2\sqrt{3}+2}, \sqrt{3/2+\sqrt{33}/6}, \frac{\sqrt{5}}{\sqrt{3}}, 2, \frac{2}{\sqrt{3}}$$

For each listed above value of d, they construct a graph. I am choosing two of their graphs, one small and another beautiful, to let you taste their proofs.

Result 58.3

$$\chi\left(E^2\{1,2\}\right) \geq 5.$$

Proof Consider the 26-vertex {1,2}-graph G whose vertices are given by the following coordinates:

$$[-2,0,0,2],[2,0,0,2],[0,0,0,0],[0,0,0,4],[0,0,-6,2],$$
$$[0,0,6,2],[-1,-3,3,3],[1,3,3,3],[-3,-3,3,1],[3,3,3,1],$$
$$[-1,-3,-3,1],[1,3,-3,1],[-4,0,0,0],[4,0,0,0],[3,-3,-3,1],$$
$$[-3,3,-3,1],[1,-3,-3,3],[-1,3,-3,3],[1,-3,3,1],[-1,3,3,1],$$
$$[-2,0,6,0],[2,0,6,0],[-2,0,-6,0],[2,0,-6,0],[0,-6,0,2],[0,6,0,2].$$

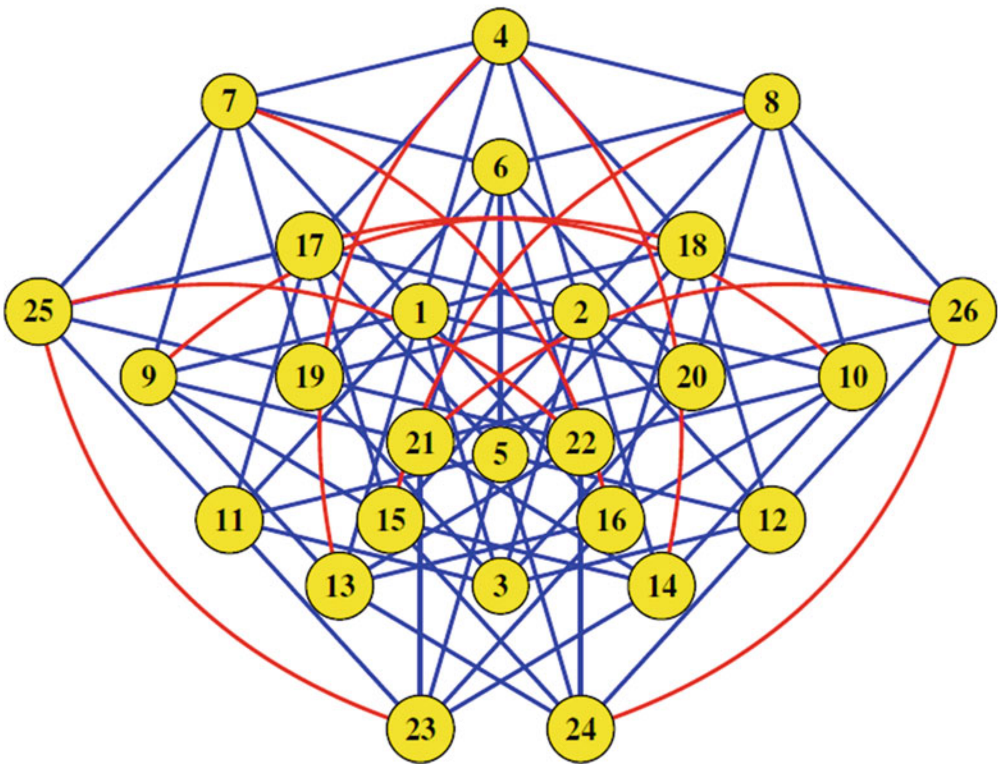


Fig. 58.1 26-vertex graph G with 75 unit edges (shown in blue) and 10 edges of length 2 (shown in red)

It can be verified that this graph (Fig. 58.1) has 75 unit edges and only 10 edges of length 2. Since the mid-point of any edge of length 2 is also a vertex of the graph, these long edges are shown in Fig. 58.1 with curved line segments.

The authors use both Maple and Sage to verify that $\chi(G\{1,2\}) \geq 5$. Given the relatively small order of this graph, a computer-free proof is certainly possible. It is a surprising fact that only 10 long edges sufficed to raise the chromatic number from 4 to 5. ■

Result 58.4

$$\chi\left(E^2\left\{1,2/\sqrt{3}\right\}\right) \geq 5.$$

Proof Consider the 103-vertex $\left\{1,2\sqrt{3}\right\}$-graph shown in Fig. 58.2.

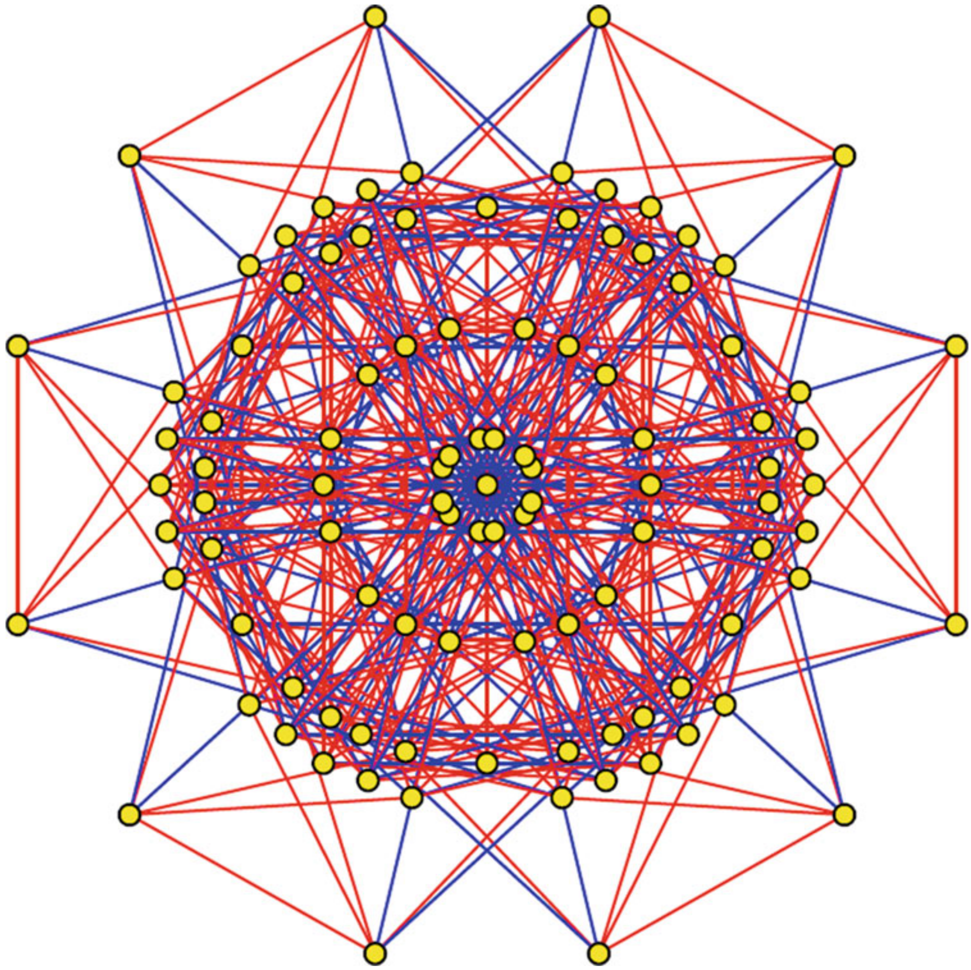


Fig. 58.2 The 103-vertex $\{1, 2\sqrt{3}\}$-Exoo–Ismailescu graph with chromatic number 5. Edges of length 1 are shown in red; edges of length $2/\sqrt{3}$ are shown in blue

This graph is too large to verify its chromatic number by hand: it has 312-unit edges, 177 edges of length $2/\sqrt{3}$, and chromatic number of at least 5. Despite its size, Sage takes only a couple of minutes to verify this. A list of vertices is available at the URL [EI4]. ■

58.3 Two-Distance Graph of Chromatic Number At Least 6

In January 2020, in *Geombinatorics*, Geoffrey Exoo and Dan Ismailescu continue their direction of studying two-distance graphs. They create the second two-distance graph of chromatic number 6 [EI5], thus strengthening their result 56.3.

Theorem 58.5 $\chi(E^2\{1, 2\}) \geq 6$.

Proof I am presenting a slightly edited proof by Exoo and Ismailescu. Let us construct a (1,2)-distance graph of the chromatic number at least 6. We build this graph in several stages, using vertices with coordinates of the form $\left(a\sqrt{3}/12 + b\sqrt{11}/12, c/12 + d\sqrt{33}/12\right)$ where *a,b,c,* and *d* are integers. We will use the following notation:

$$[a, b, c, d] := \left(\frac{a\sqrt{3}}{12} + \frac{b\sqrt{11}}{12}, \frac{c}{12} + \frac{d\sqrt{33}}{12}\right).$$

Step 1 Consider the following set *S* of 23 points in the plane:

$$\begin{aligned} S := \{&[0,0,0,0], [0,0,0,-4], [0,0,-6,-2], [0,0,-6,2], [-6,0,0,-2],\\ &[-4,0,0,0], [-4,0,-6,-2], [-4,0,-6,2], [-2,0,0,-2], [-2,0,-6,-4],\\ &[-2,0,-6,4], [0,-6,-6,0], [-5,-3,3,3], [-5,3,-3,3], [-2,-6,0,0],\\ &[-2,-6,0,-4], [-2,6,0,0], [-2,6,0,-4], [-6,-6,0,0], [-6,6,0,0],\\ &[-4,0,0,-4], [0,0,-12,0], [-8,0,0,0]\}. \end{aligned}$$

Next, consider all the points in *S* together with their reflections across the *x*-axis and across the *y*-axis, respectively. One obtains a new set *T* that has 57 points. For $k = 0, \ldots, 5$, let U_k be the image of *T* under a rotation through the angle $k\pi/3$ about the origin and define

$$V : U_0 \cup U_1 \cup U_2 \cup U_3 \cup U_4 \cup U_5.$$

It is easy to check that the image of $[a,b,c,d]$ under a rotation through $\pi/3$ about $[0, 0, 0, 0]$ is $[(a-c)/2, (b-3d)/2, (3a+c)/2, (b+d)/2]$. Let *G* be the {1,2}-graph, whose vertex set are the points in *V*.

Claim 1 The graph *G* has 205 vertices, 966 edges of length 1, 423 edges of length 2, and exactly 18 5-colorings.

Step 2 We construct a slightly larger graph *H*, by including the following nine additional vertices to the vertex set of *G*:

$$\begin{aligned} A := [-2,0,0,-6], B := [8,0,0,4], [-4,-6,-6,-4], [-4,6,6,-4], [-3,-3,-3,-5],\\ [-4,0,-12,4], [-4,0,12,4], [7,-3,3,3], [7,3,-3,3]. \end{aligned}$$

Claim 2 The graph *H* defined above has 214 vertices, 1004 edges of length 1, 446 edges of length 2, and exactly 35 5-colorings. Moreover, in each of these colorings, vertices A and B are of the same color (Fig. 58.3).

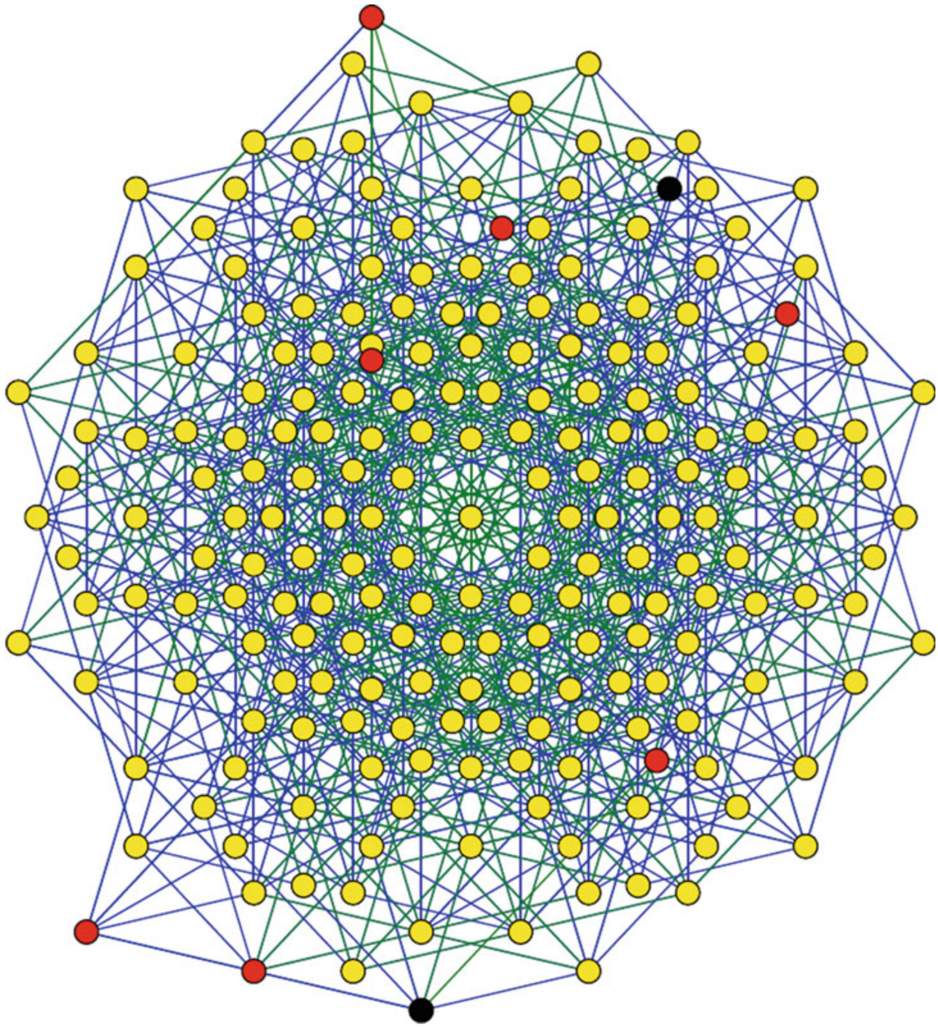


Fig. 58.3 The {1,2}-graph *H*. Vertices *A* and *B* appear in black, the other seven new vertices in red

Note that the distance between *A* and *B* is exactly 5. Rotating the vertices of *H* about vertex *A* by an angle arccos(49/50) = $\arcsin\left(3\sqrt{11}/50\right)$ creates a copy of *H*, which we denote H'. The image of vertex *B* under this rotation is a point $B' \in V(H')$, and the distance between *B* and B' is exactly 1.

Let *K* be the {1,2}-graph whose vertex set is $V(H) \cup V(H')$. It can be checked that *K* has 426 vertices, 2009 edges of length 1, and 892 edges of length 2. More importantly, every 5-coloring of *K* forces vertices *A, B,* and B' to receive the same color. Since $AB = AB' = 5$ and $BB' = 1$, it follows that $x(K) \geq 6$. This concludes the proof of Theorem 58.5.

Note 1 The final argument in the proof of Theorem 58.5 involves an isosceles triangle of side lengths 5, 5, and 1, exactly the same triangle as Matthew Huddleston used in 2008 [OTH] (see Chapter 57).

Note 2 All vertices of the graph K have coordinates in $Q\left(\sqrt{3}, \sqrt{11}\right)$ because the smallest angle of an isosceles triangle with sides 5,5 and 1 is $\arccos(49/50) = \arcsin\left(3\sqrt{11}/50\right)$.■

Computations We have two counting claims to prove. The assertions pertaining to edge counts for both of the graphs can be easily obtained by direct computation using the data available at [EI6], [EI7]. The assertions that there are exactly 18 5-colorings for graph G, and 35 5-colorings for graph H require more difficult computations.

To find all 5-colorings for these graphs, we used a simple recursive exhaustive search procedure that allowed us to divide the work across multiple processors. The outline of the search procedure is given below. Before the procedure is used, the vertices are ordered as follows.

- The vertices are partitioned into orbits based on the dihedral group generated by the transformations used in the construction (reflections in the axes and the $\pi/3$ rotation).
- Vertices within an orbit are sorted by polar angle: $0 \leq \theta \leq 2\pi$, and at all stages appear consecutively in the vertex ordering.
- Vertex orbits are sorted in descending order by degree.
- In case of ties, vertex orbits adjacent to the largest number of vertices that appear earlier in the ordering are listed first.

Then each vertex is assigned the *NC* (uncolored) value, and the following search procedure is called with vertex 0 and the list of colors as parameters.

The computations were performed using 48 threads on an AMD EPYC 7551 32-Core (64 Virtual Core) Processor and were completed in 3780 seconds of elapsed time and 81,000 seconds of total processing time for G, and 5120 seconds of elapsed time and 95,000 seconds of total processing time for H. In each case, all but three of the threads were finished halfway through the computation, which was not surprising, given that our method for splitting the work was fairly crude.

Exoo and Ismailescu include their computer program in the *Geombinatorics* paper [EI2]; please, consult it there. ■

58.4 Geoffrey Exoo

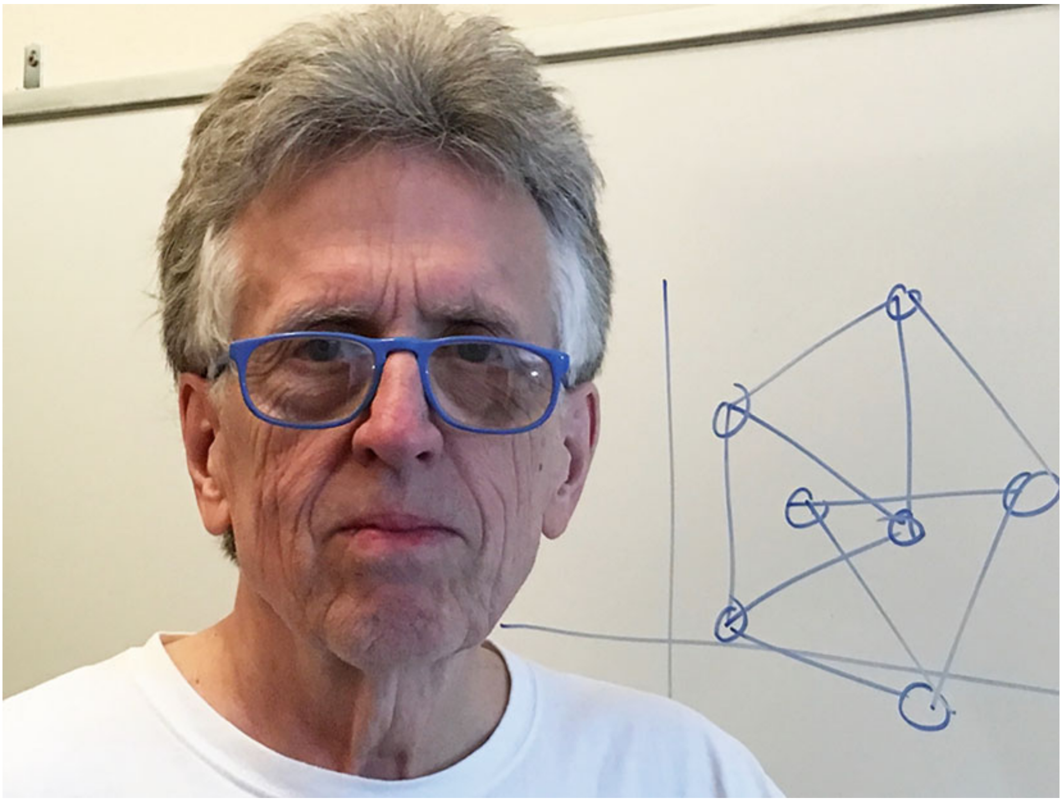
On August 31, 2020, Geoffrey Exoo answered my request to write his life story. This section is all his.

I grew up in Cleveland, Ohio, went to the University of Michigan, and landed my first real job as a Programmer/Statistician for the Michigan Employment Security Commission.

I may be the only Combinatorialist whose first publication is in the "Michigan Manpower Review Quarterly," an article on unemployment rates in the automobile industry and how they were computed in 1971.

While in Arizona later that year I tried to get a programmer job at Kitt Peak. The job was under Civil Service and one way to gain "points" for Civil Service jobs was to take the GRE exam, which I did, and listed Michigan as one of the three places to have the scores sent.

I did not get the job at Kitt Peak, though I did have a great time interviewing, but ended up returning to Michigan as a part-time graduate student and part-time programmer for few years.



Geoffrey Exoo

During that time, I had a variety of amusing programming jobs, including jobs which involved using lasers to make finger holes in bowling balls, making microporous membranes to filter Japanese beer, processing TV ratings data, and evaluating water filtration systems, to name a few. My grad studies were sort of off and on, but at one point, I learned that Graph Theory was a good dissertation topic if one wanted to land a position at Bell Labs, which I did, more because it was the center of the Unix world than anything to do with Mathematics.

When I was a grad student, I saw a few talks on using computers to make progress on Graph Theory problems. It appeared to me that the techniques being described were pretty elementary. So, I always had it in mind to try to do better.

Eventually, I took Frank Harary's Graph Theory class and wrote my dissertation with him and did manage to end up at Bell Labs. I left there shortly after the 1982 consent decree that broke up AT&T.

I applied to Indiana State, remembering a close friend who went there, always touted it as a warm, friendly, and supportive place, a view I have never questioned.

I had one brief respite from the academic world in the last few years of the twentieth century, working on transportation problems for a large trucking company. Their fundamental problem was nearly equivalent to a Traveling Salesman Problem with a sales force of 15,000 and close to 100,000 places to visit – not as easy as it sounds. Good fun though.

58.5 Dan Ismailescu

Dan Ismailescu

On September 4, 2020, Dan Ismailescu answered my request to write about himself.

I was born in Romania, two years after Ceausescu rose to power. Both my parents were school math teachers; it turned out that the apple did not fall far of the tree. Growing up during tough economic and politically oppressive times, I regarded mathematics both as a refuge and a source of entertainment. I remember mulling over a hard problem for days on end and enjoying doing so. I am grateful to my teachers and to my parents for encouraging and supporting my inclinations.

Quite predictably, I went to college and studied mathematics; after graduation, I taught high school for seven years. After the collapse of communism in Eastern Europe, I came to the US on a scholarship and completed my graduate degree at Courant Institute under the guidance of János Pach. Since 2001, I have been happily employed at Hofstra University, NY.

I always liked problems that are easy to state but difficult to solve. That is why when I discovered the Hadwiger-Nelson problem in Soifer's book, I was instantly hooked.

I first contacted Geoff in 2013 in connection to a paper of his on ε-unit distance graphs, and we started working on chromatic numbers of Euclidean spaces. Several years and joint papers later, we still enjoy our collaboration. I do feel we complement each other well mathematically and we share similar views in regard to the pace of our research. Our relationship may appear unusual to many: there are months of inactivity followed by weeks during which we may exchange 20+ emails per day. Geoff and I never met, and we spoke live exactly once (over Skype). In some sense, we are pioneers of the social distancing practice.

Chapter 59
Jaan Parts on Two-Distance 6-Coloring

The same October 2020 issue of *Geombinatorics* that carried Exoo-Ismailescu essay on 6-chromatic two-distance graphs contains Jaan Parts' paper on the same topic [Par1].

This train of thought left the station in October 2019, when Aubrey de Grey explicitly defined the Huddleston Graph G_{Hud} [G2], as reported in [Par1]:

$$G_{\text{Hud}} = P \oplus P \oplus P \oplus P \oplus P,$$

$$P = \left\{ \left(r \sin \frac{2\pi k}{5}, r \cos \frac{2\pi k}{5} \right); r = \sqrt{\frac{5+\sqrt{5}}{10}}; k \in Z \right\},$$

where P is a regular pentagon of side 1 and $\oplus$ defines the Minkowski sum. In this case, the vertices of the graph G_{Hud} are the union of all sums of the vertex coordinates of the summand-graphs, and edges connect all vertex pairs that are at a forbidden distance apart. The forbidden distances, as you recall, are 1 and the golden ratio $d = \frac{1+\sqrt{5}}{2}$ (also known as the diagonal of a regular pentagon of side 1).

The graph G_{Hud} is fairly small: its vertices count 126 and edges of lengths 1 or d count 350. As Huddleston showed, in any 5-coloring, the farthest vertices from the center of the graph G_{Hud} form monochromatic pairs with a distance 5; the central vertex must have the same color (Fig. 59.1).

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_59

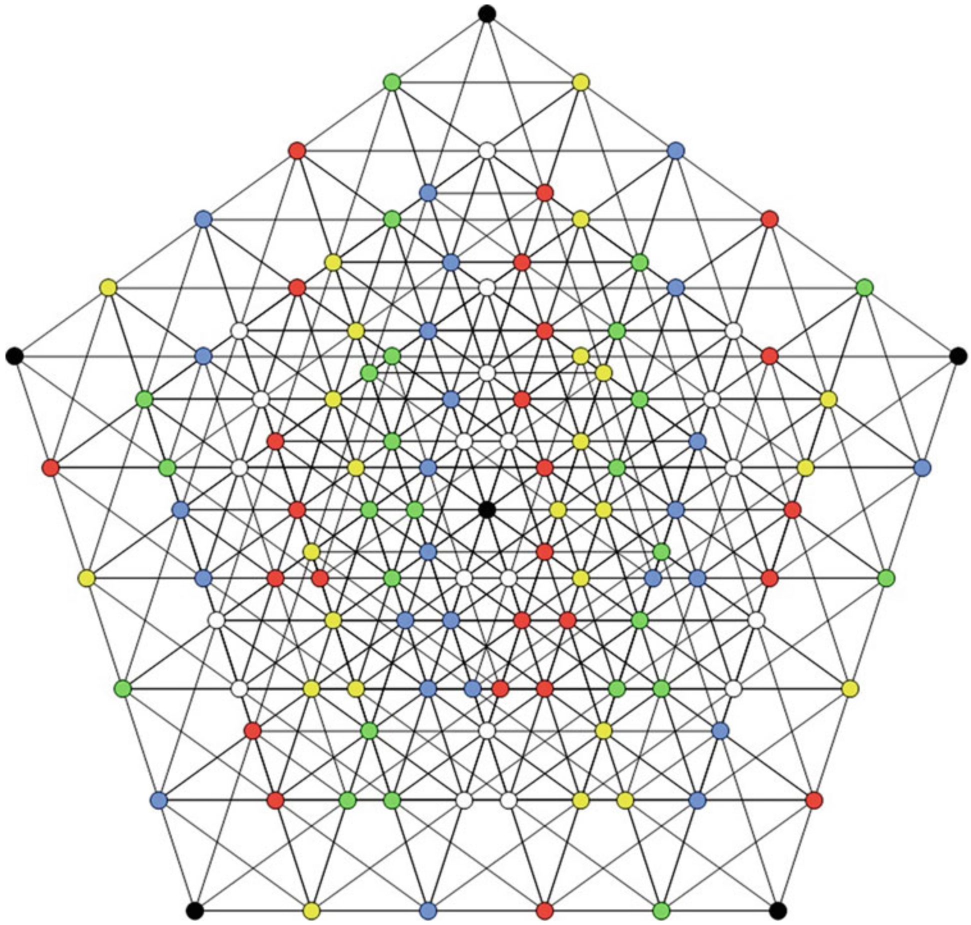


Fig. 59.1 Visualization of the Huddleston Graph with monochromatic vertices shown in black

These observations allow Parts to dramatically reduce the size of the desired graph. The graph G_{16} that he shows on the right side of Fig. 59.2 is a 16-vertex subgraph of G_{Hud}; it has a symmetry group of order 48 and 28 edges each of length 1 or d.

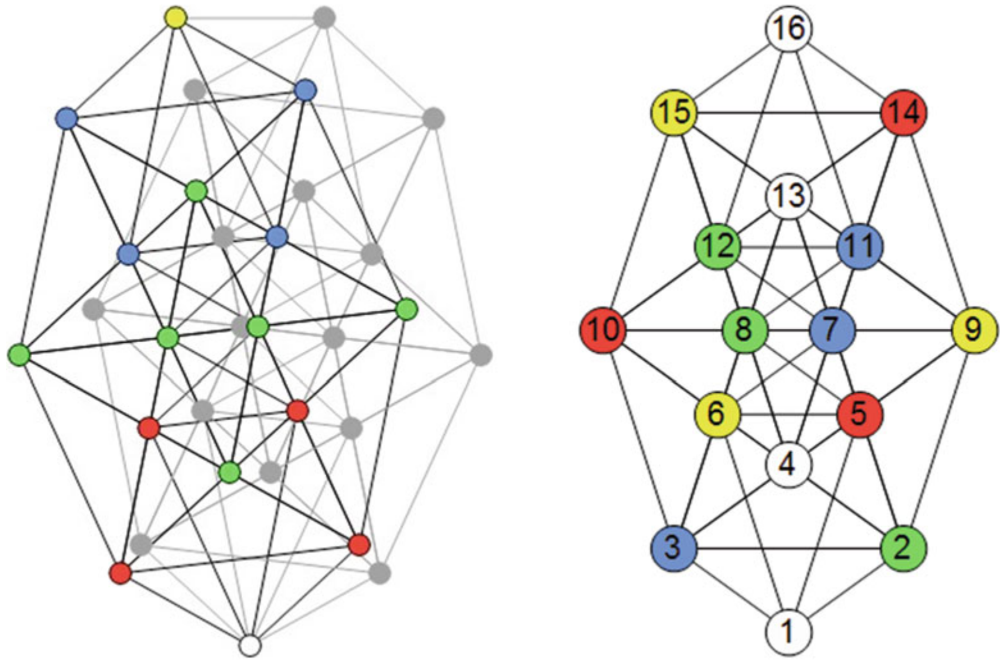
Edges of length 1:

(1,2), (1,3), (2,4), (2,5), (3,4), (3,6), (4,7), (4,8), (5,6), (5,8), (5,9), (6,7), (6,10), (7,9), (7,12), (7,13), (8,10), (8,11), (8,13), (9,11), (10,12), (11,14), (12,15), (13,14), (13,15), (14,16), (15,16).

Edges of length d (the golden ratio):

(1,5), (1,6), (2,3), (2,6), (2,7), (2,9), ((3,5), (3,8), (3,10), (4,9), (4,10), (4,11), (4,12), (5,13), (6,13), (7,10), (7,14), (8,9), (8,15), (9,13), (9,14), (10,13), (10,15), (11,15), (11,16), (12,14), (12,16), (14,15).

Now Parts is ready to construct a 6-chromatic two-distance graph G_{31} on just 31 vertices. All he needs to do is to apply spindling (see Theorem 58.1) to his graph G_{16} (Fig. 59.2 on the left). The construction and proof follow.



G_{31} G_{16}

Fig. 59.2 Jaan Parts' 31-vertex two-distance graph G_{31} and its building block G_{16}

Theorem 59.1 (Parts [Par1]). G_{31} is a 31-vertex two-distance 6-chromatic graph.

Proof Parts creates the 31-vertex graph G_{31} (left in Fig. 59.2) by spindling the graph G_{16}, i.e., rotating G_{16} around its vertex 1 by the angle $\arccos((95+\sqrt{5})/100)$, which gives an additional edge of unit length between the copies of vertex 16. This forces at least one of the two copies of vertex 16 to be of a different color than vertex 1. The graph G_{16} contains 5-clicks (thus requiring at least 5 colors), so it suffices to prove that in any 5-coloring of G_{16} vertices 1 and 16 are of the same color.

Partition the vertices of G_{16} into the following subsets: {1}, {2,3,5,6}, {4,7,8,9,10,13}, {11,12,14,15}, and{16}.

The vertices {1,2,3,5,6} form a 5-clique, thus they must be assigned different colors: 1-white, 2-green, 3-blue, 5-red, and 6-yellow.

The set {4,7,8,9,10,13} has three independent subsets (i.e., no edges inside subsets): {4,13}, {7,10}, and {8,9}. The colors green, blue, red, and yellow can be used in {4,7,8,9,10,13} only once. This means that one of {4,13}, {7,10}, {8,9} must have both of its vertices colored white. This forbids all vertices of the set {11,12,14,15} to be white, which forces vertex 16 to be white, i.e., the same color as vertex 1. ■

Parts' Theorem and William Gasarch' remark prompt me to pose the following natural open problem:

Open Problem 59.2 (Soifer 2023). Find the smallest (in the number of vertices) two-distance 6-chromatic graph.

Chapter 60
Forbidden Odds, Binaries, and Factorials

60.1 One Odd Problem

In March 1994, I arrived in Florida Atlantic University for the 25th Southeastern International Conference on Combinatorics, Graph Theory, and Computing. My main interest was as always to visit with Paul Erdős. This time he introduced me to Moshe Rosenfeld and his new problem. Good fortune preserves the image of that day:

From the left: Moshe Rosenfeld, John H. Conway, and Alexander Soifer, Florida Atlantic University, March, 1994

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_60

Definition 60.1 (Rosenfeld [Ros1]). The *odd-distance graph* E_{odd} is the graph with vertex-set E^2 in which two vertices are adjacent if and only if the distance between them is an odd integer.

Moshe showed that E_{odd} contains a subgraph K_4, i.e., its chromatic number is at least 4. He asked whether the chromatic number $\chi(E_{\text{odd}})$ of E_{odd} is finite. In fact, while the problem was new, the absence of K_4-subgraphs was not. It followed from the much more general 1974 result of Graham, Rothschild, and Straus [GRSt].

Theorem 60.2 In E^n there exist $n + 2$ points, the distance between any two of which is an odd integer if and only if $n \equiv 14 \pmod{16}$.

In the necessary part of the proof, the authors used a very old result about determinants by the leading Victorian mathematician Arthur Cayley. A natural problem comes to mind:

Problem 60.3 *Find* $\chi(E_{\text{odd}})$.

In 2009, Ardal, Manuch, Rosenfeld, Shelah, and Stacho [AMRSS] improved the lower bound to $\chi(E_{\text{odd}}) \geq 5$. Their construction started with a set S of points of a unit triangular lattice, followed by a spindling method (Theorem 58.1), i.e., rotation of S about one of its points until another point becomes distance 1 from its original position and considering the union of S and its rotated image – 21 vertices in all.

They also showed (ibid.) that the chromatic number of the graph with rational plane Q as vertex set with forbidden monochromatic odd distances $\chi(Q_{\text{odd}})$ is 2, just like in unit-distance rational plane.

Let us denote the (Lebesgue) *measurable chromatic number* of the odd-distance graph by $\chi_m(E_{\text{odd}})$. In 2009, the MIT undergraduate Jacob Steinhardt [Stein] proved the following result using tools of spectral graph theory, which may be beneficial in solving other coloring problems.

Theorem 60.4 (Steinhardt [Stein]). $\chi_m(E_{\text{odd}}) \geq \aleph_0$.

As you know, there are quite a few odd numbers :). This fact and the result in the measurable case inspired me to formulate a general case conjecture in 2009. In print it appeared in my chapter [Soi50] of the 2015 book *Open Problems in Mathematics* that John F. Nash, Jr. and Michael Th. Rassias invited me to write.

Conjecture 60.5 (Soifer, 2009). $\chi(E_{\text{odd}}) \geq \aleph_0$.

All Marijn Heule's emails are insightful and worthy of studying. On October 25, 2018, I received an email with "odd/odd-distance graphs" on the subject line:

> I observed something interesting, while trying to find a unit-distance graph with chromatic number 6: I had several promising graphs, but computing the chromatic number appeared very hard. I was actually able to find 5 colorings of several of these graphs by adding lots of non-unit edges. In particular I connected all points that are distance (2i+1)/(2j+1) apart with i, j in N. This was somewhat surprising to me as the edge density increased significantly, while the chromatic number did not increase at all (all graphs have chromatic number 5).
>
> Let an odd/odd-distance graph be a graph for which two points are connected if and only if they are exactly distance (2i+1)/(2j+1) apart with i, j in N. Now the main question that I have been thinking of:

is there a relatively small odd/odd-distance graph with chromatic number 6. So far I was not able to find any odd/odd-distance graph with chromatic number 6. The smallest odd/odd-distance graph with chromatic number 5 that I found has 21 vertices. If it is already hard to find an odd/odd-distance graph with chromatic number 6, it is not surprising that it is hard to find a unit-distance graph with chromatic number 6.

In October 2021, Marijn Heule publishes an essay in *Geombinatorics* [Heu5], which starts with his own pretty visualization of the 5-chromatic odd-distance graph (ODG) obtained in 2009 [AMRSS] (Fig. 60.1).

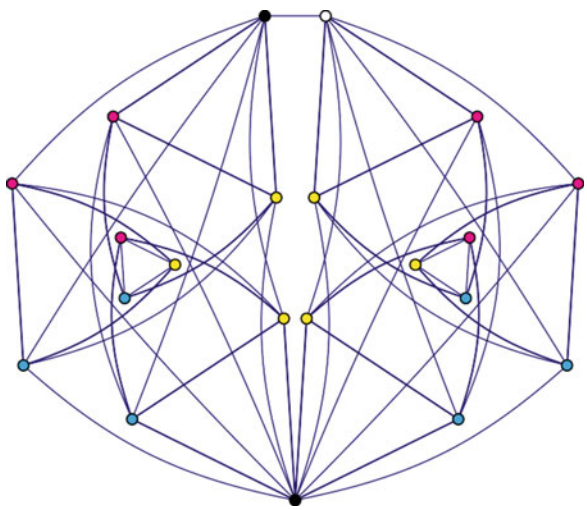


Fig. 60.1 The smallest known 5-chromatic ODG by Ardal et al. in Marijn Heule visualization. The shortest edges have length 1, while bold edges have length 3. Some edges are curved to avoid overlap in the picture

Marijn continues:

I tried to construct a 6-chromatic OD graph, but this turned out to be a challenge. Recall that the smallest known 5-chromatic OD graph has only 21 vertices. However, many large and dense OD graphs are 5-colorable. This is somewhat unexpected. Soifer conjectures that there are odd-distance graphs in the plane with infinitely large chromatic number [Soi50] and Steinhardt showed that there is no finite measurable coloring of the OD graph of the plane [Stein]. However, no OD graph with chromatic number 6 is known. I would like to pose it as a challenge with a prize.

The Heule Challenge ($500) 60.6 Construct an odd-distance graph with chromatic number 6 or prove that none exists.

In conclusion, Marijn Heule observes:

What makes OD graphs with chromatic number 6 interesting? As described above, some patterns observed in valid 4-colorings of dense UD graphs can also be observed in 4-colorings of OD graphs with significantly fewer vertices (points). Therefore, knowing which points form an OD graph with chromatic number 6 can be a big help in constructing UD graphs with chromatic number 6 (if they exist).

If no OD graph with chromatic number 6 exists, then there is clearly no UD graph with chromatic number 6. Hence the chromatic number of the plane would be 5. There are several other interesting open questions related to odd-distance graphs. One question that is potentially related to the challenge is whether all triangle-free graphs can be drawn as an odd-distance graph in the plane. As a possible first step in this direction, one could try to show whether the Mycielski graphs [Myc] are odd-distance graphs. A positive answer would solve the challenge and even imply the existence of odd-distance graphs with infinitely large chromatic number.

The next, January 2022, issue XXXI(3) of *Geombinatorics* carried Jaan Parts' essay answering Marijn Heule's challenge, where Parts constructs a 6-chromatic odd-distance graph of a relatively small order and shares his ideas about this "Odd Problem" and its relationship to unit-distance graphs [Par4]; see there details of construction and the reduction procedures. Parts' odd-distance 6-chromatic graph is visualized in Fig. 60.2.

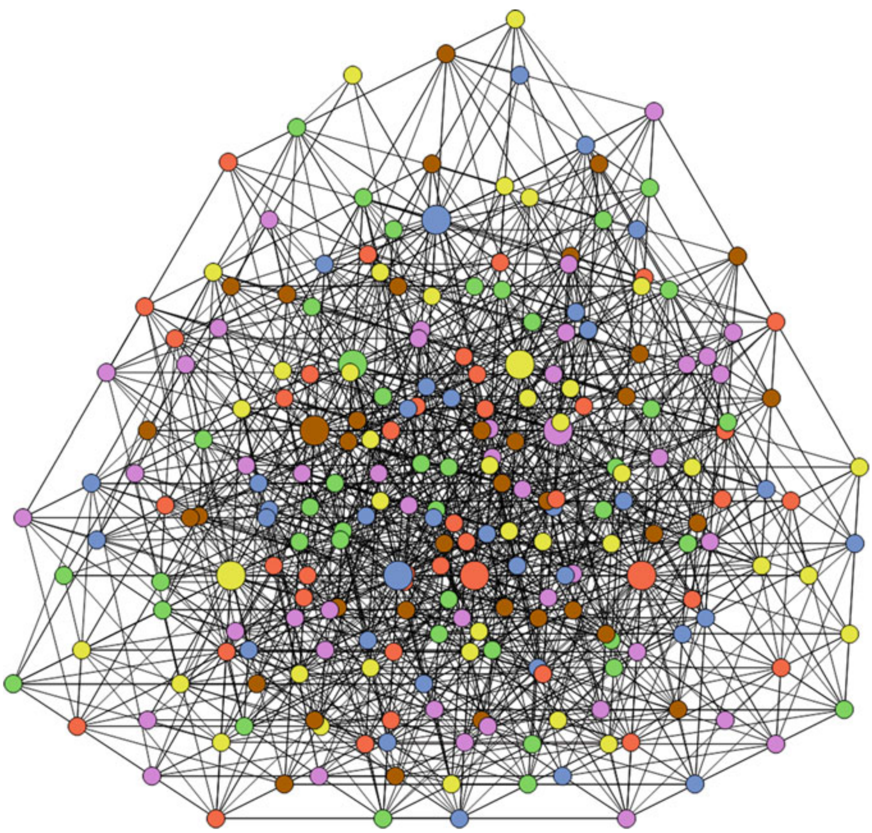


Fig. 60.2 The 6-chromatic 234-vertex odd-distance graph and its 6-coloring. The vertices of the core frame are enlarged.

Parts sums up his essay as follows.

> We confirmed the existence of 6-chromatic ODGs in the plane, determined that it is enough to use edges of four lengths {1, 3, 5, 7}, found a simple construction, changing the parameters of which we can get an infinite family of 6-chromatic 306-vertex ODGs, and even reduced the required number of vertices to 234.

Parts is waiting for Heule to offer a prize before he dives into constructing a 7-chromatic odd-distance graph (ODG):

> One could still dig here. However, for the 7-chromatic ODG, the award has not yet been announced, so this is a pointless exercise. (And here we look at Marijn inquiringly.)

Grapevine brought me news about James Davies, a graduate student at the University of Waterloo, Canada, whom I contacted on October 1, 2022. The same day, he sent me his paper [Dav], which was to appear two days later in arXiv. In it, Davies proves in the positive my Conjecture 60.5:

Theorem 60.7 ([Dav], 2022.) Every finite coloring of the plane contains a monochromatic pair of points at an odd distance from each other.

To my offer, "Welcome to submit this paper and its sequel to *Geombinatorics*," Davies replied on October 2, 2022, as follows:

> Thank you! I still need to decide where to submit, but I am certainly considering *Geombinatorics*.
>
> There is another paper that I do plan to submit to *Geombinatorics* (although admittedly it may be a while before I have the time to write it properly). It is a construction of unit distance graphs in R^d with chromatic number d+2 and arbitrarily large girth, for d=2 I think the construction is simpler than O'Donnell's as the embedding does not require the same case work.

My inquiry about a journal publication of [Dav], James Davies answered on June 5, 2023: "I ended up submitting the odd distance paper to *Geometric and Functional Analysis* back in October [2022], I am awaiting reviews still".

On December 23, 2023, James Davies informed me that this paper has finally been accepted by *Geometric and Functional Analysis* journal.

Gil Kalai, a professor emeritus at the Hebrew University of Jerusalem, while reporting Davies' Theorem, posed the following problem, in which he reproduced Fig. 7.2 that appears in this book:[1]

Challenge 60.8 What is the smallest odd distance between a monochromatic pair in the Hoffman–Soifer coloring of Fig. 7.2?

---

[1] https://gilkalai.wordpress.com/2022/10/19/james-davies-every-finite-colouring-of-the-plane-contains-a-monochromatic-pair-of-points-at-an-odd-distance-from-each-other/

In his recent lecture,[2] Davies formulated two important extensions of Theorem 60.7, by James Davies of Cambridge, Rose McCarty of Princeton, and Michał Pilipczuk of the University of Warsaw:

Theorem 60.9 Let $f(x) = a_n x^n + \cdots + a_0$ be a polynomial with integer coefficients and $a_n \geqslant 1$. Then, every finite coloring of the plane contains a monochromatic pair of distinct points at a distance of $f(x)$ from each other for some integer x.

Theorem 60.10 Every finite coloring of the plane contains a monochromatic pair of points whose distance from each other is a prime number.

As you can see, prime numbers make their first appearance in the coloring world on the plane!

On June 5, 2023, James Davies kindly sent me a present draft with these two theorems [DMP]. He expects to submit it to arXiv in a few weeks, and then to a journal.

Indeed, on December 23, 2023, Davies informed me that their trio submitted this paper to *Israel Journal of Mathematics*.

Dr. James Davies is presently the Gott Research Fellow in Mathematics at Trinity Hall, Cambridge.

60.2 Forbidden Binaries and Factorials[3]

The odd-distance graph problem reminded me of a fantastic problem, used in 2010 in the 27th Colorado (now called Soifer) Mathematical Olympiad. It was proposed by the 1990 and 1991 first prize winner and now professor at Ohio State University Matthew Kahle. There is a two-way bridge: mathematical research provides a rich source for creating original Olympiad problems, and conversely, Olympiad problems often inspire "further explorations," open problems and exciting research work.

Colorful Integers 60.11 (M. Kahle, 2008, [Soi55], [Soi40]).

A. What is the *minimum* number of colors necessary for coloring the set of positive integers so that any two integers which differ by any power of 2 are colored in different colors? (Observe that 1 is a power of two: $2^0 = 1$).
B. What is the *minimum* number of colors necessary for coloring the set of positive integers so that any two integers which differ by any factorial are colored in different colors?

Solution of 60.11.A Clearly 3 colors are necessary, since the numbers 1, 2, 3 pairwise differ by powers of 2 and thus require three distinct colors. On the other hand, coloring the positive integers cyclically modulo 3 does the trick because under this coloring the difference between two numbers of the same color is a multiple of 3, which is never equal to a power of 2. So 3 colors are also sufficient. ■

Solution of 60.11.B The first solution was found by Adam Hesterberg, who was a high school student at the time (he is now a lecturer in computer science at Harvard University) and

[2] https://video.renyi.hu/video/james-davies-odd-distances-in-colourings-of-the-plane-530

[3] The material of Section 60.2 first appeared in [Soi55].

solved the problem in one day. In order to prove the existence of a mysterious "r", Adam used nesting intervals. In mathematics we often prove existence without discovering the value of the existing object. Bob Ewell calculated the value of Adam's "r." However, on October 11, 2016, Matthew Kahle sent me a solution that explicitly determined the value of one such mysterious "r". I choose to present *Kahle's solution*.

Assume 3 colors suffice. Since $1! = 1$ and $2! = 2$, any three consecutive integers must be colored in 3 distinct colors a, b, c. Numbers 1 through 6 must be colored a, b, c, a, b, c. Accordingly, number 7 must be colored a, but this is not allowed because $7 - 1 = 3!$ – a contradiction. Thus, at least four colors are needed.

Suppose for a moment that there exists a number r, (necessarily irrational), such that $n!r$ is in the interval [1, 3] (mod 4), for every positive integer n. We will determine which of the 4 colors to use on the integer k by looking at kr (mod 4): the *color-defining-intervals* [0,1) [1,2) [2,3) [3,4) (mod 4) determine the 4-coloring of the set of integers Z.

Thus defined 4-coloring satisfies the conditions of the problem. Indeed, suppose $|i-j| = n!$ for some n. By multiplying through by r, we get $|ri - rj| = rn!$, which is between 1 and 3 (mod 4). In particular, ri and rj belong to different color-defining-intervals modulo 4, and thus i and j received different colors.

All that is left to prove is the existence of the desired r. The following lemma is useful.

Lemma 60.12 For $k \geq 1$, we have the inequality

$$\frac{1}{(4k)!} + \frac{1}{(4(k+1))!} + \frac{1}{(4(k+2))!} + \ldots \leq \frac{1}{4k-1} \times \frac{1}{(4k-1)!}$$

Proof of Lemma 60.12 Provided that $-1 < R < 1$, the formula for the sum of an infinite geometric series is

$$a + aR + aR^2 + aR^3 + \ldots. = \frac{a}{1-R}$$

Setting $a = 1/(4k)!$, $R = 1/4k$, and comparing term-by-term the factorial series in the statement of the lemma with the geometric series, the result immediately follows. ■

Solution of Problem 60.11.B We claim that the following number r,

$$\begin{aligned} r &= 1 + \frac{1}{4!} + \frac{1}{8!} + \frac{1}{12!} + \ldots \\ &= 1.0416914703416917479394211141\ldots \end{aligned}$$

has the desired property.

Let $n \geq 1$ be an integer. Suppose that k is the smallest integer such that $n < 4k$. Then

$$\begin{aligned} n!r &= n!\Big(1 + \frac{1}{4!} + \frac{1}{8!} + \frac{1}{12!} + \ldots\Big) \\ &= n! + \frac{n!}{4!} + \frac{n!}{8!} + \frac{n!}{12!} + \ldots \\ &= \underbrace{\left(n! + \frac{n!}{4!} + \ldots \frac{n!}{(4(k-1))!}\right)}_{A_n} + \underbrace{\left(\frac{n!}{(4k)!} + \frac{n!}{(4(k+1))!} \cdots\right)}_{B_n} \end{aligned}$$

Observe that A_n is a sum of integers. All but the last of these integers is a multiple of 4. The last integer summand $n!/(4(k-1))!$ is either 1 or 2 (mod 4), depending on n (mod 4). If $n \equiv 0$ or 1 (mod 4), then $n!/(4(k-1))! \equiv 1$ (mod 4), and if $n \equiv 2$ or 3 (mod 4), then $n!/(4(k-1))! \equiv 2$ (mod 4). Therefore, $A_n \equiv 1$ or 2 (mod 4) for every n.

The Lemma implies that $0 < B_n < 1$. Indeed, since $n \leq 4k-1$ we have

$$\begin{aligned} B_n &\leq (4k-1)!\left(\frac{1}{(4k)!} + \frac{1}{(4(k+1))!} + \dots\right) \\ &\leq (4k-1)! \times \frac{1}{4k-1} \times \frac{1}{(4k-1)!} \\ &= \frac{1}{4k-1} \\ &< 1. \end{aligned}$$

Summing up, we conclude that $n!r \in [1,3]$ (mod 4) for every n. ■

This problem has a lovely prehistory. In 1987, Paul Erdős posed the following problem to the well-known Israeli mathematician Yitzhak Katznelson, a Stanford professor, who recollects 14 years later [Kat]:

> In 1987 Paul Erdős asked me if the Cayley graph defined on Z [the set of integers] by a lacunary sequence has necessarily a finite chromatic number. Below is my answer [in the positive], delivered to him on the spot but never published [until 2001].

As usual in my writings, I am naming this result after both contributors, the author of the conjecture and the prover.

The 1987 Erdős–Katznelson Theorem 60.13 Let $\varepsilon > 0$ be fixed and suppose that $S = \{n_1, n_2, \dots, n_j, \dots\}$ is a sequence of positive integers such that $n_{j+1} > (1+\varepsilon)n_j$ for all $j \geq 1$.[4] Define a graph $G = G(S)$ with vertex set Z by letting the pair (n, m) be an edge if and only if $|n-m| \in S$. The chromatic number $\chi(G)$ of G is finite.

Katznelson presented the Erdős conjecture and his proof at a 1991 seminar attended by the young (at that time) Israeli mathematician Yuval Peres.

Peres, currently a professor at the University of California Berkley and a researcher at Microsoft, jointly with Wilhelm Schlag, presently at the University of Chicago, improved [PS] Katznelson's upper bound for the chromatic number in the Erdős problem. From here, I let Matthew Kahle, Professor at Ohio State University and the Colorado Mathematical Olympiad 1990 and 1991 winner, tell the rest of the story. Matt writes to me in the October 12, 2016, e-mail:

> Dear Sasha,
>
> Here is a brief history of this problem. Feel free to extract whatever is interesting for your own creative purposes. :)

[4] Such a sequence is called *lacunary*.

I saw an interesting seminar talk by Yuval Peres (about his joint work with Schlag) when I was a graduate student at University of Washington. He addressed some questions of Erdős about coloring graphs on integers like this. As long as the sequence of distances is increasing at least exponentially fast, the chromatic number is finite, and they can even get an explicit upper bound on the chromatic number. So I asked at the end of his talk about the factorial graph.

Not sure why this was my first question, but I guess it was the first sequence I thought of that grew super-exponentially fast. But the factorial graph is also particularly appealing since no periodic coloring will work, and periodic colorings are what you want to try.

Yuval said he didn't know, but he guessed that their proof would show that it was probably less than 10.

So I asked my office mate at the time, Tristram Bogart, who was also studying combinatorics, what is the chromatic number of the factorial graph. We quickly established that you need at least 4 colors, and applying what we remembered from Peres's proof, we were able to prove that the chromatic number was at most 5. So we knew that chromatic number was either 4 or 5. We bet a beer on the outcome: I bet 4 and Tristram bet 5.

There it stood for a few years, until I asked the question to some bright [high school] students at Canada/USA Mathcamp. I just defined the graph and offered $20 for figuring out what its chromatic number is. But I did not give them any hint. I did not explain the idea of Peres's proof, or what Tristram and I knew so far.

Amazingly, Adam Hesterberg came back the next day and claimed the $20, showing that the chromatic number is exactly 4. His proof of the existence of such a number was very similar to, or the same as Peres's proof, but he was a high school student, rediscovering the methods of the professionals! And he improved on whatever Tristram and I had been doing, because we were only able to find a 5-coloring, so he must have been a little more efficient!

I was happy to pay Adam the $20. I did collect on the beer from Tristram sometime later, and joked that $20 was expensive for a beer (the net result for me), but I did not mind because I was happy that my guess was right.

When I sat down to remember the proof, for your book, I remember being a little bit dissatisfied with the nested intervals. It is a subtle fact of analysis that an infinite sequence of nested closed intervals must have a point of intersection. After all, it fails if the intervals are open instead of closed! So I wondered if I could instead find an explicit r that works, perhaps as an infinite series. Playing around for a few minutes, I found the r in the proof I gave you. Somehow this is more satisfying to me, since it is slightly more elementary, and because the coloring is more explicit.

Adam Hesterberg is currently a Ph.D. student in mathematics at MIT. His research interests include graph theory, computational geometry, and theoretical computer science.

Warm regards,
Matt

The high school student in this story, Adam Hesterberg, soon after won the 2007 USA Mathematical Olympiad. He then graduated from Princeton University, was a graduate student at the Massachusetts Institute of Technology (MIT), and is now a lecturer in computer science at Harvard University.

Matt created this problems and forwarded it to me for the Colorado Mathematical Olympiad. The solution to Problem 60.11.B that you saw was written especially for my book [Soi55] by Matthew Kahle.

Imagine, two days after the Olympiad, I received a remarkable e-mail from the Olympiad judge and the first prize winner of the First (!) Colorado Mathematical Olympiad in 1984 Dr. Russel Schaffer.

Monday, April 26, 2010 3:22 PM
To: Alexander Soifer <asoifer@uccs.edu>
Cc: RWSchaffer@gmail.com
Subject: Alternate Solution to 5b
Alexander,

On the drive back to Wyoming [from Colorado Springs] on Saturday afternoon, I thought a bit about problem 5b and came up with an alternate solution.

Today, I formalized it and wrote it up.

This isn't the simple solution that you asked for on Friday. It is less elegant than the solution that you presented. It is, however, a workmanlike solution based on more straightforward intuition. No flashes of daring brilliance required. The intuition is such that a smart high school student could reasonably come up with it in the allotted time.
Russel

Let me reproduce for you Russel's email attachment.

Russel Schaffer' Traveling Solution of Problem 60.11.B Define the color for each integer $x \geq 0$ to be:

$$c(x) = \left(\sum_{i=0}^{\infty} \left[\frac{x + \sum_{j=1}^{i-1} (4j-1)!}{(4i)!} \right] \right) (mod\, 4)$$

where we use square brackets [ ] to indicate the integer part of a real number. Clearly this defines a coloring with four colors. To demonstrate that no two integers have the same color if they are separated by $n!$ for some integer n, consider $(c(d + n!) - c(d))$ $(mod\ 4)$ for some integers $d \geq 0$ and $n \geq 0$. As a notational convenience, we let $k = [n/4]$.

$$(c(d+n!)-c(d)) \equiv \sum_{i=0}^{k-1}\left(\left\lfloor\frac{d+n!+\sum_{j=1}^{i-1}(4j-1)!}{(4i)!}\right\rfloor-\left\lfloor\frac{d+\sum_{j=1}^{i-1}(4j-1)!}{(4i)!}\right\rfloor\right)$$
$$+\left\lfloor\frac{d+n!+\sum_{j=1}^{k-1}(4j-1)!}{(4k)!}\right\rfloor-\left\lfloor\frac{d+\sum_{j=1}^{k-1}(4j-1)!}{(4k)!}\right\rfloor$$
$$+\sum_{i=k+1}^{\infty}\left(\left\lfloor\frac{d+n!+\sum_{j=1}^{i-1}(4j-1)!}{(4i)!}\right\rfloor-\left\lfloor\frac{d+\sum_{j=1}^{i-1}(4j-1)!}{(4i)!}\right\rfloor\right)$$

The summand on the first line will be congruent to 0 modulo 4 because

$$4k \mid \frac{n!}{(4i)!} \text{ for all } 0 \le i \le k-1.$$

Because $\frac{n!}{4k!}$ is an integer, the summand on the second line equals:

$$\begin{array}{ll} 1 \equiv 1(mod\ 4) & \text{if } n \equiv 0\ (mod\ 4) \\ n \equiv 1(mod\ 4) & \text{if } n \equiv 1\ (mod\ 4) \\ n(n-1) \equiv 2\ (mod\ 4) & \text{if } n \equiv 2\ (mod\ 4) \\ n(n-1)(n-2) \equiv 2\ (mod\ 4) & \text{if } n \equiv 3\ (mod\ 4) \end{array}$$

Because $0 < \frac{n!}{(4i)!} < 1$, for all $i > k$, we know that each term under the summation in the third line must be 0 or 1. The entire summand on the third line is thus 0 or 1 because at most one term under the summation can be non-zero. Assume to the contrary that there are integers a and b, $k < a < b$, such that:

$$d+n!+\sum_{j=1}^{a-1}(4j-1)! \ge x_a(4a)! > d+\sum_{j=1}^{a-1}(4j-1)!$$
$$d+n!+\sum_{j=1}^{b-1}(4j-1)! \ge x_b(4\,b)! > d+\sum_{j=1}^{b-1}(4j-1)!$$

for some integers x_a and x_b. Then we would have:

$$1+\frac{d+\sum_{j=1}^{a-1}(4j-1)!}{n!}\geq\frac{x_a(4a)!}{n!}>\frac{d+\sum_{j=1}^{a-1}(4j-1)!}{n!}$$

$$1+\frac{d+\sum_{j=1}^{a-1}(4j-1)!}{n!}\geq\frac{x_b(4b)!-\sum_{j=a}^{b-1}(4j-1)!}{n!}>\frac{d+\sum_{j=1}^{a-1}(4j-1)!}{n!}$$

Now, the center quantities in both inequalities are both integers, and they are both bounded by an identical pair of values which differ by 1, therefore, they must be equal. But we cannot have

$$x_a(4a)!=x_b(4b)!-\sum_{j=a}^{b-1}(4j-1)!$$

because $(4a)!$ divides the left hand side and all terms of the right-hand side except for $(4a-1)!$. We have thus reached a contradiction and can conclude that at most one term under the summation can be nonzero.

Therefore, in the above expansion of $(c(d+n!)-c(d))\ (mod\ 4)$, we see that the first summand will always be 0, the second 1 or 2, and the third 0 or 1. We conclude that $(c(d+n!)-c(d))\ (mod\ 4)$ is always nonzero. ■

Having finished his solution, Russel continues:

Believe it or not, there is some intuition behind this solution. Consider the coloring where each integer x is assigned the color $x\ (mod\ 4)$. This works just fine for pairs of numbers whose difference is 1, 2, or 6. In fact, not only does it work for 1, 2, and 6 but it also works for 1 + 1, 2 + 1, and 6 + 1.

This gives us room to squeeze in an adjustment to make things work for pairs of numbers that differ by 24. We add 1, modulo 4, to all numbers in the block of 24 integers contained in [24, 47]. We add 2, modulo 4, to all numbers in the next block of 24 integers; add 3, modulo 4, to the following block of 24 integers, and so on.

As observed above, the adjustment did not cause problems with the coloring for pairs whose difference is 1, 2, or 6. And the adjusted coloring also works for pairs whose difference is 24, 120, 720, and 5040. We run into trouble again only when we need to compare pairs of numbers that differ by 8!.

We resolve the problems with 8! by doing another adjustment, adding increasing increments to successive blocks of 8! integers. We need only be careful that no two block boundaries get too close to the block boundaries from the first adjustment. To this end, we begin our adjustment 3! before the end of the first block of 8! integers.

As before, the adjusted coloring works until we reach the next difference of the form $4i!$. This is 12!. Again, we perform an adjustment whose block boundaries are guaranteed not to be too close to any previous block boundaries. We continue in this manner, making adjustments for all differences of the form $4i!$. The coloring given at the head of the first page formalizes an infinite sequence of these adjustments.

Let me repeat one lyrical line from Russel's solution:

Believe it or not, there is some intuition behind this solution.

Indeed, Russel possesses *some intuition*!

In fact, I believe that if Russel were not a senior when in 1984 I started the Colorado Mathematical Olympiad, he would have won as many Olympiads as he were to enter. My fault: I started the Colorado Mathematical Olympiad too late. :)

When a fabulous Problem 60.8.B gets solved, we are inspired to see better, look further, aspire a higher ground. Inspired by his Problem 60.8.B, Matthew Kahle proposes to increase the set of forbidden monochromatic distances in the plane from a singleton {1}, as in CNP, to all factorials {1!, 2!, . . ., n!, . . .}.

Open Factorial Coloring Problem in the Plane 60.14 (*M. Kahle*). Find the minimum number of colors $\chi_F(E^2)$ required for coloring the Euclidian plane E^2 in such a way that no two points of the same color are at a factorial distance (n!) apart.

We do not even know whether $\chi_F(E^2)$ is finite, so you have plenty of enjoyable research to undertake!

Of course, the dimension in this problem can be raised, and thus we find ourselves in space, in the Euclidean n-dimensional space E^n.

Open Factorial Coloring Problem in *n*-Space 60.15 Find the minimum number of colors $\chi_F(E^n)$ required for coloring the Euclidian n-space E^n in such a way that no two points of the same color are at a factorial distance (n!) apart.

Let us not forget Problem 60.11.A simply because it was trivial on the line.

Open Binary Coloring Problem in the Plane 60.16 Find the minimum number of colors $\chi_B(E^2)$ required for coloring the Euclidian plane E^2 in such a way that no two points of the same color are at a binary distance (2^n) apart.

Open Binary Coloring Problem in *n*-Space 60.17 Find the minimum number of colors $\chi_B(E^n)$ required for coloring the Euclidian n-space E^n in such a way that no two points of the same color are at a binary distance (2^n) apart.

These two open problems, perhaps, invite you to have an infinite fun, for the answers to them could be not finite but rather infinite cardinal numbers.

Davies–McCarthy–Pilipczuk [DMP] include, with credit, open Problems 60.14 and 60.16, with the following "conjecturous" comment:

> While we conjecture that both of these problems should have negative answers, due to exponential growth of the forbidden distances, it appears challenging to extend current methods to solve these two problems. Of course, it would be more exciting if either of these two problems have a positive answer [i.e., an infinite chromatic number].

They also include a promising conjecture by Boris Bukh [Buk]:

Conjecture 60.18 (Bukh). Let $A \subset R_{>0}$ be algebraically independent. Then there is a finite coloring of E^2 containing no monochromatic pair of points whose distance is contained in A.

William Gasarch's comments inspire me to pose the following two open problems. The answers are probably infinity, as in the previous four open problems – the time will tell.

Open Perfect Square Coloring Problem in the Plane 60.19 (Soifer 2023). Find the minimum number of colors $\chi_S(E^2)$ required for coloring the Euclidian plane E^2 in such a way that no two points of the same color are at a perfect square distance apart.

Open Perfect Square Coloring Problem in *n*-Space 60.20 (Soifer 2023). Find the minimum number of colors $\chi_S(E^n)$ required for coloring the Euclidian n-space E^n in such a way that no two points of the same color are at a perfect square distance apart.

There is a lively thread of research studying chromatic number $\chi(G[1, d])$ of graphs $G[1, d]$ with all distances of the segment $[1,d]$ forbidden. A recent April 2023 paper [CJW] by the Polish mathematicians Joanna Chybowska-Sokół, Konstanty Junosza-Szaniawski, and Krzysztof Węsek provides both new results and a very fine summary of this direction of inquiry. Enjoy reading it!

Chapter 61
7- and 8-Chromatic Two-Distance Graphs

This chapter is short as it awaits your research, which I hope to see in the not-too-distant future.

It is easy for me to pose the following conjecture because it is weaker than my old 2002 conjecture $\chi(E^2) = 7$. Here, as before, by "the plane" we understand the Euclidean plane.

Conjecture 61.1 (Soifer, 2022). There is a 7-chromatic two-distance graph in the plane.

And now two hard open problems:

Two-Distance Open Problem 61.2 (Soifer, 2022). Construct an 8-chromatic two-distance graph in the plane or prove that one does not exist.

And if the answer to Problem 61.2 is positive, we would like to find the answer to the following super hard problem, or obtain partial results:

Second Two-Distance Open Problem 61.3 (Soifer, 2022). Over all $d > 1$, find a two-distance graph $G = G\{1, d\}$ in the plane of maximum chromatic number $\chi(G) = \Psi$. What are the values of d in graphs that realize Ψ?

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_61

Part XII
Predicting the Future

I never think of the future – it comes soon enough.
– Albert Einstein

Prediction is very difficult, especially about the future.
– Niels Bohr

Chapter 62
What If We Had No Choice?

62.1 Prologue

On the pages of this book, we have seen a variety of approaches used in attempts to settle the chromatic number of the plane problem (CNP). Tools from graph theory (Chapter 17), topology (Chapters 8 and 26), measure theory (Chapter 9), abstract algebra (Chapter 11), and discrete and combinatorial geometry (Chapters 4, 6, and 7) have been tried – and only recently an improvement has been attained in the general case. The range for CNP still remains (too) wide: $\chi = 5, 6,$ or 7.

I wrote years ago that such a wide range was an embarrassment for mathematicians. The 4-color Map-Coloring Problem, for example, from its birth in 1852 (or a bit earlier), had a conjecture: 4 colors suffice. Since 1890, thanks to Percy John Heawood [Hea], we knew that the answer was 4 or else 5. The CNP problem is an entirely different matter. After 70+ years of very active work on the problem, we have not even been able to confidently conjecture the answer. Have mathematicians been so bad, or has the problem been so good? Have we been missing something in our assault on the CNP?

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_62

Saharon Shelah (left) and Alexander Soifer at Paul Erdős' 80th Birthday Conference, Keszthely, Hungary, July 18, 1993.

These were the questions that occupied me as I was flying cross country from Colorado Springs to the Rutgers University of New Jersey in October 2002 for a week of joint research with Saharon Shelah, who in my opinion is a genius of problem solving and a very quick learner (I knew that, for we produced two joint papers on Abelian group theory before, in 1984, when we met in Udine, Italy).[1] Per Saharon's request, I compiled a list of problems we could be interested in working on together and numbered them according to set-theorists' taste, from 0 to 12. Problem 0 read as follows:

0. *What if we had no choice?*

This was a natural question for someone who grew up in the Soviet Union with not much choice: we voted for one candidate per each office, ate whatever food was sold at the moment, and lived wherever we were allowed to live. But of course, I meant here something else that made mathematical sense. Saharon understood me. Did you? No? Let me explain.

Nicolaas G. de Bruijn and Paul Erdős reduced CNP to finite sets in the plane, as we have seen in Chapters 5 and 28. Their famous theorem, obtained, in fact, shortly before Ed Nelson posed CNP, required the Axiom of Choice. This is the choice I referred to in my problem 0 for Saharon and me to ponder:

[1] Ronald L. Graham and Joel H. Spencer [GS2] agree with me: "Shelah is widely regarded as one of the most powerful problem solvers in modern mathematics."

What if we had no Axiom of Choice?

In the absence of the Axiom of Choice, we would not have the De Bruijn–Erdős Theorem, and so CNP would not necessarily be reduced to finite plane sets. In particular, I was interested in the following questions:

What can and should we use in place of the Axiom of Choice?
What results can we prove in this alternative Set Theory?
How would "choiceless" mathematics compare to the mathematics built on choice?

And so Saharon and I met for a week in the Garden State autumn and broke some new ground. Before we look at the outcome of our meeting, I need to offer you an excursion into the Land of Choice.

62.2 The Axiom of Choice and Its Relatives

To choose one sock from each of infinitely many pairs of socks requires the Axiom of Choice, but for shoes the Axiom is not needed.
– Bertrand Russell

At present, set theory has lost its relevance.
– Lev S. Pontryagin[2]

The Axiom of Choice was used implicitly throughout the XIX century. A careful observation would uncover that it was used for proving even such a classic result as the sequential Bolzano–Weierstrass Theorem (every infinite bounded subset of reals has a sequential limit point). In 1904, while proving the Well-Ordering Principle, Ernst Friedrich Ferdinand Zermelo (1871–1953) formalized and for the first time explicitly used the Axiom of Choice [Zer]:

The Axiom of Choice (AC) Every family Φ of nonempty sets has a choice function, i.e., there is a function f such that $f(S) \in S$ for every S from Φ.

The newborn axiom prompted a heated debate in the mathematical world. In trying to defend the axiom, in a series of 1908–1909 papers, Zermelo developed a system of axioms for set theory. It was improved by Adolf Abraham Halevi Fraenkel (1891–1965) in his 1922 [Fra1], [Fra2], and 1925 [Fra3] papers. Finally, in 1928, John von Neumann named it the *Zermelo–Fraenkel Set Theory*, or **ZF** [Neu]. ZF with the addition of the Axiom of Choice was naturally denoted by **ZFC** and named the *Zermelo–Fraenkel–Choice system of axioms*.

The historian of the Axiom of Choice, Gregory H. Moore, opens his remarkable book about the Axiom of Choice as follows [Moo]:

> David Hilbert once wrote that Zermelo's Axiom of Choice was the axiom "most attacked up to the present [1926] in mathematical literature..." To this Abraham Fraenkel later [1958] added that "the axiom of choice is probably the most interesting and, in spite of its

[2] [Pon]. L. S. Pontryagin wrote his *Life of Lev Semenovich Pontryagin ...* with the title modeled after the autobiography of the famous Italian Renaissance sculptor Benevento Cellini. It contains this ridiculous statement that was an attack on A. N. Kolmogorov, who highly valued Set Theory.

late appearance, the most discussed axiom of mathematics, second only to Euclid's axiom of parallels which was introduced more than two thousand years ago."

The Axiom postulated the existence of a choice function, without giving any clue of how to find it. It came, therefore, as no surprise that the Axiom was opposed by constructivists, intuitionists, and other mathematicians, who viewed nonconstructive existence results with great suspicion. Moore [Moo] observes that "Despite this initial widespread distrust, today the vast majority of mathematicians accepts the axiom without hesitation and utilize it in algebra, analysis, logic, set theory, and topology." Yes, I agree: vast majority accepts the Axiom of Choice, and consequently, **ZFC** is the standard foundation of set theory – but is it a good thing for mathematics? A majority – any majority – political, social, mathematical often loses sensitivity that is often so naturally preserved among a minority. We will later look into the consequences of the near universal acceptance of the Axiom of Choice as a part of the foundation of mathematics. Here I will introduce other axioms, and first of all, some weaker versions of the Axiom of Choice.

Many results in mathematics really need just a countable version of choice:

The Countable Axiom of Choice ($\mathbf{AC}_{\aleph_0}$) Every countable family of nonempty sets has a choice function.

Much later, in 1942, Paul Isaac Bernays (1888–1977) introduced the following axiom [Bern]:

The Principle of Dependent Choices (DC) If E is a binary relation on a nonempty set A, and for every $a \in A$ there exists $b \in A$ with aEb, then there is a sequence $a_1, a_2, \ldots, a_n, \ldots$ such that a_nEa_{n+1} for every $n < \omega$.

AC implies **DC** (see, for example, Theorem 8.2 in [Jec]), but not conversely. In turn, **DC** implies $\mathbf{AC}_{\aleph_0}$, but not conversely. **DC** is slightly stronger than $\mathbf{AC}_{\aleph_0}$, but it is a sufficient addition to **ZF** for creating a foundation for the classical Lebesgue Measure Theory. Observe that, in particular, **DC** is sufficient for Falconer's Theorem (Theorem 9.1).

One – unfortunate in my opinion – consequence of the Axiom of Choice is the existence of sets on the line that have no length (I mean, no Lebesgue measure). This "regret" must have given birth to the following axiom:

(LM) Every set of real numbers is Lebesgue measurable.

Assuming the existence of an inaccessible cardinal[3], Robert Martin Solovay (nowadays Professor Emeritus at Berkley), using Paul Cohen's forcing, constructed in 1964 (and published in 1970) a model that proved a remarkable theorem [Sol1]. Mitya Karabash and I introduced [KS] the following term in honor of Robert M. Solovay.

The Zermelo–Fraenkel–Solovay System of Axioms for set theory, which we denote by **ZFS**, is defined as follows:

$$\mathbf{ZFS} = \mathbf{ZF} + \mathbf{AC}_{\aleph_0} + \mathbf{LM},$$

[3] A cardinal κ is called *inaccessible* if $\kappa > \aleph_0$, κ is regular, and κ is strong limit. An infinite cardinal $\aleph_\alpha$ is *regular*, if cf $\omega_\alpha = \omega_\alpha$. A cardinal κ is a *strong limit* cardinal if for every cardinal λ, $\lambda < \kappa$ implies $2^\lambda < \kappa$.

and **ZFS Plus**, or shorter, **ZFS+** stands for

$$\mathbf{ZFS+} = \mathbf{ZF} + \mathbf{DC} + \mathbf{LM}.$$

Now the Solovay Theorem formulates very concisely:

The Solovay Theorem 62.1 **ZFS+** is consistent.[4]

Solovay reports [Sol1] that "the original problem of showing **ZF+LM** consistent was suggested to the author by Paul Cohen." Here is how Paul Joseph Cohen (April 2, 1934–March 23, 2007), the man who completed Kurt Gödel's work and won Fields Medal for it in 1966, described the Solovay Theorem in 1966 [Coh2, p. 142]:

> One of the most interesting results [concerning the relationship of various forms of AC] is due to R. Solovay (as yet unpublished) which says that models *N* can be constructed in which the countable AC holds and yet every set of real numbers is Lebesgue measurable.

Indeed, this is a profound result, which in my opinion offers **ZFS+** as a viable alternative to the classical **ZFC**. In particular, **ZFS+** allows the development of the usual Lebesgue Measure Theory. On April 10, 2003, I asked Professor Solovay whether a stronger result is possible, i.e., whether **ZFS** would suffice for building the Lebesgue Measure Theory. The following day Solovay replied [Sol2]:

> I thought about this in the early 60s. The only theorem for which I needed **DC** was the Radon–Nykodim theorem. But I don't know that there isn't a clever way of getting by with just Countable Choice and proving Radon–Nykodim. I just noticed that the usual proof [found in Halmos] uses **DC**.

The Continuum Hypothesis (**CH**) states that there is no cardinal κ such that $\aleph_0 < \kappa < 2^{\aleph_0}$.

The Generalized Continuum Hypothesis (**GCH**) states that for any infinite cardinal λ there is no cardinal κ such that $\lambda < \kappa < 2^{\lambda}$.

The Axiom of Constructibility (**V=L**) introduced by Gödel in 1940 [Göd2], asserts that every set is constructible, i.e., that every set belongs to the constructible universe L.

Kurt Gödel's (1906–1978) 1940 results [Göd2] combined with Paul J. Cohen's 1963–1964 results [Coh1] prove independence of **AC** (as well as of the Continuum Hypothesis, **CH**, and the Generalized Continuum Hypothesis, **GCH**) from the rest of the axioms of **ZF** set theory.

Saharon Shelah playfully summarizes these developments in his 2003 "Logical Dreams" [She3]:

> In short: The Continuum Problem asks:
>
> How many real numbers are there?
>
> G. Cantor proved: There are more reals than rationals. (In a technical sense: "uncountable," "there is no bijection from R into Q").
>
> The Continuum Hypothesis (CH) says: yes, more, but barely so. Every set A ⊆ R is either countable or equinumerous with R.

[4] Assuming the existence of an inaccessible cardinal.

K. Gödel proved: Perhaps CH holds.
P. Cohen proved: Perhaps CH does not hold.

Kurt Gödel also showed that **ZF + V=L** implies **GCH**, while the founder of the famous Warsaw school of set theory and topology Wacław Franciszek Sierpiński (1882–1969) proved that **ZF + GCH** implies **AC**.

Finally, one can remember Lev S. Pontryagin not only as a fine mathematician and a fine anti-Semite but also as a fool, who took his fight against anything Kolmogorov's to the extreme of such a ridiculous statement as this chapter's tongue-in-cheek epigraph "*At present, set theory has lost its relevance.*" What can be more relevant to mathematics than its very foundation, Set Theory!

62.3 The First Example

The Axiom of Choice differs from other axioms of ZF by postulating the existence of a set . . . without defining it . . . Thus it is often interesting to know whether a mathematical statement can be proved without using the Axiom of Choice.

– Thomas Jech [Jec]

Theories come and go; examples live forever.

– Israel M. Gelfand

October 2002, Rutgers University. My week-long joint work with Saharon Shelah (entertainment of the mathematical kind, really) results in the first surprising example. We dedicate the paper to the memory of our teacher, friend, and coauthor Paul Erdős, on the occasion of his 90th birthday. Let us look together at our example.

Our first task is to expand the definition of the chromatic number.[5] How important is to select a productive definition? Socrates thought highly of this undertaking: "The beginning of wisdom is the definition of terms."[6] And so I took two weeks to "sleep" on the choice of the definition and consulted with my coauthors Saharon Shelah and Mitya Karabash before I chose the simplest definition, the one that came first to my mind. "Simplest" surely is not a detractor – in fact, "simple" and "natural" are attributes of definitions that survive the test of time.

Without the Axiom of Choice, the minimum, and thus the chromatic number of a graph, may not exist. In allowing a system of axioms for set theory not to include the Axiom of Choice, we need to come up with a much broader definition of the chromatic number of a graph than the one we used in Chapter 12 – if we want the chromatic number to exist. In fact, instead of the *chromatic number*, we ought to talk about the *chromatic set*. There are several meaningful ways to define it. I am choosing the following definition.

[5] It *is* the first task, but we did not think of it then, and so this definition appeared for the first time in print in the first 2009 edition [Soi44] of this book.

[6] Quoted from [Pet], p. 494.

Definition 62.1 Let G be a graph and $\boldsymbol{A}$ a system of axioms for set theory. *The set of chromatic cardinalities* $\chi^{\boldsymbol{A}}(G)$ of G is the set of all cardinal numbers $\tau \le |G|$ such that there is a proper coloring of the vertices of G in τ colors and τ is minimum with respect to this property.

As you can easily see, the set of chromatic cardinalities does not have to have just one element as was the case when $\boldsymbol{A} = \mathbf{ZFC}$. It can also be empty.

The advantage of this definition is its simplicity. Best of all, we can use inequalities on sets of chromatic cardinalities as follows.

The *inequality* $\chi^{\boldsymbol{A}}(G) > \beta$, where β is a cardinal number, means that for every $\alpha \in \chi^{\boldsymbol{A}}(G)$, $\alpha > \beta$. The inequalities $<$, $\le$, *and* $\ge$ are defined analogously. We also agree that the cardinality of the empty set is greater than or equal to the cardinality of any set.[7] Finally, if β is a cardinal number, $\chi^{\boldsymbol{A}}(G) = \beta$ means that $\chi^{\boldsymbol{A}}(G) = \{\beta\}$.

I would like to introduce simple generalizations of the notion of unit-distance graph.

A *Distance Graph* is a graph with vertex set $V \subseteq E^n$ for some n, where two vertices v_1,v_2 are adjacent if and only if the distance $|v_1v_2|$ belongs to a fixed set S of distances. In particular, when $S = \{1\}$, we get a unit-distance graph.

A *Difference Graph* is a graph with vertex set $V \subseteq E^n$ for some n, with two vertices v_1,v_2 adjacent if and only if their difference $v_1 - v_2$ belongs to a fixed set $S \subseteq E^n$ of differences. Of course, on the line distance graphs and difference graphs coincide.

As always, Z, Q, and R stand for the sets of integers, rationals, and reals, respectively. We are now ready for the first example, which will demonstrate how dramatically the chromatic number of a simple graph we construct depends upon the system of axioms for set theory: it is just 2 in **ZFC** and uncountable in **ZFS**. Let us construct this surprising example and then prove its properties.

Example 62.2 (Shelah–Soifer 2003, [SS1]). We define a graph G as follows: the set R of real numbers serves as the vertex set, and the set of edges is $\left\{(s,t)\colon s - t - \sqrt{2} \in Q\right\}$.

Result 62.3 (Shelah–Soifer 2003, [SS1]). For the distance graph G on the line, $\chi^{\mathbf{ZFC}}(G) = 2$, while $\chi^{\mathbf{ZFS}}(G) > \aleph_0$.

Claim 1 of 62.3: $\chi^{\mathbf{ZFC}}(G) = 2$.

Proof Let $S = \left\{q + n\sqrt{2}; q \in Q, n \in Z\right\}$. We define the equivalence relation E on R as follows: $sEt \Leftrightarrow s - t \in S$.

Let Y be a set of representatives for E (in choosing representatives, we are using the Axiom of Choice). For $t \in R$ let $y(t) \in Y$ be such that $tEy(t)$. We define a 2-coloring $c(t)$ as follows: $c(t)=l$, $l=0,1$ if and only if there is $n \in Z$ such that $t - y(t) - 2n\sqrt{2} - l\sqrt{2} \in Q$. ■

Without **AC** the chromatic situation changes dramatically:

Claim 2 of 62.3: $\chi^{\mathbf{ZFS}}(G) > \aleph_0$.

We will simplify the proof if we acquire a useful tool first.

[7] I know, this convention seems to be counterintuitive, but it is handy, convenient, and allows to prove meaningful results, as you will soon see.

Tool 62.4 If $A \subseteq [0,1)$ and A contains no pair of adjacent vertices of G, then A is null (of Lebesgue measure zero).

Proof Assume to the contrary that A contains no pair of adjacent vertices of G yet A has a positive measure. Then there is an interval I such that[8]

$$\frac{\mu(A \cap I)}{\mu(I)} > \frac{9}{10} \tag{62.1}$$

Choose $q \in Q$ such that $\sqrt{2} < q < \sqrt{2} + \frac{\mu(I)}{10}$.

Let $B = A - \left(q - \sqrt{2}\right) = \left\{x - q + \sqrt{2} : x \in A\right\}$. Then

$$\frac{\mu(B \cap I)}{\mu(I)} > \frac{8}{10}. \tag{62.2}$$

Inequalities (62.1) and (62.2) imply that there is $x \in I \cap A \cap B$. Since $x \in B$, we have $y = x + \left(q - \sqrt{2}\right) \in A$. So, both $x, y \in A$ and $x - y - \sqrt{2} = -q \in Q$. Thus, $\{x, y\}$ is an edge of the graph G with both endpoints in A, which is the desired contradiction. ■

Proof of Claim 2 of 62.3 Assume that the vertices of the graph G are properly colored in $\aleph_0$ colors (i.e., the adjacent vertices are colored in different colors), and $A_1^1, \ldots, A_n^1, \ldots$ are the corresponding monochromatic sets. Let $A_n = A_n^1 \cap [0,1)$ for every $n < \omega$. Since $\mu(\bigcup_{n<\omega} A_n) = \mu([0,1)) = 1$ and Lebesgue measure is a countably additive function in $\mathbf{AC}_{\aleph_0}$, there is a positive integer n such that A_n has a positive measure. By tool 62.4, A_n contains a pair of adjacent vertices of G, which contradict the assumption that the graph is properly colored. ■

Remark This example points out the circumstances in which the presence or the absence of **AC** could dictate the value of the chromatic number of the plane and many other characteristics of a variety of structures in mathematics (and consequently in physics).

62.4 Examples in the Plane

As the main object of our interest has been the good ole Euclidean plane, we aspired to construct a difference graph G_2 in the plane R^2, and thus come closer to the setting of the chromatic number of the plane problem. The chromatic number of the constructed below graph G_2 is 4 in **ZFC** and uncountable in the Zermelo–Fraenkel–Solovay system of axioms **ZFS**.

Saharon Shelah and I believe the example (and its analog G_3, presented here as well) may prove to serve as an important illumination and inspiration in this area of research.

As Thomas Jech [Jec] observes, in the Solovay model, every set of reals differs from a Borel set by a set of measure zero.

[8] The (Lebesgue) measure $\mu(S)$ of a set S is defined in Chapter 9.

Example 62.5 (Soifer–Shelah 2003, [SS2]). We define graph G_2 as follows: the set R^2 of points in the plane serves as the vertex set, and the set of edges is the union of the four sets $\{(s,t) : s,t \in R^2; s - t - \varepsilon \in Q^2\}$ for $\varepsilon = (\sqrt{2}, 0)$, $\varepsilon = (0, \sqrt{2})$, $\varepsilon = (\sqrt{2}, \sqrt{2})$, and $\varepsilon = (-\sqrt{2}, \sqrt{2})$ respectively.[9]

Result 62.6 (Soifer–Shelah 2003, [SS2]). For the difference graph G_2 in the plane, $\chi^{\mathbf{ZFC}}(G) = 4$, while $\chi^{\mathbf{ZFS}}(G) > \aleph_0$.

Claim 1 of 62.6: $\chi^{\mathbf{ZFC}}(G) = 4$.

Proof Let $S = \{(q_1 + n_1\sqrt{2}, q_2 + n_2\sqrt{2}) : q_i \in Q, n_i \in Z\}$. We define an equivalence relation E on R^2 as follows: $sEt \Leftrightarrow s - t \in S$.

Let Y be a set of representatives for E (we can choose them due to the Axiom of Choice). For $t \in R^2$ let $y(t) \in Y$ be such that $tEy(t)$. We define a 4-coloring $c(t)$ as follows: $c\ (t) = (l_1, l_2)$, $l_i \in \{0, 1\}$ if and only if there is a pair $(n_1, n_2) \in Z^2$ such that $t - y(t) - 2\sqrt{2}(n_1, n_2) - \sqrt{2}(l_1, l_2) \in Q^2$. ■

Claim 2 of 62.6: $\chi^{\mathbf{ZFS}}(G) > \aleph_0$.

Proof We create a tool similar to tool 62.4 and then prove the claim 2 similarly to its counterpart of Result 62.3. ■

We can define the edges of the graph differently.

Example 62.7 (Soifer–Shelah 2003, [SS2]). The set R^2 of points in the plane serves as the vertex set for G_3, and the set of edges is the union of the *two* sets $\{(s,t) : s,t \in R^2; s - t - \varepsilon \in Q^2\}$ for $\varepsilon = (\sqrt{2}, 0)$ and $\varepsilon = (0, \sqrt{2})$ respectively.

Result 62.8 (Soifer–Shelah 2003, [SS2]). For the difference graph G_3 in the plane, $\chi^{\mathbf{ZFC}}(G) = 2$, while $\chi^{\mathbf{ZFS}}(G) > \aleph_0$.

Claim 1 of 62.8: $\chi^{\mathbf{ZFC}}(G) = 2$.

Proof Let $S = \{(q_1 + n_1\sqrt{2}, q_2 + n_2\sqrt{2}) : q_i \in Q, n_i \in Z\}$. We define an equivalence relation E on R^2 as follows: $sEt \Leftrightarrow s - t \in S$.

Let Y be a set of representatives for E. For $t \in R^2$ let $y(t) \in Y$ be such that $tEy(t)$. We define a 2-coloring $c(t)$ as follows: $c(t) = (\varepsilon_1 + \varepsilon_2)_{mod2}$ if and only if there is a pair $(\varepsilon_1, \varepsilon_2) \in Z^2$ such that $t - y(t) - \sqrt{2}(\varepsilon_1, \varepsilon_2) \in Q^2$. ■

Claim 2 of 62.8: $\chi^{\mathbf{ZFS}}(G) > \aleph_0$.

Proof is similar to the one presented for G in Result 62.3. ■

One may wonder what is so special about $\sqrt{2}$ in our constructions. Well, $\sqrt{2}$ is the oldest known irrational number: a proof of its irrationality, apparently, comes from the Pythagoras School. Our reasoning and results would not change if we were to replace $\sqrt{2}$ everywhere with another irrational number.

[9] Q^2, of course, denotes the rational plane.

62.5 Examples in Space

Space isn't remote at all. It's only an hour's drive away if your car could go straight upwards.

– Fred Hoyle

Ideas developed above are extended here to construct difference graphs in the real n-dimensional space R^n, whose chromatic number is a positive integer in **ZFC**, and is uncountable in **ZFS**.

Example 62.9 (Soifer 2005, [Soi23]). Define a difference graph G_n as follows: the set R^n of points of the n-space serves as the vertex set, and the set of edges is $\bigcup_{i=1}^{n}\left\{(s,t) : s,t \in R^n; s-t-\sqrt{2}\varepsilon_i \in Q^n\right\}$ where ε_i are the n unit vectors on the coordinate axes forming a standard basis of R^n. For example, $\varepsilon_1 = (1,0,\ldots,0)$ – we will use this vector in the proof of *Claim* 2 below.[10]

Result 62.10 (Soifer 2005, [Soi23]). For the difference graph G_n, $\chi^{\mathbf{ZFC}}(G_n) = 2$, while $\chi^{\mathbf{ZFS}}(G_n) > \aleph_0$.

Claim 1 of 62.10: $\chi^{\mathbf{ZFC}}(G) = 2$.

Proof Let $S = \left\{q + m\sqrt{2} : q \in Q^n, m \in Z^n\right\}$. We define an equivalence relation E on R^n as follows: $sEt \Leftrightarrow s - t \in S$.

Let Y be a set of representatives for E. For $t \in R^n$ let $y(t) \in Y$be such a representative that $tEy(t)$. We define a 2-coloring $c(t)$ as follows: $c(t) = \ \| k\|_{mod2}$ iff there is $k \in Z^n$ such that $t - y(t) - \sqrt{2}k \in Q^n$, where $\|k\|$ denotes the sum of all n coordinates of k. ■

Claim 2 of 62.10: $\chi^{\mathbf{ZFS}}(G) > \aleph_0$.

The proof is similar to the one of Result 62.3 – we just need an "n-dimensional tool."

Tool 62.11 If $A \subseteq [0,1)^n$ and A contains no pair of adjacent vertices of G, then A is null (of Lebesgue measure zero).

Proof Assume to the contrary that $A \subseteq [0,1)^n$ contains no pair of adjacent vertices of G_n, yet A has positive measure. Then there is an n-dimensional parallelepiped I, with a side parallel to the first coordinate axis of length, say, a, such that

$$\frac{\mu(A \cap I)}{\mu(I)} > \frac{9}{10} \tag{62.3}$$

Choose $q \in Q$ such that $\sqrt{2} < q < \sqrt{2} + \frac{a}{10}$. Define a translate B of A as follows:

[10] Z^n is a set of integral n-tuples and Q^n is the rational n-space.

$$B = A - \left(q - \sqrt{2}\right)\varepsilon_1$$

Then

$$\frac{\mu(B \cap I)}{\mu(I)} > \frac{8}{10} \tag{62.4}$$

Inequalities (62.3) and (62.4) imply that there is $v \in I \cap A \cap B$. Since $v \in B$, we have $w = v + \left(q - \sqrt{2}\right)\varepsilon_1 \in A$. So, we have v, $w \in A$ and $v - w - \sqrt{2}\varepsilon_1 = -q\varepsilon_1 \in Q^n$. Thus, $\{v, w\}$ is an edge of the graph G with both endpoints in A, which is the desired contradiction. ■

We can certainly vary the definition of edges to get new graphs.

Example 62.12 (Soifer 2005, [Soi23]). Define graph *G*' as follows: the set R^n of points of the *n*-space still serves as the vertex set, but the set of edges is $\bigcup_{0 \le i \ne j \le n} \left\{(s, t) : s, t \in R^n; s - t - \sqrt{2}\left(\varepsilon_i - \varepsilon_j\right) \in Q^n\right\}$ where ε_i , $i = 1, \ldots, n$ are the n unit vectors on coordinate axes forming the standard basis of R^n, and $\varepsilon_0 = 0 \in R^n$.

Result 62.13 (Soifer 2005, [Soi23]). For the difference graph G_n, $\chi^{\mathbf{ZFC}}(G_n) = 2^n$, while $\chi^{\mathbf{ZFS}}(G_n) > \aleph_0$.

Claim 1 of 62.13: $\chi^{\mathbf{ZFC}}(G) = 2^n$.

Proof Indeed, the 2^n vertices of the n-dimensional unit cube generated by ε_i, $0 \le i \le n$ must all be colored in different colors, so 2^n colors are obviously needed.

Let Y be a set of representatives for E. For $t \in R^n$ let $y(t) \in Y$ be a representative such that $tEy(t)$. We define a 2^n-coloring $c(t)$ as follows: $c(t) = \left(k^1_{mod\,2}, k^2_{mod\,2}, \ldots, k^n_{mod\,2}\right)$ iff there is $k = (k^1, k^2, \ldots, k^n) \in Z^n$ such that $t - y(t) - \sqrt{2}k \in Q^n$, where $k^i_{mod2} \in \{0, 1\}$ is the remainder upon division of k^i by 2 for $i = 1, 2, \ldots, n$. ■

Claim 2 of 62.13: $\chi^{\mathbf{ZFS}}(G) > \aleph_0$.

Proof closely follows the one for claim 2 of Result 62.10. ■

Observe It is certainly possible to construct other examples of difference graphs in R^n whose chromatic number in **ZFC** is any integer between 2 and 2^n and is uncountable in **ZFS**.

These examples illuminate the influence of the system of axioms for set theory on combinatorial results. They also suggest that the chromatic number of the Euclidean space E^n may not exist "in the absolute" (i.e., in **ZF**), but depends upon the system of axioms we choose for set theory. The examples we have seen naturally pose the following open problem:

Open AC Problem 62.14 For which values of n is the chromatic number $\chi(E^n)$ of the n-space E^n is defined "in the absolute," i.e., in **ZF** regardless of the addition of the Axiom of Choice or its relative?

Chapter 63
AfterMath and the Shelah–Soifer Class of Graphs

63.1 Shelah–Soifer Graphs

In 1900s–1930s, the foundations of set theory dominated mathematicians' interests. Nowadays, the interest in the foundations in general, and in the Axiom of Choice in particular, has diminished outside set theorists and logicians. Most mathematicians have settled on the **ZFC**-based mathematics. Shelah–Soifer papers seemed to "strike the mathematical heart" [Del]. They received a remarkable critique [Del] by Jean–Paul Delahaye[1], a complimentary mention in Ronald L. Graham's articles [Gra5], [Gra6], [Gra7], and [Gra8], entered the column by Joseph O'Rourke [Oro] and were the subject of the column [Szp] by George Szpiro in the newspaper in Zürich, the city where Van der Waerden lived for 45 years. It inspired a series of works by various authors. We will look at one such paper in the next section. Another example was forwarded to me by Professor Branko Grünbaum in February 2005:

From: Janos Pach pach@CIMS.nyu.edu
Date: February 27, 2005 8:21:56 PM PST
To: eokoh@gc.cuny.edu, dlazarus@erols.com, sarioz@acm.org, aushakov@mail.ru, herr_strangelove@yahoo.com, mlaufer@gc.cuny.edu, tswaine@gc.cuny.edu, syuan@gc.cuny.edu, msilva@gc.cuny.edu, dmussa@gc.cuny.edu, jharlacher@gc.cuny.edu, Eva@Antonakos.net, mmunn@gc.cuny.edu, raghavan@cs.nyu.edu
Cc: RLandsman@gc.cuny.edu (Robert Landsman)
Subject: Combinatorial Comp. Seminar on Wednesday
SEMINAR ON COMBINATORIAL COMPUTING
March 2, Wednesday, 6:30pm
Room 6417, Graduate Center, 365 Fifth Avenue, NY

[1] Jean-Paul Delahaye is a professor of computer science and mathematics, specializing in Complexity Theory, Computational Finance, Computational and Combinatorial Games, Modeling, Simulation, etc. He is the winner of the 1998 Prix d'Alembert from the Société mathématique de France, the winner of the 2012 Kuhlmann Prize from the Society of Sciences, Agriculture and the Arts of Lille, Inspecteur général de l'éducation nationale honoraire, author of several hundred articles and a few books.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_63

INDEPENDENCE IN EUCLIDEAN RAMSEY THEORY
Jacob Fox, Massachusetts Institute of Technology

In this talk, I will present several remarkable new developments on independence in Euclidean Ramsey theory. S. Shelah and A. Soifer recently constructed a graph on the real line with chromatic number 2 in the Zermelo–Fraenkel–Choice (ZFC) system of axioms, but with uncountable chromatic number (if it exists) in a consistent system of axioms with limited choice, studied by Solovay in 1970. Motivated by these recent results, Radoicic and I discovered that the statement ``every 3-coloring of the non-zero real numbers contains a monochromatic solution to the equation x_1+2x_2–4x_3=0" is independent of the Zermelo–Fraenkel axioms for set theory. A system L: Ax=0 of linear homogeneous equations is called a-regular over R if every a-coloring of the real numbers contains a monochromatic solution to L in distinct variables. In 1943, Rado classified those L that are a-regular over R for all finite a. In ZFC, if a is an infinite cardinal, we classify those L that are a-regular. This classification depends on the cardinality of the continuum. In the Solovay model, we classify those L that are aleph_0-regular over R. We also leave several problems concerning the chromatic number of graphs on Euclidean space.

To the best of my knowledge and literature search, the 1970 fundamental work by Robert M. Solovay has been cited in set theoretic works for decades but has not been known to or used in combinatorics and Ramsey Theory before [SS1] appeared in 2003. Inspired by our surprising results, Solovay's work and what Mitya Karabash and I named Zermelo–Fraenkel–Solovay System of Axioms **ZFS**, the comparative study of **ZFC** vs. **ZFS** has entered a number of recent combinatorial works, for example, by Jacob Fox and Rados Radoicic [FR], Boris Alexeev, Jacob Fox and Ronald L. Graham [AFG], and Boris Bukh [Buk].

In February 2007, I wrote about my excitement to Ron Graham:

Hi Ron,

I downloaded and enjoyed your latest paper with Alexeev and Fox.

It is an added pleasure that apparently Shelah and I motivated you guys to compare ZFC and Solovay's axioms for sets.

His 1970 work has always been appreciated by logicians, but it seems that Saharon and I were first to use it as a tool for study of combinatorics.

I would also like to read Radoicic-Fox, which you quote in your soon-to-be published survey.

Ron responded on February 12, 2007:

Hi Sasha,

I don't have a copy of the Radoicic–Fox paper, but it should be easy to get since Fox is a grad student at Princeton and Radoicic is working in New York.

Best regards,

Ron

Ron [Gra8] summarizes this group's results as follows:

An interesting phenomenon has been recently observed by Fox, Radoicic, Alexeev and the author [FR], [AFG] which shows how the axioms of set theory can affect the outcome of some of these questions. For example, consider the linear equation E: x +

$y + z - 4w = 0$. This is certainly not partition regular, and in fact, there is a 4-coloring of the integers which prevents E from having any (nontrivial) monochromatic solution. However, suppose we changed the question and asked whether E has monochromatic solutions in reals for every 4-coloring of the reals. It can be shown that in ZFC, there exist 4-colorings of the reals for which E has no monochromatic solution. However, if we replace the Axiom of Choice (the "C" in ZFC) by LM . . ., then in the system ZF+LM (which is consistent if ZFC is), the answer is yes. In other words, in this system every 4-coloring of the reals always contains a nontrivial monochromatic solution to E. On the other hand, this distinction does not occur for the equation $x + y - z = 0$, for example.

Dmytro (Mitya) Karabash, in 2006, an undergraduate student at Columbia University and a fine mathematician, coined the term for the class of graphs Shelah and I stumbled on:

Definition 63.1 (Dmytro Karabash). *The Shelah–Soifer class S of graphs* consists of graphs G, for which $\chi^{ZFS}(G) \cap \chi^{ZFC}(G) = \varnothing$. Let S^c stand for the complement, i.e., the class of graphs which are not Shelah–Soifer graphs.

Mitya and I then looked into what causes a graph to belong to this class, and "how many" Shelah–Soifer graphs there are. The following results come from our joint 2007 paper [KS].

Definition 63.2 ([KS]). Let d be the Euclidian metric in R^n. The distance set between $A, B \subseteq R^n$ is defined as follows: $d(A, B) = \{d(x, y) : x \in A, y \in B\}$.

Definition 63.3 ([KS]). Let $D \subseteq R^+ = (0, \infty)$. The symbol G_D^n stands for the graph with the vertex set R^n and the edge set $\{(x, y) : d(x, y) \in D\}$.

The following result addresses the newly defined in this chapter notion of the set of chromatic cardinalities $\chi^{ZFS}(G_D^n)$ of the graph G_D^n.

Theorem 63.4 ([KS]). If for $D \subseteq R^+$, 0 is a limit point of D in R, then $\chi^{ZFS}(G_D^n) > \aleph_0$.

We can prove Theorem 63.4 using an argument analogous to Proof of Claim 2 in Result 62.3. For the sake of diversifying our tools, we will use instead the old result of Hugo Steinhaus (1887–1972):

The Steinhaus Lemma 63.5 ([Stei]). If $A \subseteq R$, is a set of positive Lebesgue measure, then the set $A - A = \{x - y: x, y \in A\}$ contains a ball around 0.

Proof of Theorem 63.4 Let us argue by contradiction: suppose $\chi^{ZFS}(G_D^n) \le \aleph_0$. Then there exists a countable proper coloring $c: R^n \to N$ of G_D^n. Look at the monochromatic sets $A_i = \{x \in R^n \mid c(x) = i\}$. Since all sets in R^n are measurable in ***ZFS***, we get $\sum_{i=1}^{n} \mu(A_i) = \infty$. Hence, there exists $i \in N$ such that $\mu(A_i) > 0$. Thus, there exists a set of positive measure $A \subseteq R^n$ such that $d(A, A) \cap D = \varnothing$. We reduce to case $n = 1$ by observing that there must exist a line $L \subseteq R^n$ such that $A \cap L$ has positive measure in L by the product measure theorem (see, for example, Theorem 2.36 in [Foll]). Now we apply the Steinhaus Lemma 63.5 to see that $d(A, A)$ contains some interval $[0, \varepsilon)$ and since 0 is a limit point of D, we get $d(A,A) \cap D \ne \varnothing$. ■

Definition 63.6 ([KS]). Set $D \subseteq R$ is called *integrally independent mod 2* if for any $n \in N$, $a_1, a_2, \ldots, a_n \in Z$ and $s_1, s_2, \ldots, s_n \in D$, the equality $\sum_{i=1}^{n} a_i s_i = 0$ implies $2 \mid \sum_{i=1}^{n} a_i$.

Theorem 63.7 ([KS]). *If* $D \subseteq R^+$, $|D| \le \aleph_0$, *then* $\chi^{ZFC}(G_D^1) \le \aleph_0$, i.e., for every $\alpha \in \chi^{ZFC}(G_D^1)$, $\alpha \le \aleph_0$. If in addition D is integrally independent mod 2, then $\chi^{ZFC}(G_D^1) = \{2\}$.

Proof For any $p\in R$, p lies in the connected component C of G_D^1,

$$C = C(p) = \{x : \exists n > 0, \exists \{x_i\}_{i=1}^n \subseteq R, \text{such that}$$

$$|x_i - x_{i+1}| \in D, x_0 = x, x_n = p\} = \bigcup_{i=1}^{\infty} C_i,$$

where $C_i = C_i(p)$ are defined inductively by $C_0(p) = \{p\}$ and $C_i = C_{i-1} + D = \{x+y \mid x\in C_{i-1}, y\in D\}$. Since $|D|\le\aleph_0$, for every i, $|C_i|\le\aleph_0$ and hence $|C|\le\aleph_0$. Thus, we can color the component in $\aleph_0$ colors by coloring each point in a different color. The Axiom of Choice allows us to similarly color all other components. But the chromatic number of the graph is the supremum of the chromatic numbers of its components and hence the first statement of the theorem is proved.

If D is integrally independent mod 2, then $C_i\cap C_j = \varnothing$ if and only if $i - j$ is even. Hence, coloring each C_i according to its parity is a well-defined 2-coloring. ■

Theorem 63.8 ([KS]). For the set of graphs $H = \{G_D^1 : D\subseteq E^n\}$, we have $|H \cap S| = |H \cap S^c|$, where S is the class of Shelah–Soifer graphs and S^c is its complement.

To prove this theorem, let us first prove two tools:

Tool 63.9 ([KS]). If $D\subseteq (\varepsilon, \infty)$ for some $\varepsilon > 0$, then $\chi^{ZF}(G_D^n) \le \aleph_0$.

Proof We obtain a proper $\aleph_0$-coloring of G_D^n by cutting E^n into n-cubes of side $\frac{\varepsilon}{\sqrt{n}}$ and coloring each cube into a different color. ■

Tool 63.10 ([KS]). Let G_1, G_2 be graphs with the same vertex set V and $G = G_1\cup_{edge}G_2$ be their edge union. Suppose that G_1, G_2 are countably colorable in an axiomatic system $\boldsymbol{A}$, i.e., $\chi^{\boldsymbol{A}}(G_i) \le \aleph_0$ for $i = 1,2$. Then, G is countably colorable in $\boldsymbol{A}$, i.e., $\chi^{\boldsymbol{A}}(G) \le \aleph_0$.

Proof Let c_1, c_2 be proper $\aleph_0$-colorings of G_1, G_2, respectively. Consider the coloring $c_1\oplus c_2$ of G into $\aleph_0^2$ colors defined by $c_1\oplus c_2(v) = (c_1(v), c_2(v))$. This is clearly a proper coloring of G that uses $\aleph_0^2 = \aleph_0$ colors. ■

Proof of Theorem 63.8 Let $H_D = \{G_{D\cup E}^1 : E\subseteq [1, \infty)\}$.

1) First consider $D\subseteq (0, 1)$ such that $|D| \le \aleph_0$ and 0 is a limit point of D. Theorem 63.7 implies that $\chi^{ZFC}(G_D^1) \le \aleph_0$ and Tool 63.9 implies that $\chi^{ZFC}(G_E^1) \le \aleph_0$. Hence, by Tool 63.10, $\chi(G_{D\cup E}^1) = \chi(G_D^1\cup_{edge} G_E^1) \le \aleph_0$. On the other hand, for every $G_{D\cup E}^1 \in H_D$, $\chi^{ZFS}(G_{D\cup E}^1) > \aleph_0$ by Theorem 63.4. Hence, $H_D \subseteq S$.
2) Since $|[1,\infty)| = |R^+|$ we get $|H_D| = 2^{|[1,\infty)|} = 2^{|R+|} = |S|$ and hence $|H| = |H\cap S|$.
3) Now consider $D = (0, 1]$. Since each graph $G_{D\cup E}^1$ in H_D contains a complete subgraph with the vertex set of the cardinality of continuum, we get $\chi^{ZFS}(G_{D\cup E}^1) = \chi^{ZFC}(G_{D\cup E}^1) = |c|$, where c is continuum. Hence, $H_D\subseteq S^c$ and similarly to argument 2) above, $|H|=|H\cap Sc|$. ■

Theorem 63.8 suggests that the class S is "as big as" the class S^c, whatever "as big" means. Let us make it formal.

Definition 63.11 ([KS]). Let θ be a class of all graphs and α a cardinal number. We define a set θ_α as follows:

$$\theta_\alpha = \{G \in \theta : |V(G)| \le \alpha\},$$

where $V(G)$ is the vertex set of G, i.e., θ_α be the set of graphs with cardinality of vertex set not exceeding α.

We conjecture:

Conjecture 63.12 ([KS]). For any cardinal $\alpha > \aleph_0$, $|\theta_\alpha \cap S| = |\theta_\alpha \cap S^c|$.

63.2 A Unit-Distance Shelah–Soifer Graph

On July 10, 2007, my first day home after the 3-day cross-country rally brought me from Princeton to my home in Colorado Springs, I received the following e-mail from Melbourne, Australia:

Dear Professor Soifer,

I am a student from Monash Uni[versity] in Australia and I have done some work on the chromatic number of the plane problem. I found your various publications on the topic extremely helpful. I particularly liked your recent work with Saharon Shelah and as part of my [Honours] bachelor's thesis I found another example of a graph with 'ambiguous' chromatic number. This graph is a unit-distance graph so it may be considered even further evidence that the plane chromatic number may also be ambiguous as you have suggested. It has been submitted for review but you can find a pre-print of it here http://arxiv.org/abs/0707.1177 if you are interested. As you will notice, I am greatly indebted to your work since my proof is essentially analogous to yours.

Kind regards,

Michael Payne

Indeed, the paper Michael submitted to arXiv the day before his e-mail to me contained a fabulous example. He starts with unit-distance graph G_1 whose vertex set is the rational plane Q^2 and, of course, two vertices are adjacent if and only if they are distance 1 apart.

Example 63.13 (Payne [Pay]). The desired unit-distance graph G on the vertex set R^2 is obtained by the tiling of the plane by translates of the graph G_1, i.e., its edge set is

$$\left\{(p_1, p_2) : p_1, p_2 \in R^2; p_1 - p_2 \in Q^2; |p_1 - p_2| = 1\right\}.$$

Claim 1: $\chi^{\mathbf{ZFC}}(G) = 2$.

Proof By Woodall's result 11.2, the chromatic number of the graph G_1 is equal to 2. Since the graph G consists of nonconnected components – tiles – each of which is isomorphic to G_1, the whole graph G is also 2-colorable (the Axiom of Choice is used to select "origin points" for 2-coloring of each tile). ■

Claim 2: $3 \le \chi^{\mathbf{ZFS}}(G) \le 7$.

Michael Payne shows first that any measurable set S of positive (Lebesgue) measure contains the endpoints of a path of length 3 in G. Of course, this would rule out 2-coloring of S. Payne continues: "We can then proceed in a similar fashion to Shelah and Soifer's proof in [SS1]." Let us look at his proof.

Tool 63.14 ([Pay]). For any point $p \in R^2$ and any $\varepsilon > 0$, there is $q \in Q$ with $|q| < \varepsilon$ such that there is a path of length 3 in G starting at p and ending at $p + (q, 0)$.

Proof We use the fact that the rational points are dense on the unit circle to choose an angle α such that $(\cos\alpha, \sin\alpha) \in Q^2$ and

$$\left| \cos\alpha - \frac{1}{2} \right| < \frac{\varepsilon}{3}.$$

The path starting at p and passing through the following 3 points has the desired property:

$p_1 = p + (\cos\alpha, \sin\alpha)$,
$p_2 = p + (\cos\alpha - 1, \sin\alpha)$,
$p_3 = p + (2\cos\alpha - 1, 0)$.

From the previous inequality,

$$|\,2\cos\ \alpha - 1\,| < \frac{2\varepsilon}{3} < \varepsilon,$$

and so we can simply choose $q = 2\cos\alpha - 1$. ■

Tool 63.15 ([Pay]). Any measurable set $A \subset R^2$ of positive measure $\mu(A) > 0$ contains a pair of vertices of G that are joined by a path of length 3.

Proof Michael Payne's proof of this central for the example tool is based on the use of ideas from Shelah–Soifer examples discussed in the previous chapters. Assume that $\mu(A) > 0$. Then there is a unit square S in R^2 with sides parallel to the axes such that

$$\frac{\mu(A \cap S)}{\mu(S)} > \frac{9}{10}.$$

(Since A has positive measure, it must contain points with density equal to 1. Around any such point we can find a square with the desired property.)

By Tool 63.14, we can choose a rational q such that $|q| < \frac{1}{10}$ and two points (x, y) and $(x + q, y)$ are joined by a path of length 3. Let A' be a translate of A: $A' = \{(x + q, y) : (x, y) \in A\}$. We have

$$\frac{\mu(A' \cap S)}{\mu(S)} > \frac{8}{10}.$$

since the part of A translated out of S has measure at most $\mu(S)/10 = \frac{1}{10}$. The two inequalities above imply that there exists $u \in A \cap A' \cap S$. Indeed, assume that $A \cap A' \cap S = \varnothing$, then $A \cap S$ and $A' \cap S$ are disjoint, and using additive property of measure, we get:

$$\mu(S) = \mu(A \cap S) + \mu\left(A' \cap S\right) + \mu\left(S \backslash \left(A \cup A'\right)\right).$$

Dividing by $\mu(S)$ and using the density bounds, we get

$$1 > \frac{9}{10} + \frac{8}{10} + \frac{\mu(S \backslash (A \cup A'))}{\mu(S)},$$

which is a contradiction. Therefore, there exists $u \in A \cap A' \cap S$. Since $u \in A'$, it has a preimage $v \in A$ such that $u = v + (q, 0)$. So $u, v \in A$, and u and v are connected by a path of length 3 in G. ■

Proof of Claim 2 of Payne's Example 63.13. Assume that G is 2-colored, with the corresponding monochromatic sets A_1 and A_2 which together cover the plane. At least one of the sets, say, A_1, has positive measure. Thus, by Tool 63.15, A_1 contains a pair of points connected by a path of length 3 in G. However, in a 2-coloring of a graph, points connected by paths of length 3 must have the opposite colors. Hence, no 2-coloring of G exists. ■

Chapter 64
A Glimpse into the Future: Chromatic Number of the Plane, Theorems, and Conjectures

The importance of particular axioms being used makes a surprising difference for the question of determining the chromatic number of the plane, as recently shown by Shelah and Soifer

– Ronald L. Graham ([Gra6].)

64.1 Conditional Chromatic Number of the Plane Theorem

In the previous chapters, we constructed graphs that highlight the difference between our common **ZFC** Mathematics and mathematics that could be created, such as the **ZFS** Mathematics. But what does it have to do with the main problem of this book, the chromatic number of the plane (CNP)?

Is **AC** relevant to the problem of the chromatic number χ of the plane? The answer depends upon the value of χ which we, of course, do not (yet) know. However, in 2003, Saharon Shelah and I published the following conditional result. It was 1 year after I conjectured 7 as the chromatic number of the plane and 15 years before Aubrey de Grey constructed a 5-chromatic unit-distance graph. Today, the following theorem has only a historic, or anthropological interest.

Conditional Chromatic Number of the Plane Theorem 64.1[1] (Shelah–Soifer 2003 [SS1]). Assume that any finite unit distance plane graph has chromatic number not exceeding 4. Then:

$$^{*})\ \chi^{\mathbf{ZFC}}\left(E^2\right) = 4;$$

$$^{**})\ \chi^{\mathbf{ZFS+}}\left(E^2\right) \geq 5.$$

[1] Due to the use of the Solovay’s Theorem, we assume the existence of an inaccessible cardinal.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_64

Proof The claim *) is true due to the De Bruijn–Erdős Compactness Theorem 27.1.

Claim **). In the Solovay's system of axioms **ZFS+** = **ZF** + **DC** + **LM**, every subset S of the Euclidean plane E^2 is Lebesgue measurable. Indeed, S is measurable if there is a Borel set B such that the symmetric difference $S\Delta B$ is null. Thus, every plane set differs from a Borel set by a null set. We can think of a unit segment $I = [0,1]$ as a set of infinite binary fractions and observe that the bijection $I \to I^2$ defined as $0.a_1a_2\ldots a_n\ldots \mapsto (0.a_1a_3\ldots; 0.a_2a_4\ldots)$ preserves null sets. Due to the Falconer Theorem [Fal1], [Fal2] that appeared in Chapter 9, we can now conclude that the chromatic number of the plane is at least 5 (and, of course, at most 7). ■

This conditional theorem allows for a certain historical insight. Perhaps, the problem of finding the chromatic number of the plane has withstood for over 73 years, leaving us still with a wide range for χ being 5, 6, or 7 because the answer depends upon the system of axioms we choose for set theory?

In general, the chromatic number of the Euclidean space E^n may depend upon the system of axioms we choose for set theory.

In the end of his 2007 paper, Michael Payne, the author of the important Example 63.13 in the previous chapter, remarks:

> After demonstrating the existence of graphs whose chromatic number depends on the axiomatization of set theory, Shelah and Soifer went on to formulate a conditional theorem [which essentially showed that the chromatic number of the plane may be ambiguous in a similar way to the graphs considered here [SS1]. They showed that the chromatic number of the plane may be 4 with **AC** but 5, 6 or 7 with **LM**. The fact that our new example G (Example 63.13) is a subgraph of P [unit distance plane] makes this possibility seem even more likely.

This begs a question: was the choice of the mathematical standard **ZFC** inevitable? Was this choice "best" possible?

In fact, I can formulate an unconditional theorem, which is a consequence of the same ideas as Theorem 64.1.

64.2 Unconditional Chromatic Number of the Plane Theorem

Can we get anything unconditionally, a piece of mathematical "truth-forever" result? Yes, we can, but not yet in **ZFC**.

Unconditional Chromatic Number of the Plane Theorem 64.2 $\chi^{\mathbf{ZFS+}}(E^2) \geq 5$.

64.3 The Conjecture

I trust – all living is related,
The future is my everyday,
As heretic, I end by falling
Into Simplicity, the only way.
– Boris Pasternak, *The Waves*, 1931[2]

The great Russian poet provides us with a fitting epigraph about simplicity. Indeed, much of mathematical results are surprisingly simple, as are our conjectures. Just look at Erdős–Szekeres' Happy End Conjecture 31.15! At the end of this section, I will present my very simple conjecture for the chromatic number of the Euclidean n-dimensional space. Let us start with the plane. In 2002, when I issued this conjecture, the majority view leaned toward values 4 and 5. Now, I see a good number of researchers joining me with a 7.

Chromatic Number of the Plane Conjecture 64.3 (A. Soifer, 2002).

$$\chi\left(E^2\right) = 7.$$

"OK," I hear your reply, "but then a finite unit-distance7-chromatic graph must exist in the plane!" This is true, but it would be quite large. In 1998, Dan Pritikin published the lower bound for the size of such a graph [Pri]: $|G| \geq 6198$. This lower bound stood for 22 years, when in 2020 Jaan Parts succeeded in improving it:

Lower Bound for a Unit-Distance 7-Chromatic Graph 64.4 ([Par3], 2020). Any unit-distance 7-chromatic graph G satisfies the following inequality: $|G| \geq 6993$.

Parts achieves his result by constructing a tiling of more than 99.985698% of the Euclidean plane with 6 colors, thus setting the new record in these indoor competitions. In fact, we all expect that the size of the smallest such graph to be much larger than 6993. The best we can hope to achieve at the moment is a lower bound of 6.

---

[2] [Pas]. Translated by Ilya Hoffman and Alexander Soifer. In 1931, Pasternak dedicated this poem "To N.I. Bukharin," a famous politician and philosopher, murdered on March 15, 1938, in Stalin's purges. One would hope that in the contemporary, 2004, 11-volume edition of Boris Pasternak the dedication would be restored – it was not. My personal library has the 1934 book [Pas], and it seems to be the only place where we can learn about this dedication. The original Russian text is this (I am adding four lines that follow):

В родстве со всем, что есть, уверясь
И знаясь с будущим в быту,
Нельзя не впасть к концу, как в ересь,
В неслыханную простоту.

Но мы пощажены не будем,
Когда ее не утаим.
Она всего нужнее людям,
Но сложное понятней им.

Permit me to repeat here the new \$1,000 problem of Ronald L. Graham. Ron Graham believed that every talk must include at least one proof; he also always had at least one joke. I felt that I knew Ron's style of problem posing. I also knew that we were friends, and Ron with his great sense of humor would not be angry with me for announcing *Ron Graham's New \$1000 Problem* without clearing it with him first. And so, I did it in my March 14, 2019, talk at Florida Atlantic University (with a disclaimer "subject to Ron's approval"). My audience filled the room with laughter. Immediately after the conference, on March 16, 2019, I sent Ron an email:

> Dear Ron,
>
> Unfortunately, I did not see you at my talk in Boca, where I presented *your new \$1000 problem*, of course, subject to your approval. :) The audience loved it.
>
> **The New Ron Graham's \$1,000 Problem 64.5.** Prove or disprove the existence of a 6-chromatic unit-distance graph.
>
> So . . . please, reply with a yes or a no, or your different related open problem(s) – for the inclusion in the second expanded edition of *The Coloring Book.*
>
> Yours always,
> Sasha

Ron replied the same day:

> Hi Sasha,
>
> I had to check out by 11 so unfortunately I couldn't make your talk! :(
>
> I approve of the new \$1000 problem!
>
> Ron

In 2002, I also formulated a conjecture for 3-dimesional Euclidean space E^3:

Chromatic Number of 3-Space Conjecture 64.6 (A. Soifer, 2002).

$$\chi\left(E^3\right) = 15.$$

On April 16, 2018, Aubrey de Grey emailed me for publication in *Geombinatorics* an edited paper with his breakthrough, the first 5-chromatic unit-distance graph. His email ends with the following paragraph:

> I also have some ideas for the upper bound and I suspect that there are ways to falsify you conjecture for $n = 3$. . . but nothing sufficiently concrete to be worth describing yet.

This is how I learned on April 16, 2018, that Aubrey attempted to refute my Chromatic Number of 3-Space Conjecture 64.6. The conjecture so far has withstood 5+ years since Aubrey's assault commenced. :)

My expectations for small dimensions led me to my old general conjecture for the chromatic number of the Euclidean n-dimensional space E^n:

The General Chromatic Number of E^n Conjecture 64.7 (A. Soifer, 2002). For any positive $n > 1$,

$$\chi(E^n) = 2^{n+1} - 1.$$

On March 11, 2021, I was giving a Public Lecture "Chromatic Number of the Plane and its Relatives, History, Conjectures, Results, and Credits: A Ballad in 12 Movements" at a most powerful conference on Euclidean Ramsey Theory, dedicated to the memory of Ron Graham at the Alfred Rényi Institute of Mathematics of the Hungarian Academy of Sciences. Jacob Fox asked for my rationale in creating this "bold conjecture." I replied "Simplicity, like in The Happy End Erdős–Szekeres Conjecture." Jacob was likely not satisfied, and for good reason, so I sent him a longer answer than simply claiming intuition. In the early 1990s, while "playing" with 6-colorings, I "almost" managed to color the plane with 6 colors without creating a monochromatic unit. I then began to feel that 7 was the answer to the chromatic number of the plane, even in the general case, not merely in special cases of "nice" tilings. Next step was David Coulson's paper, showing 15 as the upper bound for the chromatic number of 3-space. I thought then, in 2002, that 15 must be the exact answer for 3-space without any special circumstances. Now look at the numbers I conjectured so far: 7, 15 – which is next? I'd say 31, and this is how my conjecture was born. Moreover, we know axiomatic exponential lower and upper bounds for the chromatic number of n-spaces, and my conjectured function is exponential and lies right in the middle between the known exponential bounds. I do not expect the General Conjecure 64.7 to be proved in my lifetime.

As Paul Erdős used to say about some of his conjectures, "This conjecture will likely withstand centuries, but we shall see!"

Part XIII
Imagining the Real, Realizing the Imaginary

"What do you think of the abstract – do you believe that one should deduce one's abstraction from the forms of nature, or that one should create the form, outside of nature?" Matisse replied, "There is always a measure of reality. The rest, I agree, is imagination."

– Henry Matisse[1]

For me, the province of art and the province of nature thus became more and more widely separated, until I was able to experience both as completely independent realms.

– Wassily Kandinsky[2]

Reality is unnatural.

– Pier Paolo Pasolini

Everything you can imagine is real.

– Pablo Picasso

There is an intimate relationship between the order of Nature (which constitutes the basis of life) and the order of Art (which constitutes the basis of civilization).

– Herbert Read[3]

1 Interview with R. W. Howe [Fla], p. 186.
2 [Kan], 1913.
3 [Rea].

Chapter 65
What Do the Founding Set Theorists Think About the Foundations?

In the beginning (if there was such a thing) God created Newton's laws of motion together with the necessary masses and forces. This is all; everything beyond this follows from the development of appropriate mathematical methods by means of deduction.

– Albert Einstein, 1946 ([Ein2, p. 19].)

Kurt Gödel and Paul J. Cohen believed that we would eventually identify all the axioms of set theory and when we have done so, we will no longer be able to choose between **CH** and ¬**CH** (or, similarly, between **AC** and **DC + LM**) because the additional axioms would exclude one of the options. Cohen shares his thoughts on the subject in 1966 [Coh2, pp. 150–151, underlining is his]:

> One can feel that our intuition about sets is inexhaustible and that eventually an intuitively clear axiom will be presented which decides **CH**. One possibility is **V=L**, but this is almost universally rejected... A point of view which the author feels may eventually come to be accepted is that CH is <u>obviously</u> false.

While the majority of mathematicians are Platonists, Saharon Shelah does not share their Platonist view in his 2003 *Logical Dreams* [She3],

> Some believe that compelling, additional axioms for set theory which settle problems of real interest will be found or even have been found. It is hard to argue with hope and problematic to consider arguments which have not yet been suggested. However, I do not agree with the pure Platonic view that the interesting problems in set theory can be decided, we just have to discover the additional axiom. My mental picture is that we have many possible set theories, all conforming to **ZFC**.

Before I query the great set theorists about the state of their minds on the foundations of set theory, I read their writings on the subject matter. Robert M. Solovay in his pioneering 1970 paper [Sol1] states:

> Of course, the axiom of choice [**AC**] is true.

Saharon Shelah writes in the introduction to his 1994 classic monograph *Cardinal Arithmetic* [She2]:

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_65

> If we interpret "true" by "is provable in **ZFC**" (the usual axioms of set theory), as I do, then a large part of set theory which is done today does not deal directly with true theorems – it deals, rather, with a huge machinery for building counterexamples (forcing possible universes) or with "thin" universes (inner models). Very often the answer to "can this happen?" is "it depends." Now, I believe that this phenomenon is inevitable, and expresses a deep phase of the development of set theory, which resulted in many fascinating theorems (and also in quite a few proofs of mine). However, there is still some uneasiness about it. A way to express it is to say that if Cantor would have risen from his grave today, he would not just have problems with understanding the proofs of modern theorems – he would not understand what the theorems actually say.

And so I endeavor to ask the three of the leading set theorists, great contributors to the axiomatics of sets, the following questions:[1]

*) Has **AC** been good for mathematics?
) Ought **AC to be a part of a 'standard' system of axioms for set theory?
***) What do you think of the Solovay system of axioms (**ZFS**)?
****) How do you envision the 'standard' system of axioms for set theory?

"My opinions shift and I have no obvious candidates [for the standard system of axioms]," Paul J. Cohen replies to me [Coh3] and adds: "Solovay's result on **LM** is very nice, but hardly an axiom." Saharon Shelah in his response to me [She4] sees a certain value in using the Solovay system **ZFS** and systems weaker than **ZFC**:

> The major question is what is true, i.e., when existence tells you something more if you give an explicit construction. Now, working in **ZF**, **ZF+DC**, and also **ZF+DC+LM** and many other systems are ways to explicate the word "construct."

In Shelah's opinion [She4], **AC** has been "definitely" good for mathematics, **AC** is true and "should be in our standard system [of axioms]." Robert Solovay too believes that **AC** is true (and therefore his system, which I admire so much, **ZFS+** = **ZF+DC+LM** is false). He writes to me about **AC** [Sol3]:

(a) It is true.
(b) It plays an essential role in all sorts of Theorems, for example, the uniqueness of the algebraic closure, existence of maximal ideals, etc.

This prompts my question: *But... what is "true"*? Shelah answers it as follows [She5]:

> This is a meta-mathematical question. I will say [true is what] fits our image of set theory.
> You may say this is circularly, but this is unavoidable.
> You may be Platonist like Cantor then the meaning is clear.
> You may say [true is] what mathematicians who have not been interested in the question will accept.
> You may be a formalist and then this is a definition of **ZFC**.

[1] June 11–20, 2006, author's e-mail exchanges with Paul J. Cohen, Saharon Shelah, and Robert M. Solovay.

Shelah is clear on what is not true enough, in his 2003 paper [She3]:

> Generally I do not think that the fact that a statement solves everything really nicely, even deeply, even being the best semi-axioms (if there is such a thing, which I doubt) is a sufficient reason to say it is a "true axiom."

Not surprisingly, there is no rigorous definition of "truth," this elusive notion is subjective, and all we can hope for is to recognize a true axiom when we encounter one. For Shelah, a true axiom is "what I feel/think is self-evident."[2] This is a high bar to clear, and only **ZFC** seems to have cleared it for most creators of set theory. Even such serious candidates as **CH**, ¬**CH**, **GCH**, **V**=**L** are termed by Shelah as "semi-axioms" – because they are not sufficiently "representative" of all possibilities. Shelah elaborates [She3]:

> Still most mathematicians, even those who have worked with **GCH** [and with other semi-axioms, A. S.] do it because they like to prove theorems and they could not otherwise solve their problems (or get a reasonable picture), i.e., they have no alternative in the short run...
>
> What are our criterions for semi-axioms? First of all, having many consequences, rich, deep beautiful theory is important. Second, it is preferable that it is reasonable and "has positive measure." Third, it is preferred to be sure it leads to no contradiction...

The renown French mathematician Jean-Paul Delahaye [Del] believes that Shelah–Soifer results may have put a new emphasis on the task of finding which world of sets we think we live in:

> It turns out that knowing if the world of sets satisfies the axiom of choice or a competing axiom is a determining factor in the solution of problems that no one had imagined depended on them. The questions raised by the new results are tied to the fundamental nature of the world of sets. Is it reasonable to believe that the mathematical world of sets is real? If it exists, does the true world of sets – the one in which we think we live – allow the coloring of S. Shelah and A. Soifer in two colors or does it require an infinity of colors? ...
>
> A series of results concerning the theory of graphs, published in 2003 and 2004 by Alexander Soifer of Princeton University and Saharon Shelah, of the University of Jerusalem, should temper our attitude and invite us to greater curiosity for the alternatives offered to the axiom of choice. The observation demonstrated by A. Soifer and S. Shelah should force mathematicians to reflect on the problems of foundations: what axioms must be retained to form the basis of mathematics for physicists and for mathematicians?

[2] [She5].

Chapter 66
So, What Does It All Mean?

I know of mathematicians who hold that the axiom of choice has the same character of intuitive self-evidence that belongs to the most elementary laws of logic on which mathematics depends. It has never seemed so to me.

– Alonzo Church (Talk at the International Congress of Mathematicians in Moscow, 1966 [Chu].)

Shelah–Soifer's results we have discussed in this book seem surprising and even strange. How can the presence of the Axiom of Choice or its version affect whether we need 2 colors or an uncountable infinity of colors for coloring a particular easily understood graph? How can the chromatic number of the plane depend upon our choice of the axioms for set theory? What *do* these results mean?

Jean-Paul Delahaye opens his article about these Shelah–Soifer papers in *Pour la Science*, the French edition of *Scientific American*, as follows [Del]:

> The axiom of choice, a benign matter for the non-logician, puzzles mathematicians. Today, it manifests itself in a strange way: it takes, depending on the axiom's variants, either two or infinity of colors to resolve a coloring problem.
>
> Just as the parallels postulate seemed obvious, the axiom of choice has often been considered true and beyond discussion. The inventor of set theory, Georg Cantor (1845–1918), had used it several times without realizing it; Giuseppe Peano (1858–1932) used it in 1890, in working to solve a differential equation problem, consciously; but it was Ernst Zermelo (1871–1953), at the beginning of the 20th century, who identified it clearly and studied it.
>
> When Gödel and Cohen proved independence of **AC** from the rest of the axioms **ZF** of set theory, they created a parallel, so to speak, between **AC** and the parallels postulate. As so, when Shelah–Soifer came out, it showed that various buildings of mathematics can be constructed.

Delahaye observes, "These [Shelah–Soifer's] results mean, as with the parallels postulate, that several different universes can be considered," and continues:

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_66

In the case of geometry, the independence of the parallels postulate proved that non-Euclidian geometries deserved to be studied and that they could even be used in physics: Albert Einstein took advantage of these when, between 1907 and 1915, he worked out his general theory of relativity.

Regarding the axiom of choice, a similar logical conclusion was warranted; the universes where the axiom of choice is not satisfied must be explored and could be useful in physics.

Jean Alexandre Eugène Dieudonné (1906–1992), one of the founding members of the celebrated Nicolas Bourbaki project, describes the state of the foundations in 1976 as follows:

Beyond classical analysis (based on the Zermelo–Fraenkel axioms supplemented by the Denumerable Axiom of Choice), there is an infinity of different possible mathematics, and for the time being no definitive reason compels us to choose one of them rather than another.[1]

The Solovay System of Axioms **ZFS+** is stronger than the system referred to by Dieudonné. It allows us to build classical analysis, including the complete Lebesgue measure theory, and it eliminates such counter-intuitive objects, existing in **ZFC**, as non-measurable sets of reals.

Using the axiom of choice in their 1924 paper [BT], two celebrated Polish mathematicians Stefan Banach (1892–1945) and Alfred Tarski (1902–1983) decomposed the 3-dimensional closed unit ball into finitely many pieces and moved those pieces through rotations and translations (pieces were allowed to move through one another) in such a way that the pieces formed 2 copies of the original ball. Since the measure of the union of two disjoint measurable sets is the sum of their measures, and measure is invariant under translations and rotations, we can conclude that there is a piece in Banach–Tarski decomposition that has no measure (i.e., volume). Having LM in the system of axioms for set theory would eliminate this and a good number of other paradoxes.

ZFC allows us to create imaginary objects – or shall I say, unimagined objects – such as sets on the line that have no length, sets in the plane that have no area, etc. Are we not paying a high price for the comfort of having a powerful tool?

Having lived most of my life in **ZFC** and having enjoyed using transfinite induction in my group theory works, in the course of my work with Shelah I came to a realization that I prefer **ZFS+** = **ZF** + **DC** + **LM** over **ZFC**: **LM** assures that every set of reals is measurable (which is consistent with my intuition: every point set on the line ought to have length), while **DC** seems to give us as much choice as is consistent with **LM**.

Of course, by downgrading **AC** to **DC**, we would lose such tools as the transfinite induction and the well ordering of uncountable sets, as would consequently lose some important theorems, such as the existence of basis for a vector space. However, new theorems would be found when mathematicians spend as much time building on the Solovay foundation **ZFS+** as they have on **ZFC.**

[1] Quoted from [Moo, p. 4].

Chapter 67
Imagining the Real or Realizing the Imaginary: Platonism Versus Imaginism

As far as the propositions of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality.

– Albert Einstein, 1921 [Ein1]

Reality is merely an illusion, albeit a very persistent one.

– Albert Einstein, Annalen der Physik, Sep. 27, 1905

Physical concepts are free creations of the human mind, and are not, however it may seem, uniquely determined by the external world.

– Albert Einstein, The Evolution of Physics

We all know that Art is not truth. Art is a lie that makes us realize the truth.

– Pablo Picasso, 1923 [Pic]

The mathematician is an inventor, not a discoverer.

– Ludwig Wittgenstein, 1937–1944 [Witt, p. 47e]

Undoubtedly, a vast majority of mathematicians are Platonists.[1] They believe that mathematical objects exist "out there" independently of the human mind, and mathematicians merely discover them. The Platonists believe that a mathematical statement, such as **AC**, is objectively either true or false – we simply do not know which it is, although in a poll of mathematicians, "**AC** is true" would win hands down. Likewise, a question, what is the chromatic number of the first Shelah–Soifer graph *G*, surely, must have a definitive answer; it cannot be "2 or uncountable infinity." Therefore, for the Platonists either **ZFC** or **ZF+DC +LM** is true, we just do not know which. *Platonists imagine the real.*

I find it contradictory that most mathematicians subscribe to Platonism, and at the same time believe that mathematics is an art. In my opinion,

Science reflects what is outside the Man, in Nature, whereas Art reflects what is within.

[1] Jim Holt reminisces [Hol]: "Some years ago, while giving a lecture to an international audience of elite mathematicians in Berkeley, I asked how many of them were Platonists. About three-quarters raised their hands."

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_67

Let me ask great thinkers to assess the goals of Art, while keeping in mind that Mathematics is an art. In 1859, the 39-year-old great French poet Charles Pierre Baudelaire (he lived sadly only 46 years, 1821–1867) observes:

> In recent years we have heard it said in a thousand and different ways, "Copy nature; just copy nature. There is no greater delight, no finer triumph than an excellent copy of nature." And this doctrine (the enemy of art) was alleged to apply not only to painting but to all the arts, even to the novel and to poetry. To these doctrinaires, who were so completely satisfied by Nature, a man of imagination would certainly have the right to reply: "I consider it useless and tedious to represent what exists, because nothing that exists satisfies me."

The 19-year-old future Nobel Laureate Albert Camus (another tragically short life of 46 years, 1913–1960) digs deeper in his 1932 *Essay on Music* [Cam1]:

> According to [realism], Art ought to concern itself exclusively with the imitation of Nature and the exact reproduction of Reality. This is a definition that not only demeans Art, but, further, destroys it. To reduce Art to a servile imitation of Nature is to condemn it to produce only the imperfect.
>
> The greater part of the aesthetic emotion, in fact, is a product of our personality. The beautiful is not in Nature; it is we who put it there. The sense of beauty we feel in front of landscape does not come from the landscape's aesthetics perfection. It comes from the fact that the look of things is in perfect agreement with our instincts, our propensities, with everything that makes up our unconscious personality... The greater part of an aesthetic emotion is therefore produced by ourselves, and Amiel's saying "A landscape is a state of the soul" will always be true.

Indeed, "*The beautiful is not in Nature; it is we who put it there*." How shall we describe those who hold Albert Camus' view, the view dual to that of Plato? I propose to call them – us – *Imaginists*. This name was once used in Russia in the early XX century: there were several poets, Sergei Esenin, Vadim Shershenevich, Anatoly Mariengof, and others, who valued imagination and called themselves "Imaginists." I trust they would not object to our application of this term to mathematicians and scholars, who view their work essentially as *Imaginism*. By marrying Picasso to Wittgenstein, I arrive at the following proverb:

Mathematics is an invention that makes us realize reality.

Mathematics is certain only as an invention, or as Einstein put it, "*As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality*."

I believe that *mathematicians do not only imagine the real but also realize the imaginary*. Just like all artists, mathematicians create objects that challenge reality in every aspect: beauty, simplicity, intuitiveness, and counter-intuitiveness. I believe that

Nature is but one of many inspirations for creating mathematics.

Mathematics of **ZFC** *is* the house that Jack built. Has Jack built the only possible house? Has he built the best possible house? Must we give up a village for a house, as Richard III offered his kingdom for a horse?

I believe that mathematicians put all their eggs in one **ZFC** basket, and thus missed out on many alternative mathematical universes that can be built on many alternative foundations, one of which is Solovay's **ZFS+**. Saharon Shelah too thinks that we ought to build on many foundations, but he puts his main emphasis on building *up* from **ZFC**. It seems, we have been too comfortable, too nonchalant about seeing problems with **ZFC** and doing nothing about them. Mathematical results presented in the previous chapter may be not important by themselves, but they illuminate the existence of many buildings of mathematics that could be built. I hope this is a wake-up call. Delahaye [Del] concludes his analysis of Shelah–Soifer's series of papers with the possibility of the emerging "revolution of set theories":

> In set theory, as in geometry, all axiomatic systems are not created equal. Thinking carefully about their meaning and the consequences of each one of them, and asking ourselves (as it is done in geometry) what the particular usefulness of this or that axiom is in expressing and addressing issues of mathematical physics, may be relevant once again and could lead – why not – to a revolution of set theories, similar to the revolution in non-Euclidian geometries.

You, my reader, and I are engaged in the art of mathematics. We also ought to be citizens of our small, endangered planet. Let us leave the Ivory Tower and get involved with the world, involved like the Little Flutist in the great French chansonnier Georges Brassens' song "*Le Petit joueur de flûtiau*." I am translating for you the lines opening this ballad and those concluding it.

A flutist played in the royal court.
The sound of his songs was sweeter than words.
The flutist's play delighted the court:
"The noble coat of arms for such a perform!"
But the musician refused the reward:
My King, I don't need your royal award!
From noble titles and royal feasts
The sound of my flute will fret a bit.
In land that so dear and sweet to my heart,
Each one will say: "The flutist's corrupt."
………………………………………………..
The flutist paid bow to the King as he stormed
From the royal castle to the place called home,
Where everyone ready accepting him
Without noble titles and royal feasts.
He went to his family, friends, and his lover,
To village of his, to his own barrack.
In land that so dear and sweet to his heart,
No one will say, "The flutist's corrupt."

Part XIV
Farewell to the Reader

Chapter 68
Two Celebrated Problems

Mathematics, rightly viewed, possesses not only truth, but supreme beauty... capable of a stern perfection such as only the greatest art can show.

– Bertrand Russell

Unfinished, a picture remains alive, dangerous. A finished work is a dead work, killed.

– Pablo Picasso

Histories of two beautiful coloring problems, The Four-Color Problem (4CP) and The Chromatic Number of the Plane Problem (CNP) have been strikingly similar in many ways.

Both problems are easy to enjoy, easy to formulate, and hard to solve.

Both problems were created by young students, ages 20 and 18, respectively, born in the year '32:

4CC by Francis Guthrie, born in 1832;
CNP by Edward Nelson, born in 1932.

4CC inspired and motivated the development of much of graph theory. CNP inspired a lot of mathematical work and results in a great variety of fields in which solutions were sought: combinatorics, graph theory, topology, measure theory, abstract algebra, geometry, combinatorial geometry, etc.

As we have seen on the pages of this book, CNP and 4CP have an essential nonempty intersection: The Townsend-Woodall 5-Color Theorem.

Each of these two problems had a chief promoter. For 4CP this was Augustus De Morgan, who had kept the problem alive for decades. Paul Erdős' contribution to keeping CNP alive is even greater. First of all, as De Morgan did for 4CP, Erdős kept the flaming torch of the problem lit. Even though the problem was not his, Paul made CNP well known by posing it in his countless problem talks and many publications, for example, we see it in [E61.22], [E63.21], [E75.24], [E75.25], [E76.49], [E78.50], [E79.04], [ESi], [E80.38], [E80.41], [E81.23], [E81.26], [E85.01], [E91.60], [E92.19], [E92.60], and [E94.60]. Moreover, Paul Erdős created a good number of fabulous related problems, some of which we have discussed in this book.

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1_68

Both problems required a very long time to be conquered. Victor Klee and Stan Wagon [KW], observing that solving 4CP took 124 years, suggested that CNP might require as long for its solution:

> If a solution of CNP takes as long as 4CC, then we will have a solution by the year 2084.

CNP is a classic problem of mathematics. As Ron Graham and Eric Tressler write [GT] in a volume I edited [Soi49],

> The unit distance graph in the plane … is simple enough to describe to a nonmathematician, and so enigmatic that finding its chromatic number is a new four-color map problem for graph theorists.

The interest in CNP is growing. I was invited to write a chapter about CNP for the book "[Famous] Open Problems in Mathematics" edited by "A Beautiful Mind" (John F. Nash, Jr.) and Michael Th. Rassias [Soi50] for Springer, and another chapter for "Topics in Chromatic Graph Theory" edited by Lowell W. Beineke and Robin J. Wilson with academic consultant Bjarne Toft [Soi35] for Cambridge University Press. Alfréd Rényi Institute of Mathematics of the Hungarian Academy of Sciences invited my "Public Lecture" about CNP and related problems [Soi56] at their spectacular 2021 three-day *Workshop on Euclidean Ramsey Theory* dedicated to the memory of our Captain, Ron Graham. As you have witnessed in this book, CNP attracted research not only by mathematicians, but also by a gerontologist, a computer scientist, and an engineer.

Will we completely solve CNP by 2084? I would say, partially yes, completely, probably not. Paul Erdős would say, "We shall see!" Arnold Schoenberg believes that faith can move mountains. Erdős urges us to believe that the transfinite Book of all theorems and their best proofs exists. Such belief led Appel and Haken to succeed at the breaking point of available computing. Such belief is needed to conquer my favorite open problem of mathematics, the chromatic number of the plane. We shall overcome!

I concluded the first edition of this book quite prophetically:

> Thank you for inviting my book into your home and holding it in your hands, or on your screen. Your feedback, problems, conjectures, solutions would be welcome in my home. Who knows, maybe they will inspire a new edition in the future, and we will meet again!

Indeed, new brilliant contributions inspired me to write and Springer to publish this new much expanded edition. Paul Erdős on some occasions said, "I hope some of my theorems will live forever." I hope the conjecture of my life, prophesying the value of the chromatic number of the Euclidean n-space E^n will not live forever, but will be settled within the next, say, 200 years, but, in Erdős' words, we shall see:

$$\boldsymbol{\chi(E^n) = 2^{n+1} - 1}.$$

The time has come to let my book live her own life. I hope she will add colors to your life and inspiration to your work. Reach out, share your conjectures, theorems, and ideas about things you are passionate about, such as Art, Mathematics, and Freedom!

Bibliography

[Abb] Abbott, H.L.: Some problems in combinatorial analysis. Ph.D. thesis, University of Alberta, Canada, March 1965

[AH] Abbott, H.L., Hanson, D.: A problem of Schur and its generalizations. Acta Arithmetica. **20**, 175–187 (1972)

[AM] Abbott, H.L., Moser, L.: Sum-free sets of integers. Acta Arithmetica. **11**, 392–396 (1966)

[AZ] Abbott, H.L., Zhou, B.: On small faces in 4–critical graphs. Ars Combin. **32**, 203–207 (1991)

[Abe1] Abelin-Schur, H., talk at Schur–Gedenkkolloquium at the Humboldt University, Berlin, November 15, 1991, manuscript

[Abe2] Abelin-Schur, H., phone conversation of May 18, 1995

[Abb] Abbott, T.K.: Catalogue of the Manuscripts in the Library of Trinity College, Dublin to which Is Added a List of Fagel Collection of Maps in the Same Library. Hodges, Figgis, & Co, Dublin (1900)

[Aks] Aksionov, V.A.: On continuation of 3-coloring of planar graph (Russian). Diskret. Analiz. **26**, 3–19 (1974)

[AMe] Aksionov, V.A., Mel'nikov, L.S.: Some counter examples associated with the three-color problem. J. Combin. Theory Series B. **28**, 1–9 (1980)

[AFG] Alexeev, B., Fox, J., Graham, R.L.: On minimal colorings without monochromatic solutions to a linear equation. In: Combinatorial Number Theory: Dedicated to Professor Ron Graham on the Occasion of His Seventieth Birthday, pp. 1–22. Walter de Gruyter, Berlin (2007)

[AnM1] Angeltveit, V., McKay, B.D.: $R(5, 5) \leq 48$. J. Graph Theory. **89**, 5–13 (2018)

[AnM2] Angeltveit, V., McKay, B.D.: Personal communication to Stanisław P. Radziszowski (2020)

[Ant] Antonioni, M.: Reality and Cinéma-Vérité, Cinema Nuovo, 1965 (Spanish). English translation: Antonioni, M., Blow-Up. A Film by Michelangelo Antonioni, Simon and Schuster, New York, 1971, 11–13

[AH0] Appel, K., Haken, W.: The existence of unavoidable sets of geographically good configurations. Ill. J. Math. **20**, 218–297 (1976)

© Alexander Soifer 2024

[AH1] Appel, K., Haken, W.: The solution of the Four-Color map problem, Sci. Am., **237**(4), 108–121 (Oct. 1977)

[AH2] Appel, K., Haken, W.: Every planar map is four Colorable. Part I. Discharging. Ill. J. Math. **21**, 429–490 (1977)

[AH3] Appel, K., Haken, W., Koch, J.: Every planar map is four Colorable. Part II. Reducibility. Ill. J. Math. **21**, 491–567 (1977)

[AH4] Appel, K., Haken, W.: Every Planar Map Is Four Colorable. American Mathematical Society, Providence (1989)

[Arc] Archdeacon, D.: Coupled coloring of planar maps. Congr. Numer. **39**, 89–94 (1986)

[AMRSS] Ardal, H., Manuch, J., Rosenfeld, M., Shelah, S., Stacho, L.: The odd-distance plane graph. Discrete Comput. Geom. **42**(2), 132–141 (2009)

[ABPR] Arman, A., Bondarenko, A.V., Prymak, A., Radchenko, D.: Upper bounds on chromatic number of En in low dimensions, February 12, 2022., arXiv:2112.13438v2

[AT] Arman, A., Tsaturian, S.: A result in asymmetric Euclidean Ramsey theory, 15 Feb. 2017, arXiv:1702.04799v1

[Arn] Arnautov, V.I.: Nediskretnaya Topologizuemost Schetnykh Kolets, Dokl. Akad. Nauk SSSR 191 1970 747–750 (Russian). English Translation: Nondiscrete topologizability of countable rings, Soviet Math. Dokl. 11 (1970), 423–426

[Arr] Arrias, E.: In Memoriam Prof. P. J. H. Baudet. Eigen Haard **48**(5), 92–94, January 28, 1928

[BaO] Baker, H., Oliver, E.: Ericas in Southern Africa. Purnell, Cape Town

[BT] Banach, S., Tarski, A.: Sur la décomposition des ensembles de points en parties respectivement congruentes. Fundam. Math. **6**, 244–277 (1924)

[Bau1] Baudet, P.J.H.: Groepentheoretische onderzoekingen. Doctoral dissertation, Martinus Nijhoff, The Hague (1918)

[Bau2] Baudet, P.J.H.: Het Limietbegrip (Inaugural Address at Delft). P. Noordhoff, Groningen (1919)

[Bau3] Baudet, P.J.H.: Een nieuwe theorie van onmeetbare getallen. Christiaan Huygens. **I**, 33–47

[Bau4] Baudet, P.J.H.: Een stelling over rekenkundige reeksen van hoogere orde. Christiaan Huygens **I**, 1921–1922, 146–149 (published posthumously)

[Bau5] Baudet, P.J.H.: Gelijktijdige invoering der negatieve en imaginaire getallen. Christiaan Huygens **I**, 226–231 (published posthumously by J. Teixeira de Mattos and F. Schuh on the basis of a lecture and notes by Baudet)

[BII1] Baudet: Letter to A. Soifer of August 24, 1995

[BII2] Baudet, H.: Letter to A. Soifer of October 17, 1995

[BII3] Baudet, H.: Letter to A. Soifer of November 27, 1995

[BII4] Baudet, H.: Letter to A. Soifer of December 18, 1995

[BII5] Baudet, H.: Letter to A. Soifer of January 29, 1996

[BII6] Baudet, H.: Letter to A. Soifer of February 15, 1996

[BII7] Baudet, H.: Letter to A. Soifer of March 6, 1996

[BII8] Baudet, H.: Letter to A. Soifer of March 23, 1996

[BII9] Baudet, H.: Letter to A. Soifer of April 7, 1996
[BII10] Baudet, H.: Letter to A. Soifer of April 18, 1996
[BII11] Baudet, H.: Letter to A. Soifer of May 10, 1996
[BII12] Baudet, H.: Letter to A. Soifer of May 27, 1996
[BII13] Baudet, H.: Letter to A. Soifer of June 12, 1996
[BG] Baumert, L.D., Golomb, S.W.: Backtrack programming. J. Ass. Comp. Machinery. **12**, 516–524 (1965)
[Beh] Behzad, M.: Graphs and their chromatic numbers, Ph.D. Thesis, Michigan State University (1965)
[BC1] Behzad, M., Chartrand, G.: An Introduction to total graphs. In: Rosenstiehl, P. (ed.) Theory of Graphs. International Symposium, Rome, July 1966, pp. 31–33. Gordon and Breach, New York (1967)
[BC2] Behzad, M., Chartrand, G.: Introduction to the Theory of Graphs. Allyn and Bacon, Boston (1971)
[BCL] Behzad, M., Chartrand, G., Lesniak-Foster, L.: Graphs & Digraphs. Prindle, Weber & Schmidt, Boston (1979)
[BO] Belov, A.Ya., Okhitin, S.B.: Ob Odnoj Kombinatornoj Zadaze (On one combinatorial problem). Uspekhi Mat. Nauk **48**(2), 169–170 (1993) (Russian)
[BP] Benda, M., Perles, M.: Colorings of metric spaces. Geombinatorics. **IX**(3), 113–126 (2000)
[Ber] Bergelson, V.: Multiplicatively large sets and ergodic Ramsey Theory. Israel J. Math. **148**, 23–40 (2005)
[BL] Bergelson, V., Leibman, A.: Polynomial extension of van der Waerden's and Szemerédi's theorems. J. Am. Math. Soc. **9**, 725–753 (1996)
[Berl] Berlekamp, E.R.: A construction for partitions which avoid long arithmetic progressions. Can. Math. Bull. **11**, 409–414 (1969)
[Bern] Bernays, P.: A system of axiomatic set theory III. J. Symb. Log. **7**, 65–89 (1942)
[Ber1] Berhhart, A.: Six–rings in minimal five–color maps. Am. J. Math. **69**, 391–412 (1947)
[Ber2] Berhhart, A.: Another reducible edge configuration. Am. J. Math. **70**, 144–146 (1948)
[Bert] Bertsch, S.M.: Nobel Prize Women in Science: Their Lives, Struggles, and Momentous Discoveries. Carol Pub., Secaucus (1993)
[Big] Biggs, N.L.: De Morgan on map colouring and the separation axiom. Archive Hist. Exact Sci. **28**, 165–170 (1983)
[BLW] Biggs, N.L., Lloyd, E.K., Wilson, R.J.: Graph Theory 1736–1936. Clarendon Press, Oxford (1976) Latest version: reprinted with corrections, 1998
[BN1] Bikov, A., Nenov, N.: The edge Folkman number $Fe(3; 3;4)$ is greater than 19. Geombinatorics. **XXVIII**(1), 5–13 (2017)
[BN2] Bikov, A., Nenov, N.: On the vertex Folkman numbers $F_v(a_1,...,a_s; m-1)$ when $\max\{ a_1,...,a_s \}= 6$ or 7, arXiv:1512.02051v2, 8 April 2017
[BN3] Bikov, A., Nenov, N.: Lower bounding the Folkman numbers $F_v(a_1,...,a_s; m-1)$, 5 Nov. 2017, arXiv:1711.01535v1

[BN4] Bikov, A., Nenov, N.: On the independence number of (3,3)-Ramsey graphs and the Folkman number $F_e(3,3;4)$. Australas. J. Comb. **77**, 35–50 (2020)

[Bir] Birkhoff, G.D.: The reducibility of maps. Am. J. Math. **35**, 114–128 (1913)

[BRa] Bogolyubsky, L.I., Raigorodskii, A.M.: A remark on lower bounds for the chromatic numbers of spaces of small dimension with metrics ℓ_1 and ℓ_2. Math Notes. **105**, 180–203 (2019)

[BS] Boltyanski, V., Soifer, A.: Geometric Etudes in Combinatorial Mathematics. Center for Excellence in Mathematical Education, Colorado Springs (1994)

[BT] Bóna, M., Tóth, G.: A Ramsey–type problem on right–angled triangles in space. Discret. Math. **150**, 61–67 (1996)

[Bor1] Borodin, O.V.: Structural properties of plane graphs without adjacent triangles and an application to 3-colorings. J. Graph Theory. **21**(2), 183–186 (1996)

[Bor2] Borodin, O.V.: Colorings of plane graphs: a survey. Discret. Math. **313**(4), 517–539 (2013)

[Bor3] Borodin, O.V.: Solution of the Ringel problem on vertex-face coloring of planar graphs and coloring of 1-planar graphs. Metody Diskret. Analiz. **41**, 12–26 (1984) (Russian)

[Bor4] Borodin, O.V.: A new proof of the 6-color theorem. J. Graph Theory. **19**, 507–521 (1995)

[Bor5] Borodin, O.V.: Structure of neighborhoods of edges in planar graphs and simultaneous coloring of vertices, edges and faces. Math. Notes. **53**, 483–489 (2013)

[BGJR] Borodin, O.V., Glebov, A.N., Jensen, T. R., Raspaud, A.: Planar graphs without triangles adjacent to cycles of length from 3 to 9 are 3-colorable. Siberian Electron. Math. Rep. **3**, 428–440 (2006)

[BGRS] Borodin, O.V., Glebov, A.N., Raspaud, A., Salavatipour, M.R.: Planar graphs without cycles of length from 4 to 7 are 3-colorable. J. Combin. Theory Ser. B. **93**, 303–311 (2005)

[BGMR] Borodin, O.V., Glebov, A.N., Montassier, M., Raspaud, A.: Planar graphs without 5- and 7-cycles and without adjacent triangles are 3-colorable. J. Combin. Theory Ser. B. **99**, 668–673 (2009)

[BGl] Borodin, O.V., Glebov, A.N.: Planar graphs without 5-cycles and with minimal distance between triangles at least 2 are 3-colourable. J. Graph Theory. **66**(1), 1–31 (2011)

[BoR] Borodin, O.V., Raspaud, A.: A sufficient condition for planar graphs to be 3-colorable. J. Combin. Theory Ser. B. **88**, 17–27 (2003)

[Boza5] Boza, L.: Personal communication to Stanisław P. Radziszowski (2013)

[BMP] Brass, P., Moser, W., Pach, J.: Research problems in discrete geometry. Springer, New York (2005)

[Bra1] Brauer, A.: Über Sequenzen von Potenzresten, Sitzunberichte de Preussischen Akademie der Wissenschaften, Physicalisch-Mathematische Klasse, pp. 9–16. Preussische Akademie der Wissenschaften, Berlin (1928)

[Bra2] Brauer, A.: Combinatorial methods in the distribution of *k*th power residues, combinatorial mathematics and its applications. Proceedings of the Conference, University of North Carolina, pp. 14–37 (1969)

[Bra3] Brauer, A.: Gedenkrede auf Issai Schur. In: Schur, I. (ed.) Gesammelte Abhandlungen, vol. 1, pp. v–xiv. Berlin, Springer (1973)

[BBH] Brauer, A., Brauer, R., Hopf, H.: Über dies Irreduzibilität einigen spezieller Klassen von Polynomen. Jahresber. Deutsch. Math.–Verein. **35**, 99–112 (1926)

[BraH] Brauer, H.: Memoires, 114 pages, manuscript shared with A. Soifer

[BraR] Brauer, R.: Collected Papers, vol's I, II, and III. MIT Press, Cambridge, MA (1980)

[Bro] Brooks, R.L.: On Colouring the nodes of a network. Proc. Cambridge Philos. Soc. 37, pp. 194–197 (1941)

[BrR] Brown, J.I.: Rödl, Vojtěch Folkman numbers for graphs of small order. Ars Combin. **35**(A), 11–27 (1993)

[BDP1] Brown, N., Dunfield, N., Perry, G.: Colorings of the plane I. Geombinatorics. **III**(2), 24–31 (1993)

[BDP2] Brown, N., Dunfield, N., Perry, G.: Colorings of the plane II. Geombinatorics. **III**(3), 64–74 (1993)

[BDP3] Brown, N., Dunfield, N., Perry, G.: Colorings of the plane III. Geombinatorics. **III**(4), 110–114 (1993)

[Bru1] Bruijn, N. G. de: Letter to P. Erdős, August 18, 1947

[Bru2] Bruijn, N. G. de: Letter to B. L. van der Waerden, March 29, 1977

[Bru3] de Bruijn, N.G.: Commentary (on B. L. Van der Waerden's article [Wae2]). In: Berin, E.M.J., et al. (eds.) Two Decades of Mathematics in the Netherlands, 1920–1940. Part I, pp. 116–124. Mathematical Centre, Amsterdam (1978)

[Bru4] Bruijn, N. G. de: e-mail to A. Soifer, May 9, 1995

[Bru5] Bruijn, N. G. de: e-mail to A. Soifer, June 21, 1995

[Bru6] Bruijn, N. G. de: e-mail to A. Soifer, January 27, 2004

[Bru7] Bruijn, N. G. de: e-mail to A. Soifer, May 26, 2004

[Bru8] Bruijn, N. G. de: e-mail to A. Soifer, June 1, 2004

[Bru9] Bruijn, N. G. de: e-mail to A. Soifer, June 15, 2004

[Bru10] Bruijn, N. G. de: e-mail to A. Soifer, June 23, 2004

[Bru11] Bruijn, N. G. de: e-mail to A. Soifer, November 1, 2005

[Bru12] Bruijn, N. G. de: e-mail to A. Soifer, November 3, 2005

[BE1] de Bruijn, N.G., Erdős, P.: On a combinatorial problem. Indagationes Math. **10**, 421–423 (1948)

[BE2] de Bruijn, N.G., Erdős, P.: A colour problem for infinite graphs and a problem in the theory of relations. *Indagationes Math.* **13**, 369–373 (1951)

[BFS] Brüning, J., Ferus, D., Siegmund-Schultze, R.: Terror and Exile: Persecution and Expulsion of Mathematicians from Berlin between 1933 and 1945. Deutsche Mathematiker–Vereinigung, Berlin (1998)

[Buk] Bukh, B.: Measurable sets with excluded distances. Geom. Funct. Anal. **18**, 668–697 (2008)

[BM] Burkill, H., Mirsky, L.: Monotonicity. J. Math. Anal. Appl. **41**, 391–410 (1973)

[CET] Calkin, N.J., Erdős, P., Tovey, C.A.: New Ramsey bounds from cyclic graphs of prime order. SIAM J. Discret. Math. **10**, 381–387 (1997)

[CGMS] Campos, M., Griffiths, S., Morris, R., Sahasrabudhe, J.: An exponential improvement for diagonal Ramsey. arXiv:2303.09521v1

[Cam1] Camus, A.: Essay on music, 1932. In: Youthful Writings by Albert Camus. Alfred A. Knopf, New York (1976)

[Cam2] Camus, A.: [The Artist and His Age], Lecture at Uppsala University, December 14, 1957, in the anthology Speaking Out: Lectures and Speeches 1937–1958. Vintage International, New York (2022)

[Can1] Cantwell, K.: Finite euclidean Ramsey theory. J. Combin. Theory Ser. A. **73**(2), 273–285 (1996)

[Can2] Cantwell, K.: All regular polytopes are Ramsey. J. Combin. Theory Ser. A. **114**, 555–562 (2007)

[Cay1] Cayley, A.: Report of meeting, 13 June 1878. In: Proceedings of the London Mathematical Society, IX, p. 148. C. F. Hodgson & Son, Gough Square, London (1878)

[Cay2] Cayley, A.: Report of meeting, 13 June 1878. In: Nature, XVIII, p. 204. MacMillan and CO., London (1878)

[Cay3] Cayley, A: On the colouring of maps. In: Proceedings of the Royal Geographical Society, London, pp. 259–261 (1879)

[CL] Chartrand, G., Lesniak, L.: Graphs & Digraphs, 4th edn. Chapman & Hall/CRC, Boca Raton (2005)

[CLZ] Chartrand, G., Lesniak, L., Zhang, P.: Graphs & Digraphs. Chapman & Hall/CRC, Boca Raton (2011)

[CRW] Chen, M., Raspaud, A., Wang, W.: Three-coloring planar graphs without short cycles. Inf. Process. Lett. **101**, 134–138 (2007)

[CKR] Cherkashin, D., Kulikov, A., Raigorodskii, A.: On the chromatic numbers of small-dimensional Euclidean spaces. Discret. Appl. Math. **243**, 125–131 (2018)

[Che] Chetwynd, A.G.: Total colourings of graphs – a progress report, graph theory, combinatorics, and applications. In: Alavi, Y., et al. (eds.) Proceedings of the Sixth Quadrennial Conference on Graph Theory and Applications of Graphs, 1988, vol. 1, 1991, pp. 233–244, Wiley, New York

[Chi1] Chilakamarri, K.B.: Unit-distance graphs in rational n-spaces. Discret. Math. **69**, 213–218 (1988)

[Chi2] Chilakamarri, K.B.: On the chromatic number of rational five-space. Aequationes Math. **39**, 146–148 (1990)

[Chi3] Chilakamarri, K.B.: Unit-distance graphs in Minkowski metric spaces. Geom. Dedicata. **37**, 345–356 (1991)

[Chi4] Chilakamarri, K.B.: The unit-distance graph problem: a brief survey and some new results. Bull. ICA. **8**, 39–60 (1993)

[Chi5] Chilakamarri, K.B.: A characterization in Z^n of finite Unit-distance graphs in R^n, (manuscript)

[Chi6] Chilakamarri, K.B.: A 4-chromatic unit-distance graph with no triangles. Geombinatorics. **IV**(3), 64–76 (1995)

[CG] Chung, F., Graham, R.L.: Forced convex *n*-gons in the plane. Discret. Comput. Geom. **19**, 3677–3371 (1998)

[Cho] Chow, T.: Distances forbidden by two-colorings of Q^3 and A_n (manuscript)

[Chow] Chowla, S.: There exists an infinity of 3–combinations of primes in A.P. Lahore Philos. Soc. **6**(2), 15–16 (1944)

[Chu] Church, A., Paul, J.: Cohen and the continuum problem. In: Petrovsky, I.G. (ed.) Proceedings of International Congress of Mathematicians (Moscow – 1966), pp. 15–20. Mir, Moscow (1968)

[Chv] Chvátal, V.: Some unknown van der Waerden numbers. In: Combinatorial Structures and their Applications, pp. 31–33. Gordon and Breech, New York (1969)

[CH] Chvátal, V., Harary, F.: Generalized Ramsey theory for graphs. Bull. Am. Math. Soc. **78**, 423–426 (1972)

[CJW] Chybowska-Sokół, J., Junosza-Szaniawski, K., Węsek, K.: Coloring distance graphs on the plane. Discret. Math. **346**, 113441 (2023)

[Cib] Cibulka, J.: On the chromatic number of real and rational spaces. *Geombinatorics*. **XVIII**(2), 53–65 (2008)

[Coh1] Cohen, P.J., The independence of the continuum hypothesis: Proceeding of the National Academy of Science, USA, Part I: 50 (1963), 1143–1148; Part II: 51, pp. 105–110 (1964)

[Coh2] Cohen, P.J.: Set Theory and the Continuum Hypothesis. Benjamin/Cummings Reading, Massachusetts (1966)

[Coh3] Cohen, P.J.: e-mail to A. Soifer, June 12, 2006

[CS] Conway, J.H., Sloane, N.J.A.: Sphere Packings, Lattices and Groups, 3rd ed., with contrib. by Bannai, E., Borcherds, R.E., Leech, J., Norton, S.P., Odlyzko, A.M., Parker, R.A., Queen, L., Venkov, B.B. Springer, New York (1999)

[Cou1] Coulson, D.: An 18-colouring of 3-space omitting distance one. Discret. Math. **170**(1–3), 241–247 (1997)

[Cou2] Coulson, D.: A 15–colouring of 3-space omitting distance one. Discret. Math. **256**(1–2), 83–90 (2002)

[Cou3] Coulson, D.: Tilings and colourings of 3-space. Geombinatorics. **XII (3)**, 102–116 (2003)

[Cour] Courant, R.: Reminiscences from Hilbert's Göttingen. Math. Intell. **3**, 154–164 (1981)

[Cox] Coxeter, H.S.M.: The four-color map problem, 1840–1890. Math. Teach. **52**, 283–289 (1959)

[CoxJ] Cox, J.: Letter to A. Soifer dated November 21, 1991

[Cox1] Coxeter, H.S.M.: Map-coloring problems. Scripta Math. **23**, 11–25 (1957)

[Cox2] Coxeter, H.S.M.: The Four-Color map problem, 1840–1890. Math. Teach. **52**, 283–289 (1959)

[Cro] Croft, H.T.: Incidence incidents. Eureka (Cambridge). **30**, 22–26 (1967)

[CFG] Croft, H.T., Falconer, K.J., Guy, R.K.: Unsolved Problems in Geometry. Springer, New York (1991)
[CT] Csizmadia, G., Tóth, G.: Note on a Ramsey–type problem in geometry. J. Combin. Theory Ser. A. **65**(2), 302–306 (1994)
[Dal1] van Dalen, D.: Mystic, Geometer, and Intuitionist: The Life of L. E. J. Brouwer, Volume I: The Dawning Revolution. Oxford University Press, Oxford (1999)
[Dal2] van Dalen, D.: Mystic, Geometer, and Intuitionist: The Life of L. E. J. Brouwer, Volume II: Hope and Disillusion. Oxford University Press, Oxford (2005)
[DGK] Danzer, L., Grünbaum, B., Klee, V.: Helly's Theorem and its Relatives. American Mathematical Society, Providence (1963)
[Dav] Davies, J.: Odd distances in colourings of the plane, revised 6 October 2022., arXiv:2209.15598v2, accepted by *Geometric and Functional Analysis* in January 2024
[DMP] Davies, J., McCarty, R., Pilipczuk, M.: Prime and polynomial distances in colourings of the plane, submitted to *Israel Journal of Mathematics*
[Del] Delahaye, J.-P.: Coloriages irréals. Pour la Science, Février, 88–93 (2005). English translation: Imaginary coloring, Geombinatorics **XV**(3), 101–119 (2006)
[DR] Delannoy, H., Ramsey, A.S.: Réponse à question 51. Interméd. Math. **1**, 192 (1894)
[DeM1] De Morgan, A.: Letter to W. R. Hamilton, October 23, 1852; TCD MS 1493, 668; Trinity College Dublin Library, Manuscripts Department
[DeM2] De Morgan, A.: Letter to W. Whewell, December 9, 1853, Trinity College Cambridge Whewell Add. mss., a.202[125]
[DeM3] De Morgan, A.: Letter to R. L. Ellis, June 24, 1854, Trinity College Cambridge, Whewell Add. mss., c.67[111]
[DeM4] [De Morgan, A]: Review of the philosophy of discovery..., The Athenaeum 1694, London, pp. 501–503 (1860)
[DeM5] De Morgan, A: Correspondence to Sir Herschel, in Memoir of Augustus De Morgan, by his wife S. E. De Morgan, Longmans, Green, and Co., London, p. 333 (1882)
[Des1] Descartes, B., A three–colour problem, Eureka **9**, April 21, 1947
[Des2] [Descartes, B.]: A three–colour problem, solutions to problems in Eureka No. 9. Eureka **10,** March 24, 1948
[Des3] Descartes, B.: Advanced problems and solutions: solution of problem 4526. Am. Math. Mon. **61**(5), 352–353 (1954)
[Dick] Dick, A.: Emmy Noether 1882–1935. Birkhäuser, Boston (1981)
[Dic1] Dickson, L.E.: On the last theorem of Fermat. Q. J. Pure Appl. Math. **40**, 27–45 (1908)
[Dic2] Dickson, L.E.: History of the Theory of Numbers, vol.'s 1 (1919), 2 (1920), 3 (1923). Carnegie Institution of Washington, Washington, DC
[Die] Diestel, R.: Graph Theory, 2nd edn. Springer, New York (2000)

[Dir] Dirac, G.A.: A property of 4-chromatic graphs and some remarks on critical graphs. J. Lond. Math. Soc. **27**, 85–92 (1952)
[Dol1] Dold-Samplonius: Bartel Leendert van der Waerden befragt von Yvonne Dold-Samplonius. NTM Int. J. Hist. Ethics Nat. Sci. Technol. Med. **2**, 129–147 (1994). English translation: Interview with Bartel Leendert van der Waerden, Notices Amer. Math. Soc. **44**(3), pp. 313–320 (1997)
[Dol2] Dold-Samplonius: In memoriam: Bartel Leendert van der Waerden (1903–1996). Hist. Math. **24**, 125–130 (1997)
[DuR] Dudek, A., Rödl, V.: On the Folkman number *f*(2, 3, 4). Exp. Math. **17**, 63–67 (2008)
[Dum] Dumitrescu, A.: A Remark on the Erdős-Szekeres Theorem. Amer. Math. Monthly **112**(10), 921–924 (2005)
[DT] Dummit, D.S., Foot, R.M.: Abstract Algebra, 2nd edn. Wiley, New York (1999)
[Dup] Duparc, H.J.A.: Handwritten recollection notes for A. Soifer, Delft, September 1996
[DU] Dynkin, E.B., Uspensky, V.A.: Matematicheskie Besedy, Gosudarstvennoe Izdatel'stvo Tekhniko-Teoreticheskoy Literatury, Moscow, 1952 (Russian). Partial English translation (part 1 of three only): Multicolor Problems. D. C. Heath & Co., Boston, 1963
[Dye] Dye, D.S.: A Grammar of Chinese Lattice. Harvard-Yenching Institute, Harvard University Press, Cambridge, MA (1937)
[DHJP] Polymath, D.H.J.: Polymath 16. https://dustingmixon.wordpress.com/2018/04/14/polymath16 first-thread-simplifying-de-greys-graph/
[Eck0] Eckmann, B.: e-mail to A. Soifer of December 7, 2004
[Eck1] Eckmann, B.: e-mail to A. Soifer of December 23, 2004
[Eck2] Eckmann, B.: e-mail to A. Soifer of December 27, 2004
[Eck3] Eckmann, B.: e-mail to A. Soifer of December 30, 2004
[Ein1] Einstein, A.: Geometry and experience, a lecture before the Prussian Academy of Sciences, January 27, 1921. In: Seelig, C. (ed.) Ideas and Opinions by Albert Einstein. Crown, New York (1954)
[Ein2] Einstein, A. (author), Schillp, P.A. (ed.) Albert Einstein: Philosopher–Scientist. Open Court, La Salle (1970)
[Eis] Eisenreich, G.: B. L. Van der Waerden Wirken von 1931 bis 1945 in Leipzig. In: Beckert, H., Schumann, H. (eds.) 100 Jahre Mathematisches Seminar der Karl–Marx Universität Leipzig. VEB Deutscher Verlag der Wissenschaften, Berlin (1981)
[E42.06] Erdős, P.: Some set–theoretical properties of graphs. Revista de la Univ. Nac. de Tucumán, Ser A. Mat. y Fis. Teór. **3**, 363–367 (1942)
[ER] Erdős, P., Rado, R.: Combinatorial theorems on classification of subsets of a given set. Proc. London Math. Soc. **2**(3), 417–439 (1952)
[Epp] Epple, M.: Email to Alexander Soifer, April 23, 2014
[E47/8/4ltr] Erdős, P.: Letter to N. G. de Bruijn, August 4, 1947
[E47.09] Erdős, P.: Some remarks on the theory of graphs. Bull. Amer. Math. Soc. **53**, 292–294 (1947)

[E50.07] Erdős, P.: On integers of the form $2^k + p$ and some related problems. Summa Brasil. Math. **2**, 113–124 (1950)

[E51.01] Erdős, P.: A colour problem for infinite graphs and a problem in the theory of relations. Proc. Konink. Nederl. Akad. Wetensch. Amsterdam. **54**, 371–373 (1951)

[E52.03] Erdős, P.: Egy kongruenciarendszerekröl szóló problémáról (On a problem concerning covering systems). Mat. Lapok. **3**, 122–128 (1952) (Hungarian)

[E57.13] Erdős, P.: Some unsolved problems. Michigan Math. J. **4**, 291–300 (1957)

[E59.06] Erdős, P.: Graph theory and probability. Can. J. Math. **11**, 34–38 (1959)

[E60.09] Erdős, P.: On sets of distances of n points in Euclidean space. MTA MKI Közl. **5**, 165–169 (1960)

[E61.05] Erdős, P.: Graph theory and probability II. Can. J. Math. **13**, 346–352 (1961)

[E61.22] Erdős, P.: Some Unsolved Problems. *Magyar tudomanyos akademia mathematikai kutato intezetenek közlemenyel* 6 ser. A (1–2), 221–254 (1961)

[E63.21] Erdős, P.: Nekotorye Nereshennye Problemy (Some unsolved problems). Matematika. **7**(4), 109–143 (1963) (A Russian translation of [E61.22])

[E71.13] Erdős, P.: On the application of combinatorial analysis to number theory, geometry and analysis. In: Actes du Congrès International des Mathematiciens (Nice, 1970), Tome 3, pp. 201–210. Gauthier–Villars, Paris (1971)

[E73.21] Erdős, P.: Problems and results on combinatorial number theory. In: A Survey of Combinatorial Theory (Proc. Internat. Sympos., Colorado State Univ., Fort Collins, Colorado, 1971), pp. 117–138. North–Holland, Amsterdam (1973)

[E75.24] Erdős, P.: On some problems of elementary and combinatorial geometry. Annali Matematica Pura Applicata. **103**(4), 99–108 (1975)

[E75.25] Erdős, P.: On some problems of elementary and combinatorial geometry. Ann. Mat. Pura Appl. **103**(4), 99–108 (1975)

[E75.33] Erdős, P.: Problems and results on finite and infinite graphs. In: Recent Advances in Graph Theory (Proc. Second Czechoslovak Sympos., Prague, 1974), pp. 183–192. Academia, Prague (1975)

[E76.35] Erdős, P.: Problems and results in combinatorial analysis. Colloquio Internacionale sulle Teorie Combinato*rie* (Rome, 1973), Tomo II, Atti dei Convegni Lincei, No. 17, pp. 3–17. Accad. Naz. Lincei, Rome 1976

[E75.40] Erdős, P.: Some problems on elementary geometry. Austr. Math. Soc. Gaz. **2**, 2–3 (1975)

[E76.49] Erdős, P.: Problem (p. 681); P. Erdős, Chairman, Unsolved Problems. In: Nash-Williams, C.St.J.A., Sheehan, J. (eds.) Proceedings of the Fifth British Combinatorial Conference 1975, University of Aberdeen, July 14–18, 1975, Congr. Numer. **XV**, Winnipeg, Utilitas Mathematica (1976)

[E77.26] Erdős, P.: Problems and results on combinatorial number theory, II. J. Indian Math. Soc. (N.S.). **40**(1–4), 285–298 (1976)

[E77.28] Erdős, P.: Problems in number theory and combinatorics. In: Proceedings of the Sixth Manitoba Conference on Numerical Mathematics (Univ. Manitoba, Winnipeg, Man., 1976). Congr. Numer. **XVIII**, 35–58, Utilitas Math., Winnipeg, Man. (1977)

[E78.50] Erdős, P.: Problem 28 in "Exercises, Problems and Conjectures,". In: Bollobás, B. (ed.) Extremal Graph Theory, London Mathematical Society Monographs, p. 285. Academic (1978)

[E79.04] Erdős, P.: Combinatorial problems in geometry and number theory. In: Relations between combinatorics and other parts of mathematics, Proc. Sympos. Pure Math., Ohio State Univ., Columbus Ohio (1978); Proc. Sympos. Pure Math., XXXIV. American Mathematical Society, Providence, pp. 149–162 (1979)

[E80.03] Erdős, P.: A survey of problems in combinatorial number theory. In: Combinatorial Mathematics, Optimal Designs and their Applications (Proc. Sympos. Combin. Math. and Optimal Design, Colorado State Univ., Fort Collins, Colo., 1978). Ann. Discret. Math. 6 (1980), 89–115

[E80.38] Erdős, P.: Some combinatorial problems in geometry. In: Geometry and Differential Geometry (Proc., Haifa, Israel, 1979). Lecture Notes in Math., 792. Springer, pp. 46–53 (1980)

[E80.41] Erdős, P.: Some combinational problems in geometry. In: Geometry and Differential Geometry (Proc. Conf., Univ. Haifa, Haifa, 1979), Lecture Notes in Math., 792, pp. 46–53. Springer, Berlin (1980)

[E81.16] Erdős, P.: On the combinatorial problems which I would most like to see solved. Combinatorica. **1**, 25–42 (1981)

[E81.20] Erdős, P.: Problems and results in graph theory. In: The Theory and Application of Graphs (Kalamazoo, Mich., 1980), pp. 331–341. Wiley, New York (1981)

[E81.23] Erdős, P.: Some applications of graph theory and combinatorial methods to number theory and geometry. Algebraic Methods in Graph Theory (Colloq. held in Szeged, Hungary 1978), Vol. I, Coll. Math Soc. J. Bolyai. **25**, 137–148 (1981)

[E81.26] Erdős, P.: Some new problems and results in graph theory and other branches of combinatorial mathematics. In: Combinatorics and Graph Theory, Proc. Symp. Calcutta 1980, Lecture Notes in Mathematics, 885, pp. 9–17. Springer (1981)

[E81.32] Erdős, P.: My Scottish book "problems". In: The Scottish Book. Mathematics from the Scottish Café, pp. 35–43. Birkhäuser, Boston (1981)

[E83.01] Erdős, P.: Some combinatorial problems in geometry. In: Geometry and Differential Geometry, Lecture Notes in Mathematics, 792, pp. 46–53. Springer, Berlin (1983)

[E83.03] Erdős, P.: Combinatorial problems in geometry. Math. Chron. **12**, 35–54 (1983)

[E85.01] Erdős, P.: Problems and results in combinatorial geometry. In: Discrete Geometry and Convexity, Annals of the New York Academy of Sciences, 440, pp. 1–11. The New York Academy of Sciences, New York (1985)

[E85.33] Erdős, P.: Some problems and results in number theory. In: Number Theory and Combinatorics, Japan 1984 (Tokyo, Okayama and Kyoto, 1984), pp. 65–87. World Scientific Publishing, Singapore (1985)

[E85.34] Erdős, P.: Some solved and unsolved problems of mine in number theory. In: Topics in Analytic Number Theory (Austin, Tex., 1982), pp. 59–75. University of Texas Press, Austin (1985)

[E87.12] Erdős, P., My joint work with Richard Rado, Surveys in Combinatorics 1987 (New Cross, 1987), London Mathematical Society Lecture Note series 123, Cambridge University Press, Cambridge/New York, 1987, 53–80

[E88.28] Erdős, P.: Problems and results in combinatorial analysis and graph theory. Proceedings of the First Japan Conference on Graph Theory and Applications (Hakone, 1986). Discret. Math. **72**(1–3), 81–92 (1988)

[E89.27] Erdős, P.: Problems and results on extremal problems in number theory, geometry, and combinatorics. Proceedings of the 7th Fischland Colloquium, (Wustrow, 1988), Rostock. Math. Kolloq. **38**, 6–14 (1989)

[E89.35] Erdős, P.: Some problems and results on combinatorial number theory. In: Graph theory and its applications: East and West (Jinan, 1986). Annals of the New York Academy of Sciences 576, pp. 132–145. New York Academy of Sciences, New York (1989)

[E89.60] Erdős, P.: Video recording of the talk "Some of My Favorite Problems I". University of Colorado at Colorado Springs, March 16, 1989

[E89.61] Erdős, P.: Video recording of the talk "Some of My Favorite Problems II". University of Colorado at Colorado Springs, March 17, 1989

[E89CS] Erdős, P.: An interview in the documentary Colorado Mathematical Olympiad (1989)

[E90.23] Erdős, P.: Problems and results on graphs and hypergraphs: similarities and differences. In: Mathematics of Ramsey Theory, Algorithms and Combinatorics 5, pp. 12–28. Springer, Berlin (1990)

[E91.31] Erdős, P.: Problems and results in combinatorial analysis and combinatorial number theory. In: Graph theory, Combinatorics and Applications, vol. 1 (Kalamazoo, MI, 1988), pp. 397–406. Wiley, New York (1991)

[E91.60] Erdős, P.: Video recording of the talk "Some of My Favorite Problems I". University of Colorado at Colorado Springs, December 23, 1991

[E91/7/12ltr] Erdős, P.: Letter to A. Soifer of July 12, 1991

[E91/7/16ltr] Erdős, P.: Letter to A. Soifer of July 16, 1991

[E91/8/10ltr] Erdős, P.: Letter to A. Soifer of August 10, 1991

[E91/8/14ltr] Erdős, P.: Letter to A. Soifer of August 14, 1991

[E91/10/2ltr] Erdős, P.: Letter to A. Soifer, received on October 2, 1991 (marked "1977 VII 25")

[E92.14] Erdős, P.: Obituary of My Friend and Coauthor Tibor Gallai. Geombinatorics. **II**(1), 5–6 (1992) corrections: II(2), 1992, 37

[E92.15] Erdős, P.: In memory of Tibor Gallai. Combinatorica. **12**(4), 373–374 (1992)

[E92.19] Erdős, P.: On some unsolved problems in elementary geometry. Mat Lapok (N.S.). **2**(2), 1–10 (1992) (Hungarian)

[E92.60] Erdős, P.: Video recording of the talk “Some of My Favorite Problems II”. University of Colorado at Colorado Springs, January 10, 1992

[E93.20] Erdős, P.: Some of my favorite solved and unsolved problems in graph theory. Quaest. Math. **16**(3), 333–350 (1993)

[E94.21] Erdős, P.: Problems and results in discrete mathematics, trends in discrete mathematics. Discret. Math. **136**(1–3), 53–73 (1994)

[E94.22] Erdős, P.: Problems and results on set systems and hypergraphs. In: Extremal Problems for Finite Sets (Visegr˙ád, 1991), Bolyai Society Mathematical Studies 3, pp. 217–227. János Bolyai Mathematical Society, Budapest (1994)

[E94.26] Erdős, P.: Some problems in number theory, combinatorics and combinatorial geometry. Math. Pannon. **5**(2), 261–269 (1994)

[E94.60] Erdős, P.: Video recording of the talk “Twenty Five Years of Questions and Answers”. In: 25th South–Eastern International Conference on Combinatorics, Graph Theory and Computing. Florida Atlantic University, Boca Raton, March 10, 1994

[E95.20] Erdős, P.: Some of my favourite problems in number theory, combinatorics, and geometry. Combinatorics Week (Portuguese), São Paulo, 1994. Resenhas 2, 165–186 (1995)

[E95.32] Erdős, P.: On the 120-th anniversary of the birth of Schur. Geombinatorics. **V**(1), 4–5 (1995)

[E95BR] Erdős, P.: Interview with A. Soifer, March 7, 1995, Boca Raton, Florida

[E97.18] Erdős, P.: Some of my favorite problems and results. In: Graham, R.L., Nešetřil, J. (eds.) The Mathematics of Paul Erdős, vol. I, pp. 47–67. Springer, Berlin (1997)

[E97.21] Erdős, P.: Some of my favourite unsolved problems. Math. Japon. **46**(3), 527–537 (1997)

[EBC] Erdős, P., Bollobás, B., Catlin, P.A.: Hadwiger’s conjecture is true for almost every graph. Eur. J. Comb. **1**(3), 195–199 (1980)

[EB] Erdős, P., de Bruijn, N.G.: A colour problem for infinite graphs and a problem in the theory of relations. Indag. Math. **13**, 371–373 (1951)

[EHT] Erdős, P., Harary, F., Tutte, W.T.: On the dimension of a graph. Mathematika, London. **12**, 118–122 (1965)

[EG0] Erdős, P., Graham, R.: Old and new problems and results in combinatorial number Theory: van der Waerden’s theorem and related topics. L’Enseignement Mathématique. **25**(3–4), 325–344 (1979)

[EG] Erdős, P., Graham, R.: Old and New Problems and Results in Combinatorial Number Theory. L’Enseignement Mathématique, Université de Genève (1980)

[EGMRSS] Erdős, P., Graham, R.L., Montgomery, P., Rothschild, B.L., Spencer, J., Straus, E.G.: Euclidean Ramsey theorems I, II, and III. J Combin. Theory Ser. A. 14, 341–363 (1973); Coll. Math. Soc. Janos Bolyai. 10 (1973). North Holland, Amsterdam, 529–557 and 559–583 (1975)

[EH1] Erdős, P., Hajnal, A.: On chromatic numbers of graphs and set-systems. Acta Mathematica Acad. Sci. Hungar. **17**(1–2), 61–99 (1966)

[EH2] Erdős, P., Hajnal, A.: Research problem 2–5. J. Combin. Theory. **2**, 104 (1967)

[EP1] Erdős, P., Purdy, G.: Some extremal problems in geometry, III. In: Proc. Sixth Southeastern Conf. on Combinatorics, Graph Theory and Computing (Florida Atlantic Univ., Boca Raton, FL, 1975), Congr. Numer. XIV, pp. 291–308. Utilitas Mathematica, Winnipeg (1975)

[EP2] Erdős, P., Purdy, G.: Some extremal problems in combinatorial geometry. In: Graham, R., Grötschel, M., Lovász, L. (eds.) Handbook of Combinatorics. North Holland, Amsterdam (1991)

[EP3] Erdős, P., Purdy, G.: Some Extremal Problems in Discrete Geometry (manuscript, shared by Erdős with A. Soifer)

[ESS] Erdős, P., Sárközy, A., Sós, V.T.: On conjecture of Roth and some related problems. I. Irregularities of Partitions (Fertöd, 1986). In: Algorithms Combin. Study Res. Texts 8, pp. 47–59. Springer, Berlin (1989)

[ESi] Erdős, P., Simonovits, M.: On the chromatic number of geometric graphs. Ars Comb. **9**, 229–246 (1980)

[ESu1] Erdős, P., Surányi, J.: Válogatott Fejezetek a Számelméletbol, Polygon, Szeged, 2nd ed. (1996) (Hungarian)

[ESu2] Erdős, P., Surányi, J.: Topics in the Theory of Numbers, trans. from the Hungarian, 2nd edn. Springer, New York (2003)

[ES1] Erdős, P., Szekeres, G.: A combinatorial problem in geometry. Compos. Math. **2**, 463–470 (1935)

[ES2] Erdős, P., Szekeres, G.: On some extremum problems in elementary geometry. Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae, Sectio Math. **3**, 53–62 (1960/61)

[ET] Erdős, P., Turán, P.: On some sequences of integers. J. Lond. Math. Soc. **11**, 261–264 (1936)

[EW] Erdős, P., Wilson, R.J.: On the chromatic index of almost all graphs. J. Comb. Theory Ser. B. **23**, 255–257 (1977)

[Err] Errera, A.: Du Coloriage des Cartes et de Quelques Questions d'Analysis Situs. Thèse pour le grade de Docteur Special en Sciences Physiques et Mathematiques, Gauthier–Villars & Co, Paris (1921)

[Eri] Erickson, M.: An upper bound for the Folkman number F(3,3;5). J. Graph Theory. **17**(6), 679–681 (1993)

[Ex3] Exoo, G.: Applying optimization algorithm to Ramsey problems. In: Alavi, Y. (ed.) Graph theory, combinatorics, algorithms, and applications, pp. 175–179. SIAM, Philadelphia (1989)

[Ex4] Exoo, G.: A lower bound for $R(5, 5)$. J. Graph Theory. **13**, 97–98 (1989)

[Ex5] Exoo, G.: On two classical Ramsey numbers of the form $R(3,n)$. SIAM J. Discret. Math. **2**, 488–490 (1989)

[Ex9] Exoo, G.: Announcement: On the Ramsey Numbers $R(4,6)$, $R(5,6)$ and $R(3,12)$. Ars Combinatorica 35 (1993) 85. The construction of a graph proving $R\ (4,6) \geq 35$ is presented in detail at https://cs.indstate.edu/ge/RAMSEY (2001)

[Ex15] Exoo, G.: Some applications of pq-groups in graph theory. Discuss. Math. Graph Theory. **24**, 109–114 (2004) Constructions available at https://cs.indstate.edu/ge/RAMSEY

[Ex17] Exoo, G., Personal communication to S. P. Radziszowski (2005–2006). Constructions at https://cs.indstate.edu/ge/RAMSEY

[Ex18] Exoo, G.: A lower bound for Schur numbers and multicolor Ramsey numbers of K_3. Electron. J. Combinatorics **1**, No. 1, R8, 1–3, 1994. http://www.combinatorics.org/Volume_1/Abstracts/v1i1r8.html

[Ex19] Exoo, G.: ε-unit distance graphs. Discrete Comput. Geom. **33**(1), 117–123 (2005)

[EI1] Exoo, G., Ismailescu, D.: Small order triangle-free 4-chromatic unit distance graphs. Geombinatorics. **XXVI**(2), 49–64 (2016)

[EI2] Exoo, G., Ismailescu, D.: The Hadwiger-Nelson problem with two forbidden distances. Geombinatorics. **XXVIII**(1), 51–68 (2018)

[EI3] Exoo, G., Ismailescu, D.: The chromatic number of the plane is at least 5 – a new proof, arXiv:1805.00157, May 2018; and also Discrete Comput. Geom. (2019). https://doi.org/10.1007/s00454-019-00058-1

[EI4] http://cs.indstate.edu/ge/ExooIsmailescuData

[EI5] Exoo, G., Ismailescu, D.: A 6-chromatic two-distance graph in the plane. Geombinatorics. **XXIX**(3), 97–103 (2020)

[EI6] Exoo, G., Ismailescu, D..: http://cs.indstate.edu/ge/ ExooIsmailescuData/vertices.1.2.graph.txt

[EI7] Exoo, G., Ismailescu, D..: http://cs.indstate.edu/ge/ExooIsmailescuData/matrix.1.2.graph.txt

[EI8] Exoo, G., Ismailescu, D.: On the Chromatic Number of R^n for Small Values of n (2014). https://arxiv.org/abs/1408.2002

[EIL] Exoo, G., Ismailescu, D.P., Lim, M.: On the chromatic number of R^4. Discrete Comput. Geom. **52**, 416–423 (2014)

[Fal1] Falconer, K.J.: The realization of distances in measurable subsets covering R^n. Com. Theory (A). **31**, 187–189 (1981)

[Fal2] Falconer, K.J.: 8-page manuscript written for this book, dated February 17, 2005

[Fis1] Fischer, K.G.: Additive k-colorable extensions of the rational plane. Discret. Math. **82**, 181–195 (1990)

[Fis2] Fischer, K.G.: The connected components of the graph $Q\left(\sqrt{N_1}, \ldots, \sqrt{N_d}\right)^2$. Congr. Numer. **72**, 213–221 (1990)

[Fla] Flam, J.: Matisse on Art. University of California Press, Berkeley (1995)

[Foll] Folland, G.: Real Analysis: Modern Techniques and their Applications, 2nd edn. Wiley, New York (1999)

[Fol] Folkman, J.: Graphs with monochromatic complete subgraphs in every edge coloring, Rand Corporation Memorandum RM–53–58–PR, October 1967. Also appeared under the same title in SIAM J. Appl. Math. **18**, 19–24 (1970)

[FR] Fox, J., Radoicic, R.: The axiom of choice and the degree of regularity of equations over the reals, unpublished

[Fra1] Fraenkel, A.: Zu den Grundlagen der Cantor–Zermeloschen Mengenlehre. Math. Ann. **86**(3–4), 230–237 (1922) (German)

[Fra2] Fraenkel, A.: Axiomatische Begründung der transfiniten Kardinalzahlen. I. Math. Zeitschr. **13**(1), 153–188 (1922) (German)

[Fra3] Fraenkel, A.: Untersuchungen über die Grundlagen der Mengenlehre. Math. Zeitschr. **22**(1), 250–273 (1925) (German)

[FR1] Frankl, P., Rödl, V.: All triangles are Ramsey. Trans. Am. Math. Soc. **297**, 777–779 (1986)

[FR2] Frankl, P., Rödl, V.: Large triangle-free subgraphs in graphs without K_4. Graphs Combin. **2**(2), 135–144 (1986)

[FR3] Frankl, P., Rödl, V.: A partition property of simplices in Euclidean space. J. Am. Math. Soc. **3**, 1–7 (1990)

[FW] Frankl, P., Wilson, R.M.: Intersection theorems with geometric consequences. Combinatorica. **1**, 357–368 (1981)

[F] Fredrickson, H.: Schur numbers and the Ramsey numbers N(3,3,...,3;2). J. Combin. Theory Ser A. **27**, 371–379 (1979)

[FS] Fredrickson, H., Sweet, M.M.: Symmetric sum-free partitions and lower bounds for Schur numbers. Electron. J. Combin. **7**(1), R32, 1–9, (2000). http://www.combinatorics.org/Volume_7/Abstracts/v7i1r32.html

[Fre] Frei, G.: Helmut Hasse (1898–1979). Expo. Math. **3**, 55–69 (1985)

[FLR] Frei, G., Lemmermeyer, F., Roquette, P. (eds.): Emil Artin and Helmut Hasse: the Correspondence 1923–1958. Springer, Basel (2014)

[FR] Frei, G., Roquette, P. (eds.): Helmut Hasse – Emil Artin: Die Korrespondenz 1923–1934. Universitätsverlag Göttingen (2008). http://www.rzuser.uni-heidelberg.de/~ci3/hasse-artin/hasart-web-080922.html

[FTW] Frei, G., Top, J., Walling, L.: A short biography of B. L. van der Waerden. Nieuw Archief voor Wiskunde, Vierde serie. **12**(3), 137–144 (1994)

[Fri] Friedman, H., Strong undecidable $\prod_2^0$ statements (in preparation)

[FF] Fritsch, R., Fritsch, G.: The Four-Color Theorem. Springer, New York (1998)

[Fur1] Furstenberg, H.: Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. J. D'Analyse Math. **31**, 204–256 (1977)

[Fur2] Furstenberg, H.: A polynomial Szemerédi theorem, manuscript dated January 1994

[Gal] Galperin, G.: A letter to A. Soifer, undated, post–stamped March 1, 2005

[Gars1] Garsia, A.: blue-lines manuscript sent to A. Soifer

[Gars2] Garsia, A.: e-mail to A. Soifer, February 28, 1995

[Gar1] Gardner, M.: The celebrated Four-Color map problem of topology. Sci. Am. **206**, 218–226 (Sept. 1960)

[Gar2] Gardner, M.: A new collection of "brain teasers". Sci. Am. **206** 172–180 (Oct. 1960)

[Gar3] Gardner, M.: The coloring of unusual maps leads into Uncharted territory. Sci. Am. **242**, 14–21 (1980)

[Gas] Gasarch, W.: IMU's non-controversial changing the name of the Nevanlinna prize. Computational Convexity, June 4 (2019). https://blog.computationalcomplexity.org/2019/06/imus-non-controversial-changing-name-of.html

[GS] Gethner, E., Springer II, W.M.: How false is Kempe's proof of the four-color theorem? Congr. Numer. **164**, 159–175 (2003)

[GS] Gismondi, S.J., Swart, E.R.: A new type of 4–colour reducibility. Congr. Numer. **82**, 33–48 (1991)

[Gof] Goffman, C.: And what is your Erdős number? Am. Math. Mon. **76**(7), 761 (1969)

[Göd1] Gödel, K.: Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I. Monatshefte für Mathematik und Physik. **38**, 173–198 (1931)

[Göd2] Gödel, K.: The Consistency of the Axiom of Choice and of the Generalized Continuum Hypothesis with the Axioms of Set Theory. Princeton University Press (1940)

[Gol1] Golomb, S.W.: Letter to A. Soifer, September 10, 1001

[Gol2] Golomb, S.W.: Letter to A. Soifer, September 25, 1001

[Gon] Gonthier, G.: A computer–checked proof of the Four Colour Theorem. http://research.microsoft.com/~gonthier/4colproof.pdf

[GO] Goodman, J.E., O'Rourke, J.: Handbook of Discrete and Computational Geometry. CRC Press, Boca Raton (1997)

[Gou] Goudsmit, S.A.: Alsos. Henry Schuman, New York (1947)

[Gow] Gowers, W.: A new proof of Szemerédi's theorem. GAFA, Geom. funct. anal. **11**, 465–588 (2001)

[Gra0] Graham, R.L.: On edgewise 2-colored graphs with monochromatic triangles and containing no complete hexagon. J. Combin. Theory. **4**, 300 (1968)

[Gra1] Graham, R.L.: Rudiments of Ramsey Theory. American Mathematical Society, Providence (1981) revised 1983

[Gra2] Graham, R.L.: Recent developments in Ramsey theory. In: Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Warsaw, 1983), 1555–1567. PWN, Warsaw (1984)

[Gra3] Graham, R.L.: Recent trends in Euclidean Ramsey Theory. Discret. Math. **136**, 119–127 (1994)

[Gra4] Graham, R.L.: A problem talk for ca. 200 finalists of the USA Mathematical Olympiad, Massachusetts Institute of Technology, May 4, 2002

[Gra5] Graham, R.L.: Talk at the Mathematical Sciences Research Institute, Berkeley, August 2003

[Gra6] Graham, R.L.: Open problems in Euclidean Ramsey Theory. Geombinatorics. **XIII**(4), 165–177 (2004)

[Gra7] Graham, R.L.: Some of my favorite problems in Ramsey theory. In: Landman, B., et al. (eds.) Proceedings of the 'Integers Conference 2005' in Celebration of the 70th Birthday of Ronald Graham, Carrolton, Georgia, USA, October 27–30, 2005, pp. 229–236. Walter de Gruyter, Berlin (2007)

[Gra8] Graham, R.L.: Old and new problems and results in Ramsey theory. In: Horizon of Combinatorics (Conference and EMS Summer School), Budapest and Lake Balaton, Hungary, 10–22 July 2006, to appear in 2007

[GN] Graham, R.L., Nešetřil, J.: Ramsey theory and Paul Erdős. In: Halász, G., et al. (eds.) Paul Erdős and his Mathematics II, pp. 339–365. Springer, Budapest (2002)

[GR0] Graham, R.L., Rothschild, B.L.: Rota's geometric analogue to Ramsey's theorem. In: Proceedings of Symposia in Pure Mathematics, vol. 19, pp. 101–104. American Mathematical Society, Providence (1971)

[GR1] Graham, R.L., Rothschild, B.L.: Ramsey's theorem for n-parameter sets. Trans. Am. Math. Soc. **159**, 257–292 (1971)

[GR2] Graham, R.L., Rothschild, B.L.: A survey of finite Ramsey theorems. In: Proceedings of the Second Louisiana Conference, on Combinatorics, Graph Theory and Computing, pp. 21–40 (1971)

[GR3] Graham, R.L., Rothschild, B.L.: Some recent developments in Ramsey theory. In: Combinatorics (Proc. Advanced Study Inst., Breukelen, 1974), Part 2: Graph theory; foundations, partitions and combinatorial geometry, 61–76. Math. Centre Tracts, No. 56, Math. Centrum, Amsterdam (1974). http://www.ams.org/mathscinet-getitem?mr=373948

[GRS1] Graham, R.L., Rothschild, B.L., Spencer, J.H.: Ramsey theory, 1st edn. Wiley, New York (1980)

[GRS2] Graham, R.L., Rothschild, B.L., Spencer, J.H.: Ramsey theory, 2nd edn. Wiley, New York (1990)

[GRSt] Graham, R.L., Rothschild, B.L., Straus, E.G.: Are there $n + 2$ points in E^n with odd integral distances? Am. Math. Mon. **81**, 21–25 (1974)

[GS1] Graham, R.L., Spencer, J.H.: On small graphs with forced monochromatic triangles. In: Recent Trends in Graph Theory (Proc. Conf., New York, 1970), 137–141. Lecture Notes in Mathematics, vol. 186. Springer, Berlin (1971)

[GS2] Graham, R.L., Spencer, J.H.: Ramsey theory. Sci. Am. 112–117 (July 1990)

[GT] Graham, R.L., Tressler, E.: Open problems in Euclidean Ramsey theory. In: Ramsey Theory Yesterday, Today, and Tomorrow by Alexander Soifer (ed.) [Soi49], Springer Birkhäuser, Progress in Mathematics 285, New York, 2011, 115–120

[GY0] Graver, J.E., Yackel, J.: An upper bound for Ramsey numbers. Bull. Am. Math. Soc. **72**, 1076–1079 (1966)

[GY] Graver, J.E., Yackel, J.: Some graph theoretic results associated with Ramsey's theorem. J. Combin. Theory. **4**, 125–175 (1968)

[Grav] Graves, R.P.: Life of Sir William Rowan Hamilton, Knt., LL. D., D. C. L., M. R. I. A., Andrews Professor of Astronomy in the University of Dublin, and Royal Astronomer of Ireland, etc.: including selections from his poems, correspondence, and miscellaneous writings, vol. 1 (1882), 2 (1885) and 3 (1889), Hodges, Figgis, & Co., Dublin

[GT] Green, B., Tao, T.: The primes contain arbitrarily long arithmetic progressions, 6th revision (2007). arXiv:math/0404188v6

[GG] Greenwood, R.E., Gleason, A.M.: Combinatorial relations and chromatic graphs. Can. J. Math. **7**, 1–7 (1955)

[G1] de Grey, A.D.N.J.: The Chromatic Number of the Plane Is at least 5. Geombinatorics. **XXVIII**(1), 18–31

[G2] Grey, A.D.N.J. de: Comments in Polymath16 1, thread 14, comment 24439, October 20, 2019. http://michaelnielsen.org/polymath1/index.php?title=Hadwiger-Nelson-problem

[G3] de Grey, A.D.N.J.: A small 6-chromatic unit-distance graph in R^3. Geombinatorics. **XXX**(1), 5–13

[GR] Grinstead, C., Roberts, S.: On the Ramsey numbers *R* (3,8) and *R(3,9)*. J. Combin. Theory Ser. B. **33**, 27–51 (1982)

[Gro] Groves, L.R.: Now it Can Be Told. Harper, New York (1962)

[Grö] Grötzsch, H.: Ein Dreifarbensatz ftir dreikreisfreie Netze auf der Kugel. Wissenschaftliche Zeitschrift der Martin–Luther–Universität Halle–Wittenberg. Mathematisch–Naturwissenschaftliche Reihe, 8, 109–120 (1958–1959)

[Grü] Grünbaum, B.: Grötzsch's theorem on 3-colorings. Mich. Math. J. **10**, 303–310 (1963)

[Gut1] The Late Professor Guthrie, *Cape Times*, October 23, 1899, Cape Town, South Africa

[Gut2] A biography of Francis Guthrie in *Dictionary of South African Biography*, vol. II, Tafelberg-Uitgewers LTD., Editor–in–chief D.W. Krüger, 279–80

[Gut3] Ritchie, W.: The History of the South African College 1829–1918. T. Maskew Miller, Cape Town (1918)

[GutFr] Guthrie, F.: Note on the colouring of maps. Proc. R. Soc. Edinb. **10**, 727–728 (1880)

[Guy] Guy, R.K.: Unsolved Problems in Number Theory, 2nd edn. Springer, New York (1994)

[Had1] Hadwiger, H.: Über eine Klassification der Steckenkomplexe. Vierteljschr. Naturforsch. Ges. Zürich. **88**, 133–142 (1943)

[Had2] Hadwiger, H.: Ein Überdeckungssätz für den Euklidischen Raum. Port. Math. **4**, 140–144 (1944)

[Had3] Hadwiger, H.: Uberdeckung des euklidischen Raum durch kongruente Mengen. Port. Math. **4**, 238–242 (1945)

[Had4] Hadwiger, H.: Ungelöste Probleme, Nr. 11. Elem. Math. **16**, 103–104 (1961)

[Hal] Halmos, P. R.: Has Progress in Mathematics Slowed Down? Amer. Math. Monthly **98**, 561–578 (1990)

[HD1] Hadwiger, H., Debrunner, H.: Ausgewählte einzelprobleme der kombinatorishen geometrie in der ebene. L'Enseiguement Mathematique. **1**, 56–89 (1955)

[HD2] Hadwiger, H., Debrunner, H.: Kombinatorischen Geometrie in der Ebene. L'Enseiguement Mathematique, Geneva (1959)

[HD3] Hadwiger, H., Debrunner, H.: Combinatorial Geometry of the Plane edited by Yaglom, I.M., Nauka, Moscow (1965). (extended by the editor Russian Translation version of [HD2])

[HDK] Hadwiger, H., Debrunner, H., Klee, V.: Combinatorial Geometry in the Plane. Holt, Rinehart and Winston, New York (1964)

[HN1] Hadziivanov, N.G., Nenov, N.D.: On the Graham–Spencer number. C. R. Acad. Bulg. Sci. **32**, 155–158 (1979) (Russian)

[HN2] Hadziivanov, N.G., Nenov, N.D.: Every (3, 3)–Ramsey graph without 5-cliques has more than 11 vertices (Russian). Serdica. **11**, 341–356 (1985)

[HJ] Hales, A., Jewett, R.: Regularity and positional games. Trans. Am. Math. Soc. **106**, 222–229 (1963)

[Ham] Hamilton, W.R.: Letter to A. De Morgan dated October 26, 1852; TCD MS 1493, 669; Trinity College Dublin Library, Manuscripts Department

[Har0] Harary, F.: Graph Theory. Addison Wesley, Reading (1969)

[Har1] Harary, F.: The Four-Color Conjecture and other graphical diseases. In: Harary, F. (ed.) Proof Techniques in Graph Theory, pp. 1–9. Academic Press, New York (1969)

[Har2] Harary, F., A Tribute to Frank P. Ramsey: J. Graph Theory. **7**(1983), 1–7 (1903–1930)

[Har3] Harary, F.: Letter to A. Soifer, February 19, 1996

[Harb] Harborth, H.: Konvexe Fünfecke in ebenen Punktmengen. Elem. Math. **33**, 116–118 (1978)

[HaKr] Harborth, H., Krause, S.: Ramsey numbers for circulant colorings. Congr. Numer. **161**, 139–150 (2003)

[HJNRX] Hassan, Z.R., Jiang, Y., Narváez, D.E., Radziszowski, S., Xu, X.: On some generalized vertex Folkman numbers. Graphs and Combinatorics (to appear). https://arxiv.org/abs/2110.03121

[HR] Hatfield, L., Ringel, G.: Pearls in Graph Theory: A Comprehensive Introduction, Academic Press, Boston (1994)

[Hav] Havel, I.: On a conjecture of B. Grünbaum. J. Combin. Theory. **7**, 184–186 (1969)

[Hea1] Heawood, P.J.: Map–Colour theorem. Q. J. Pure Appl. Math. **XXIV**, 332–338 (1890)

[Hea2] Heawood, P.J.: On the Four–Colour map theorem. Q. J. Pure Appl. Math. **XXIX**, 270–285 (1898)

[Head1] Headmaster of Clifton College, [Untitled], J. Educ. **IX**, 11–12 (January 1, 1887)

[Head2] Headmaster of Clifton College, [Untitled], J. Educ. **XI**, 277 (June 1, 1889)

[Hee1] Heesch, H.: Untersuchungen zum Vierfarbenproblem (German), B. I. Hochschulscripten, 810/810a/810b, Bibliographisches Institut, Mannheim–Vienna–Zürich, 1969

[Hee2] Heesch, H.: Chromatic reduction of the triangulations T_e, $e = e_5 + e_7$. J. Combin. Theory Ser. B. **13**, 46–55 (1972)

[Hei1] Heisenberg, W.: Die aktive und die passive Opposition im Dritten Reich, unpublished 4-page manuscript with the cover letter, dated November 12, 1947

[Hei2] Heisenberg, W.: Physics and Beyond: Encounters and Conversations. Harper & Row, New York (1971)

[Hei3] Heisenberg, W.: Across the Frontiers. Harper & Row, New York (1974)
[Hei4] Heisenberg, W.: Encounters with Einstein and Other Essays on People, Places, and Particles. Princeton University Press, Princeton (1989)
[HeiE] Heisenberg, E.: Inner Exile: Recollections of a Life with Werner Heisenberg. Birkhäuser, Boston (1984)
[Het1] Editorial, Die?? Neen, die niet! Het Parool No. 312, January 16, 1946, p. 3
[Het2] Editorial, Prof. Van der Waerden nog niet Benoemd. Het Parool No. 313, January 17, 1946, p. 1
[Het3] Editorial, Prof. dr. B. L. van der Waerden. Het Parool No. 320, January 25, 1946, p. 1
[Het4] Red. (Editors), Untitled commentary on Van der Waerden's letter. Het Parool No. 326, February 1, 1946, p. 3
[Het5] Editorial, Rondom Van der Waerden. Het Parool No. 336, February 13, 1946, p. 3
[Heu1] Heule, M.: Schur number five. arXiv:1711.08076 [cs.LO], (21 Nov 2017)
[Heu2] Heule, Marijn J. H.: Schur number five. Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18), 6598–6606 (2018)
[Heu3] Heule, M.J.H.: Computing small unit-distance graphs with Chromatic Number 5. Geombinatorics. **XXVIII**(1), 32–50 (2018)
[Heu4] Heule, M.J.H.: Private communication to A. Soifer, February 15, 2019
[Heu5] Heule, M.J.H.: Odd-distance virtual edges in unit-distance graphs. Geombinatorics. **XXXI**(2), 68–76 (2021)
[HHW] Heule, M.J.H., Hunt Jr., W.A., Wetzler, N.D.: Trimming while checking clausal proofs. In: Formal Methods in Computer-Aided Design, pp. 181–188. IEEE (2013)
[HKM] Heule, M.J.H., Kullmann, O., Marek, V.W.: Solving and Verifying the Boolean Pythagorean Triples problem via Cube-and-Conquer, arXiv:1605.00723 [cs.DM], (3 May 2016)
[Hil] Hilbert, D.: Ueber die Irreducibilität ganzer rationaler Functionen mit ganzzahligen Coefficienten. J. Reine Angew. Math. **110**, 104–129 (1892)
[HI] Hill, R., Irving, R.W.: On group partitions associated with lower bounds for symmetric Ramsey numbers. Eur. J. Comb. **3**(1), 35–50 (1982)
[HH] Hilton, A.J.W., Hind, H.R.: The total chromatic number of graphs having large maximum degree. Discret. Math. **17**(1–3), 127–140 (1993)
[Hin1] Hind, H.R.: An upper bound for the total chromatic number. Graphs Combin. **6**(2), 153–159 (1990)
[Hin2] Hind, H.R.: An upper bound for the total chromatic number of dense graphs. J. Graph Theory. **16**(3), 197–203 (1992)
[HMR] Hind, H., Molloy, M., and Reed, B., Total coloring with Δ + poly(logΔ) colors, SIAM J. Comput. 28(3), 1999, 816–821
[Hin] Hindman, N.: Finite sums from sequences within cells of a partition of N. J. Combin. Theory Ser. A. **17**, 1–11 (1974)
[HOx] Hobbs, A.M., Oxley, J.G., William, T.: Tutte, 1917–2002. Not. Am. Math. Soc. **51**, 320–330 (2004)

[HO] Hochberg, R., O'Donnell, P.: Some 4-chromatic unit-distance graphs without small cycles. Geombinatorics. **V**(4), 137–114 (1996)

[Hod] Hodge, Sir W.: Solomon Lefschetz 1884–1972, Biographical Memoirs of Fellows of the Royal Society **19**, 433–453 (December 1973)

[Hof] Hofer, H.: Letter to A. Soifer of March 3, 1995

[HS1] Hoffman, I., Soifer, A.: Almost chromatic number of the plane. Geombinatorics. **III**(2), 38–40 (1993)

[HS2] Hoffman, I., Soifer, A.: Another six–coloring of the plane. Discret. Math. **150**, 427–429 (1996)

[Hol] Holt, J.: Irreligion, a review of the book by Paulos, J. A., A Mathematician Explains Why the Arguments for God Just Don't Add Up; The New York Times Book Review, January 13, 2008, 18

[Hor] Horgan, J: The death of proof. Sci. Am. 92–103 (October 1993)

[Hort] Horton, J. D.: Sets with no Empty Convex 7-Gons. Canad. Math. Bull. **26**(4), 482–484 (1983)

[HUA1] Archive of Humboldt University at Berlin, documents UK–Sch 342, Bd. I, B1.1., 1R, 2R, 3

[HUA2] Archive of Humboldt University at Berlin, document UK–Sch 342, Bd. I, B1.25

[HZ1] Huang, Y.R., Zhang, K.M.: An new upper bound formula for two color classical Ramsey numbers. J. Combin. Math. Combin. Comput. **28**, 347–350 (1998)

[HWSYZH] (also abbreviated as HW+) Huang, Y.R., Wang, Y., Sheng, W., Yang, J., Zhang, K.M., Huang, J.: New upper bound formulas with parameters for Ramsey numbers, Discret. Math., 307 (2007) 760–763

[IMU1] Report of the 18th General Assembly of the International Mathematical Union (IMU) São Paulo, Brazil, July 28–29, 2018, final January 10, 2020. chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.mathunion.org/fileadmin/IMU/Organization/GA/GA_2018/18%20GA%20Report%20Final%20180729%20Sao%20Paulo.pdf

[IMU2] IMU Announcement of the Abacus Medal, http://www.union-matematica.org.ar/archivo/wp-content/uploads/2019/04/IMU_Abacus_Medal_v2.pdf

[Irv0] Irving, R.W.: An extension of Schur's theorem on sum-free partitions. Acta Arithmetica. **25**, 55–64 (1973)

[Irv1] Irving, R.W.: On a bound of Graham and Spencer for a graph–coloring constant. J. Combin. Theory Ser. B. **15**, 200–203 (1973)

[Irv2] Irving, R.W.: Generalised Ramsey numbers for small graphs. Discret. Math. **9**, 251–264 (1974)

[Isb1] Isbell, J.: Letter to A. Soifer of August 26, 1991

[Isb2] Isbell, J.: Letter to A. Soifer of September 3, 1991

[Iva] Ivanov, L.L.: Оценка хроматического числа пространства R^4. *УМН* **61**:5 (371), 181–182 (2006); Russ. Math. Surv. **61**(5), 984–986 (2006)

[Jec] Jech, T.J.: The Axiom of Choice. North–Holland, Amsterdam (1973)

[Jef] Jeffrey, R.C.: The Logic of Decision. McGraw–Hill, New York (1965)

[JefB] Jeffreys, B.: Letter to the editor. Math. Gazette. **63**, 126–127 (1979)

[JKSV] Jelínek, V., Kyncl, J., Stolar, R., Valla, T.: Monochromatic triangles in two-colored plane, *Combinatorica* (submitted March 2006) Vol. 29, Issue 6, November 2009, 699–718, https://doi.org/10.1007/s00493-009-2291-y arXiv:math.CO/0701940v1, Jan. 31, 2007

[JT] Jensen, T.R., Toft, B.: Graph Coloring Problems. Wiley, New York (1995)

[Jes] Jessop, J.: Some botanists and botanical collectors associated with Southern Africa. Lantern, 80–85 (1963)

[Joh1] Johnson Jr., P.D.: Coloring abelian groups. Discret. Math. **40**, 219–223 (1982)

[Joh2] Johnson Jr., P.D.: Simple product colorings. Discret. Math. **48**, 83–85 (1984)

[Joh3] Johnson Jr., P.D.: Two-coloring of real quadratic extensions of Q^2 that forbid many distances. Congr. Numer. **60**, 51–58 (1987)

[Joh4] Johnson Jr., P.D.: Two-colorings of a dense subgroup of Q^n that forbid many distances. Discret. Math. **79**, 191–195 (1989/90)

[Joh5] Johnson Jr., P.D.: Product colorings. Geombinatorics. **I**(3), 11–12 (1991)

[Joh6] Johnson Jr., P.D.: Maximal sets of translations forbidable by two-colorings of abelian groups. Congr. Numer. **91**, 153–158 (1992)

[Joh7] Johnson Jr., P.D.: About two definitions of the fractional chromatic number. Geombinatorics. **V(3)**, 11–12 (1996)

[Joh8] Johnson Jr., P.D.: Introduction to "Colorings of metric spaces" by Benda and Perles. Geombinatorics. **IX**(3), 110–112 (2000)

[Joh9] Johnson Jr., P.D.: Coloring the rational points to forbid the distance one – a tentative history and compendium. Geombinatorics. **XVI**(1), 209–218 (2006)

[Juh] Juhász, R.: Ramsey type theorems in the plane. J. Combin. Theory (A). **27**, 152–160 (1979)

[JRW] Jungreis, D.S., Reid, M., Witte, D.: Distances forbidden by some 2-coloring of Q^2. Discret. Math. **82**, 53–56 (1990)

[KaT] Kahle, M., Taha, B.: New lower bounds for $\chi\ (R^d)$ for d = 8, …, 12. Geombinatorics. **XXIV**(3), 109–116 (2015)

[KK] Kahn, J., Kalai, G.: A counterexample to Borsuk's conjecture. Bull. Am. Math. Soc. **29**(1), 60–62 (1993)

[Kai] Kainen, P.: Is the four-color theorem true? Geombinatorics. **III**(2), 41–56 (1993)

[Ka1] Kalbfleisch, J.G.: Construction of special edge–chromatic graphs. Can. Math. Bull. **8**, 575–584 (1965)

[Ka2] Kalbfleisch, J.G.: Chromatic graphs and Ramsey's theorem. Ph.D. thesis, University of Waterloo (January 1966)

[Ka3] Kalbfleisch, J.G.: On an unknown Ramsey number. Mich. Math. J. **13**, 385–392 (1966)

[KM] Kakeya, S., Morimoto, S., On a theorem of M. Bandet [sic] and van der Waerden: Jpn. J. Math. **7**, 163–165 (1930)

[Kan] Kandinsky, W.: Reminiscences, Berlin, 1913. In: Kandinsky: Complete Writings on Art. Da Capo Press, New York (1994)

[KS] Karabash, D., Soifer, A.: On Shelah-Soifer class of graphs. Geombinatorics. **XVI**(4), 363–366 (2007)

[Kat] Katznelson, Y.: Chromatic numbers of Cayley graphs on Z and recurrence. Combinatorica. **21**, 211–219 (2001)

[Kei] Keizer, M.d.: Het Parool 1940–1945: Verzetsblad in Oorlogstijd, 2nd edn. Otto Cramwinckel Uitgever, Amsterdam (1991)

[KK] Kelly, J.B., Kelly, L.M.: Paths and circuits in critical graphs. Am. J. Math. **76**, 786–792 (1954)

[Kem1] Kempe, A.B.: How to colour a map with four colours. Nature **20**, 275 (July 17, 1879)

[Kem2] Kempe, A.B.: On the geographical problem of the four colours. Am. J. Math. **II**, 193–200 (1879)

[Kem3] Kempe, A.B.: Untitled. Proc. London Math. Soc. **10**, 229–231 (1878–1879)

[Kem4] Kempe, A.B.: How to colour a map with four colours. Nature **21**, 399–400 (Feb. 26, 1880)

[Kem5] Kempe, A.B.: Untitled abstract. Proc. Lond. Math. Soc. **22**, 263–264 (1890–1891)

[Ker] Kersten, I.: Biography of Ernst Witt (1911–1991). Contemp. Math. **272**, 155–171 (2000)

[Kery] Kéry, G.: On a theorem of Ramsey. Matematikai Lapok. **15**, 204–224 (1964) (Hungarian)

[KS] Ketonen, J., Solovay, R.: Rapidly growing Ramsey functions. Ann. Math. **113**, 267–314 (1981)

[Key] Keynes, J.M.: Essays in Biography. W. W. Norton & Co., New York (1963)

[KSz] Khalfalah, A., Szemerédi, E.: On the number of monochromatic solutions of $x + y = z^2$. Comb. Probab. Comput. **15**(1–2), 213–227 (2006)

[Khi1] Khinchin, A.Y.: Tri Zhemchuzhiny Teorii Chisel. Gos. Izd-vo Tekhn.-teoret. lit-ry, Moskva (1947) (Russian)

[Khi2] Khinchin, A.Y.: Tri Zhemchuzhiny Teorii Chisel. Gos. Izd-vo Tekhn.-teoret, lit-ry, 2nd ed., Moskva (1948) (Russian)

[Khi3] Chintschin, A.J.: Drei Pearlen der Zahlentheorie. Akademie Verlag, Berlin (1951)

[Khi4] Khinchin, A.Y.: Three Pearls of Number Theory. Graylock, Rochester (1952)

[Kip] Kipling, R.: Just So Stories, 1st edn. Macmillan Publishers Ltd., London (1902)

[KW] Klee, V., Wagon, S.: Old and New Unsolved Problems in Plane Geometry and Number Theory, 2nd edn. Mathematical Association of America (1991) (with Addendum)

[Kle] Kleinberg, E.M.: The independence of Ramsey's theory. J. Symb. Log. **34**, 205–206 (1969)

[KP] Kleitman, D.J., Pachter, L.: Finding convex sets among points in the plane. Discret. Comput. Geom. **19**, 405–410 (1998)

[Kne1] Knegtmans, P.J.: Socialisme en Democratie: De SDAP tussen klasse en natie (1929–1939), Cip–Gegevens Koninklijke Bibliotheek, Den Haag (1989)

[Kne2] Knegtmans, P.J.: Een Kweetsbaar Centrum van de Geest: De Universiteit van Amsterdam tussen 1935 en 1950. Amsterdam University Press (1998)
[Kne3] Knegtmans, P.J.: e-mail to A. Soifer, May 25, 2004
[Kne4] Knegtmans, P.J.: e-mail to A. Soifer, May 26, 2004
[Kne5] Knegtmans, P.J.: e-mail to A. Soifer, May 28, 2004
[Kne6] Knegtmans, P.J.: e-mail to A. Soifer, June 7, 2004
[Kne7] Knegtmans, P.J.: e-mail to A. Soifer, June 10, 2004
[Kne8] Knegtmans, P.J.: e-mail to A. Soifer, October 4, 2004
[Knu] Knuth, D.E.: The Art of Computer Programming, vol. 4, fascicle 6: Satisfiability (2015)
[KoP] Kouril, M., Paul, J.L.: The Van der Waerden number W(2,6) is 1132. Exp. Math. **17**(1), 53–61 (2008). https://doi.org/10.1080/10586458.2008.10129025
[Kri1] Křiž, I.: Permutation groups in Euclidean Ramsey Theory. Proc. Am. Math. Soc. **385**, 899–908 (1991)
[Kri2] Křiž, I.: All trapezoids are Ramsey. Discret. Math. **108**, 59–62 (1992)
[KMi] Kronk, H.V., Mitchem, J.: The entire chromatic number of a normal graph is at most seven. Bull. Am. Math. Soc. **78**(5), 799–800 (September 1972)
[KT] Kuhn, H.W., Tucker, A.W.: John von Neumann's work in the theory of games and mathematical economics. Bull. Am. Math. Soc. **64**, 100–122 (1958)
[Kun] Kundera, M.: The Curtain: An Essay in Seven Parts, 1st edn. HarperCollins, New York (2006)
[Lam] Lamb, E.: Maths proof smashes size record: supercomputer produces a 200-terabyte proof – but is it really mathematics? Nature **534**, 17–18 (June 2016)
[LR] Landman, B.M., Robertson, A.: Ramsey Theory on the Integers. American Mathematical Society, Providence (2004)
[Lan1] Lang, S.: Algebra. Addison-Wesley, Reading (1965)
[Lan2] Lang, S.: Letter to the editor. Mitteilungen DMV **1**, 5 (2002). Republished in Notices Am. Math. Soc., June–July, 2005, p. 620
[LRX] Lange, A., Radziszowski, S., Xu, X.: Use of MAX-CUT for Ramsey arrowing of triangles. J. Combin. Math. Combin. Comput. **88**, 61–71 (2014)
[Lar] Larman, D.G.: A note on the realization of distances within sets in Euclidean space. Commentarii Math. Helvetici. **53**, 529–535 (1978)
[LR] Larman, D.G., Rogers, C.A.: The realization of distances within sets in Euclidean space. Mathematika. **19**, 1–24 (1972)
[Lau] von Laue, M.: Bull. At. Sci. **4**(4), 103 (1948)
[Led1] Ledermann, W.: Issai Schur and his school in Berlin. Bull. London Math. Soc. **15**, 97–106 (1983)
[Led2] Ledermann, W.: Ledermann's St Andrews interview, conducted by O'Connor, J. J., and Robertson, E. F., September 2000. http://wwwgap.dcs.stand.ac.uk/~history/HistTopics/Ledermann_interview.html#s69

[LN] Ledermann, W., Neumann, P.: The life of Issai Schur through letters and other documents. In: Joseph, et al. (eds.) Memory of Issai Schur, pp. xiv–xci. Birkhäuser, Boston (2003)

[Leh1] Lehto, O.: On the life and work of Lars Ahlfors. Math. Intell. **20**(4), 4–8 (1999)

[Leh2] Lehto, O.: Nevanlinna, Rolf (1895–1980), http://www.kansallisbiografia.fi/english/?id¼47111

[LS] Lewin Sime, R.: Lise Meitner: A Life in Physics. University California Press, Berkeley (1997)

[Lin] Lin, S.: On Ramsey numbers and K_r-colorings of graphs. J. Combin. Theory Ser. B. **12**, 82–92 (1972)

[Lit] Litten, F.: Die Carathéodory–Nachfolge in München 1938–1944. Centaurus. **37**, 154–172 (1994)

[Lov] Lovász, L.: Three short proofs in graph theory. J. Combin. Theory Ser. B. **19**, 269–271 (1975)

[Lov1] Lovász, L.: On chromatic number of finite set-systems. Acta Math. **19**, 59–67 (1968)

[Lov2] Lovász, L.: Combinatorial Problems and Exercises. North Holland, Amsterdam (1979)

[LSS] Lovász, L., Saks, M., Schrijver, A.: Orthogonal representations and connectivity of graphs. Linear Algebra Appl. **114/115**, 439–454 (1989)

[LV] Lovász, L., Vesztergombi, K.: Geometric representations of graphs. In: Halás, G., et al. (eds.) Paul Erdős and his Mathematics II, pp. 471–498. Springer (2002)

[Lu] Lu, L.: Explicit construction of small Folkman graphs. SIAM J. Discrete Math. **21**, 1053–1060 (2008)

[Luk] Lukomskaya, M.A.: New proof of Van der Waerden's theorem about arithmetic progression and some generalization of this theorem. Uspehi Matematicheskikh nauk. **3**(6(28)), 201–204 (1948)

[LCW] Luo, X., Chen, M., Wang, W.: On 3-colorable planar graphs without cycles of four lengths. Inf. Process. Lett. **103**, 150–156 (2007)

[Mac] Mackey, J.: Combinatorial remedies. Ph.D. thesis, Department of Mathematics, University of Hawaii (1994)

[MV] Mattheus, S., Verstraete, J.: The asymptotics of $r(4,t)$. arXiv:2306.04007v3

[Man1] Mann, M.: A new bound for the chromatic number of the rational five-space. Geombinatorics. **XI**(2), 49–53 (2001)

[Man2] Mann, M.: Hunting unit-distance graphs in rational n-spaces. Geombinatorics. **XIII**(2), 86–97 (2003)

[May] May, K.: The origin of the four-color conjecture. Isis **56**(3), 346–348 (Fall 1965)

[MR4] McKay, B.D., Radziszowski, S.P.: $R(4,\ 5) = 25$. J. Graph Theory. **19**, 309–322 (1995)

[MR5] McKay, B.D., Radziszowski, S.P.: Subgraph counting identities and Ramsey numbers. J. Combin. Theory Ser B. **69**, 193–209 (1997)

[MZ] McKay, B.D., Zhang, K.M.: The value of the Ramsey number *R (3,8)*. J. Graph Theory. **16**, 99–105 (1992)

[Meh] Mehrtens, H., Irresponsible purity: the political and moral structure of mathematical sciences in the National Socialist state, Science, Technology and National Socialism, by Renneberg, M. and Walker, M. (ed's), Cambridge University Press, 1994., 324–338 and 411–413

[Mel] Mellor, D.H.: Cambridge philosophers I: F. P. Ramsey. Philosophy. **70**, 243–262 (1995)

[Me] Mel'nikov, L.S.: Problem 9. In: Fiedler, M. (ed.) Recent Advances in Graph Theory. Proceedings of the International Symposium, Prague, 1974, p. 543. Academia, Praha (1975)

[Mil] Milosz, C.: The Captive Mind. Octagon, New York (1981)

[Min] Minnaert, M.G.J.: Light and Color in the Outdoors. Springer, New York (1993)

[Mir] Mirsky, L.: The combinatorics of arbitrary partitions. Bull. Inst. Math. **11**, 6–9 (1975)

[MMRS] Montejano, A., Montejano, L., Roldán-Pensado, E., Pablo, S.: About an Erdős–Grünbaum conjecture concerning piercing of non-bounded convex sets. http://arxiv.org/abs/1407.0642

[Moo] Moore, G.H.: Zermelo's Axiom of Choice: Its Origins, Development, and Influence. Springer, New York (1982)

[Mor] Mordell, L.J.: On the rational solutions of the indeterminate equations of the third and fourth degrees. Proc. Camb. Philos. Soc. **21**, 179–182 (1922)

[Mor1] Morrison, P.: Alsos: the story of German science. Bull. At. Sci. **3**(12) (1947) pp. 354, 365

[Mor2] Morrison, P.: Bull. At. Sci. **4**(4), 104 (1948)

[MM] Moser, L., Moser, W.: Solution to problem 10. Can. Math. Bull. **4**, 187–189 (1961)

[MP] Moser, W., Pach, J.: Research Problems in Discrete Geometry. McGill University (1986)

[Mül] Müller, T.: A counterexample to a conjecture of Grünbaum on piercing convex sets in the plane. Discret. Math. **313**(24), 2868–2871 (2013)

[Myc] Mycielski, J.: Sur le coloriage des graphes. Colloq. Math. **3**, 161–162 (1955)

[Nec] Nechushtan, O.: On the space chromatic number. Discret. Math. **256**, 499–507 (2002)

[Nel1] Nelson, E.: Letter to A. Soifer of August 23, 1991

[Nel2] Nelson, E.: Letter to A. Soifer of October 5, 1991

[NW] Nelson, R., Wilson, R.J. (eds.): Graph Colourings. Longman Scientific and Technical, Essex (1990)

[Nen1] Nenov, N.D.: A new lower bound for the Graham–Spencer number (Russian). Serdica. **6**, 373–383 (1980)

[Nen2] Nenov, N.D.: An example of a 15-vertex (3,3)-Ramsey graph with clique number 4 (Russian). C. R. Acad. Bulg. Sci. **34**, 1487–1489 (1981)

[NR] Nešetřil, J., Rödl, V.: The Ramsey property for graphs with forbidden complete subgraphs. J. Combin. Theory Ser. B. **20**, 243–249 (1976)

[Neu] von Neumann, J.: Über die Definition durch transfinite Induktion und verwandte Fragen der allgemeinen Mengenlehre. Math. Annalen. **99**, 373–391 (1928)

[Nil] Nilli, A.: On Borsuk's problem. Contemp. Math. **178**, 209–210 (1994)

[NM] von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press, Princeton (1944)

[Odo1] O'Donnell, P.: A triangle-free 4-chromatic graph in the plane. Geombinatorics. **IV(1)**, 23–29 (1994)

[Odo2] O'Donnell, P.: A 40 vertex 4-chromatic triangle-free unit-distance graph in the plane. Geombinatorics. **V**(1), 30–34 (1995)

[Odo3] O'Donnell, P.: High Girth unit-distance graphs. Ph. D. dissertation, Rutgers University, May 25, 1999

[Odo4] O'Donnell, P.: Arbitrary girth, 4-chromatic unit-distance graphs in the plane part I: graph description. Geombinatorics. **IX**(3), 145–150 (2000)

[Odo5] O'Donnell, P.: Arbitrary girth, 4-chromatic unit-distance graphs in the plane part II: graph embedding. Geombinatorics. **IX**(4), 180–193 (2000)

[Oro] O'Rourke, J.: Computational geometry column 46. Int. J. Comput. Geom. Appl. **14**(6), 475–478 (2004)

[OSe] Ó Seanór, S.: Letter to A. Soifer from March 21, 1997

[OED] Oxford English Dictionary, May 24, 2023. https://www.oed.com/view/Entry/73969?redirectedFrom=four-colour+problem#eid3604483

[OTH] Owings, J., Tetiva, M., Huddleston, M.: Coloring the plane. Am. Math. Mon. **115**(2), 170–172 (2008)

[PH] Paris, J., Harrington, L.: Mathematical incompleteness in Peano arithmetic. In: Barwise, J. (ed.) Handbook of Mathematical Logic, pp. 1133–1142. North-Holland, Amsterdam (1977)

[Pash] Pash, B.T.: The Alsos Mission. Award House, New York (1969)

[Par1] Parts, J.: A small 6-chromatic two-distance graph in the plane. Geombinatorics. **XXIX**(3), 111–115 (2020)

[Par2] Parts, J.: Graph minimization, focusing on the example of 5-chromatic unit-distance graphs in the plane. Geombinatorics. **XXIX**(4), 137–166 (2020)

[Par3] Parts, J.: What percent of the plane can be properly 5- and 6-colored? Geombinatorics. **XXX**(1), 25–39 (2020)

[Par4] Parts, J.: A 6-chromatic odd-distance graph in the plane. Geombinatorics. **XXXI**(3), 124–137 (2022)

[Par5] Parts, J.: On the plane and its coloring. Geombinatorics. **XXXI**(4), 189–195 (2022)

[Pas] Pasternak, B.L.: Волны (Waves), dedicated "To N.I. Bukharin," 1931, in Vtoroe Rozhdenie (Second Birth), Sovetskij Pisatel', Moscow, 1934, 7–21. See it also in Complete Works in 11 Volumes, vol. 2, 50–59, but without the dedication to Nikolai I. Bukharin, Slovo, Moscow, 2004 (Russian)

[Pay] Payne, M.S.: A Unit-distance graph with ambiguous chromatic number. arXiv: 0707.1177v1[math.CO] 9 July 2007

[Per] Peremans, W.: Van der Waerden Day in Groningen. Nieuw Archief voor Wiskunde, Vierde serie. **12**(3), 135–136 (1994)

[PS] Peres, Y., Schlag, W.: Two Erdős problems on lacunary sequences: chromatic number and Diophantine approximation. Bull. Lond. Math. Soc. **42**(2), 295–300 (2010)
[Pet] Peter, L.J.: Peter's Quotations: Ideas for Our Time. William Morrow, New York (1977)
[Pete] Peters, P.: Minjnheer de Red. Propria Cures, 56th year, No. 22, February 8, 1946
[Pic] Picasso, P., Ashton, D. (eds.): Picasso on Art: A Selection of Views. Da Capo, New York (1992)
[PRU] Piwakowski, K., Radziszowski, S.P., Urbánski, S.: Computation of the Folkman number Fe(3, 3; 5). J. Graph Theory. **32**, 41–49 (1999)
[Poi] Poincaré, H.: The Foundations of Science. The Science Press, Lancaster (1946)
[Poisk] Materialy Konferenzii "Poisk–97", Moscow, 1997, (Russian)
[Pol] Pólya, G., Alexanderson, G.L.: The Pólya Picture Album: Encounters of a Mathematician. Birkhäuser, Boston (1987)
[Poly] Polymath, D.H.J.: On the chromatic number of circular disks and infinite strips in the plane. Geombinatorics. **XXX**(4), 190–201 (2020)
[Pon] Pontryagin, L.S.: Life of Lev Semenovich Pontryagin, mathematician, compiled by himself. Born 1908, Moscow, Moscow, 1998, (Russian)
[Pow] Powers, T.: Heisenberg's War. Little, Brown and Co. (1993)
[Pre] Prévert, J.: Paroles. Le Calligraphe, Paris (1945) [copyright Éditions du Point du Jour, 1946], (French)
[Pri] Pritikin, D.: All unit-distance graphs of order 6197 are 6-colorable. J. Combin. Theory Ser. B. **73**(2), 159–163 (1998)
[Pr1] Prömel, H.J.: Complete disorder is impossible: the mathematical work of Walter Deubner. Comb. Probab. Comput. **14**, 3–16 (2005)
[Pr2] Prömel, H.J.: Ramsey Theory for Discrete Structures. Springer, Cham (2013)
[PR] Prömel, H.J., Rödl, V.: An elementary proof of the canonizing version of Gallai–Witt's theorem. J. Combin. Theory Ser. A. **42**, 144–149 (1983)
[PR] Prömel, H.J., Voigt, B.: Aspects of Ramsey–Theory, chapters I, II, III, and V, Universität Bonn, unpublished report with chapter IV not supplied by the authors (1989)
[PP] Purdy, C., Purdy, G.: Minimal forbidden distance one graphs, proceedings of the nineteenth southeastern conference on combinatorics, graph Theory, and computing, Baton Rouge, LA, 1988. Congr. Numer. **66**, 165–171 (1988)
[Rad1] Rado, R.: Studien zur Kombinatorik. Math. Zeitschrift. **36**, 424–480 (1933)
[Rad2] Rado, R.: Note on combinatorial analysis. Proc. London Math. Soc. **48**(2), 122–160 (1943)
[RT] Radoicic, R., Tóth, G.: Note on the chromatic number of the space. In: Aronov, B., Basu, S., Pach, J., Sharir M. (eds.) Discrete and Computational Geometry: the Goodman–Pollack Festschrift, pp. 695–698, Springer, Berlin (2003). http://www.ams.org/mathscinet/search/mscdoc.html?code=05D10,%2852A37%29

[Radz1] Radziszowski, S.P.: Small Ramsey numbers. Electron. J. Combin., Dynamic Survey, DS1, revision #16, January 15, 2021, 116 pages, http://www.combinatorics.org/; see it also here: https://www.combinatorics.org/ojs/index.php/eljc/article/view/DS1/versions

[Radz2] Radziszowski, S.P.: On the Most wanted Folkman graph. Geombinatorics. **XVI**(4), 367–381 (2007)

[RK2] Radziszowski, S.P., Kreher, D.L.: Upper bounds for some Ramsey numbers $R(3, k)$. J. Combin. Math. Combin. Comput. **4**, 207–212 (1988)

[Raig1] Raigorodskii, A.M.: On a bound in Borsuk's problem. Russ. Math. Surv. **54**(2), 453–454 (1999)

[Raig2] Raigorodskii, A.M.: On the chromatic number of a space. Russ. Math. Surv. **55**(2), 351–352 (2000)

[Raig3] Raigorodskii, A.M.: Borsuk's problem and the chromatic numbers of some metric spaces. Russ. Math. Surv. **56**(1), 103–139 (2001)

[Raig4] Raigorodskii, A.M.: The Borsuk and Hadwiger problems and systems of vectors with restrictions on scalar products. Russ. Math. Surv. **57**(3), 606–607 (2002)

[Raig5] Raigorodskii, A.M.: The Borsuk problem for integral polytopes. Sb. Math. **193**(10), 1535–1556 (2002)

[Raig6] Raigorodskii, A.M.: Chromatic Numbers (Russian). Moscow Center of Continuous Mathematical Education, Moscow (2003)

[Rai] Raiskii, D.E.: Realizing of all distances in a decomposition of the space R^n into $n+1$ Parts. Mat. Zametki **7**, 319–323 (1970) [Russian]; Engl. transl., Math. Notes **7**, 194–196 (1970)

[Ram1] Ramsey, F.P.: Mathematical logic. Math. Gazette. (1926)

[Ram2] Ramsey, F.P.: On a problem of formal logic. Proc. London Math. Soc. Ser 2. **30**(part 4), 338–384 (1930)

[Ram3] Ramsey, F.P., Braithwaite, R.B. (ed. and Introduction), Moore, G.E. (Preface): The Foundations of Mathematics, Kegan Paul, Trench, Trubner & Co., London 1931

[Ram4] Ramsey, F.P.: Truth and probability, 1926. In: Ramsey, F.P., Mellor, D.H. (eds.) Philosophical Papers. Cambridge University Press, Cambridge (1990)

[Ram5] Ramsey, F.P., Galavotti, M.C. (eds.): Notes on Philosophy and Mathematics. Bibliopolis, Napoli (1991)

[Rea] Read, H.: Gabo, Naum Gabo: March 2–April 1968, Albright–Knox Art Gallery. Buffalo, New York, p. 5 (1968)

[Reid] Reid, C.: Hilbert. Allen & Unwin/Springer-Verlag, London/Berlin (1970)

[Rei] Reid, M.A.: Undergraduate Algebraic Geometry. Cambridge University Press (1988)

[Rit] Ritchie, W.: The History of the South African College, vol. 1. T. Maskew Miller, Cape Town (1918)

[Rob1] Robertson, N.: E-mail to A. Soifer, May 9, 1994

[RSST] Robertson, N., Sanders, D.P., Seymour, P.D., Thomas, R.: The four-colour theorem. J. Combin. Theory Ser. B. **70**(1), 2–44 (1997)

[RST] Robertson, N., Seymour, P.D., Thomas, R.: Hadwiger's conjecture for K_6-free graphs. Combinatorica. **13**(3), 279–361 (1993)
[Ros1] Rosenfeld, M.: Odd integral distances among points in the plane. Geombinatorics. **V**(4), 156–159 (1996)
[Ros2] Rosenfeld, M.: The odd-distance plane graph. Geombinatorics. **XIX**(2), 62–66 (2009)
[Rot] Roth, K.F.: On certain sets of integers. J. Lond. Math. Soc. **28**, 104–109 (1953)
[Saa1] Saaty, T.L.: Remarks on the four color problem; the Kempe catastrophe. Math. Magazine. **40**(1), 31–36 (1967)
[Saa2] Saaty, T.L.: Thirteen colorful variations on Guthrie's four-color conjecture. Am. Math. Mon. **79**(1), 2–43 (1972)
[SK] Saaty, T. L., Kainen, P. C.: The Four-Color Problem: Assaults and Conquest. McGraw-Hill, New York (1977)
[San1] Sanders, J.H.: A generalization of Schur's theorem. Ph.D. dissertation, Yale University (1968)
[San2] Sanders, J.H.: A generalization of Schur's theorem. arXiv:1712.03620v1, Dec. 11, 2017
[SZ1] Sanders, D.P., Zhao, Y.: A note on the three color problem. Graphs Combin. **11**, 91–94 (1995)
[SZ2] Sanders, D.P., Zhao, Y.: On total 9-colouring planar graphs of maximum degree seven. J. Graph Theory. **31**, 67–73 (1999)
[SZ3] Sanders, D.P., Zhao, Y.: On simultaneous edge-face colorings of plane graphs. Combinatorica. **17**, 441–445 (1997)
[Sav] Savage, L.J.: The Foundation of Statistics. Wiley, New York (1954)
[Schi] Schiffer, M. M., Issai Schur: some personal reminiscences, in Mathematik in Berlin: Geschichte und Documentation, by Begehr, H. (ed.), v. 2 (1998), 177–181
[Scho] Schoenberg, A.: Preface. In: Anton Webern, 6 Bagatellen für Streichquartett, op. 9, Partitur, Universal Edition
[Schuh] Schuh, F.: In memoriam (Pierre Joseph Henry Baudet). Christian Huygens. **I**, 145 (1921–1922)
[Sch] Schur, I.: Über die Kongruenz $x^m + y^m \equiv z^m$ (mod. p). Jahresbericht der Deutschen Mathematiker–Vereinigung. **25**, 114–117 (1916)
[I] Irving, R.W.: An extension of Schur's theorem on sum-free partitions. Acta Arithmetica. **25**, 55–64 (1973)
[Seg] Segal, S.L.: Mathematicians under the Nazis. Princeton University Press, Princeton (2003)
[Sei] Seidenberg, A.: A simple proof of Erdős and Szekeres. J. Lond. Math. Soc. **34**, 352 (1959)
[Sha] Shader, L.E.: All right triangles are Ramsey in E^2! J. Combin. Theory Ser. A. **20**, 385–389 (1976)
[She1] Shelah, S.: Primitive recursive bounds for van der Waerden numbers. J. Am. Math. Soc. **1**, 683–697 (1988)
[She2] Shelah, S.: Cardinal Arithmetic. Clarendon Press, Oxford (1993)

[She3] Shelah, S.: Logical dreams. Bull. Am. Math. Soc. **40**, 203–228 (2003)

[She4] Shelah, S.: e-mail to A. Soifer, June 12, 2006

[She5] Shelah, S.: e-mail to A. Soifer, June 21, 2006

[SS1] Shelah, S., Soifer, A.: Axiom of choice and chromatic number of the plane. J. Combin. Theory Ser. A. **103**, 387–391 (2003)

[SS2] Shelah, S., Soifer, A.: Chromatic number of the plane & its relatives part III: its future. Geombinatorics. **XIII**(1), 41–46 (2003)

[Sim] Simmons, G.J.: The chromatic number of the sphere. J. Austral. Math. Soc. Ser. A. **21**, 473–480 (1976)

[Sko] Skolem, T.: Ein Kombinatorische Satz mit Anwendung auf ein Logisches Entscheidungsproblem. Fundamenta Math. **20**, 254–261 (1933)

[Soa] Soare, R.I.: Recursively Enumerable Sets and Degrees. Springer (1987)

[SS3] Soifer, A., Shelah, S.: Axiom of choice and chromatic number: an example in the plane. J. Combin. Theory Ser. A. **105**(2), 359–364 (2004)

[Soi0] Soifer, A.: Mathematics as Problem Solving, Center for Excellence in Mathematical Education, Colorado Springs, 1987, 2nd edn. Springer, New York (2008)

[Soi1] Soifer, A.: How Does One Cut a Triangle? Center for Excellence in Mathematical Education, Colorado Springs, 1990. Second expanded edition. Springer, New York (2009)

[Soi2] Soifer, A.: Triangles in a two-colored plane. Geombinatorics **I**(1), 6–7 and **I**(2) 13–14 (1991)

[Soi3] Soifer, A.:Triangles in a three-colored plane. Geombinatorics **I**(2), 11–12 (1991) and **I**(4), 21 (1992)

[Soi4] Soifer, A.: Chromatic number of the plane: a historical essay. Geombinatorics. **I**(3), 13–15 (1991)

[Soi5] Soifer, A.: Relatives of chromatic number of the plane I. Geombinatorics. **I**(4), 13–15 (1992)

[Soi6] Soifer, A.: A six–coloring of the plane. J. Comb. Theory Ser. A. **61**(2), 292–294 (1992)

[Soi7] Soifer, A.: Six–realizable set X_6. Geombinatorics. **III**(4) (1994)

[Soi8] Soifer, A.: An infinite class of 6-colorings of the plane. Congr. Numer. **101**, 83–86 (1994)

[Soi9] Soifer, A.: The Colorado Mathematical Olympiad and Further Explorations: From the Mountains of Colorado to the Peaks of Mathematics. Springer, New York (2011)

[Soi10] Soifer, A.: Issai Schur: Ramsey theory before Ramsey. Geombinatorics. **V (1)**, 6–23 (1995)

[Soi11] Soifer, A.: Pierre Joseph Henry Baudet: Ramsey theory before Ramsey. Geombinatorics. **VI**(2), 60–70 (1996)

[Soi12] Soifer, A.: The Baudet–Schur conjecture on monochromatic arithmetic progressions: an historical investigation. Congr. Numer. **117**, 207–216 (1996)

[Soi13] Soifer, A.: On the border of two years: an interview with Paul Erdős. Geombinatorics. **VI(3)**, 99–110 (1997)

[Soi14] Soifer, A.: And you Don't even have to believe in G'd, but you have to believe that the book exists. Geombinatorics. **VI**(3), 111–116 (1997)

[Soi15] Soifer, A.: Map coloring in Victorian age: problems, history, results. Math. Comp. **10**(1), 20–31 (1997)

[Soi16] Soifer, A.: Competitions, mathematics, life. Math. Comp. **11**(2), 20–41 (1998)

[Soi17] Soifer, A.: 50th anniversary of one problem: chromatic number of the plane. Math. Comp. **16**(1), 9–41 (2003)

[Soi18] Soifer, A.: Chromatic number of the plane & its relatives part I: the problem & its history. Geombinatorics. **XII**(3), 131–148 (2003)

[Soi19] Soifer, A.: Chromatic number of the plane & its relatives part ii: polychromatic number and 6–coloring. Geombinatorics. **XII**(4), 191–216 (2003)

[Soi20] Soifer, A.: In search of Van der Waerden. Leipzig and Amsterdam, 1931–1951. Part I: Leipzig. Geombinatorics. **XIV**(1), 21–40 (2004)

[Soi21] Soifer, A.: In search of Van der Waerden. Leipzig and Amsterdam, 1931–1951. Part II: Amsterdam, 1945. Geombinatorics. **XIV**(2), 72–102 (2004)

[Soi22] Soifer, A.: A journey from Ramsey theory to mathematical Olympiad to finite projective planes. Math. Comp. **17**(2), 8–16 (2004)

[Soi23] Soifer, A.: Axiom of choice and chromatic number of R^n. J. Combin. Theory Ser. A. **110**(1), 169–173 (2005)

[Soi24] Soifer, A.: In search of Van der Waerden. Leipzig and Amsterdam, 1931–1951. Part III: Amsterdam, 1946–1951. Geombinatorics. **XIV**(3), 124–161 (2005)

[Soi25] Soifer, A.: On Stephen P. townshend's 1976 Proof. Geombinatorics. **XVI**(4), 181–183 (2005)

[Soi26] Soifer, A.: In search of Van der Waerden: the early years. Geombinatorics. **XIV**(3), 305–342 (2007)

[Soi27] Soifer, A.: Chromatic number of the plane theorem in Solovay's set Theory. Geombinatorics. **XVII**(2), 85–87 (2007)

[Soi28] Soifer, A.: Editorial dedicated to Paul Erdős' 80th birthday. Geombinatorics. **II**(3), 43–45 (1993)

[Soi29] Soifer, A.: On the border of two years: an interview with Paul Erdős, December 24, 1991–January 9, 1992. Geombinatorics. **VI**(3), 99–110 (1997)

[Soi30] Soifer, A.: And you don't even have to believe in G'd but you have to believe that the book exists. Geombinatorics. **VI**(3), 111–116 (1997)

[Soi31] Soifer, A.: To leave, to die, or to compromise? A review of Constantin Carathéodory: mathematics and politics in turbulent time by Georgiadou, M. Geombinatorics. **XIV**(1), 41–46 (2004)

[Soi32] Soifer, A.: Map coloring in the Victorian age: problems and history. Math. Comp. **10**, 20–31 (1997)

[Soi33] Soifer, A., Ramsey Theory before Ramsey, prehistory and early history: an essay in 13 Parts, in Ramsey Theory Yesterday, Today and Tomorrow, by A. Soifer (ed.), "Progress in Mathematics" series, Birkhäuser/Springer, New York, 2010, 1–26

[Soi34] Soifer, A.: Book Review of Emil Artin and Helmut Hasse. The correspondence 1923–1958, by Frei, G., Lemmermeyer, F., Roquette, P.J. (eds.). Springer, Basel (2014)

[Soi35] Soifer, A., Geometric graphs. In: Beineke, L.W., Wilson, R.J. (eds.), and Toft, B. (acad. Consult) Topics in Chromatic Graph Theory. Cambridge University Press, Cambridge, pp. 161–180 (2015)

[Soi37] Soifer, A.: Letter of the Executive Committee of the World Federation of National Mathematics Competitions to the Executive Committee of the International Mathematics Union

[Soi38] Soifer, A.: Letter to the Executive Committee of the International Mathematics Union

[Soi39] Soifer, A.: Van der Waerden: Thoughts about Life and Fate. Moscow Center of Continuing Mathematical Education, Moscow (2008) (Russian)

[Soi40] Soifer, A.: Between the line and the plane: chromatic étude in 6 movements. Math. Comp. **23**(2), 30–45 (2010)

[Soi41] Soifer, A.: An open letter to all nine Rolf Nevanlinna's prize laureates, Geombinatorics XXVII(2), 84–90 (October 2017)

[Soi42] Soifer, A.: The international Mathematics union abacus medal. Geombinatorics. **XXIX**(1), 33–36 (2019)

[Soi43] Soifer, A.: A taste for duality: Erdős, Dilworth, Gallai, Milgram, and Mirsky. Geombinatorics. **XXVI**(1), 30–41

[Soi44] Soifer, A.: The Mathematical Coloring Book: Mathematics of Coloring and the Colorful Life of its Creators, 1st edn. Springer, New York (2009)

[Soi45] Soifer, A.: Progress in my favorite open problem of mathematics, chromatic number of the plane: an étude in five movements. Geombinatorics. **XXVIII**(1), 5–17

[Soi46] Soifer, A.: Chromatic problems and prizes. Geombinatorics. **XXVIII**(4), 206–210 (2019)

[Soi47] Soifer, A.: The Scholar and the State: In Search of Van der Waerden. Springer Birkhäuser, Basel (2015)

[Soi48] Soifer, A.: Remembering Ronald Lewis Graham. Geombinatorics. **XXX**(2), 56–62 (2020)

[Soi49] Soifer, A. (ed.): Ramsey Theory: Yesterday, Today, and Tomorrow Progress in Mathematics 285. Springer Birkhäuser, New York (2011)

[Soi50] Soifer, A.: Hadwiger–Nelson problem. In: Nash Jr., J.F., Rassias, M.T. (eds.) Open Problems in Mathematics, pp. 439–457. Springer, Cham (2016)

[Soi52] Soifer, A.: Truth, like water, will find its way out: my response to Günter Ziegler and the *Deutsche Mathematiker-Vereinigung*. Geombinatorics. **XXV**(2), 76–86 (2015) https://geombina.uccs.edu/truth-water-2

[Soi53] Soifer, A.: Taking sides: *geombinatorics*' response to the war and boycotts. Geombinatorics. **XXXI**(4), 148–155 (2022)

[Soi54] Soifer, A.: There is no censorship in the Western World, or is there? An essay in six movements. Geombinatorics. **XXXII**(2), 87–95 (2022)

[Soi55] Soifer, A.: The Colorado Mathematical Olympiad, The Third Decade and Further Explorations: From the Mountains of Colorado to the Peaks of Mathematics. Springer, New York (2017)

[Soi56] Soifer, A.: Chromatic Number of The Plane and Its Relatives: History, Conjectures, Results And Credits. Invited Public Lecture of The Workshop On Euclidean Ramsey Theory dedicated to the memory of Ron Graham, Hungarian Academy of Sciences, Alfréd Rényi Institute of Mathematics, Budapest, Hungary, 2021, March 10–12. https://video.renyi.hu/video/chromatic-number-of-the-plane-and-its-relatives-history-conjectures-results-and-credits-public-lecture-215

[Soi57] Soifer, A.: Executive Committee of the International Mathematics Union; it takes thunder to end cover-up: an essay in 6 movements. Geombinatorics. **XXXIII**(1), 39–54 (2023)

[Soi91/7/6ltr] Soifer, A., 3-page handwritten letter to Paul Erdős

[SL] Soifer, A., Lozansky, E.: Pigeons in every pigeonhole. Quantum, 24–28 and 32 (January 1990)

[Sol1] Solovay, R.M.: A model of set theory in which every set of reals is Lebesgue measurable. Ann. Math. 92 Ser. **2**, 1–56 (1970)

[Sol2] Solovay, R.M.: e-mail to A. Soifer, April 11, 2003

Sol3] Solovay, R.. e-mail to A. Soifer, June 12, 2006

[Sp1] Spencer, J.H.: Introductory Commentary. In: Erdős, P. (ed.) The Art of Counting, p. xxiii. MIT Press, Cambridge, MA (1973)

[Sp2] Spencer, J.H.: Ramsey theory and Ramsey theoreticians. J. Graph Theory. **7**, 15–23 (1983)

[Sp3] Spencer, J.H.: Three hundred million points suffice. J. Combin. Theory Ser. A. **49**, 210–217 (1988)

[Sp4] Spencer, J.H.: Erratum: "Three hundred million points suffice". J. Combin. Theory Ser. **A50**(2), 323 (1989)

[Sp5] Spencer, J.H.: Ramsey's theorem – a new lower bound. J. Combin. Theory, Ser. A. **18**, 108–115 (1975)

[Spe3] Spencer, T.: University of Nebraska at Omaha, personal communication with S. P. Radziszowski (1993), and, Upper Bounds for Ramsey Numbers via Linear Programming, manuscript, (1994)

[Ste] Steinberg, R.: The state of the three color problem, in Gimbel, J., Kennedy, J. W., and Quintas, L. V. (eds.), Quo Vadis, Graph Theory? Annals of Discrete Math. **55**, 211–248 (1993)

[Stein] Steinhardt, J.: On coloring the odd-distance graph. Electron. J. Combin. **16**(12), 1–7 (2009)

[Stei] Steinhaus, H.: Sur les distances des points des ensembles de measure positive. Fundam. Math. **1**, 93–104 (1920)

[Ster] Stern, M.: A Review of the Book by Olli Lehto, Korkeat Maailmat. Rolf Nevanlinnan elämä. Otava, Helsinki (2001)

[StSh] Stevens, R.S., Shanturam, R.: Computer generated van der Waerden partitions. Math. Comput. **17**, 635–636 (1978)

[Sto] Story, W.E.: Note on the preceding paper. Am. J. Math. **II**, 201–204 (1879)

[Str] Struick, D.: Letter to A. Soifer, March 3, 1995

[Sve1] Sved, M.: Paul Erdős – portrait of our new academician. Austral. Math. Soc. Gaz. **14**(3), 59–62 (1987)

[Sve2] Sved, M.: Old snapshots of the young. Geombinatorics. **II**(3), 46–52 (1993)

[Sza] Szabó, Z.: An application of Lovász' local lemma – a new lower bound for the van der Waerden number. Random Struct. Algoritm. **1**(3), 343–360 (1990)

[Sze1] Székely, L.A.: Remarks on the chromatic number of geometric graphs. In: Graphs and other Combinatorial Topics, Proceedings of the Third Czechoslovak Symposium on Graph Theory, Prague, August 24–27, 1982, ed. Fiedler, M., Teubner Verlaggesellschaft, Leipzig, 1983, 312–315

[Sze2] Székely, L.A.: Measurable chromatic number of geometric graphs and sets without some distances in Euclidean space. Combinatorica. **4**, 213–218 (1984)

[Sze3] Székely, L.A.: Erdős on unit-distances and the Szemerédi–Trotter theorems. In: Halász, G., et al. (eds.) Paul Erdős and his Mathematics II, pp. 649–666. János Bolyai Mathematical Society/Springer, Budapest/Berlin (2002)

[SW] Székely, L.A., Wormald, N.C.: Bounds on the measurable chromatic number of R^n. Discret. Math. **75**, 343–372 (1989)

[Szek] Szekeres, G.: A combinatorial problem in geometry: reminiscences. In: Erdős, P. (ed.) The Art of Counting, pp. xix–xxii. MIT Press, Cambridge, MA (1973)

[SP] Szekeres, G., Peters, L.: Computer solution to the 17-point Erdős-Szekeres problem. ANZIAM J. **48**, 151–164 (2006)

[Sz1] Szemerédi, E.: On sets of integers containing no four elements in arithmetic progression. Acta Math. Acad. Sci. Hungar. **20**, 89–104 (1969)

[Sz2] Szemerédi, E.: On sets of integers containing no k elements in arithmetic progression. Acta Arithmetica. **XXVII**, 199–245 (1975)

[Szp] Szpiro, G.: Das Auswahlaxiom und seine Konsequenzen: Wie viele Farben braught man zur Kolorierung der Ebene? – Je nachdem. . . . Neue Zürcher Zeitung, Nr. 85, 65 (April 13, 2005)

[Tai1] Tait, P. G., On the colouring of maps, Proc. Roy. Soc. Edinb. 10 (1878–80), 501–503

[Tai2] Tait, P. G., Remarks on the previous communication, Proc. Roy. Soc. Edinb. 10 (1878–80), 729

[Tai3] Tait, P.G.: Note on a theorem in the geometry of position. Trans. Roy. Soc. Edinb. **29**, 657–660 & plate XVI (1880). Reprinted: Scientific Papers **1**(54), 1898–1900, 408–411 & plate X

[Tar] Tarkovsky, A.: Time within Time: The Diaries 1970–1986, transl. by Hunter-Blair, K., Verso, London (1993)

[Tie1] Tietze, H.: Über das Problem der Nachbargebiete im Raume. Monatschefte für Mathematik und Physik. **16**, 211–216 (1905)

[Tie2] Tietze, H.: Gelöste und ungelöste mathematische Probleme aus alter und neuer Zeit; vierzehn Vorlesungen für Laien und für Freunde der Mathematik, München, Biederstein, 1949. English translation from the 2nd rev. German ed: Famous Problems of Mathematics: Solved and Unsolved Mathematical Problems, from Antiquity to Modern Times, Graylock Press, Baltimore, 1965

[Tho] Thomassen, C.: On the Nelson unit distance coloring problem. Am. Math. Mon. **106**, 850–853 (1999)

[Time] Whoosh, Time, December 26, 1949, p.30

[TV1] Tóth, G., Valtr, P.: Note on the Erdős–Szekeres theorem. Discret. Comput. Geom. **19**, 457–459 (1998)

[TV2] Tóth, G., Valtr, P.: The Erdős-Szekeres theorem, upper bounds and generalizations. In: Goodman, J.E., et al. (eds.) Discrete and Computational Geometry – Papers from the MSRI Special Program MSRI Publications 52, pp. 557–568. Cambridge University Press, Cambridge (2005)

[Tow1] Townsend, S.P.: Every 5-colouring map in the plane contains a monochrome unit. J. Combin. Theory Ser. A. **30**, 114–115 (1981)

[Tow2] Townsend, S.P.: Colouring the plane with no monochrome unit. Geombinatorics. **XIV**(4), 181–193 (2005)

[Tro] Trotsky, L.: What is national socialism? (June 10, 1933). In: The Struggle Against Fascism in Germany. Pathfinder Press, New York (1971)

[Tym] Tymozcko, T.: The four-color problem and philosophical significance. J. Philos. **76**(2), 57–83 (1979)

[Ung] [Ungar, P.]: Problem 4526, advanced problems and solutions. Amer. Math. Monthly **60**, 123 (1953); correction 60, 336 (1953)

[UPE] Upsilon Pi Epsilon: International Honor Society for the Computing and Information Disciplines. http://upe.acm.org/newsbrief_jan_2015.html

[VE] Vasiliev, N.B., Egorov, A.: Zadachi Vsesoyuznykh Matematicheskikh Olimpiad (Problems of Soviet National Mathematical Olympiads). Nauka, Moscow (1988) (Russian)

[VGR] Villarino, M.B., Gasarch, W., Regan, K.W.: Hilbert's proof of his irreducibility theorem. Am. Math. Mon. **125**(6), 513–530 (2017)

[Viz1] Vizing, V.G.: Ob ozenke khromaticheskogo klassa *p*-grafa (On an estimate of the chromatic class of a *p*-graph). Diskrenyi Analiz. **3**, 25–30 (1964) (Russian)

[Viz2] Vizing, V.G.: Kriticheskie grafy s khromaticheskim klassom (Critical graphs with the given chromatic class). Diskrenyi Analiz. **5**, 9–17 (1965) (Russian)

[Viz3] Vizing, V.G.: Nekotorye Nereshennye Zadachi v teorii grafov. Uspekhi Mat. Nauk. **XXIII**(6), 117–134 (1968) (Russian). English Translation: Some unsolved problems in graph theory. Russ. Math. Surv. **23**, 125–141 (1968)

[VND] Voronov, V.A., Neopryatnaya, A.M., and Dergachev, E.A.: Constructing 5-chromatic unit distance graphs embedded in the Euclidean plane and two-dimensional spheres (2022). arXiv:2106.11824v4

[Wae1] Waerden, B.L. van der: De algebraiese Grondslagen der meetkunde van het aantal [Ph.D. thesis]. Zutphen – W. J. Thieme & Cie (1926)

[Wae2] van der Waerden, B.L.: Beweis einer Baudetschen Vermutung. Nieuw Archief voor Wiskunde. **15**, 212–216 (1927)

[Wae3] Waerden, B.L. van der: Moderne Algebra, vol. 1(1930) and 2(1931), 1st ed.. Verlag von Julius Springer, Berlin

[Wae4] van der Waerden, B.L.: Die Gruppentheoretische Methode in der Quantenmechanik. J. Springer (1932)

[Wae5] van der Waerden, B.L.: Nachruf auf Emmy Noether. Math. Annalen. **111**, 469–476 (1935)

[Wae6] Waerden, B. L. van der: Moderne Algebra, vol. 1(1937) and 2 (1940), 2nd ed., Verlag von Julius Springer, Berlin

[Wae7] Waerden, B. L. van der: Letter to Behnke, May 10, 1940, Erich Hecke's Papers, Private German Collection

[Wae8] Waerden, B. L. van der: Letter to Hecke, May 16, 1940, Erich Hecke's Papers, Private German Collection

[Wae9] Waerden, B. L. van der: Prof. Van der Waerden verweert zich, Het Parool No. 326, February 1, 1946, p. 3

[Wae10] Waerden, B. L. van der: Rechtzetting, Proria Cures, February 1, 1946

[Wae11] Waerden, B. L. van der: Moderne Algebra, vol. 1 and 2 (1950), 3rd ed. Springer, Berlin

[Wae12] Waerden, B. L. van der: Over de Ruimte, Rede, 4 December 1950, P. Noordhoff N.V., Groningen–Djakarta (1950)

[Wae13] Waerden, B. L. van der: Einfall und Überlegung in der Mathematik, 1. Mitteilung. Elem. Math. **VIII**(6), 121–144 (1953); 2. Mitteilung. Elem. Math. **IX**(1), 1–24 (1954); 3. Mitteilung: Der Beweis der Vermutung von Baudet. Elem. Math. **IX**(3), 49–72 (1954), (German)

[Wae14] Waerden, B. L. van der: Einfall und Überlegung: Drei Kleine Beiträge zur Psychologie des Mathematischen Denkens (Idea and Reflection: Three Little Contributions to the Psychology of the Mathematical Thinking. In 3 parts; 3rd part: Der Beweis der Vermutung von Baudet), Birkhäuser, Basel (1954) (German)

[Wae15] van der Waerden, B.L.: Science Awakening, P. Noordhoff, Groningen, 1954. Expanded edition. Oxford University Press, New York (1961)

[Wae16] van der Waerden, B.L.: Wie der Beweis der Vermutung von Baudet gefunden wurde. Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität. **28**, 6–15 (1965)

[Wae17] Waerden, B. L. van der: Letter to Henry Baudet, October 20, 1965

[Wae18] van der Waerden, B.L.: How the proof of Baudet's conjecture was found. In: Mirsky, L. (ed.) Studies in Pure Mathematics, pp. 251–260. Academic, London (1971)

[Wae19] Waerden, B. L. van der: Letter to N. G. de Bruijn, April 5, 1977

[Wae20] van der Waerden, B.L.: On the Sources of My Book Moderne Algebra. Hist. Math. **2**, 31–40 (1975)

[Wae21] Waerden, B. L. van der: Beweis einer Baudetschen Vermutung, Bertin, E.M. J. et al. (eds.) Two Decades of Mathematics in the Netherlands 1920–1940: A Retrospection on the Occasion of the Bicentennial of the Wiskundig Genootschap, Part I, Mathematical Center, Amsterdam, 1978, 110–115 (with Commentary by N. G. de Bruijn, 116–124)
[Wae22] van der Waerden, B.L.: A History of Algebra. Springer, Berlin (1985)
[Wae23] Waerden, B. L. van der: Letter to Soifer, A., March 9, 1995
[Wae24] Waerden, B. L. van der: Letter to Soifer, A., April 4, 1995
[Wae25] Waerden, B. L. van der: Undated letter to Soifer, A., (Swiss Post Office seal April 24, 1995)
[Wae26] van der Waerden, B.L.: Wie der Beweis der Vermutung von Baudet gefunden wurde. Elem. Math. **53**, 139–148 (1998)
[WaD1] Waerden, D. van der: e-mail to A. Soifer, February 22, 2004
[WaD2] Waerden, D. van der: e-mail to A. Soifer, April 29, 2004
[WaH1] Waerden, H. van der: Letter to A. Soifer, June 20, 2004
[WaH2] Waerden, H. van der: Letter to A. Soifer, September 10, 2010
[Wag] Wagner, K.: Über eine Eigenschaft der ebenen Komplexe. Math. Ann. **114**, 570–590 (1937)
[Wal1] Walker, M.: Uranium machines, nuclear explosives, and national socialism: the German quest for nuclear power, 1939–1949. Ph. D. dissertation in the History of Science, Princeton University, October 1987
[Wal2] Walker, M.: Nazi Science: Myth, Truth, and the German Atomic Bomb. Plenum, New York (1995)
[Wa] Waller, A.O.: Simultaneously colouring the edges and faces of plane graphs. J. Combin. Theory Ser. B. **69**, 219–221 (1997)
[Wal] Walters, M.: Combinatorial proofs of the polynomial van der Waerden theorem and the polynomial Hales-Jewett theorem. J. Lond. Math. Soc. **61**(1), 1–12 (2000)
[WZ] Wang, W., Zhu, X.: Entire colouring of plane graphs. J. Combin. Theory Ser. B. **101**(6), 490–501 (2011)
[Wer] Wernicke, P.: Über den kartographischen Vierfarbensatz. Math. Ann. **58**(3), 413–426 (1904)
[W] Whitehead, E.G.: The Ramsey number N(3,3,3,3;2). Discret. Math. **4**, 389–396 (1973)
[WT] Whitney, H., Tutte, W.T.: Kempe chains and the four colour problem. Utilitas Math. **2**, 241–281 (1972)
[Wiki] https://en.wikipedia.org/wiki/IMU_Abacus_Medal
[Wil1] Wilson, R.J.: New light on the origin of the four-color conjecture. Hist. Math. **3**, 329–330 (1976)
[Wil2] Wilson, R.J.: Four Colours Suffice. Princeton University Press (2002)
[WilJ] Wilson, J.: New light on the origin of the Four-Color conjecture. Hist. Math. **3**, 329–330 (1976)
[WilJC] Wilson, J. C., On a supposed solution of the ‘four-colour problem’. Math. Gaz. 3 (1904–6), 338–340

[Wit] Witt, E.: Ein kombinatorischer Satz der Elementargeometrie. Math. Nachr. **6**, 261–262 (1952)

[Witt] Wittgenstein, L.: Remarks on the Foundation of Mathematics. Macmillan, New York (1956)

[Woo1] Woodall, D.R.: Distances realized by sets covering the plane. J. Combin. Theory Ser. A. **14**, 187–200 (1973)

[Woo2] Woodall, D.R.: The four-colour conjecture is proved. Manifold. **19**, 14–22 (1977)

[Woo3] Woodall, D.R.: The four-colour theorem. Bull. Inst. Math. & Appl. **14**, 245–249 (1978)

[WW] Woodall, D.R., Wilson, R.J.: The Appel-Haken proof of the Four-Color theorem. In: Beineke, L., Wilson, R.J. (eds.) Selected Topics in Graph Theory, pp. 83–101. Academic, London (1978)

[Wor] Wormald, N.C.: A 4-chromatic graph with a special plane drawing. J. Austral. Math. Soc. Ser. A. **28**, 1–8 (1979)

[Xu1] Xu, B.: On 3-colorable plane graphs without 5- and 7-cycles. J. Combin. Theory Ser. B. **96**, 958–963 (2006)

[Xu2] Xu, B.: A 3-color theorem on plane graphs without 5-circuits. Acta Math. Sin.. (in press)

[Xu3] Xu, B.: A sufficient condition on 3-colorable plane graphs without 5- and 6-circuits. Acta Math. Appl. Sin. **30**(3), 765–772 (2014)

[XR] Xu, X., Radziszowski, S.P.: On some open questions for Ramsey and Folkman numbers. In: Gera, R., Hedetniemi, S., Larson, C. (eds.) Graph Theory, Favorite Conjectures and Open Problems Problem Books in Mathematics, vol. 1, pp. 43–62. Springer (2016)

[XXER] Xu, X., Xie, Z., Exoo, G., Radziszowski, S.P.: Constructive lower bounds on classical multicolor Ramsey numbers. Electron. J. Combin. **#R35**, 11 (2004) http://www.combinatorics.org/

[XXR] Xu, X., Xie, Z., Radziszowski, S.P.: A constructive approach for the lower bounds on the Ramsey numbers $R(s, t)$. J. Graph Theory. **47**, 231–239 (2004)

[Zak1] Zaks, J.: On the chromatic number of some rational spaces. Combin. Optim. Res. Rep. Corr. **88–26**, 1–3 (1988)

[Zak2] Zaks, J.: On four–colourings of the rational four–space. Aequationes Math. **27**, 259–266 (1989)

[Zak3] Zaks, J.: On the connectedness of some geometric graphs. Combin. Theory (B). **49**, 143–150 (1990)

[Zak4] Zaks, J.: On the chromatic number of some rational spaces. Ars Combin. **3**, 253–256 (1992)

[Zak5] Zaks, J.: On the connectedness of some geometric graphs. Combin. Theory (B). **49**, 143–150 (1990)

[Zak6] Zaks, J.: Uniform distances in rational unit-distance graphs. Discret. Math. **109**(1–3), 307–311 (1992)

[Zak7] Zaks, J.: On odd integral distances rational graphs. Geombinatorics. **IX**(2), 90–95 (1999)

[Zer] Zermelo, E.: Beweis dass jede Menge wohlgeordnet werden kann. Math. Ann. **59**, 514–516 (1904)

[Zieg] Ziegler, G.M.: Günter M. Ziegler, "Alexander Soifer: 'The Mathematical Coloring Book. Mathematics of Coloring and the Colorful Life of Its Creators'. Springer, 2009, xxx+607 pp." Published online: 18 September 2014; Deutsche Mathematiker-Vereinigung and Springer-Verlag Berlin Heidelberg 2014**;** Jahresber Dtsch Math-Ver (2014) 116:261–269. (English) Zbl 1306.00028

[Zyk1] Zykov, A.A.: On some properties of linear complexes, (Russian). Mat. Sbornik N.S. **24**(66), 163–188 (1949). On some properties of linear complexes. Am. Math. Soc. Translation 1952, (1952). no. 79

[Zyk2] Zykov, A.A.: Problem 12 (by V. G. Vizing), Beiträge zur Graphentheorie, Vorgetragen auf dem Internationalen Kolloquium in Manebach (DDR) vom 9.–12. Mai 1967, p. 228. B. G. Teubner Verlagsgesellschaft, Leipzig, 1968

[Zyk3] Zykov, A.A.: Theory of Finite Graphs I, (Russian). Nauka, Novosibirsk (1969) (volume II has not been published). [Ein1]

Name Index

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1

C

D

E

I

J

K

T

U

V

W

X

Y

Z

Subject Index

© Alexander Soifer 2024
A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1

Index of Notations

© Alexander Soifer 2024

A. Soifer, *The New Mathematical Coloring Book*, https://doi.org/10.1007/978-1-0716-3597-1