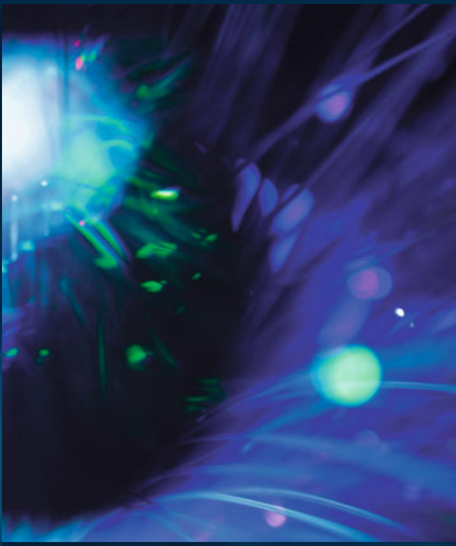
THE FRONTIERS COLLECTION



Shyam Wuppuluri
Giancarlo Ghirardi (Eds.)

SPACE, TIME AND THE LIMITS OF HUMAN UNDERSTANDING

Foreword by John Stachel and
Afterword by Noam Chomsky

Springer

THE FRONTIERS COLLECTION

THE FRONTIERS COLLECTION

The books in this collection are devoted to challenging and open problems at the forefront of modern science, including related philosophical debates. In contrast to typical research monographs, however, they strive to present their topics in a manner accessible also to scientifically literate non-specialists wishing to gain insight into the deeper implications and fascinating questions involved. Taken as a whole, the series reflects the need for a fundamental and interdisciplinary approach to modern science. Furthermore, it is intended to encourage active scientists in all areas to ponder over important and perhaps controversial issues beyond their own speciality. Extending from quantum physics and relativity to entropy, consciousness and complex systems—the Frontiers Collection will inspire readers to push back the frontiers of their own knowledge.

More information about this series at http://www.springer.com/series/5342

For a full list of published titles, please see back of book or springer.com/series/5342

Shyam Wuppuluri · Giancarlo Ghirardi
Editors

Space, Time and the Limits of Human Understanding

Springer

Editors
Shyam Wuppuluri
R. N. Podar School (CBSE)
Podar Educational Complex
Mumbai
India

Giancarlo Ghirardi
Abdus Salam International Centre for Theoretical Physics
Trieste
Italy

ISSN 1612-3018 ISSN 2197-6619 (electronic)
THE FRONTIERS COLLECTION
ISBN 978-3-319-44417-8 ISBN 978-3-319-44418-5 (eBook)
DOI 10.1007/978-3-319-44418-5

Library of Congress Control Number: 2016950405

© Springer International Publishing AG 2017
Chapter 15 is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/). For further details see license information in the chapter.
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.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Vox audita perit litera scripta manet,
The thing heard perishes but the written word remains

—Latin Maxim.

Foreword

First of all, let me thank Mr. Shyam Wuppuluri for his years of effort devoted to inspiring, assembling, and editing this volume. Rather than attempting to comment on the papers in it, I shall try to outline my approach to some issues that the reader may find helpful to bear in mind when reading these papers. Unfortunately, the concepts needed to discuss these issues are so interrelated that I have not found it possible to provide a simple, sequential introduction to these concepts. Indeed, they are inextricably intermingled in my presentation.

Every human community is based on an interrelated complex of labor processes that enable the community not just to survive but that, if successful, enable it to grow and thrive. Each such labor process involves three elements: the *labor* of a group of people using some *tools* to act upon the *initial objects* of labor (the "raw materials"). The goal of this process is action upon these initial objects in such a way as to produce *final objects* (the "finished products") that will benefit at least some members of the community.

But not all tools and initial objects are external to the members of the community. Language—and other symbolic systems that language enables people to create—allows them to form conceptual systems, some of which in turn become intellectual tools that can be used to modify the initial conceptual systems and to create new ones.

In other words, not all labor need to be manual. Just as important is *intellectual labor*. Like all forms of labor, intellectual labor starts from some object—in this case a conceptual object—and works on this object with the aim of modifying it to produce a new conceptual object (see my paper "Problems not Disciplines"). There are many types of intellectual tools (art, literature, religion, etc.) but most important for the questions discussed in this volume are the scientific disciplines, including both the so-called natural and social sciences (better called the human sciences—the *Geisteswissenschaften* in German).

Why include the social sciences? As in all other forms of labor, a successful intellectual labor process is ultimately the result of the collective labor of more than one individual: *knowledge* is our name for such results. Rather than being an individual problem, as it has been treated in Western philosophy since the time of Descartes, the problem of knowledge is a social problem (see my paper Where is Knowledge?). And it should always be borne in mind that the ultimate goal of all intellectual labor is to provide additional resources to some human community for action upon the material world and, intended or not, to produce changes in the community itself.

As noted already, in the case of intellectual labor the objects of labor are *complexes of concepts*; and the collective labor acting upon these complexes is that of some *intellectual community*. When it is successful, *new knowledge* is our name for the results of such a process of intellectual labor. And it is important to bear in mind that knowledge itself is indeed a *process* that will never end as long as intellectual communities continue to exist. So one should be suspicious of any proposed axiomatization of some realm of human knowledge, especially if it is accompanied by claims of the closure of what should be an open, unending process.

An important philosophical issue is the distinction between the *conceptual objects* that constitute our knowledge and the *objects of that knowledge*, and the relation between these two. Such philosophical doctrines as conflicts as empiricism, rationalism, naturalism, materialism and idealism depend on the answer to this question. Suffice it to say that I regard the goal of action on the objects of knowledge as primary. The introduction of a new conceptual object, or the modification of an existing one, must make clear its relation to some object(s) of knowledge. In the case of physical concepts, such as space and time, this means that their definition must be accompanied by some account of the means by which an ideal measurement of this quantity can be carried out.

Let me emphasize: I am not maintaining that it is meaningful because it is measurable; rather that *because* it is meaningful it must in principle be measurable. No one has stated this lesson more eloquently than Gaston Bachelard:

> In order to embody new experimental evidence, it is necessary to deform the original concepts, study the conditions of applicability of these concepts, and above all incorporate the conditions of applicability of a concept into the very meaning of the concept. … The classic division that separates a theory from its application ignores this necessity to incorporate the conditions of applicability into the very essence of the theory (*La formation de l'esprit scientifique*. Paris: Vrin. 1938, pg. 61; transl. by J.S.).

Like all tools, intellectual tools themselves may become the object of a labor process aimed at improving their effectiveness—indeed one tool may even be set aside totally in favor of some new tool. The prime example is relation between language, the first intellectual tool, and logic. Logic is a tool created for the improvement of language when it is applied to certain subjects. Indeed, one should rather say "logics," since more than one consistent logical system may be employed in the critical reconstruction of a language. So one might say that logic is about language and language is about the world. Often this two-step relation is reduced to a one-step relation: logic is about the world—a view I have criticized for decades

> [A] logic always has some language as its object. The more formalized one wants the logic to be, the more formalized the language must be. … Since a logic presupposes a language,

> the objects logic studies must be linguistic objects. I believe in the existence of other objects which are quite independent of language; but such objects, in contrast to linguistic references to them, cannot have a logic. Lest this be thought a mere verbal quibble, let me point out that someone who believed that all reality was fundamentally conceptual in nature could meaningfully and non-metaphorically speak of the logic of the world.
>
> The doctrine that all relations are fundamentally logical relations and that there is hence no basic distinction between logic and ontology is not unknown in the philosophical literature. … It has even been given a name, "panlogism" … However, if one does adopt this position, it should be done with full awareness, and not tacitly through acceptance of a certain approach to quantum logic" (excerpts from "Do Quanta Need a New Logic?").

I maintain that, in this respect, mathematics is similar to logic. Mathematical structures do not apply directly to the world, but to other conceptual structures that have been created to apply to some aspects of the world. Again this two-step relation is often short-circuited with the assertion that mathematical structures apply directly to the world (see Where is Knowledge?). In the case of physical structures, the easiest way to see the fallacy of such an assertion is by looking the question of units. In order to quantify any physical concept, a system of appropriate units must be introduced; and it is only the ratio of some physical quantity to its unit that is a pure number, and the usual mathematical relation between two such physical concepts can only be applied to pure numbers. To define the distance between two points, one must first specify some unit of distance, and the numerical value of the distance between these points is then the ratio of that distance specified in these units divided by this unit of distance. Given a set of such primary units, secondary units may then be computed. For example, in order to specify the velocity of some object one must divide a spatial distance by a temporal interval; so units of space and time must first be specified.

Finally, let me turn to the concepts of space and time. If we now consider the objects of the dynamical natural sciences, their objects themselves are processes (e.g., mechanical, electromagnetic, quantum, etc.). As long as the Newtonian worldview prevailed, it made sense to consider such processes as composed of a sequence of states, one for each value of the absolute time. But with the advent of the special theory of relativity, and even more the general theory of relativity, a viewpoint based on states is no longer tenable. The processes *are primary*; and any introduction of a temporal sequence of states—to the extent that it is possible—is quite secondary and dependent on additional definitions. A number of theoretical physicists have emphasized the primacy of process, including Lee Smolin:

> Relativity theory and quantum theory each. tell us—no, better, they scream at us—that our world is a history of processes. Motion and change are primary. Nothing is, except in a very approximate and temporary sense. How something is, or what its state is, is an illusion. It may be a useful illusion for some purposes, but if we want to think fundamentally we must not lose sight of the essential fact that 'is' is an illusion. So to speak the language of the new physics we must learn a vocabulary in which process is more important than, and prior to stasis (*Three Roads to Quantum Gravity.* Basic Books, 2001).

To put it another way, the four-dimensional concept of space–time is primary, its possible division into three-dimensional space and one-dimensional time is quite secondary and depends on the introduction of additional concepts.

What about time, the reader may object? Surely this is a primary concept. Here one must make a distinction between two concepts of time that are often conflated: between what I have called *local time* and *global time* (see "Albert Einstein: A Man for the Millennium"). To clarify what I mean by the concept of the local time between two events, I often draw an analogy between this concept and the concept of the spatial distance between two places. It is well known that the distance between two different places depends on the spatial path between them. The *shortest* distance is along the *straightest* path between them. The local time between two non-coincident events in space–time depends on the spatiotemporal path between them. That is the analogy; the difference is that the *longest local time* interval between the two events is along the *straightest* path in space–time has. As alluded to above in the discussion of units, it is the local time that is measured by clocks and the spatial distance that is measured by rulers ("measuring rods").

Clearly, when one speaks of two non-coincident events as "taking place at the same time," it cannot be the concept of local time that is involved. Indeed it is here that the concept of *global time* enters the story. When the Newtonian worldview prevailed, the additional concept of the absolute time provided such a definition of the global time. Two non-coincident events either occurred at the same absolute time or one preceded the other, and clocks measured this absolute time in a path-independent way. Once the concept of absolute time was abandoned, clocks measured only the path-dependent local time, and two non-coincident events might or might not be connected by a temporal path. Given an initial event, all events not connected to it by a timeline path form an entire region of space–time that has been called the *elsewhere* of the event. The concept of global time amounts to some definition of a three-dimensional slice of this four-dimensional elsewhere; and in general nothing physically significant can depend on this definition.

The great divide between general relativity and all previous dynamical theories is that they were all based on fixed and given space–time structures. All dynamical processes took place on the fixed and given stage provided by these structures. In contrast, general relativity has no such fixed structures. All space–time structures are themselves dynamical entities, subject to field equations. So that until a solution to these field equations is specified, no answers to questions about space–time structures can be given. As I have put it elsewhere: in all the rest of physics, "where" and "when" are parts of the question; in general relativity, "where" and "when" are parts of the answer.

I do not expect the reader to uncritically accept my viewpoint on these issues; but hope that he or she will bear in mind the need for adopting some standpoint about them when confronting many of the questions discussed in this volume.

John Stachel

Appendix

Those interested in further exploring my approach to these questions may want to start with: Where is Knowledge? in Joseph Kouneiher et al, eds., *Frontiers of Fundamental Physics/The Eleventh International Symposium, AIP Conference Proceedings*, vol. 1446 (American Institute of Physics, 2012), pp. 312–334. Some other relevant papers of mine (many of them may be found at this website: http://www.bu.edu/cphs/ces/research/papers-by-stachel/) are:

- "Do Quanta Need a New Logic?" In *From Quarks to Quasars/ Philosophical Problems of Modern Physics,* edited by Robert Colodny, pp. 229–347. Pittsburgh: U. of Pittsburgh Press, 1986.
- "Quantum Logic," in *The Philosophy of Science/An Encyclopedia*, vol. 2, Sahotra Sarkar and Jessica Pfeffer, eds. (New York/Abington: Routledge), pp. 633–644 (2005).
- "Structure, Individuality and QG," in D.P. Rickles, S. French and J. Saatsi, eds., *Structural Foundations of Quantum Gravity* (Oxford University Press), pp. 53–82 (2006).
- Albert Einstein: A Man for the Millenium?, in L. Mornas and J. Diaz Alonzo, eds., *A Century of Relativity Physics/ERE 2005 XXVII Spanish Relativity Meeting: AIP Conference Proceedings, vol. 841* (American Institute of Physics, 2006).
- Problems not Disciplines, in J. Renn and K. Gavroglu, eds, *Positioning the History of Science* (Springer 2007) pp. 163–167.
- The Story of Newstein or: Is Gravity Just Another Pretty Force?, in J. Renn (ed.) *The Genesis of General Relativity*, Vol. 4, *Theories of Gravitation in the Twilight of Classical Physics: The Promise o f Mathematics and the Dream of a Unified Theory* (Springer 2007), pp. 421–458.
- Prolegomena to any future quantum gravity, in D. Oriti (ed.), *Approaches to Quantum Gravity*, (Cambridge Univ. Press, 2009).
- John Stachel and Kaća Bradonjić, Quantum Gravity: Meaning and Measurement, in *Studies in the History and Philosophy of Modern Physics*, 2013.
- The Hole Argument and Some Physical and Philosophical Implications, *Living Reviews in Relativity*, vol. 17, 2014, pp. 1–66.

Preface

This book offers, to a diverse nonspecialist audience, a panorama of contextual perspectives on the topics of space and time. Almost every field has in its language, a notion of space and time. This duo is intertwined into the fabric of our existence. If you are an aficionado of classical music, you would see that absolute music and programmed music differ not only in their nature but also in purpose. While the former is composed for its own sake, the latter is composed keeping in mind a context. Akin to this, space and time, have been analyzed either for their own sake or to fit a context and thereby to serve some purpose. To each his own.

When this collection was planned, it seemed natural to organize the articles, written by towering figures belonging to seven diverse fields, into various sections to facilitate a better understanding. And with the hope that a global meaning would emerge from such contextual viewpoints when the dots are connected. We thus segregated this volume into various fields, namely philosophy, physics, mathematics, biology/cognitive science, logic/computer science and a section 'Miscellaneous' which includes literature, space–time geography and art.

Einstein once remarked, "Space and Time are modes by which we think and not a condition in which we live." The Philosophy section is therefore devoted to some foundational and metaphysical aspects of Space and Time, which are intuitive in nature. Articles contain views of philosophers like Descartes, Newton, Leibniz, Hume, Kant, Poincare, Cassirer, Sartre, Husserl, etc., on space and time. In addition to the western philosophy, two articles are dedicated to elaborating on the views of space and time of the Indian schools of thought—Vedic and *Nyaya*. It also contains a treatise on the nature of space and time.

Thinking often gets translated into a theory after it passes the toll gate of the scientific method, which not only verifies the claims experimentally but also checks their consistency. The Physics section is aimed at providing a clear and precise account of empirical/physical status of space and time. Its articles cover topics ranging from relativity theory to quantum theory, ideas of Minkowski, Einstein, and Hermann Weyl, and others, and some more recent advances in our understanding of Space and Time.

The Book of Nature is written in the language of mathematics, as Galileo remarked. The next section deals with the mathematical foundations of space and time. Articles address themes such as the nature of metric, manifolds, spatial relationships, and the nature of the continuum. They draw from the works of Euclid, Leibniz, Einstein, Hermann Weyl, Grothendieck, and others, thereby trying to analyze and model the notions of Space and Time from a mathematical setting. These contributions are both pedagogical and expository in nature.

Space and time are intangible. But our consciousness paints them for us all the time. Be it when our house is packed with guests or when a deadline is pressing. But what are these notions produced by our brains? The section Biology/Cognitive Science deals with spatial and temporal perceptions and how they are grounded in our cognition, space in biological systems, time in biological systems (circadian rhythms) and time from an evolutionary perspective.

Merrick Furst once remarked, “The biggest difference between time and space is that you can’t reuse time.” Within the computational context, we concern ourselves with storage (space), computing time (time) and their associated complexities. This section, Logic/Computer Science, deals with topics such as computational complexity, space–time tradeoffs in analytic engines and topics in classical and modern logic that pertain to the second part of the title, limits of human understanding/thinking. It offers a concise account of the logician Kurt Godel’s work which includes his incompleteness theorems, their impact on theoretical physics and economics and his views on time.

The last section, ‘Miscellaneous’, contains handpicked articles which elucidate the notions of space and time with in the fields of literature, space–time geography and visual arts.

An intriguing foreword by Prof. John Stachel and a captivating afterword by Prof. Noam Chomsky add icing to the cake. All the 44 authors have taken utmost care to keep the articles as nontechnical as possible and self-contained. They have included an appendix with their articles (when needed) to furnish the technicalities for those keen to explore deeper.

During my undergraduate days, I, like many others, was deeply influenced by the fresh worldview found in Douglas Hofstadter’s books. He is in a sense the ‘Pied Piper’, that Hermann Weyl once addressed David Hilbert as, who seduced many rats to follow him into the deep river of thinking. His emphasis that analogy making is at the core of our cognition, motivated me and eventually led to inception of this volume.

An original thinker belonging to any domain of science should be Sherlock’ian in many aspects. Just as Holmes decodes the mysterious stick figures inscribed on the walls (The adventure of the Dancing Men) by finding patterns and thereby extracting meanings, the true pursuit of science begins with decoding such analogical stick figures painted by Mother Nature on the canvas of perceived reality. They can be tracks on bubble chambers or patterns within numbers. In such a pursuit, one greatly benefits from knowledge of diverse fields which can trigger cross-domain analogies. We sincerely hope, through this erudite volume, that a symphony of patterns and a tapestry of intuitions will emerge, providing a holistic insight into the questions: ‘What is Space?’ and ‘What is Time?’

Juhu, Mumbai Shyam Wuppuluri

Acknowledgements

Of course Rome was neither built in a day nor by a single person! I am very fortunate to find a lot of people who went out of their way to support me in this endeavor. Chronologically, I would like to thank Prof. John Stachel from the bottom of my heart for agreeing to write a foreword and guiding me patiently all through since the project's inception. His support kick-started the project. My decision to proceed with Springer owed much to the wonderful association with the editor, Dr. Angela Lahee, who has played a role not less than that of a guardian angel concerning this project. Her unparalleled expertise, enthusiasm, and patience, coupled with timely and valuable advice are gratefully acknowledged. The project could not have been in better hands than hers.

While the overall structure and contents were in place, with 38 authors on board but still moored in the harbor, I sought a co-editor, who would provide me with feedback. Little did I know that Prof. Giancarlo Ghirardi, a close associate of Prof. John Stewart Bell, would be the one to join me in this journey as a co-editor. Despite his health issues he took out time from his busy academic life and dealt with the project with care and patience. I have greatly benefited from his association, humility, wisdom, and support. Language definitely fails to translate the gratitude I have towards him. My heartfelt thanks to Prof. Noam Chomsky, who does not need an introduction. I am yet to find an adjective which has not been used previously to describe his multifaceted personality. Though I feel guilty that I cannot take the name of each author, I would like to thank every author, wholeheartedly, for their efforts and especially those who generously and patiently collaborated, incorporating the feedback and customizing the article. I am fortunate to have such associations. Sadly I must mention that Prof. Jonathan Borwein, who not only contributed to this volume but also leaves behind a wealth of other intellectual achievements, died on 2nd August, 2016. Mourning his passing, I would like to warmly acknowledge his efforts and support. The fact that he will continue to live through his works offers great solace to all his admirers.

I would also like to thank people, who though are not directly connected to this project, have influenced me, positively, in many ways. After all action at a distance has its place in physics!

I would like to thank from the bottom of my heart—Prof. Sujatha Ramdorai, who has been a beacon of support, not just to me but to many others who are passionate about science. My existence owes a lot to the solacing support she offered at various times. I would also like to thank Mr. M. G. Subramanian for teaching me the art of online communication, during my formative years. I would like to express my gratitude towards the director-principal Mrs. Avnita Bir and Podar management for always supporting me.

Lastly I would like to thank Prof. R. Sridharan, who, though he never taught me directly, nevertheless has inspired me deeply through his writings, humanity, humility, intellect and emphasis on quality. The relationship that I share with him is that of an arbitrary pebble with an ocean *ingente*—whose deep bond is both unbeknownst to the passersby and undisturbed by the chaotic wind.

Such is the stance of unsung bonds and friendships whose existence forms a fragrant and radiant twine which is stitched alongside with the other dull cords of humdrum routines into the robe of hermit's life—which is what makes his life both interesting and worth living!

Juhu, Mumbai Shyam Wuppuluri

Contents

Part I Philosophy

Part II Physics

Part III Mathematics

Part IV Biology/Cognitive Science

Part V Logic/Computer Science

Part VI Miscellaneous

About the Editors

Shyam Wuppuluri is working as a research associate for R. N. Podar institute. A computer science graduate, he has a long-standing interest in various areas of mathematics, theoretical physics, philosophy and cognitive science. E-mail: shyam.wuppuluri@gmail.com

Giancarlo Ghirardi served as the director of the Department of Theoretical Physics at University of Trieste for more than two decades. He is now the President of the Consortium for Physics of the University of Trieste and Consultant and Member of the Academic Board of the ICTP, Trieste. Giancarlo made important contributions to the foundations of quantum mechanics, particularly through the GRW interpretation that was proposed jointly with Rimini and Weber. He has authored many articles and books, and organized many conferences. An intellectual leader, he is a recipient of many honors, accolades and international awards. E-mail: ghirardi@ictp.it

Part I
Philosophy

Chapter 1
Space as a Source and as an Object of Knowledge: The Transformation of the Concept of Space in the Post-Kantian Philosophy of Geometry

Francesca Biagioli

1.1 Introduction

This paper deals with the transformation of the concept of space in the post-Kantian philosophy of geometry from the second half of the nineteenth century to the early twentieth century. Kant famously characterized space and time as forms of intuitions, which lie at the foundations of the apodictic knowledge of mathematics. The success of his philosophical account of space was due not least to the fact that Euclidean geometry was widely considered to be a model of apodictic certainty at that time. However, such later scientific developments as non-Euclidean geometries and the general theory of relativity called into question the certainty of Euclidean geometry and posed the problem of reconsidering space not so much as a source of knowledge, but as an open question for empirical research.

The first section offers a discussion of the main objections against Kant's view of space as a source of knowledge. The opposed view of space as an object of knowledge emerged in geometrical empiricism, a tradition that can be traced back to such mathematicians as Carl Friedrich Gauss, Nikolai Lobachevsky, Bernhard Riemann, Richard Dedekind, and Felix Klein, and that found one of its clearest expressions in the epistemological writings of the physiologist and physicist Hermann von Helmholtz. The second section provides a general introduction to the new phase of this debate inaugurated by Einstein's general theory of relativity of 1915. On the one hand, Einstein relied on geometrical empiricism for the view that geometry has an empirical meaning. On the other hand, he distanced himself from the received view of space by claiming that the general covariance of his field equations removed from space and time the last remnant of physical objectivity. In the third section,

---

F. Biagioli (✉)
University of Konstanz, Zukunftskolleg, P.O. Box 216, 78457 Konstanz, Germany
e-mail: francesca.biagioli@uni-konstanz.de

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_1

I discuss two different strategies for a philosophical account of the transformation of the concept of space from a Kantian perspective: (1) Hermann Weyl vindicated the foundational role of the Kantian concept of space by observing that any coordinate assignment, even in Einstein's relativistic space-time theory, presupposes the ideal perspective of the transcendental subject for its setting; (2) Ernst Cassirer emphasized the heuristic aspect of the concept of space as a hypothetical system of mathematical relations. Although both strategies offer a possible philosophical account of space informed by the sciences, I argue that Cassirer's focus on the structure of spatial notions, rather than their subjective origin, had the advantage of reflecting a variety of uses of spatial concepts in human culture and art, which he considered to be no less essential than scientific concepts to a philosophical account of the concept of space.

1.2 Three Roads to Geometrical Empiricism in the Nineteenth Century

1.2.1 The Philosophical

The first motivation for geometrical empiricism was to overcome some of the philosophical difficulties of Kant's account of space in the *Critique of Pure Reason*. Kant characterized space and time by distinguishing these concepts from sensations, on the one side, and general concepts, on the other. The first distinction is in terms of form and matter of appearance: "Since that within which the sensations can alone be ordered and placed in a certain form cannot itself be in turn sensation, the matter of all appearance is only given to us a posteriori, but its form must all lie ready for it in the mind a priori, and can therefore be considered separately from all sensation" (Kant [1], A20/B34). More specifically, space and time are forms of intuition, according to Kant, insofar as the order of appearance is directly present to the mind in the localization of objects in space and time. Therefore, he claimed that space is a necessary representation, which lies at the foundation of all outer intuitions (A24/B38). Kant went on to argue that space differs from general concepts in the way in which it is related to its parts: whereas general concepts contain a finite collection of possible instantiations under them, any limitation of space, including infinite division, lies in a single concept of space as one of its parts. Therefore, he characterized space as an intuition that lies at the origin of knowledge concerning external reality. The principles of (Euclidean) geometry offered the first example of conceptual knowledge derived from intuition with apodictic certainty (A25/B39).

One classical objection to this view is that it presupposed the syllogistic logic of Kant's time, according to which logical reasoning is restricted to finite domains and construction in intuition is necessary to justify existential assumptions concerning infinite domains (e.g., the parallel to a given line from a point outside it and incommensurable magnitudes).[1] After the emergence of mathematical logic in the nineteenth century, it became possible to account for the same distinction in terms of two different means of logical proof or rules of quantification. This approach to Kant's intuition, which is known as "logical," goes back to Cassirer [3] and was given a modern formulation by Hintikka [4]. Alternative approaches include the phenomenological approach advocated by Parsons [5], among others, based on the fact that immediacy is no less essential than singularity to the Kantian conception of space. But how to accommodate the former feature of intuition with the discovery of non-Euclidean geometry in the nineteenth century? Not only does such a discovery contradict the view that Euclidean geometry is evidently true, but the question arises whether geometric knowledge can be true at all.

In order to overcome this difficulty, Friedman [6] argued for mediating between these two approaches in line with Helmholtz's geometrical empiricism. On the other hand, Helmholtz ruled out the view of geometrical axioms as evident truths by explaining how basic geometric notions are derived by observation and experience of rigid motions. On the other hand, he inferred a naturalized form of intuition by considering the possible changes of perspective of a perceiving subject. As Helmholtz [7, p. 162] put it, "Kant's doctrine of the a priori given forms of intuition is a very fortunate and clear expression of the state of affairs; but these forms must be devoid of content and free to an extent sufficient for absorbing any content whatsoever that can enter the relevant form of perception." Helmholtz believed that the form/content distinction deserved a new formulation after the emergence of experimental psychology of vision, on the one hand, and non-Euclidean geometry, on the other.

1.2.2 The Natural

Helmholtz's naturalization of the Kantian theory of perception goes back to his 1855 lecture "Über das Sehen des Menschen." On that occasion—Helmholtz was delivering the Kant Memorial Lecture as a Professor of Physiology at the University of Königsberg—he maintained that Kant's view that the perception of physical objects presupposes some subjective forms of intuition received an empirical confirmation in Johannes Müller's theory of specific nerve energies. Müller showed that sensuous qualities depend not so much on the perceived object as on the constitution of our nerves in the case of perceptions usually associated with light. His theory accounted for the fact that the same visual sensation can have different causes (e.g., an electric current or a blow to the eye). Vice versa, light does not necessarily cause visual sensations, but, for example, ultraviolet rays cause only chemical reactions.

---

[1] Several examples are discussed in Friedman [2, Ch. 1].

Helmholtz was Müller's student in Berlin and relied on the same experimental methodology. However, he developed an original approach to visual perception in his *Handbuch der physiologischen Optik* (1867). Helmholtz advocated a theory of local signs, according to which sensations are signs for their stimuli. According to him, visual perception depends on our capacity to interpret those signs by drawing unconscious inferences from nerve stimuli to existing objects. Such a capacity deserves an empirical explanation in terms of psychological—rather than purely physiological—processes. Therefore, Helmholtz called his approach "empirist" and contrasted it with nativist views.[2] In this connection, he distanced himself from Kant's conception of space and time as pure intuitions, whose form can be defined independently of any empirical content. Not only did Kant overlook the empirical conditions for the formation of these concepts, but his theory of pure sensibility led his followers to assume that there are innate laws grounded in the form of spatial intuition, that is, the axioms of (Euclidean) geometry (Helmholtz [9], p. 456).

As it emerges most clearly in Helmholtz's later paper "Die Tatsachen in der Wahrnehmung" (1878), Helmholtz's criticism did not rule out the possibility of generalizing the Kantian notion of form of spatial intuition to all possible combinations of contents. After the passage quoted above, Helmholtz went on to say that "the axioms of geometry limit the form of intuition of space in such a way that it can no longer absorb every thinkable content, if geometry is at all supposed to be applicable to the actual world. If we drop them, the doctrine of the transcendentality of the form of intuition is without any taint. Here Kant was not critical enough in his critique; but this is admittedly a matter of theses coming from mathematics, and this bit of critical work had to be dealt with by the mathematicians" (Helmholtz [7], p. 162).[3] Furthermore, Helmholtz believed that even his empiricist epistemology relied ultimately on the assumption of a lawful course of nature, which sometimes he presented as a condition of the possibility of experience in Kant's sense.[4]

I turn back to the latter issue in connection with Cassirer's theory of the symbolic forms of experience. The following section offers a brief overview of different positions on geometrical empiricism in nineteenth-century mathematics.

---

[2] See [8, Ch. 5]. Helmholtz addressed two different questions. The first concerned the two-dimensionality of vision. At the time Helmholtz was writing, the dominant view endorsed by Johannes Müller, among others, was that a two-dimensional, spatial representation is primitively given in vision. In this view, only the perceptions of depth and of distance (i.e., the kind of perceptions that presuppose three-dimensionality) have to be learned. By contrast, Helmholtz sought to derive all spatial representations from the association of nonspatial sensations. The second question concerned the singularity of vision. Helmholtz called nativist Müller, Ewald Hering, and all those who derived the singularity of vision from the supposition of an anatomical connection between the two retinas.

[3] For a discussion of Helmholtz's claims about the "transcendental" status of space, see Biagioli [10].

[4] See, e.g., Helmholtz [7, p. 142]. Notice, however, that there were important turning points in Helmholtz's relation to Kant in this regard. Cf. Hatfield [8] and Hyder [11].

1.2.3 The Mathematical

The third road to geometrical empiricism goes back to the first attempts to rethink the question concerning the form of physical space after the discovery of non-Euclidean geometry.[5] Different hypotheses about the existence and the number of parallel lines to a given line had been explored in the eighteenth-century by such mathematicians as Girolamo Saccheri, Johann Heinrich Lambert, and Adrien-Marie Legendre. However, it was only in the 1820s that János Bolyai and Nikolai Lobachevsky, independently of each other, developed a system of geometry based on the denial of Euclid's parallel postulate. Lobachevsky was also one of the first to address the question whether non-Euclidean theorems (e.g., the proposition that the sum of the angles in a triangle is less than 180°, which followed from the denial of the parallel postulate) can be tested empirically by using astronomical measurements.

Although such an experiment proved to be impractical, much of the debate that followed from Riemann to Einstein concerned the possibility to explore the link between geometry and experience. As Gauss (in [13], p. 87) put it in a letter to Bessel dated April 9, 1830, "the theory of space has an entirely different place in knowledge from that occupied by pure mathematics. There is lacking throughout our knowledge of it the complete persuasion of necessity (also of absolute truth) which is common to the latter; we must add in humility that if number is exclusively the product of our mind, space has a reality outside our mind and we cannot completely prescribe its laws." On the one hand, Gauss's remark is reminiscent of Newton's view of geometry as the part of mechanics that deals with measurement in the *Philosophiæ Naturalis Principia Mathematica* [14]. On the other hand, by 1830, Gauss had enough knowledge of non-Euclidean geometry to see that there might be different possible hypotheses when it comes to the laws of space.[6]

A further development of this tradition is found in Gauss's student's, Bernhard Riemann, habilitation lecture of [17] "Über die Hypothesen, welche der Geometrie zu Grunde liegen" and in a series of papers published by Helmholtz between 1868 and 1878 and later collected by Paul Hertz and Moritz Schlick in the centenary edition of Helmholtz's *Epistemological Writings* (1977).[7] Both Riemann and Helmholtz started with a general notion of space as a manifold, comparable with such empirical manifolds as color and tone systems, in order to then pose the question of the necessary and sufficient conditions for introducing metrical relations. In this regard, Riemann showed the possibility of formulating infinitely many

---

[5] For an introductory account of the discovery of non-Euclidean geometry and its prehistory, see Engel and Stäckel [12].

[6] Sartorius von Waltershausen [15, p. 81] reported that even Gauss made an attempt to test the Euclidean hypothesis during his geodetic work. However, this interpretation is controversial and it was only after Bolyai's and Lobachevsky's works that the question arose whether the geometry of space could be non-Euclidean (see [16]).

[7] For a comprehensive account of nineteenth-century philosophy of geometry, see Torretti [18].

geometrical hypotheses based on Gauss's geometry of surfaces. Riemann foreshadowed Einstein's general theory of relativity by considering even the hypothesis of variably curved spaces and by articulating the conjecture of a metric, whose coefficients would depend on the local distribution of matter and forces.

The greater generality of Riemann's inquiry notwithstanding (Helmholtz restricted his consideration to manifolds of constant curvature), Helmholtz was one of the first to draw attention to the possibility of a physical interpretation of non-Euclidean geometry. He identified the fundamental precondition for the possibility of measurement as the requirement that the points of a system in motion remain fixedly linked or the free mobility of rigid bodies. He considered this to be a fact ascertained by observation and experiment, beginning with our ability to localize objects in space by performing congruent displacements. He then showed that the metrical geometry that is implicit in spatial perception and in measurement includes both Euclidean and non-Euclidean cases as possible mathematical specifications. He agreed with the conclusion of Riemann's inquiry, insofar as he believed his argument to prove the empirical—rather than a priori—origin of geometrical axioms.

The standard formulation of geometrical empiricism, in this sense, is found in Helmholtz's 1870 public lecture "Über den Ursprung und die Bedeutung der geometrischen Axiome," which initiated a philosophical and mathematical debate on the status of geometrical axioms.[8] A further development of Helmholtz's view of axioms as empirical generalizations is found in Pasch [20], which contains one of the first axiomatic treatments of geometry. The notion of axiom was progressively weakened with the emergence of the axiomatic method. Both mathematicians and philosophers called axioms "postulates" to emphasize the conceptual nature of the axioms and the possibility of formulating different hypotheses when it comes to physical space (Klein, some of the positions advocated by the Peano School, and Cassirer). In the twentieth century, it became common usage to refer to axioms as definitions in disguise (Poincaré) or implicit definitions (Hilbert, Schlick) of geometrical concepts as objects, whose existence depends solely on the coherence of the system of relations established by the axioms.

To sum up, with the advancement of the axiomatic method, geometrical empiricism lost its significance for a clarification of the questions concerning the origin and meaning of geometrical axioms. It became ever more clear that an appropriate understanding of how axioms work in modern mathematics presupposes a sharp separation between geometries as axiomatic systems and interpreted geometry or the theory of space and space-time. However, the tradition discussed above remained an important reference in the debate concerning physical and philosophical interpretations of the mathematical structures under consideration.

In order to highlight this point, the following section discusses the related problem of establishing geometrical and physical invariants.

---

[8] On the discussion of Helmholtz's view in neo-Kantianism, see Biagioli [19].

1.3 Invariants and Symmetries

The first characterization of geometrical objects in terms of invariants of transformation groups goes back to Felix Klein's "Vergleichende Betrachtungen über neuere geometrische Forschungen" [21], which is best known as "Erlanger Programm," after Klein's appointment as Professor at the University of Erlanger in the same year. An essential contribution to the implementation of such a research program is found in Sophus Lie's *Theorie der Transformationsgruppen*, which appeared in three volumes between 1888 and 1893.[9] The third volume contains a critique of Helmholtz's mathematical reasoning, along with the theorem that became the standard derivation of a metric of constant curvature from a set of necessary and sufficient assumptions about infinitesimal free mobility.

The fundamental ideas of the group-theoretical approach emerged from the observation that some geometrical properties are invariant under specific types of geometrical transformations. Such congruent displacements as translations, for example, leave invariant parallelism, lengths and measure of angles. In modern terminology, these are called the invariants of the Euclidean group. However, the same invariants might not be preserved by other transformations. Collineations in projective (and non-Euclidean) geometry, for example, leave invariant such properties as, of three or more points: to lie on the same line; of curves: to be a conic. But such transformations are known to alter such invariants of the Euclidean group as parallelism and absolute measurements.

Klein arrived at the idea of a group-theoretical classification of geometry by applying the basic notions of the algebraic theory of groups introduced by Evariste Galois and Camille Jordan to projective geometry—which flourished in the nineteenth century after the works of Jean-Victor Poncelet, Jakob Steiner, Christian von Staudt, and Arthur Cayley, among others.[10] Transformations form a group if: (i) the product of any two elements of the group also belongs to the group; (ii) there is in the group a neutral element (i.e., an element that leaves the other elements unchanged when combined with them); (iii) for every transformation in the group, there is in the group the inverse transformation.

Such mathematicians as Klein, Lie, and Poincaré showed that the same notions offered a general point of view for the comparison of geometrical researches. The significance of their classificatory ideas lies not least in the fact that the same ideas offered a point of comparison between different physical theories. Klein [25], for example, used the same approach to account for the shift from classical mechanics to special relativity. Whereas Galilean transformations preserve the invariant quantities of Newtonian mechanics (i.e., acceleration, force, mass, and therefore time, length, and simultaneity), the Lorentz transformations preserve the velocity of light, but not length and time (simultaneity).

---

[9] On Klein's relationship to Lie and the reception of the Erlanger Programm, see Hawkins [22] and Rowe [23].

[10] On the development of group theory from Galois to Klein, see Wussing [24].

The shift from special to general relativity seems to mark a break with this tradition, insofar as the space-time of general relativity is variably curved, the local value of curvature depending on the distribution of mass and energy, as first suggested by Riemann's conjecture. In the context of general relativity, such a correlation between space, time, and matter followed from Einstein's theory of gravitation. It is worth noting that, nevertheless, Einstein attached much importance to Helmholtz's interpretation of non-Euclidean geometry via the free mobility of rigid bodies for the development of his own ideas. In Einstein [26] he claimed that, without the view of geometry set forth by Helmholtz, he would have been unable to formulate the theory of relativity. But how are we to understand this comparison? Not only did Einstein's general relativity presuppose a different mathematical approach for the determination of the measure of curvature, but it implied a different concept of space. By considering hypotheses other than Euclidean geometry in the characterization of physical space, Helmholtz conceived of space as an object of research, whose properties can be determined empirically, to the required degree of approximation. In general relativity it became impossible to ask about the geometrical structure of space or space-time per se, insofar as geometrical properties depend on the distribution of matter. In classical mechanics and special relativity, the idea of space and time as objects endowed with a particular structure depends on the fact that inertial systems provide a privileged frame of reference. By contrast, Einstein's principle of relativity—in its various formulations—presupposes that the laws of physics apply to systems in any kind of motion. One of the conditions for this is that the form of natural laws remains unchanged under arbitrary changes of space-time values or general covariance. As Einstein [27, p. 117] put it, general covariance "takes away from space and time the last remnant of physical objectivity." Space-time coincidences, on the other hand, become fully objective in the sense of independence from the observer's perspective.

Einstein's claim has been much discussed at that time and in more recent literature.[11] The concluding section of this paper deals with Weyl's and Cassirer's interpretations as two different ways to account for the transformation of the concept of space from geometrical empiricism to general relativity and to provide new answers to the question concerning subjective and objective aspects of knowledge.

1.4 Subjectivity and Objectivity

1.4.1 Hermann Weyl

Hermann Weyl was both a great mathematician, whose contributions ranged from analysis, algebra, topology, differential geometry, and fundamental physics, and one of the few philosopher-scientists of the twentieth century to defend a Kantian

---

[11] For a discussion of different positions, see Norton [28] and Ryckman [29].

conception of space in line with Husserl's phenomenological approach to the a priori.[12] In his main work on the general theory of relativity, *Raum-Zeit-Materie*, which appeared in four editions between 1918 and 1921, Weyl [30, p. 98]) accounted for the shift to the new theory by saying that the space of general relativity "is nothing more than a three-dimensional manifold devoid of all form; it acquires a definite form only through the advent of the material content filling it and determining its metric relations." With reference to his metrical infinitesimal geometry, nevertheless, Weyl argued that the empirical determination of metrical relations did not rule out the possibility of a priori knowledge, insofar as the borderline between a priori and a posteriori was set somewhere else. In Weyl's account: whereas the metric at a point P, along with the metrical relation of P to its neighboring points, is everywhere the same and retains the status of an a priori condition of experience, the mutual orientation of the metrics in different points is a posteriori.

In *Raum-Zeit-Materie*, Weyl advocated a phenomenological conception of the a priori in order to deal with problem of coordination. Referring to Einstein's remark about general covariance, Weyl [30, p. 8] wrote: "[T]he objectivity of things conferred by the exclusion of the ego and its data derived directly from intuition, is not entirely satisfactory; the co-ordinate system which can only be specified by an individual act (and then only approximately) remains as an inevitable residuum of this elimination of the percipient." Whereas Einstein focused on the objectivity achieved in general relativity when it comes to determine space-time coincidences, Weyl's emphasis is on the fact that a residuum of subjectivity still plays the role of a necessary presupposition for the setting of the coordinate system. In other words, Weyl vindicated the Kantian conception of space as a source of knowledge by identifying the minimal phenomenological structure that is required for space-time coordination.

1.4.2 Ernst Cassirer

Ernst Cassirer started his career as one of the leading figures of the Marburg School of neo-Kantianism and became known as the founder of the philosophy of symbolic forms.[13] This contained one of the most promising attempts in twentieth-century philosophy to develop a unified perspective on human culture, including natural and social sciences, the humanities, and the arts as different expressions of the subject/object relation. Cassirer identified a range of different symbolic forms in which such a relation can be articulated. The broader scope of this view notwithstanding, one of Cassirer's starting point is found in his 1921 book *Zur*

---

[12] See Ryckman [29, Chaps. 5 and 6].

[13] On the development of Cassirer's thought from neo-Kantianism to the philosophy of symbolic forms, see Ferrari [31].

Einstein'schen Relativitätstheorie. Commenting on the significance of general covariance for the formulation of Einstein's principle of relativity, Cassirer observed that general relativity confirmed the tendency of every physical theory to abstract from what is immediately given in perception in order to gain the appropriate conceptual expression and understanding of the facts of experience. As Duhem [32, p. 322] put it, in order to provide the basis for the development of physical theory, empirical facts have to be transformed and put into a "symbolic form."

Regarding the spatial order of experience as a symbolic form, even before general relativity Cassirer emphasized that the mathematical concept of space provides a combination of different hypotheses. He referred, in particular, to the classic cases of manifolds of constant curvature investigated by Helmholtz, Klein, and Lie. Despite the difference between these cases and the variably curved space-time of general relativity, Cassirer argued for a further generalization of the previous system of hypotheses in order to attain to complete objectivity in facing the problem of coordination. His interpretation focused on the fact that, by generalizing the principle of relativity, Einstein needed not single out any privileged systems of reference.

Without entering into the details of Cassirer's interpretation, for my present purpose I limit myself to point out that the example of Einstein's space-time theory played an important role in Cassirer's conception of his own philosophical project. He wrote in his concluding chapter on the problem of redefining physical reality in symbolic terms rather than as absolute reality: "It is the task of systematic philosophy, which extends far beyond the theory of knowledge, to free the idea of the world from this one-sidedness. It has to grasp the whole system of symbolic forms, the application of which produces for us the concept of an ordered reality, and by virtue of which subject and object, ego and world are separated and opposed to each other in definite form, and it must refer each individual in this totality to its fixed place" (Cassirer [33], p. 447).

The relevant point for my brief comparison with Weyl is that Cassirer articulated a different strategy for vindicating the core ideas of the Kantian theory of space. Instead of identifying the a priori of space as a subjective, but necessary source of knowledge, Cassirer investigated the different ways in which the concept of a spatial order is used to articulate the subject/object relation in anthropological, scientific, and artistic contexts. A number of examples are found in Cassirer's main works on the philosophy of symbolic forms. A very clear characterization of this approach for a wider audience is found in his 1931 paper [34] "Mythischer, ästhetischer und theoretischer Raum," which was delivered during the fourth *Congress für Ästhetik und allgemeine Kunstwissenschaft* in Hamburg. In this paper, he argued that the variety of spatial concepts notwithstanding, space as a symbolic form is characterized by its purely relational nature, which is a necessary precondition for the determination of things and meanings. Cassirer referred, for example, to the mythical significance of spatial orientation in ancient myths and in religious representations of the afterlife. The example of spatial and temporal concepts in the arts shows even more clearly the potential of the symbolic forms in inventing a new constellation of meanings.

In Cassirer's view, the clarification of the symbolic function of space and time in modern physics opened the door to looking at these concepts from different perspectives, while deepening Kant's insight into the nature of space and time as orders of appearance rather than perceived objects. However, Cassirer's proposal is not to return to Kant's original view of space as a source of knowledge, but to investigate how thought can anticipate experience in virtue of the creative and hypothetical character of our systems of concepts. My suggestion is that the philosophy of symbolic forms offers an original synthesis between the two opposed ways to consider space as a source and an object of knowledge. As Cassirer [33], p. 426) put it, empiricism and idealism meet in certain presuppositions with regard to the doctrine of empirical space and of empirical time: "Both here grant to experience the decisive role, and both teach that every exact measurement presupposes universal empirical laws."

References

1. Kant, I. (1998). *Critique of pure reason*. (P. Guyer & A.W. Wood, trans.). Cambridge: Cambridge Univ. Press.
2. Friedman, M. (1992). *Kant and the exact sciences*. Cambridge, Mass.: Harvard Univ. Press.
3. Cassirer, E. (1907). Kant und die moderne Mathematik. *Kant-Studien, 12*, 1–49.
4. Hintikka, J. K. (1973). *Logic, language-games and information: Kantian themes in the philosophy of logic*. Oxford: Clarendon Press.
5. Parsons, C. (1992). The transcendental aesthetic. In P. Guyer (Ed.), *The Cambridge companion to Kant* (pp. 62–100). Cambridge: Cambridge University Press.
6. Friedman, M. (2000). Geometry, construction, and intuition in Kant and his successors. In G. Sher & R. L. Tieszen (Eds.) *Between logic and intuition: Essays in honor of Charles Parsons* (pp. 186–218). Cambridge University Press.
7. von Helmholtz, H. (1977). *Epistemological writings*. Dordrecht: Reidel.
8. Hatfield, G. C. (1990). *The natural and the normative: Theories of spatial perception from Kant to Helmholtz*. Cambridge, Mass.: MIT Press.
9. von Helmholtz, H. (1867). *Handbuch der physiologischen Optik*. Leipzig: Voss.
10. Biagioli, F. (2014). What does it mean that "space can be transcendental without the axioms being so"? Helmholtz's claim in context. *Journal for General Philosophy of Science, 45*, 1–21.
11. Hyder, D. J. (2009). *The determinate world: Kant and Helmholtz on the physical meaning of geometry*. Berlin: de Gruyter.
12. Engel, F., & Stäckel, P. E. (Eds.). (1895). *Die Theorie der Parallellinien von Euklid bis auf Gauss: Eine Urkundensammlung zur Vorgeschichte der nichteuklidischen Geometrie*. Leipzig: Teubner.
13. Kline, M. (1980). *Mathematics: The loss of certainty*. Oxford: Oxford University Press.
14. Newton, I. (1687). *Philosophiæ Naturalis Principia Mathematica*. London: Streater.
15. von Waltershausen, S. W. (1856). *Gauss zum Gedächtnis*. Leipzig: Hirzel.
16. Gray, J. J. (2006). Gauss and non-Euclidean geometry. In A. Prékopa & E. Molnár (Eds.). *Non-Euclidean geometries: János Bolyai memorial volume* (pp. 61–80). New York: Springer. .

17. Riemann, B. (1854/1867). Über die Hypothesen, welche der Geometrie zu Grunde liegen. *Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen, 13*, 133–152.
18. Torretti, R. (1978). *Philosophy of geometry from Riemann to Poincaré*. Dordrecht: Reidel.
19. Biagioli, F. (2014). Hermann Cohen and Alois Riehl on geometrical empiricism. *HOPOS: The Journal of the International Society for the History of Philosophy of Science, 4*, 83–105.
20. Pasch, M. (1882). *Vorlesungen über neuere Geometrie*. Leipzig: Teubner.
21. Klein, F. (1872). *Vergleichende Betrachtungen über neuere geometrische Forschungen*. Erlangen: Deichert.
22. Hawkins, T. (1984). The *Erlanger Programm* of Felix Klein: Reflections on its place in the history of mathematics. *Historia Mathematica, 11*, 442–470.
23. Rowe, D. E. (1992). Klein, Lie, and the Erlanger Programm. In L. Boi, D. Flament, & J.-M. Salanskis (Eds.), *1830–1930: A century of geometry, epistemology, history and mathematics* (pp. 45–54). Berlin: Springer.
24. Wußing, H. (1969). *Die Genesis des abstrakten Gruppenbegriffes*. Berlin: VEB Deutscher Verlag der Wissenschaften.
25. Klein, F. (1911). Über die geometrischen Grundlagen der Lorentzgruppe. *Physikalische Zeitschrift, 12*, 17–27.
26. Einstein, A. (1921). *Geometrie und Erfahrung*. Berlin: Springer.
27. Einstein, A. (1916). Die Grundlagen der allgemeinen Relativitätstheorie. *Annalen der Physik, 49*, 769–822.
28. Norton, J. D. (1989). Coordinates and covariance: Einstein's view of spacetime and the modern view. *Foundations of Physics, 19*, 1215–1263.
29. Ryckman, T. A. (2005). *The reign of relativity: Philosophy in physics 1915–1925*. New York: Oxford University Press.
30. Weyl, H. (1950). *Space, time, matter* (H.L. Brose, trans.). New York: Dover.
31. Ferrari, M. (2003). *Ernst Cassirer—Stationen einer philosophischen Biographie*. Hamburg: Meiner.
32. Duhem, P. (1914). *La Théorie physique, son objet—sa structure* (2nd ed.). Paris: Rivière.
33. Cassirer, E. (1923). *Substance and function and Einsteins theory of relativity*. (M.C. Swabey & W.C. Swabey, Trans.). Chicago: Open Court.
34. Cassirer, E. (1931). Mythischer, ästhetischer und theoretischer Raum. *Zeitschrift für Ästhetik und allgemeine Kunstwissenschaft, 25*, 21–36.

Chapter 2
Time in Physics and Time in Awareness

E.C.G. Sudarshan

2.1 Introduction

Time is a part of our immediate awareness. We come across it as duration without necessarily attributing a measure to it; as sequencing with the notion of “before” and “after” between pairs of “events”; as a cyclic variable related to some natural process like the diurnal cycle—morning, noon, evening, midnight or like the pulse or heartbeat or the ticking of the clock. The time labelling astronomical events looks the same whether things are going forward or backward while time labelling the germination, growth, maturity, decline and perishing of a plant or animal does look different and strange backwards.

This megalomorphian zoology of time is reflected both in our language and in our science. We talk of “this time tomorrow” or “whenever I am afraid I become confused” in which we embed a cyclic process in a well ordered time. We will return to cyclic time later.

2.2 Durational Time in Mechanics; Linear and Cyclic Time

In mechanics processes are described in time and the laws of motion pertain to rates of change with respect to time. While time *ab initio* is defined in a bizarre fashion by Newton as “flowing uniformly” it is definition without content since flowing requires something else with respect to which this flow is measured! Further how does one

---

E.C.G. Sudarshan (✉)
Physics Department, University of Texas Austin, Austin, TX, USA
e-mail: bhamathig@gmail.com

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_2

recognize the flow of time? It is like asking “Where does this road go?” meaning thereby where do people, vehicles and objects traversing the road go—the road does not go anywhere!

A practical measure of time is obtained by observing the change in configuration of a physical system that is believed to be “moving” uniformly. It becomes acceptable since many such physical systems exist and the time as measured by each of these coincide more or less. It could be the pulsations of a quartz crystal or of radio frequency electromagnetic wave. For longer measures of time we could use the change of seasons, the ebb and flow of tides, sunrise and sunset. We could use as a rough and ready reckoning our heartbeats or the hunger or sleep cycle. For precision measurements we could use a precise system like quartz crystal or an isolated atom—to the extent that these are uniform we could expect that time as measured by them would all give a measure of “true time”.

The first law of motion asserts that an isolated body has a constant rate of change of position, an uniform motion. So, to the extent that we have isolated bodies and to the extent that we can keep it in the view, the measurement of time is reduced to a measurement of distance. The time so measured is a linear time or, rather, only a segment of it since the uniformly moving body was not in our vicinity either for a very early time or for very late time. By extending our method of length measurement to greater distances we could measure a longer segment of time: longer but still finite.

If the motion is not uniform but still the departure from uniformity is known we can correct for the variation and determine the “uniformly flowing” time. Use a sundial shadow to measure time on a sunny day is a familiar example. A cheap clock gains time in cold weather and loses time in hot weather; if we know the temperature and thermal expansion of the pendulum suspender we could determine the correct time.

In contrast to linear time most clocks are cyclic. The first example is the motion of planets around the sun, where the line joining the planet to the sun describes equal areas in equal times (that is, constant areal velocity). So if we measure the area described (for a purely circular orbit this is proportional to the angular displacement) we can measure the time elapsed.

For times longer than a complete revolution we need to count the number of complete revolutions together with the angular displacement. The same is true if we are using the oscillations of a pendulum (or the days spent in confinement). Now the counters that we have are usually cyclic, though the complete cycle is a long time. Most wrist chronometers indicate time in seconds (or fractions thereof), minutes, hours, days and months. We could build a calender that goes over a century or a millennium. But eventually, like the odometer of a Volvo, the counter will recycle! We can extend the period as much as we want by having more digits on the counter; but any finite system configuration reading time must be cyclic. Linear time going from minus infinity to plus infinity is an abstraction. When quantum theory is invoked, this cyclicity can be proved based on the positivity of the energy of any isolated system.

2.3 Metric Time and Its Relation to Symmetry Principles

The use of linear and of angular motion for measuring time by comparison of configurations (linear position and angular orientation respectively) is not accidental but due to certain symmetry requirement on physical laws for an "isolated" system. The quantity of linear motion ("momentum") and of angular motion ("angular momentum") of an isolated body should be unchanged with respect to time if the physical laws governing the system are to be unchanged by change in spatial location or by change in angular orientation. These are part of relativity laws of Newton-Galileo Mechanics. They are based on our ideas of homogeneity and isotropy of our space.

To give a measure of time ("metrizing" time) we need to invoke yet another physical quantity that is also invariant stemming from another symmetry, that is the irrelevance of the epoch in which we study a physical system. For a known "particle" (an isolated body whose structure can be disregarded for describing the motion) given its mass and its energy we can determine the momentum which is also mass times velocity (in non-relativistic theory). Similarly for angular motion, given the "moment of inertia" of the rotating system and its energy we could determine the angular velocity. In terms of knowing the angular velocity and the angular displacement we can measure the time elapsed.

It is interesting to recognize that the determination of time elapsed can be even more intimately related to the principle of relativity if we have also the requirement that laws of physics should be unchanged when one goes from one inertial frame to another inertial frame; including the one which moves uniformly with respect to another.

The quantity that is associated with this invariance (and in a technical sense, generating this transformation) is a quantity called the mechanical moment. This quantity increases linearly with respect to time, the coefficient of increase being the momentum. Thus time elapsed is the ratio of the change in the mechanical moment divided by the momentum. This general form of metrization is valid even when Einstein's relativity has to be used. We can work out a similar relation between the ratio of the generator of changes in angular velocity and the angular momentum. The times so defined are the same for all (isolated) systems.

There is thus an intimate relation between elapsed time measured by the comparison of configurations and the basic principles governing the laws of motion. Since configurations (spatial location or angular orientation) involves comparison between something fixed and something moving, there must be parts, that is structure in a physical system used as a measuring instrument of time in addition to the instrument for measuring position or orientation. The time measured in all these cases is "duration" which is metrized as described above. Nothing has "happened": the motion could be reversed without our finding anything different from the forward motion. Forwards and backwards do not have anything intrinsically different though by conventions we could make some distinction as for example a clock that goes anti clockwise. [Is a clock that goes anticlockwise a clock or an anticlock?] This mechanical time has no "sense".

2.4 Irreversibility and Historical Time

Our usual experience (in the waking state!) is with a “sense” of time. In fact much more than metric time, our time is a directed time in which a temporal order is imposed on our experiences. We have memory of the past but not of the future; and we distinguish the distant past from our recent past. Our internal measure of time is not very reliable but our temporarily sequencing is generally very good. When we look outside for evidence of this sequencing we find many processes with such time sense. Trees grow, rivers and wind erode land, ice melts and hot coffee cools. Light bulbs and candles shed light, broadcasting antennas and stars radiate, our bodies age.

There is a curious connection with a class of physical processes involving heat. We know that mechanical energy can be converted into heat: after a cold shower on a cool day we rub ourselves vigorously with a towel! But if we use heat to do mechanised work as in a steam engine we can convert only a part of the heat. Of all such “heat engines” the most efficient are reversible engines and even they are only partially efficient. The irreversible processes seem to declare that time is some intrinsic change, a degradation of energy. This is the second law of thermodynamics. The irreversible changes increase a quantity called entropy. Thermal equilibrium states have the maximum entropy under these conditions. All natural evolutions seem either to increase entropy (irreversible processes) or keep it unchanged (reversible processes). Natural evolution is towards thermal equilibrium, towards increasing entropy.

Irreversible processes have a sense of time and run backwards they look bizarre and whether it would be a splash in pond converging towards a point from which a stone is ejected or a dead body becoming alive and growing younger by the day. Candles don’t grow absorbing candlelight converging on it (though sometimes we feel rejuvenated). Time as history relates to irreversible processes and the “second law of thermodynamics”.

There is now a fundamental problem of relating historical directed time to durational (mechanical) undirected time. Undoubtedly irreversible process not only take place but are the very mechanism of metabolism and nourishment for organisms and lead to pattern formation and structure in physical systems. But if mechanics is at the basis of all physical laws how can we obtain the observable irreversibility based on a reversible mechanics? This is an old problem of physics. There has been repeated attempts at solution, the latest being based on the “mixing” or “chaotic” behavior of complex dynamical systems.

In an irreversible system as a rule transient structures arise, which are called “dissipative structures” by Prigogine. Familiar examples are singing telegraph wires (aeolian tones) arising out of eddy viscosity as wind blows past a highly stretched wire, the dancing brilliance of gas jet (or candle) that is about to get extinguished or the wild oscillations of a bridge about to collapse. But the most familiar one is the generation of metastable states and a definite lifetime: like the excited states of an atom or a radioactive nucleus. In the domain of high energy physics most particles are unstable. The regularity of decay law tells us that this law itself could be used as a clock—the fraction of the surviving particles depends experimentally on the time elapsed. Historical time and durational time can be compared.

2.5 The Birth of Time

Does time have a beginning? In Newtonain physics time has no beginning and no end. Somewhat like Melchi-Zadek (Who came to bless Abraham and to whom Abraham paid tithe) had no beginning and no genealogy! But like Newtonian time, Melchi-Zadek was also an idealization. We must think of Jacob, Abraham's grandson who had to operate in the phenomenal world. What about time around us?

All around us we see irreversible processes. Candles burn, food gets eaten and digested; and most of the phenomena around us are dictated by the generous supply of energy from the sun. Energy is being continually degraded; entropy is increasing and all the processes around us are maintained by this steady flow of solar energy.

Has it always been so and will it always be so? Surely not, if we recognize that the sun will burn down in a finite time. Even without knowing the mechanism of the sun's energy we can conclude, along with Olbers, that if sun and the stars are uniformly distributed throughout space and they had all existed for an infinitely long time then the night sky couldn't be dark. There would be equilibrium. The sun and stars would not emit or absorb. This simple but profound observation ties time to Cosmology.

As soon as we talk of cosmology we have to ask are the principles of homogeneity and isotropy, of time invariance and equivalence of uniformly moving physical systems really true? Could not, for example, physical laws depend on the cosmological epoch? Did the universe have a beginning of time? What does it mean? The present view generally held by cosmologists is that universe is evolving.

If things are changing in time the measurement of time is going to get tricky. Instead of considering time we can consider its square, its logarithm or any function that suits our fancy. Milne, in his kinematic relativity chooses the logarithm so that the origin of the new time has gone to minus infinity. More generally in relativistic cosmology space and time are not fixed but become dynamical variables: the space-time manifold obeys equations of motion; and the origin of our present version of the universe could be explosive singularity called the "Big Bang".

How does quantum mechanics alter all this? In quantum dynamics also one distinguishes the reversible evolution for isolated systems and the irreversible processes of external interference on the quantum system. The measurement of time becomes somewhat more subtle but it can be done. One conclusion was already mentioned only cyclic time can be strictly measured by a quantum clock, but survival probabilities do give a measure of linear time in an approximate manner.

2.6 Causation and Time Sequence

Closely related to the idea of temporal succession and temporary duration is the notion of causation. Causes and effects are interrelated; this interrelation is the essence of dynamical dependence, of the very notion of describability. But while this relation is symmetric in reversible processes it is asymmetric and directed in irreversible processes; the effect is dependent on the cause and later while the cause is earlier and alterable. It is this scheme of describing nature and the correspondence with our memory in normal waking state, that makes it natural for us to view time as being directed and historical.

In this vein, let us ask if our awareness alters the structure of time, clearly when we are in deep sleep we have no components to our awareness and hence no configuration. Therefore time itself is not! In normal waking state we are generally aware of historical time but there are contexts in which we are aware of duration but not of historicity. Times of intense creativity or total absorption in the task, of deep peace in which we do things like breathing or walking or even driving without the instruction of historical time, being immersed in music or times of meditative awareness in which texture of time seems to alter historical and durational times seem to coexist. And in dreams our experience of time is such that linear and cyclic times, reversible and irreversible time all coexist. Sometimes it is as if the irreversible sequence of events have been reversed.

There are some people who say that the "arrow of time" (the directed sense of time) stems from the expanding universe. The "Big Bang" and the consequent expansion of spacetime dominates the functioning of the physical law and all the universe and hence each part of it feels this all-pervading time sense. Since the creation of the perceived physical structures and even the synthesis of chemical elements is subsequent to "Big Bang" we may state this as saying the arrow of time is as old as matter itself; that rather than irreversible time originating in the interactions of complex systems they all had the same origin, the "Big Bang". It is reversible mechanics and its durational time that is the idealization.

2.7 Perception of Time: Archetypes and Artifacts

So far we have talked about the time in the physical sciences. When we come to the perception of time our archetypes and the very nature of our awareness are relevant. Three of these archetypes have already been alluded to: That of a "uniformly flowing" linear time and both durational and historical processes embedded in it. This is at the base of our waking awareness and is the one used in physical science. The second is closed cyclic loops of time which contain aspects of the waking state but in which time loses its "sense". The topology of time and space is altered. This is the time of our dreams and of our poetry and the fine arts. Third is the structureless duration of deep (dreamless) sleep in which there are no configurations to demarcate time. But there is a transcendent fourth, in which all the previous "times" coexist without clashing and all seen as artifacts. This is the creative time of contemplative awareness in which the rising and dissolution of historical time, the multiple cycles of cyclic time and duration itself maybe comprehended. While this is part of every person's awareness it has not yet found a place in physical science. To freely adapt a stanza from Ramayana:

> Transcendence of the limited renders time a chariot;
> In limitation, are causes and effects;
> Know that in harmony is pristine existence;
> Go forth in well being and joy!

References

1. Eddington, A. S. (1928). *Nature of the Physical World (1928)*. New York, Cambridge: The Macmillan Company, The University Press Year.
2. Harrison, E. R. (1989). *Darkness at night*. Harvard University Press.
3. Gamow, G. Big Bang Theory. https://en.wikipedia.org/wiki/George_Gamow#Big_Bang_nucleosynthesis
4. *Mandukya upanishad* (on the 3 states—Waking, sleep and dreaming).
5. Venkatesananda, S. (1993). *Yoga Vasistha*. State University of New York Press.

Chapter 3
Time and Space in Ancient India: Pre-philosophical Period

Nataliya Yanchevskaya and Michael Witzel

Abbreviations

RV	Ṛgveda
AV	Atharvaveda
BĀU	Bṛhadāraṇyaka-Upaniṣad
MU	Maitrī (aka Maitrāyaṇa, Maitrāyaṇī or Maitrāyaṇīya)-Upaniṣad
JUB	Jaiminīya-Upaniṣad-Brāhmaṇa
MBh	Mahābhārata
YV	Yogavāsiṣṭharāmāyaṇa

The earliest source of our knowledge of Ancient Indian perceptions of space and time is the Ṛgveda, one of the most ancient monuments of Indo-European liturgical poetry, a collection of hymns dedicated to different deities of the Vedic pantheon, and the oldest of the 4 Vedas (c. 1200–1000 BCE). In this paper we attempt to reconstruct Vedic views on time and space (information found in different layers of Vedic texts is fragmentary and therefore needs to be extracted and systematized) and then show historical development of these ideas, from the Ṛgveda to the latest texts of the Vedic tradition (Upaniṣads, c. 500 BCE) and then to the epics (first of all, the Mahābhārata).

N. Yanchevskaya
Princeton University, Princeton, NJ, USA
e-mail: nay@princeton.edu

N. Yanchevskaya · M. Witzel (✉)
Department of South Asian Studies, Harvard University, 1 Bow Street, Cambridge, MA 02138, USA
e-mail: witzel@fas.harvard.edu

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_3

3.1 Time in the Ṛgveda

We do not find any mention of time *in abstracto* in the Ṛgveda. The word *kāla* that is the most common term for denoting the abstract notion of time in the later tradition is used in the Ṛgveda only once, in RV 10.42.9,[1] a hymn that belongs to a later stratum of the text. The word *kāla* (***kāle***—Loc. Sg. from *kāla,* "in [a moment of] time") in this stanza denotes the proper moment to throw the dice to win the perfect score (*kṛta*), i.e., it is rather a specific "proper instant" than a general/abstract notion. Another word that means "time" in the later texts—*samaya*—is not used in the Ṛgveda at all.

Although the Ṛgveda has no unifying term for time as a general notion, it seems that all major Indian ideas and concepts of time developed in the later tradition originate there: first and foremost, it is the idea of two types of time or *two times*—the first is time-eternity, and the second one is transitory, "concrete", or "profane" time. The Ṛgveda also has the seeds of a view that time is the Lord of the universe that regulates and even creates it.

A word that denotes this all-powerful time-eternity in the Ṛgveda is the year (*saṃvatsara*). In one of the cosmogonic hymns of the 10th maṇḍala (RV 10.190) it is stated that "the Year (*saṃvatsara*) that dispenses days and nights" was produced "out of the foaming ocean".[2] It is clear from the hymn that this foaming or undulating ocean is, in fact, *cosmic ocean* that symbolizes the universe in its unordered and chaotic state (thus, according to Schayer, the year is "the first product out of Cosmic Waters"[3]). Before the cosmic ocean there was only "the kindled heat" (*tapas*) that gave birth to the Cosmic Order (*ṛta*) and Truth/Reality (*satya*), and then "the night (*rātri*) was born".[4] From the night the cosmic ocean is produced, and then—the year. The night here is not simply a period of day but rather the impenetrable primordial darkness before creation.

The Year in this hymn is called "the Lord of everything that blinks",[5] i.e., of the living. After the year is born, the universe changes its state from chaotic into organized. It becomes ordered; elements of creation—the sun and the moon, heaven and earth, interspace (*antarikṣa*) and sunlight appear and become properly arranged "according to the order" by someone who is called "the Arranger" (*dhātṛ*).[6] This Arranger might be the Year itself or an epithet of one of the Vedic deities, such as Prajāpati, however, it is clear that only with the creation of the year there appears a force that organizes the universe and, functionally, creates it.

---

[1] *uta prahām atidīvyā jayāti kṛtaṃ yac chvaghnī vicinoti* ***kāle*** | *yo devakāmo na dhanā ruṇaddhi sam it taṃ rāyā sṛjati svadhāvān* ||.

[2] RV 10.190.2.

[3] See Schayer [11], p. 5.

[4] RV 10.190.1.

[5] *samudrād arṇavād adhi saṃvatsaro ajāyata* | *ahorātrāṇi vidadhad viśvasya miṣato vaśī* ||.

[6] RV 10.190.3.

This hymn is probably one of the earliest mythological expressions of an idea that time organizes space and arranges phenomena of the created universe.

In the Ṛgveda we also observe the origin of another important concept found in various later texts, such as the epics and the Purāṇas, that of the wheel of time (*kāla-cakra*). In different verses of the famous riddle hymn RV 1.164 (the *"asya vāmasya"* hymn) there is a vivid description of a single-wheeled chariot[7] of the sun whose wheel "rolls around heaven" (*varvarti … pari dyām*). It is stated that the wheel is "undecaying" (or "unaging"[8]) and "unobstractable",[9] therefore it simultaneously covers both temporal and spatial domain and is unbound in either of them. It is "twelve-spoked"[10] or "five-spoked"[11]; it is called "the one [with] twelve segments or fellies"[12] in a different stanza of the hymn; it has "three naves"[13]; it bears "pairs of seven hundred twenty sons" and is compared to their "father who has five feet and twelve parts, the overflowing one in the upper half of heaven".[14] All living beings (or, in a different interpretation, all worlds—*bhuvanāni viśvā*) "stand on it" and are "fixed (or dependent) upon it"[15]; it also has "three hundred … and sixty pegs".[16]

Although the quoted hymn is a riddle and nothing is stated directly, it can be deduced what different elements of the wheel mean and what the whole wheel symbolizes. Twelve spokes of the wheel, as well as its twelve segments or fellies are the twelve months. Five spokes and five feet signify the five seasons (Vedic tradition distinguishes either five or six seasons); the "three naves" are the three major seasons (every 4 months), associated with solemn Vedic rituals; "three hundred … and sixty pegs" are the days of the year, whereas "pairs of seven hundred twenty sons" are days and nights of the year. The whole wheel hence represents a year. It is a *wheel of the year* that rolls in heaven and carries all beings or worlds. It is closely associated with the sun and the moon ("overflowing one in the upper half of heaven") and its path is the path of the sun.

Remarkably, this wheel is called in the hymn "the wheel of Cosmic Order (*ṛta*)",[17] therefore the year is equated to *ṛta*—a cornerstone notion of the Vedic worldview that we provisionally translate as "Cosmic Order". This term is rather difficult to translate due to its richness and importance for the Vedic culture: not only it signifies the Cosmic Order and organizing principle of the universe, but also

---

[7] *ratham ekacakram.*

[8] *ajaram.*

[9] *anarvaṃ.*

[10] RV 1.164.11.

[11] RV 1.164.13.

[12] RV 1.164.48.

[13] RV 1.164.02, 48.

[14] RV 1.164.12.

[15] RV 1.164.13, 14.

[16] RV 1.164.48.

[17] *cakram… ṛtasya* (RV 1.164.11).

the Truth[18] and (to some extent) social organization and religious law. If we consider the functions of *ṛta* in the Ṛgveda, we can see that, first of all, it regulates and arranges everything and serves as a foundation of the universe. It is said that Mitra and Varuṇa "govern the universe by means of *ṛta"* (RV 5.63.7); the gods are "strong through ṛta" (RV 1.106.3 and in many other verses); rivers flow to Varuṇa's *ṛta* (RV 2.28.4); Varuṇa, the guardian of *ṛta*, "spread out the threefold earth by means of *ṛta*" (RV 4.42.4); Uṣas, the dawn, appears in the sky every day because she "doesn't violate *ṛta*" (RV 1.123.9), she "follows the path of *ṛta*" (RV 1.124.3); by *ṛta* "the nourishments are moved far", for *ṛta* "the earth [and heaven] are wide and deep" (RV 4.23.9-10); "Agni… stretched heaven and earth by means of *ṛta*" (RV 5.1.7), and so on. By *ṛta* the rivers flow, the sun goes up and down, the wind blows; by means of *ṛta* it rains; generations of people are born and destroyed; *ṛta* is the means to overcome obstacles and achieve success; the day follows the night and one year follows another one, etc. To sum up, the universe functions because of *ṛta*. It establishes the order of events and governs the space-time continuum.

It is germane at this point to note that the word *ṛta* etymologically stems from the Indo–European root **-ar* that has a wide range of meanings, including "to connect, to build, to order"; "to fit together".[19] From this root originates a plethora of words used (usually, in their secondary meaning) to denote various aspects of time in different Indo-European languages, for example, Avestan *ratu*—"a period of time", Latin *articulus*—"moment" and *artus*—"narrow in space or time", Old Greek *ἄρτι* and Armenian *ard*—both meaning "just now, currently", Greek *ἀμαρτήας*—"simultaneous",[20] as well as another Vedic word related to time—*ṛtu* that in the later tradition means mostly a season but in the Ṛgveda it is rather "proper time/proper moment" or, specifically, a "[proper] time of a sacrifice". Significantly, from the same root stems the Sanskrit word *ara*—a spoke of a wheel or of an altar formed like a wheel, in the epics—a spoke of the time-wheel, and one of the divisions of time in the Jaina tradition. All the aforementioned words have meanings associated with fitting something together (i.e., *articulus* and *artus* mean "a joint, connection", etc.) and some of them have connotations related to rightness (i.e., Greek *ἄρτι*). From these parallels we can conclude that a relationship of Vedic *ṛta* with time and the wheel of time (the wheel of year in this case) is not accidental. Although we cannot say that *ṛta is* time, it is definitely on some older Indo–European level related to time both etymologically and functionally. It organizes the universe, creates a sequence of events, and exerts its power over all living beings. As we will see, in the later tradition these are the functions of time. Also, just like the year in the discussed creation hymn, *ṛta* is very much connected to the Cosmic Waters.[21] Both the Year and Ṛta, the Cosmic Order, represent time-eternity that regulates and arranges the universe in the Ṛgveda.

---

[18] This meaning was proposed by Lüders; see [5], II: 402–406.

[19] See Watkins [19]: 5.

[20] See Pokorny [8]: 174–178.

[21] See Lüders [5], II: 402–406.

Apart from the time-eternity, there are several terms in the Ṛgveda that denote the "concrete" transitory time and its units. The Ṛgveda contains multiple references to the progress of time: again, the year (*saṃvatsara*, *vatsara*, etc.), the seasons (*śarad, varṣa, hemanta, vasanta,* etc.), month (*mās*), day, night, lunar asterisms (*nakṣatra*) that are used to calculate the lunar months of 27/28 days. People obviously were very well aware of the passing of time with designations for dawn, dusk, days and ominous nights (those that are "weaving their cloth alternately"); they were even aware of (undetermined) previous and future "ages", the *yuga*s (RV 7.86), and of personal time, including stages of life, especially aging: one is wished to live "a 100 years," and indeed, *one* person is mentioned as having lived 116 years. Personal time is connected to the "objective" time: for example, it is said that Uṣas, the dawn, whose daily appearance in the sky signifies the passage of time, "diminishes human generations"[22] and "shortens the lifetime of the mortals".[23]

Accurate observations of the passage of time over the year were necessary to carry out rituals properly: the technical term for a priest was *ṛtv-ij* "the one who offers according to season/at a proper moment"—a derivation from the above-mentioned *ṛtu*. The term *ṛtu* is the closest that the Ṛgveda has to the general definition of transitory time (interestingly, Sāyaṇa, the commentator of the Ṛgveda, in his commentary explains *ṛtu* by using the word *kāla*, time[24]): it can mean a [proper] moment, or a proper time, a season, a time fit for the sacrifice, and so on. *Ṛta* and *ṛtu* being cognates that stem, as explained above, from the Indo-European verbal root **-ar*, make an interesting pair: from one of them (*ṛta*) probably originates the idea of time-eternity, while from the second one (*ṛtu*) likely does the concept of transitory time whose passage can be observed and calculated.

One of the earliest in the Indian tradition mythological descriptions of creation of time-space continuum (although, of course, it is not conceptualized in these terms in the text) is also found in the Ṛgveda, in the famous Puruṣasūkta hymn (RV 10.90) of the Xth maṇḍala. Cosmic giant Puruṣa is sacrificed to himself and/or to the gods (who are, however, created as a result of this sacrifice), and different elements of the universe are produced out of his body. Puruṣa contains all space (mythological equivalent of space in this hymn is earth) and *even more than that:* "Having covered the earth from all sides, he extended beyond [it] by ten fingers"[25]; "he surpassed the earth from front and from behind".[26] Consequently, the whole universe—all parts of space and all beings in that space—is created from his sacrificed body: the sun and the moon; the wind and the cardinal points; heaven, earth, and interspace; the gods (Indra, Agni, etc.) and living beings. Vedic hymns, poetic meters, sacrificial formulas, and the four social classes (*varṇa*) are also born from

---

[22] *praminatī manuṣyā yugāni* (RV 1.124.2).

[23] RV 1.92.10; 1.179.1.

[24] Sāyaṇa [10], Vol. 1: 962, 969, etc.

[25] RV 10.90.1c-d: *sa bhūmiṃ viśvato vṛtvāty atiṣṭhad daśāṅgulam* ||.

[26] RV 10.90.5c-d: *sa jāto aty aricyata paścād bhūmim atho puraḥ* ||.

Puruṣa's body. Not only Puruṣa covers all space, but also he encompasses the two types of time:

> Puruṣa alone is all this – what has been and what is to be. And he is lord of immortality, because he overgrows [the universe] by means of food.[27]

This stanza demonstrates that Puruṣa contains the past and the future ("what has been and what is to be"), i.e., all changeable/transitory time. He is also called "the lord of immortality", i.e., he belongs to the realm of eternity and transcends passing time of this world "by means of food". This expression is a way to show that as an "eater" of the universe Puruṣa dominates it. Food in this case can be understood as sacrificial food that feeds the ancestors: Puruṣa is the ultimate ancestor who never runs out of sacrificial food, because his food is the whole world and he himself is the whole world, thus he is immortal. Puruṣa is the first ancestor, and all future generations will sacrifice and be sacrificed to him always while this world exists. Interestingly, the past and the future in the stanza are clearly contrasted with immortality: they are mentioned together as items that are related, however, it is emphasized that they are different and belong to separate realms. In other words, although there is no clear "philosophical" statement of a concept of two types of time (transitory time and time-eternity), here we observe its formation by mythological and poetic means.

Next stanza states that Puruṣa is not only the lord, but also a receptacle of immortality:

> One quarter of him is all beings. Three quarters of him – immortality in heaven.[28]

Later in the hymn we find additional temporal terms (time-divisions) that play a role in the sacrifice of Puruṣa. The seasons are elements of sacrifice (sacrificial butter, wood and oblation) and, because Puruṣa contains everything, they are also parts of Puruṣa:

> When the gods put forth the sacrifice with Puruṣa as the offering, spring was its melted butter, summer was the firewood, autumn was the oblation.[29]

Thus, time (in the form of the seasons) is seen as a means of Puruṣa's sacrifice and, consequently, a means in the act of creation, a power of creation, so to speak.

To sum up, in the form of mythological motifs and poetic expressions, this hymn introduces the following principal ideas concerning time and space:

- Time and space are interconnected, there is time-space continuum that is embodied by the Cosmic Giant and created from his sacrificed body;

---

[27] RV 10.90.2: *puruṣa evedaṃ sarvaṃ yad bhūtaṃ yac ca bhavyam| utāmṛtatvasyeśāno yad annenātirohati* ||.

[28] RV 10.90.3c: *pādo 'sya viśvā bhūtāni tripād asyāmṛtaṃ divi* ||.

[29] RV 10.90.6: *yat puruṣeṇa haviṣā devā yajñam atanvata* | *vasanto asyāsīd ājyaṃ grīṣma idhmaḥ śarad dhaviḥ* ||.

- In accordance with what was shown in our other examples, there are indications that an idea of two types of time is being developed. Those two are time-eternity, related to immortality, and transitory time manifested as various time-units and states-in-time (i.e., past, present, and future);
- Time is necessary for creation. It is possible that time is a means of creation, a power that creates.

Overall, we can see from the materials of the Ṛgveda that the text has no designation for a unifying notion of time. However, it can be concluded from the analysis of the hymns that the Ṛgvedic people were aware of “eternity” and “temporality” and envisioned the two as both connected and distinct. From the study of poetic metaphors and mythological motifs and images related to the “temporal” and the “eternal”, from their position and relationship in the text, it can be concluded that some kind of a notion that could unify “eternity” and “temporality” and relate “temporality” to different units and states was emerging, although probably had not solidified enough to be used in the Ṛgveda to designate explicitly a power or a god of ETERNITY/TEMPORALITY; they construed it as connected with the movement of the sun and as a power that regulates the universe. Later, in the Atharvaveda, when a unified designation of time (*kāla*) is introduced, archaic myths and ideas expressed in the Ṛgveda become associated with this designation, and from that period on an elaborate mythology of time rises that is shaped into various “doctrines of time” during the philosophical period.

3.2 Time-Space Calculations and Ritual Time During the Vedic Period

From the Vedic texts it is clear that the many solemn Vedic rituals are strictly time-bound and liminal: they have to be carried out at daybreak, sunset, at new/full moon, at the three major seasons (every 4 months), and yearly. Some of them follow the yearly course of the sun from its rising point in the southeast in winter (*mahāvrata*) to its rise in the northeast in summer (*viṣuvant*),[30] thus at the time of the two solstices. This apparent movement of the sun along the eastern horizon is likened to a voyage of a ship, with resting points, “islands”, in between.

Temporal and spatial calculations (including that of true east) could be performed by very simple means, such as marking the two solstice points by sticks in the ground and taking half its distance as true east. This is important as the Vedic sacrificial ground, symbolizing sun (square), earth (round) and moon (halfmoon) as the three sacred fires, is oriented towards true east. Sacrificial time and space are thus closely linked.

---

[30] On the path of the sun see, for example, RV 5.81.4, cf. 1.115.3, AB 3.44.

Close observations of both sun and moon were necessary to calculate when various rituals are to be performed throughout the year. The solar year from one winter solstice to the next, of c. 360 + 5 days with its 12 × 30-day months, does not correspond to the lunar year of 336 days (28 × 12), with its lunar months observed from one full moon to the next. A complex system of intercalary months was soon created, that led to a system of 5-year cycles (*samvatsra, idvatsara, parivatsara, anuvatsara,* and *vatsara*), first systematically established in the Jyotiṣa (date unclear, as it is composed in post-Vedic language, although it is often claimed to date from around 1300 BCE). Post-Ṛgvedic texts contain standard sentences such as "the year has 12 months" (of 30 days each) but also "…13 months" (that is the short "month" of 5 days at the end of the 12 months in the solar year of 365 days.) The concurrent lunar year is divided into months, each of which has 27/28 named Nakṣatra constellations.

The human life cycle is equally bound to ritual time. A boy is initiated in his 7th year, then studies for 12 years before getting married. Then, he has to perform, along with his wife, a simplified set of rituals, notably the New and Full Moon rites, but also those for the spring and fall harvests (of wheat, rice), others for pregnancy and the birth of a (preferably) male child, followed by a fixed set of rituals until initiation.

Most prominent is a long series of death and anniversary rituals (*śrāddha*); they begin with the rituals of the first 12 days after death and extend to annual ones for three generations of male and female ancestors. Individual ancestors are no longer remembered as individuals, though certain famous persons are well recalled and immortalized in verses. Reference to older and more recent deities, kings, ancestors, and teachers is often made.

Clearly—different from what is often alleged—a sense of (distant) times is well established, even though these are not set within a fixed time frame, a (local) era, like our BCE/CE. Such eras were established only late in the first millennium BCE, probably under Hellenistic influence, when weekdays were imported as well.

3.3 Space Organization in Vedic Cosmography

In the worldview of the oldest (Vedic) texts the earth is flat, an ancient view that is reinforced by the almost featureless North Indian landscape, stretching from Afghanistan to, roughly, Benares. the earth is construed as a flat disk (as easily visible from a hill or mountain) with heaven (vyoman, rocana, sānu, viṣṭap, pṛṣṭha —"ridge") above. Between both there is a wide expanse of the "interspace" (antarikṣa). While more ancient mythology had a binary division of Father Heaven/Mother Earth (dyāus pitar, pṛthivī mātar), this tripartite structure appeared, according to a myth, because heaven and earth had been separated from their primordial union by a god (Indra) or a mountain (like Atlas in Greek mythology) or (an offering) pole, or by the world tree (Plakṣa) that connects the two at zenith near the current North Pole, the 'navel/top' of sky around which the stars revolve.

Replicating the ancient Northern Eurasian pattern favoring a division of the Universe into 9 levels, the Vedic texts often speak of 3 heavens (highest, middle, and lowest),[31] that is where the gods and ancestors reside, further the three layers of the Earth, and the three 'interspaces' (*antarikṣa*) between heaven and earth. The gods are said to have "resorted to heaven and earth, the waters and the sun (*svar*)… and the wide interspace".

Heaven and earth meet at the rim of the horizon: there is a small gap "as thin as a wing of an insect", through which one can go outside, mount the back of the sky, and reach the eternal light beyond it. The sky is thought (as in Iran) to consist of stone, having holes through which the light beyond shines: the stars.[32] A number of them are mentioned in the Ṛgveda, such as the ancient Seven Bears (as in Greece), however generally substituted by the Seven Sages (Ṛṣis). Other early Vedic stars are the Wolf, the Milky Way, and later also the Dolphin. Some of the major constellations are mentioned as well, such as the old Eurasian *Ursa Maior* (Great Bear, Wagon, Dipper) constellation. The pole star (*dhruva*—"the firm one") was at first not yet in its present position, due to precession but it appears in late Vedic texts. Its position in the nighttime sky is taken by the top of the world tree, around which the stars revolve.

The earth is circular, like a wheel (RV 10.89.4). However, one's own territory is always imagined as occupying its center (as the "Middle country", like in China). Though round, the earth is also divided by the 4 quarters (and their additional four intermediate regions).

The most remarkable feature in the flat lands of Northern India is the rivers that separate the tribal areas: the seven rivers of the Panjab, to which later the Ganges & Jamna (Yamunā) were added, along with their intermediate (doab) regions (also called *antarvedi*). Another major feature is the "Snow Mountains" (Himālaya) in the north, and in the post-Ṛgvedic texts also the Indian Ocean (*samudra*) in the east and west of the subcontinent. Mythologically, the Earth is perceived as surrounded by an ocean, and there are also oceans above and below it. In the post-Ṛgvedic texts it is said that clouds rise from the ocean. There also are mythological references to a "northern" or mythical "upper" ocean (AV) as well. Toward the end of the Vedic period, the central Indian mountain ranges (Pariyātra, Vindhya) and some lands south of it came into view.

The Netherworld—not a hell—appears as some undefined space (*nirṛti*) below the "three earths", is thought of as "deepest darkness", where only extraordinary miscreants go after death. Later, the netherworld was thought to be inhabited by Nāgas (and Piśācas) and some of the human ancestors. Each night the sun passes through this dark space to merge again in the east:

---

[31] RV 5.60.6.

[32] JUB 1.25, 4.51, cf. 1.3.1.; see Witzel [20]: 620

> You travel, o Savitar, through the three shining realms, you live with the rays of the Sun, you circumambulate the Night on both sides…[33]

Very rarely the sun is said to have a bright and a dark side: it flips over at sunset and moves back through the sky to the east while showing its dark side; in the morning it flips over again.

In the later Vedic period, exemplified by the Brāhmaṇa texts (before c. 500 BCE), the area of Vedic culture had expanded all over Northern India, up to the borders of Bengal. As mentioned, the traditional center of the Vedic universe (after the RV) was the "Middle Country", roughly extending from Kurukṣetra west of Delhi to Allahabad. It is regarded as the center of the Vedic peoples (Āryāvarta), the "middle direction," or "middle country" (*madhyamā dīś, madhyadeśa*). This core area of the Kuru and Pañcāla tribes was surrounded by other, often smaller tribes: in the west, the Mahāvṛsa, Bahīka, Saindhava, Kamboja, and even the Balhika (Balkh, in N. Afghanistan), and the Kosala, Videha, and Aṅga in the east. The south, even the area north of the Vindhya mountains (Cedi, Śabara, etc.), is hardly mentioned and South India (Āndhra) basically still is out of view. Interestingly, these various tribes or peoples are described as arranged clockwise in a spatial framework and enumerated while following the 4 directions. This is an ancient Indo-Iranian scheme, probably already an Indo-European one.

3.4 Time in the Later Vedic Tradition: Atharvaveda and Upaniṣads

Time *in abstracto* (*kāla*) has been attested since the second oldest Indian text, the Atharvaveda. In AV 19.53–54 we find an elaborate hymn specifically dedicated to time. It says that time is the creator of the universe, the highest God, encompassing everything that exists: "Time created the Earth… in Time are all beings… Time is the lord of everything… Time created the creatures…" (AV 19.53).[34] "This Time goes on as the first god" (AV 19.53.2d).[35] According to these hymns, Time creates the world, puts it in motion, and sustains its existence: "Through Time the sun rises… through Time the wind blows… upon Time the worlds [are founded]" (AV 19.54).[36] Here we can clearly see the parallels with the year—"the Lord of everything that blinks"—from the hymns of the Ṛgveda discussed above. Additionally, time, as described in the Atharvaveda, creates, regulates, and organizes the universe just like *ṛta*, the Cosmic Order, in the Ṛgveda.

---

[33] RV 5.081.04a-c: *uta yāsi savitas trīṇi rocanota sūryasya raśmibhiḥ sam ucyasi* | *uta rātrīm ubhayataḥ parīyasa* …||.

[34] Quoted as translated by M. Bloomfield; see Bloomfield, M., trans. [2].

[35] *kālaḥ sa īyate prathamo nu devaḥ* ||.

[36] Quoted as translated by M. Bloomfield; see Bloomfield, M., trans. [2].

Time in these two hymns is also called a horse that “drives seven wheels”[37] and “carried all beings/worlds”,[38] which definitely corresponds to an image of the wheel of the year (and of *ṛta*) in RV 1.164 that “rolls around heaven” and also carries all beings or worlds (*bhuvanāni viśvā*).

Interestingly, these hymns give us what probably is the earliest direct indication of the unified time’s division into two types. The first one—“fluid”, empirical Time, perhaps of cyclic nature—actively creates and sustains the universe and flows around the worlds (“this [time] did flow around the worlds”—AV 19.53.4b[39]). The second form—Time-eternity, the uncreated Absolute—is shown in AV 19.53.3: “A full jar [likely, of *amṛta,*[40] the drink of immortality] is placed upon Time. We see it existing in many forms. It faces all these worlds (beings). It is called Time in the highest heaven”[41]. Time-eternity here is symbolized by a full jar (of *amṛta*); “eternity”, because *amṛta* is immortality. The words “we see it existing in many forms” signify the created world, i.e. *our* world (the only one we can see), the world of fluid, empirical time that unfolds from eternity, from the Time in the highest heaven. This might be another parallel to RV 1.164, where the wheel of the year is said to be the father of “seven hundred twenty sons” (days and nights of the year) who “has five feet and twelve parts, *the overflowing one in the upper half of heaven*”.[42] Apparently, there is the same mythological motif in these two cases: the one of time-eternity understood as immortality and a drink of immortality (sometimes, like in RV 1.164, associated with the moon: in post-Vedic mythology the moon is considered to be a receptacle of soma, Vedic sacrificial libation, that—also in the later tradition—is interchangeable with *amṛta*) placed somewhere “in the highest heaven”.

Thus, in the Atharvarveda we see further development of the views that were in formation in the Ṛgveda. First, it is the idea of all-powerful time that creates and regulates the universe. Second, a conceptual division of time into two types: one is beyond this world—time-eternity, located in the realm of immortality; another one—transitory, fluid, and cyclic time that “flows around the worlds”. Remarkably, it is the second, “fluid” time that is actively evolved in creation.

Time is also discussed in the Upaniṣads. Mentions are fragmentary and few, and do not represent a consistent and integral system. A summary can be given as follows:

---

[37] AV 19.53.2.

[38] AV 19.53.4a.

[39] *sa eva saṃ bhuvanāni pary ait* |.

[40] In the previous stanza (AV 19.53.2): *sapta cakrān vahati kāla eṣa saptāsya nābhīr* ***amṛtaṃ*** *nv akṣaḥ* | (This Time rides/drives with seven wheels, it has seven naves, amṛta/immortality is its axle).

[41] *pūrṇaḥ kumbho 'dhi kāla āhitas taṃ vai paśyāmo bahudhā nu santam* |.
sa imā viśvā bhuvanāni pratyaṅ kālaṃ tam āhuḥ parame vyoman ||

[42] RV 1.164.12.

1. Time (*kāla*) or the Year (*saṃvatsara*) is considered to be a unifying notion of all measurable, empirical intervals of time (year, day, hours, etc.) and all time-states (states-in-time), i.e. past, present, and future (sometimes mentioned as "three times" or "three divisions of time").[43]
2. The Year (*saṃvatsara*) is said to be the body of the sacrificial horse. Sacrificial horse is described in this passage as an embodiment of the space-time continuum: the whole body (*ātman*) of the sacrificial horse is said to be the year (*saṃvatsara*),[44] whereas particular body parts are identified either with smaller divisions/intervals of time or with segments of space. Its limbs are identified with seasons (*ṛtu*),[45] "its joints are the months and fortnights; its feet are the days and nights",[46] "its forequarter is the rising sun; and its hindquarter is the setting sun".[47] Its other body parts are identified with the sun (eye), the dawn (head), the sky (back), the intermediate region (abdomen), the earth (underbelly), the primary and intermediate space quarters (flanks and ribs), etc. Grammatically it is interesting that the horse in the passage is described without verbs and without any markers of time ("then" and such), the passage uses a verbless series of equivalences between parts of the horse and parts of the universe. The horse is described timelessly, it is eternal. When it comes to describing behavior and activities of the horse, verbs do appear, but the horse's behavior and natural processes are juxtaposed without any relation to "external" time, from a point of view of "internal" or relative time, i.e., in relation to the horse. This relation is emphasized syntactically by the usage of the relative–correlative clauses "when-then"/"this-that" (*yad-tad*): when the horse does A, then B happens in the world. All verbs are in the present tense, i.e., it is the "timeless present":

   > When it yawns, lightning flashes; when it shakes itself, it thunders; and when it urinates, it rains. Its neighing is speech itself.[48]

   Therefore, the sacrificial horse, described here as eternal and timeless, constitutes the whole universe, the realms of space and time, and its physical activities are understood as natural phenomena, specifically, the thunderstorm, one of most mythologized natural events in the Indo–European and Indian cultures.
   Identification of the body of the sacrificial horse with a year has its roots in the conduct and duration of the *aśvamedha* (horse sacrifice) ritual: the whole ritual

---

[43] For the Year (Time) consisting of seasons, periods, and states-in-time see, for example, MU 6.14, MU 6.33, BĀU 3.8.4, BĀU 3.8.6, etc. See also Māṇḍūkya Upaniṣad 1.1, Śvetāśvatara Upaniṣad 6.5, etc. on the three times or the threefold division of time.

[44] *saṃvatsara ātmāśvasya medhyasya* (BĀU 1.1.1).

[45] *ṛtavo 'ṅgāni* (ibid).

[46] *māsāś cārdhamāsāś ca parvāṇy ahorātrāṇi pratiṣṭhā* (ibid); translated as in Olivelle [6]: 37.

[47] (Ibid); translated as in Olivelle [6]: 37.

[48] ***yad*** *vijṛmbhate* ***tad*** *vidyotate* | ***yad*** *vidhūnute* ***tat*** *stanayati* | ***yan*** *mehati* ***tad*** *varṣati* | *vāg evāsya vāk* || BĀU 1.1.1 || Translation is quoted as in Olivelle [6]: 37.

takes a year during which a sacrificial horse after preparatory rites is allowed to roam freely (escorted by an army that guards the horse against possible attacks of neighbors and other incidents). While the horse wanders, “daily offerings were made to Savitar, the horse being associated or identified with the sun and the solar year.”[49] At the end of the year the horse returns at the beginning of the king’s consecration (*aśvamedha* being a part of the king’s consecration rites) to be sacrificed with elaborate rituals that included the soma pressings, etc. Another significant detail is that during the ritual the sacrificial horse is tied by a twelve cubits (*aratni*) long *muñja* grass rope that represents a year (12 months). In sum, the considered passage from BĀU is a continuation of an idea that first appears in the Puruṣasūkta: time-space continuum is contained within and created out of the body of the sacrificed.

3. Bṛhadāraṇyaka-Upaniṣad (BĀU 3.8.3-8) states that the time-space continuum of the universe (“the things above the sky, the things below the earth, and the things between the earth and the sky, as well as all those things people here refer to as past, present and future”—BĀU 3.8.3[50]) is “woven back and forth” on *ākāśa*, the infinite space (BĀU 3.8.4[51]); and *ākāśa* in turn is “woven” on the Imperishable Brahman (BĀU 3.8.8).[52] This idea might constitute the origin of a later view developed by philosophers of the Sāṃkhya school that time is produced by means of infinite space/ether (*ākāśa*) out of the ultimate reality (in the case of the Sāṃkhya—from Prakṛti) and somehow is “part” of *ākāśa*.
4. It is said in the Bṛhadāraṇyaka-Upaniṣad 1.2.4 that Time as the Year (*saṃvatsara*) is the product of an intercourse between the ultimate being (functioning as Death and Hunger) and Speech. It was also the ultimate being’s sperm, which was transformed into the Year: “He desired: ‘may my second atman be born’. By means of [his] mind he, hunger [and] death, copulated with Speech. What was semen became the Year. For there had been no year before.”[53] Schayer notes: “Time created as a child of Death is itself Death.”[54]
5. Time is said to be so thin that it cannot be observed directly, but can be measured by the movement of the Sun: “because of thinness, a measure is this, by this [i.e., the movement of the Sun] Time is measured” (Maitrī-Upaniṣad, 6.14[55]).
6. The idea of Time being equal to the Absolute, the eternal Brahman, is expressed in Maitrī-Upaniṣad, 6.15–16: “Indeed, there are two forms of Brahman: Time and Non-Time. Then, the one that is before the Sun is non-Time with no parts. And the

---

[49] Stutley [12]: 257.

[50] quoted as translated by P. Olivelle, see Olivelle [6]: 44.

[51] Ibid, p. 44.

[52] Ibid, p. 45.

[53] *So 'kāmayata dvitīyo ma ātmā jāyeteti |sa manasā vācaṃ mithunaṃ samabhavad aśanāyā mṛtyuḥ | tad yad reta āsīt sa saṃvatsaro 'bhavat | na ha purā tataḥ saṃvatsara āsa |.*

[54] See Schayer [11], p. 6.

[55] *saukṣmyatvādetatpramāṇamanenaiva pramīyate hi kālo.*

one from the Sun is Time with parts. The year is a form of the one with parts. These creatures are born from the year".[56] The two forms of Brahman—embodied or formed (*mūrta*) and non-embodied or formless (*amūrta*)—are discussed in BĀU 2.3.1, MU 6.3, etc. Significantly, in MU 6.14 two forms of Time—embodied/formed and non-embodied/formless (*kālo mūrtiramūrtimān*)—are also mentioned. Thus, as there are two forms of Brahman, there are two corresponding Times: the absolute non-embodied Time (Eternity-Time, Time which is "Non-Time", without parts) and revealed, embodied Time ("concrete" Time of the phenomenal world, "fluid" and divided into parts—years, days, minutes, etc.). The Eternity-Time, which parallels the "before-the-Sun", "Non-time" Brahman, clearly corresponds to the Atharvaveda's "Time in the highest sky".

7. Additionally, in Maitrī-Upaniṣad there are speculations about transient and transitory nature of everything, which corresponds to one of the major tenets of Buddhism.

3.5 Time and Space in the Epics

In addition to the above ideas developed in the Upanishads, there existed an ancient doctrine specifically dedicated to Time. It is mentioned both in the epics and in the oldest Buddhist and Jain literature as Kālavāda, i.e., "The Doctrine of Time". Because it is referred to quite often, we can say that Kālavāda was widely known, and the existence of handbooks dedicated to Kālavāda has been reported.[57] Unfortunately, information that has survived is very fragmentary and dispersed throughout texts of different traditions and periods. Thus, we can only reconstruct the system. According to fragments found in the epics (first of all, the Mahābhārata) and epico-philosophical texts (such as Yogavāsiṣṭha), time is considered the primary cause of creation, existence and destruction of all beings; it is the absolute beginning of everything; it is something eternal, all-pervading and permanent:

> …everything is born from Time and then is taken away by it (MBh 13.1.49[58])
>
> …having pervaded everything, [Time] abides (YV 1.23.7[59])
>
> Time, having devoured all [the universe] with no remainder, is the essence (*ātman*) of everything (YV 1.23.6[60])

---

[56] *dve vāva brahmaṇo rūpe kālaścākālaścātha yaḥ prāgādityāt so 'kālo 'kalo 'tha ya ādityādyaḥ sa kālaḥ sakalaḥ sakalasya vā etadrūpaṃ yat saṃvatsaraḥ saṃvatsarāt khalvevemāḥ prajāḥ prajāyante* |.

[57] See [13], vol. 2: 45.

[58] quoted as in Vassilkov [16]: 20.

[59] … *sarvamākramya tiṣṭhati* ||.

[60] *kālaḥ kavalitānantaviśvo viśvātmatāṃ gataḥ* ||.

> Like the wind turning blades of grass in all directions, all beings are in the power of Time… (MBh 11.2.7[61])

> Cold, heat and rainy season turn around by [the will of] Time, as well as happiness and suffering in [the world of] men [do]. (MBh 12.28.34[62])

According to the MBh, time controls death and old age. It is compared with the great ocean that hides these two monsters. It also "cooks" or "ripens" living beings and devours everything—all worlds and all creatures:

> Nobody knows that this world sinks into the Ocean of Time where two monsters are (hiding): Old Age and Death. (MBh 12.28.43[63])

> Time cooks beings, Time takes away every living thing, and Time itself that bums living beings is, in its turn, tamed by Time. (MBh 1.1.188[64])

As the ultimate cause of everything in the world, "time establishes everything and time cooks it".[65]

The epics (especially the MBh) describe time, often compared to Yama, the god of death, as a force demolishing the universe at the end of a world period (*kalpa*) and as a devourer of things:

It takes away, destroys … devours and kills (YV 1.23.11[66])

Additionally, some texts mention empirical time, which, quite contrary to the permanent time discussed above, is changeable and consists of intervals. In texts related to Kālavāda, time can paradoxically possess the opposite qualities (indivisibility and being divided into intervals, permanence and impermanence, etc.) simultaneously, for example:

> Isn't [Time] the most cunning deceiver, o sage? It is divided though indivisible, is being burnt though never burns out, is seen though invisible (YV 1.23.16[67])

One of the most important and typical images of Kālavāda is that of the Wheel of Time (*kāla-cakra*). It is described in different parts of the MBh as well as in various other texts. In MBh 1.3 in the story of Uttaṅka, sage Gautama's a disciple who descends to the realm of the *nagas* (serpents) in the underworld, the wheel of time is depicted as a part of a loom where two deities—Dhātṛ and Vidhātṛ—spin threads of human destiny:

---

[61] quoted as in Vassilkov [16]: 21.

[62] Ibid, p. 19.

[63] Quoted as in Vassilkov [16]: 21.

[64] Ibid, p.22.

[65] *kālaḥ sthāpayate sarvaṃ kālaḥ pacati vai tathā* | (MBh 12.217.39 cd).

[66] *haratyayaṃ nāśayati karotyatti nihanti ca* |.

[67] *bhidyate nāvabhgno 'pi dagdho 'pi hi na dahyate* | *dṛśyate nāpi dṛśyo 'pi dhūrtacūḍāmaṇirmune* ||.

> Then he saw two women at a loom weaving cloth having placed [it on a loom]. And threads in that loom were black and white. And also he saw a wheel turned by six boys and saw a handsome-looking man.[68]

> Three hundred and sixty [spokes] are fixed in the middle – in this eternal ever-turning wheel furnished with 24 parts (*parvan*s) that is turned by six boys. And two young women [are] weaving [on] this loom always turning around the threads, spinning white and black [threads], [and] worlds and beings perpetually.[69]

> Those two women are Dhātṛ and Vidhātṛ. And those black and white threads are nights and days. The twelve spokes are the 12 months. As for the twelve spoke wheel turned by six boys, those are the six seasons (*ṛtu*) [and] the wheel is a year…[70]

Although certain details and numbers differ, this wheel is strikingly similar to the wheel of the year (and of *ṛta*) in RV 1.164 as well as to the wheel(s) of time in the Atharvaveda and constitutes a development of the same idea.

During the epic period an elaborate system of world ages and cycles of creation and destruction of the universe, the so-called system of *yugas*, took the finite shape. This system is connected to the major tenets of Hinduism—the doctrine of karma, dharma, and transmigration. Sometimes it is called the doctrine of cyclical time. In this framework, there are three major units of cyclical time: the *kalpas*, the *manvantaras*, and the *yugas*. *Kalpa* is a cycle of creation and destruction of the universe, also called a day of Brahmā (the creator god). When Brahmā wakes up from his sleep, the world is created; when he falls asleep it is destroyed. *Manvantara* ("a period of Manu") is a period when a particular progenitor of humans, Manu, rules. Each *manvantara* a new Manu is born. A new Indra, king of the gods, is also born in each *manvantara.* The *yugas* are associated with the increase and decrease of dharma (moral law and rules of conduct for the four social classes). There are 4 *yugas*: Kṛta, Tretā, Dvāpara, and Kali. Each of them is related to a particular throw of a dice: Kṛta is the best throw, and Kali is the worst. Each *yuga* dharma diminishes by one forth. Consequently, Kṛta is considered the best age "when dharma stands on four legs", and Kali is the worst "when dharma stands on one leg" (similarly to ages in Greek mythology—from the golden age to the iron one). At the end of Kali-yuga dharma is at its lowest point, therefore people become immoral, neglect their religious and social duties, etc. After Kali-yuga, Kṛta-yuga starts again. Between the *yugas* there are transitional periods called the *saṃdhis*. A cycle of the 4 *yugas* (including all the *saṃdhi* periods) constitutes a *mahāyuga*. Each *mahāyuga*

---

[68] … *athāpaśyat striyau tantre adhiropya paṭaṃ vayantyau* || *tasmiṃś ca tantre kṛṣṇāḥ sitāś ca tantavaḥ* | *cakraṃ cāpaśyat ṣaḍbhiḥ kumāraiḥ parivartyamānam* || *puruṣaṃ cāpaśyad darśanīyam* | *sa tān sarvāṃs tuṣṭāva ebhir mantravādaślokaiḥ* || (MBh 1.3.147–149).

[69] *trīṇy arpitāny atra śatāni madhye ṣaṣṭiś ca nityaṃ carati dhruve 'smin* | *cakre caturviṃśatiparvayoge ṣaḍ yat kumārāḥ parivartayanti* || *tantraṃ cedaṃ viśvarūpaṃ yuvatyau vayatas tantūn satataṃ vartayantyau* | *kṛṣṇān sitāṃś caiva vivartayantyau bhūtāny ajasraṃ bhuvanāni caiva* || (MBh 1.3.150–151).

[70] *ye te striyau dhātā vidhātā ca* | *ye ca te kṛṣṇāḥ sitāś ca tantavas te rātryahanī* || *dvādaśāraṃ dvādaśa māsāḥ* | *yad api tac cakraṃ dvādaśāraṃ ṣaṭ kumārāḥ parivartayanti te ṛtavaḥ ṣaṭ saṃvatsaraś cakram* | *yaḥ puruṣaḥ sa parjanyaḥ* | (MBh 1.3.172b–173b).

lasts 12 thousand “years of the gods”, one year of the gods being 360 human years. A *kalpa* (one day of Brahmā) lasts 100 *mahāyugas*, i.e., 12 million “years of the gods” or 4,320,000 human years. From these calculations we can see that this system is an elaboration of old Vedic model where the year represents time-eternity and a cycle of time.

The organization of space (cosmography) in the epics is also different from that in the Vedas. Generally, cosmological system and space organization as described in the epics is very close to the accounts found in the Purāṇas. In the Epic-Purāṇic mythology, a continent or an island called Jambūdvīpa (“Rose Apple continent”) in the Purāṇas and Sudarśana (“Beautiful”) in the Mahābhārata is situated in the center of the flat earth. This continent (or an island) is envisioned as circular or having the form of four-petaled lotus, the four continents spreading in each direction. In the Purāṇas the Southern continent is called Bharata, whereas in the Bhīṣma Parvan of the Mahābhārata we find two names for it—Bharata and alternatively Jambūdvīpa. The continent is a disk whose diameter according to the Purāṇas is 100,000 *yojanas*[71] (1000 yojanas in the MBh). In the center, in the seed cup of the lotus, there stands Mount Meru (or Sumeru) whose height according to both MBh and the Purāṇas is 84,000 *yojanas*. High above the Mount Meru there is *Dhruva,* the Pole Star; celestial bodies move around it in circles. Gaṅgā (the Ganges river) falls on Meru and runs in four directions. On the top on Mount Meru there is a city of Brahmā in the center, surrounded by the eight cities of the Lokapālas (guardians of the cardinal directions)—each of the cities marking one of the main or intermediate cardinal directions. The continent of Jambūdvīpa (or Sudarśana) is surrounded by a series of seven concentric oceans (first the salt ocean, then oceans of sugar-cane juice, wine, clarified butter, milk, whey, and fresh water) and seven “continents” (*dvīpa*), each larger than the previous one. At the edge of the world (*samā* in the MBh) there is a great mountain range *Lokāloka* that separates *loka* (world) from *aloka* (un-world).[72] The Mahābhārata describes four world elephants that dwell on the world-edge. Each of the elephants marks one of the cardinal points.

Different peoples that inhabit the world and their arrangement in the epic landscape are also described in the MBh. By the time when the Sanskrit epics were composed by the wandering bards of Northern India, the geographical horizon had expanded as to include almost the entire subcontinent. Therefore the epics mention its actual geographical features and various populations, even the distant Hūṇa (Huns) and Cīna (Chinese).

Vertically, space is organized into 3 worlds—heaven (the upper world), earth, and underworld. In different sources it is mentioned that heaven and the netherworld are divided into seven layers. However, an interesting feature of the epic

---

[71] The exact length of a *yojana* differs in various sources, however, generally, it is a distance traversed in one harnessing or without unyoking.

[72] Kirfel [4]: 121.

worldview is that the vertical axis of the universe is much less important and less described than the horizontal space organization.[73] Heroes who are at the center of attention in the epics mostly act in the world of men (although sometimes they ascend to heaven or descend to the netherworld), therefore the world of men and the horizontal domain of space is of the foremost importance. For the epics, the principal place is home of the heroes, therefore it is considered to be the "central country" (*madhyadeśa*). The farther a place from the central country, the more dangerous and "demonic" it is depicted in the epics.[74]

To conclude, in the epic texts we find a developed cosmography that has its roots in the Vedic and Indo-European mythology and a very elaborate doctrine of time. Kālavāda, the doctrine of time found in the epics, absorbed archaic myths, poetic metaphors, and pre-philosophical and religious ideas expressed in the Ṛgveda, Atharvaveda, and the Upaniṣads and supplemented them with other—newer, or maybe even more archaic—myths and ideas. Later, many of the concepts and even images of Kālavāda found their way to different philosophical systems of Classical India. This paper is an attempt to better understand the origins of the notion of time and structure of space in Ancient India and show the development of the concepts on all stages from the Ṛgveda to the epics.

References

1. Āgāśe, K. (Ed.). (1896). *The Bṛhad Āranyaka Upaniṣad. The Brihadāraṇyakopanishanmitāksharā by Nityānanda Muni*. Āpte: Poona [BĀU].
2. Bloomfield, M. (trans.) (1897). *Hymns of Atharva-Veda*. Oxford.
3. Cowell, E. B. (Ed.). (1935). Upanishads. Maitrayaniyopanisad. English & Sanskrit. The Maitrī or Maitrāyaṇīya Upanishad/with commentary of Rāmatīrtha; edited with an English translation by E. B. Cowell. Re-issue. Calcutta, Asiatic Society of Bengal [MU].
4. Kirfel, W. (1920). *Die Kosmographie Der Inder*. Nach Den Quellen Dargestellt: Kurt Schroeder, Bonn & Leipzig.
5. Lüders, H. (1959). *Varuṇa* (vol. I and II). Göttingen. Vandenhoeck & Ruprecht.
6. Olivelle, P. (trans.). (1996). Upaniṣads. Oxford, New York: Oxford University Press.
7. Pansikar, V. L. S. (Ed.). (1984). *Yogavāsiṣṭhaḥ. The Yogavāsiṣṭha of Vālmiki with the commentary*. Delhi: Motilal Banarsidass [YV].
8. Pokorny, J. (2007). Proto-Indo-European etymological dictionary. A Revised Edition of Julius Pokorny's Indogermanisches Etymologisches Wörterbuch. Indo-European Language Association—http://dnghu.org/.
9. Roth, R., & Whitney, W. D. (Eds.). (1856). *Vedas. Atharvaveda. Atharva veda sanhita/herausgegeben von R. Roth und W.D. Whitney*. Berlin: Ferd. Dümmler [AV].
10. Sāyaṇa. (1935–1959). *Ṛgveda-Samhitā with commentary of Sāyaṇāchārya* (Vols. 1–5). Poona. Vaidika Samohodhana Mandala.
11. Schayer, S. (1938). *Contributions to the problem of time in Indian philosophy*. Krakow.
12. Stutley, M. (1969). The Aśvamedha or Indian Horse Sacrifice. In *Folklore* (Vol. 80, No. 4, pp. 253–261). Taylor & Francis, Ltd.

---

[73] Vassilkov [17]: 12.

[74] Details in Vassilkov [17] and Vassilkov [18]

13. Stcherbatsky, Th. (1995). *Teoriya poznaniya i logika po ucheniyu pozdnejshih buddistov*. Saint Petersburg: Asta-Press (in Russian).
14. The Mahābhārata for the first time critically edited by Vishnu S. Sukthankar [and others]. Poona, Bhandarkar Oriental Research Institute, 1933 [MBh].
15. van Nooten, B. A., & Holland, G. B. (Eds.). (1994). *Vedas. Ṛgveda. Rig Veda: a metrically restored text with an introduction and notes*. Cambridge, Mass: Dept. of Sanskrit and Indian Studies, Distributed by Harvard University Press, Harvard University [RV].
16. Vassilkov, Ya. V. (1999). Kālavāda (the doctrine of Cyclical Time) in the Mahābhārata and the concept of Heroic Didactics. In *Composing a tradition. Concepts, Technics, and Relationships. Proceedings of the First Dubrovnik International Conference on the Sanskrit Epics and Purānas*. Zagreb.
17. Vassilkov, Ya. V. (2006). Kul'turnoe prostranstvo eposa: spetsifika "Mahabharaty". In *Radlovskie chteniya, 2006: Tezisy dokladov* (pp. 11–13). Saint Petersburg (in Russian).
18. Vassilkov, Ya. V. (2008). Ob odnom sposobe organizatsii sakral'nogo prostranstva v indijskoj praktike palomnichestva. In *Radlovskij Sbornik: nauchnye issledovaniya i muzejnye proekty MAE RAN v 2007g.* Saint Petersburg, pp. 126–133 (in Russian).
19. Watkins, C. (Ed.). (2000). *The American heritage dictionary of Indo-European roots*. Boston: Houghton Mifflin.
20. Witzel, M. (2004). The Ṛgvedic Religious System and its Central Asian and Hindukush Antecedents. In A. Griffiths & J. E. M. Houben (Eds.), *The Vedas: Texts, language and ritual* (pp. 581–636). Forsten: Groningen.

Chapter 4
Śrīharṣa on the Indefinability of Time

Jonathan Duquette and Krishnamurti Ramasubramanian

Abstract The conception of time as an absolute, eternal and imperishable entity is commonplace in several religious and philosophical systems. In the context of classical Indian philosophy, this position was advocated by the Nyāya school of logic and epistemology. This article presents an outline of the critique of the Nyāya concept of time put forward by Śrīharṣa, a 12th-century scholar in the Advaita Vedānta school of philosophical theology. In his philosophical treatise, the *Khaṇḍanakhaṇḍakhādya*, Śrīharṣa dismantles the Nyāya position based on a critical examination of its definition of causality and time-forms. The dismissal of the ontological reality of time is also discussed with reference to the works of two later Advaitins, namely Citsukha and Madhusūdana Sarasvatī.

Abbreviations

KKK	*Khaṇḍanakhaṇḍakhādya*
TP	*Tattvapradīpikā*

4.1 Introduction

Since Antiquity, thinkers of all civilizations, cultures and scholarly backgrounds have pondered on questions such as: What is time? How does it relate to the physical world? It is real or merely a human construct? Is it an object of knowledge, and if so, how do we apprehend it? What purpose does it serve in the scheme of

---

J. Duquette (✉)
Newton International Fellow, University of Oxford, Oxford, UK
e-mail: jonathan.duquette@orinst.ox.ac.uk

K. Ramasubramanian
Cell for Indian Science and Technology in Sanskrit,
Department of Humanities and Social Sciences, IIT Bombay, Mumbai, India
e-mail: ram@hss.iitb.ac.in

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_4

reality? Various answers have been given to such questions in different areas of knowledge ranging from cosmology, physics and mathematics to philosophy, psychology and theology. The Indian tradition offers a wide spectrum of well-formulated views about time (*kāla*[1]). Brahmanical, Buddhist and Jain schools all endeavoured to understand the multifaceted nature of time in its metaphysical, logical and epistemological aspects, and for long debated with each other over such issues. In this article, we will focus on how one school of thought—the Advaita Vedānta school of philosophical theology, whose views lean towards metaphysical idealism—refutes the concept of time advocated by another school, the Nyāya school of logic and epistemology, which defends a version of realistic pluralism. Our aim is to provide a first systematic outline of the incisive critique of the Nyāya concept of time put forward by the 12th-century scholar of Advaita Vedānta, Śrīharṣa, in his celebrated treatise the *Khaṇḍanakhaṇḍakhādya* ("Edible Pieces of Refutation," hereafter KKK). We believe that an examination of Śrīharṣa's critique is not only of historical and cultural interest, but also philosophically significant insofar as it stands as one of the most sophisticated attempts in the history of Indian philosophy to dismiss the ontological reality of time.

In the scholarly tradition known as Advaita (non-dualist) Vedānta, which flourished most richly after Śaṅkara, a philosopher and theologian who lived around the 8th century CE, philosophical inquiry starts from the metaphysical premise that Brahman—the immanent and transcendent "absolute" revealed in the *Upaniṣad*s, sacred texts of the brahmanical tradition—is the sole reality of everything. In this view, the individual self is identified with the transcendental self, Brahman, which is pure consciousness; as to the world in which dwells the self, it is considered an apparent manifestation of that unchanging Brahman. In other words, ontological plurality does not exist in a fundamental sense. It is experienced because of error in judgements and mistaken cognitions, and vanishes in the wake of the knowledge of the true nature of the self. Thus the Advaita tradition stands in stark contrast with the Nyāya tradition as the latter defends a robust realism including a plurality of selves, universals and substances. Historically, their metaphysical divergences have led these two schools to engage one another on several philosophical issues; the ontological nature of time is just one among them.

It is with this dialectical background in mind that one has to approach Śrīharṣa's unique contributions to Indian philosophy. One of the greatest scholars and vitally original thinkers of Advaita Vedānta, Śrīharṣa stands out in the history of Indian philosophy because of his uncompromising critique of the logical positions defended by Naiyāyikas (the philosophers of the Nyāya school). In his polemical treatise, the KKK—the title indicates that the main purpose of the work is a refutation (*khaṇḍana*)—Śrīharṣa aims to show that the definitions (*lakṣaṇa*) and means of knowledge (*pramāṇa*) with the help of which the Nyāya school constructs its own ontological categories of existence (*padārtha*), are defective. By showing that all concepts and categories of human reason are indefinable, Śrīharṣa is able to

---

[1] All technical terms used in this article are in Sanskrit.

dismantle the metaphysical realism of Nyāya and thereby urge that an Advaita philosophy based on non-distinctness (*abheda*) is alone coherent and logically defendable.[2] Before examining in detail the crux of Śrīharṣa's critique of time (Sect. 4.3), we first provide a brief summary of Nyāya views on time and causality and explain how the reality of time is vital to its entire metaphysical framework. In a way to complement our treatment of Śrīharṣa, we conclude the article (Sect. 4.4) with a brief discussion of views on time held by later Advaitins, namely Citsukha and Madhusūdana Sarasvatī.

4.2 Time in Nyāya

The philosophical position of Nyāya regarding the problem of time derives from the specific way in which it understands the world—its constitution, structure and functioning.[3] Nyāya defends a robust realism based on a pluralist metaphysics framed in terms of "categories of existence" (*padārtha*, lit. a thing denoted by a word).[4] Substances (one of the *padārtha*) include perishable (*kṛtaka*) entities such as pots and trees, as well as imperishable (*akṛtaka*) or eternal (*nitya*) entities such as atoms, selves and others. Time (*kāla*) counts among the imperishable substances. Naiyāyikas define it so because time is assumed to possess properties that can only be attributed to an eternal substance. Among those is the fact that time does not presuppose any substratum for its own existence; it is also all-pervading, uniform and unchanging. In terms of its function, time is regarded as the locus (*ādhāra*) where all contingent (*kādācitka*) entities come to existence. This function of time pertains to contingent entities alone because eternal substances do not exist "in time" according to Nyāya: they are timeless and as such beyond the realm of things conditioned by processes of production, change and destruction. Hence time itself, insofar as it is an eternal substance, is timeless. Not unlike the Newtonian absolute time, time in Nyāya constitutes the static background in which events happen and from which they derive their chronological order. Since no change in terms of action, movement or modification is conceivable without time, time must be an objective category of existence. It is not a subjective construction (*puruṣabuddhiprabhava*) but ontologically real.

---

[2] For the view that Śrīharṣa's method of argument was essentially deconstructionist, i.e., aimed at showing that the opponents' doctrines contradicted themselves, see Granoff [9]; for the argument that he had a positive program involving a personal defense of Advaita teachings, see Phillips [15].

[3] For a general treatment of time according to Nyāya(-Vaiśeṣika), see: Balslev [6]: 25–43, [7, 8] [10]: 205–228, [12–14]. The following discussion is based especially on the first two references.

[4] Generally speaking, Nyāya accepts seven such categories: 1. substance (*dravya*); 2. quality (*guṇa*); 3. activity (*karman*); 4. generality (*sāmānya*); 5. individuator (*viśeṣa*); 6. inherence (*samavāya*); 7. absence (*abhāva*).

Another fundamental property of time is its unity (*ekatva*): time is a unitary, undivided entity. The immediate question is: if time is indeed one single objective real, how do we account for the experienced plurality of time-forms and temporal distinctions? For Naiyāyikas, time-forms such as past, present and future, and temporal distinctions such as posteriority (*paratva*), anteriority (*aparatva*), simultaneity (*yaugapadya*), succession (*ayaugapadya*), etc., are purely conventional and can be explained on the basis of external factors—that is, factors that are not inherent to the substance of time. Time itself is not inherently divided into past, present and future; it is the particular events taking place "in time" that enable us to speak of time as past, present and future, and this is determined through external factors such as the motion of the sun, moon, etc. Technically speaking, time-forms and temporal distinctions result from limiting conditions (*upādhi*) that are imposed on the unitary substance of time. Naiyāyikas thus distinguish between two kinds of time: absolute time (*mahākāla*), which is eternal and of course real; and empirical or "conventional" time (*khaṇḍakāla*), which arises from the association of absolute time with external limiting conditions.

According to Naiyāyikas, temporal distinctions are the "marks" or "indicators" (*liṅga*) on the basis of which they *infer* the existence of time. In their view, perception (*pratyakṣa*) cannot reveal the existence of time and much less its properties. Inference (*anumāṇa*) alone is the means of knowledge (*pramāṇa*) to cognize time as ontologically real. The conceptual problem that naturally poses itself is that temporal distinctions seem *a priori* to be *based on* time—for what means "anteriority" if not that something precedes something else *in time*? If this is the case, it is difficult to understand how one can infer the existence of time on the basis of temporal distinctions without falling into circular reasoning. To avoid this problem, Naiyāyikas explain anteriority and other such distinctions without direct reference to time. For instance, anteriority is regarded as a quality (*guṇa*) that inheres in every substance (recall that every quality *must* inhere in a substance in the Nyāya metaphysical framework). In turn, a substance possesses this quality by virtue of its relation, for instance, to a relatively large or small number of solar revolutions. To give an example: we cognize that an old man is "anterior" to a young man because the birth of the former precedes that of the latter by many revolutions of the sun. The point is that the movement of the sun need not involve any temporal reference according to Naiyāyikas.

However, how are we to infer the existence of time on the basis of a notion such as anteriority? If anteriority is a quality that inheres in a particular object A (let's say an old man) by virtue of its relation to solar motion, how is the object A in question connected with solar motion? This is not a trivial question for Naiyāyikas. Since the motion of the sun is in principle an activity (*kriyā*, one of the seven categories mentioned earlier), it must inhere in a substance: namely in the sun, not in the object A which is located far from it. In other words, there is no relation of inherence (*samavāya*) between the motion of the sun and the object A. The way to resolve this problem is to assume that both entities (the motion and the object A) are indirectly connected to each other via a mediating substance. Technically, this indirect connection is termed *svasaṃyuktasaṃyuktasamavāya* in Nyāya: inherence (*samavāya*) in a substance which is in conjunction (*saṃyukta*) with something that is in conjunction with the individual in question (*svasaṃyukta*), where the term "conjunction" (*saṃyoga*) refers to the particular relation that stands between two substances (like a table and a book standing on it). That is to say: the object A is assumed to be in conjunction with a (mediating) substance that is itself in conjunction with the sun in which solar motion inheres. In this manner, the object A and the solar motion are connected together. Naiyāyikas aim to prove that the mediating substance in question can only be *time*, and nothing else.[5] Hence absolute time is established, through inference, as the ontological basis that makes possible and meaningful such conventional notions as posteriority, anteriority, etc. and other temporal expressions. Time accounts for the sequential character of things and events, yet it is in no way defined or affected by such things and events.

4.3 Śrīharṣa's Refutation of Time (*kālakhaṇḍana*)

The Naiyāyika concept of time as a real, unitary and all-pervading substance was challenged by Buddhist, Jain as well as brahmanical philosophers. In the fourth chapter (*caturthapariccheda*) of his KKK, Śrīharṣa offers an incisive critique of this concept from the perspective of Advaita Vedānta.[6] This critique occurs in the context of Śrīharṣa's broader refutation of the Naiyāyika definition of causality (*kāraṇalakṣaṇa*). The philosophical realism advocated in Nyāya rests to a great extent on its view of causality, and causality itself depends upon the implicit postulate that time is real. By demonstrating that the ontological reality of time cannot be logically ascertained, Śrīharṣa provides an additional proof that the time-based definition of causality advocated by Naiyāyikas is incorrect. As we shall see (Sect. 4.4), the Advaita defense of the transactional reality (*vyāvahārikasattā*) of the empirical world is not reconcilable with a conception of causality that implies the reality of time.

---

[5] For a more detailed discussion of this proof, see [7]: 184–187.

[6] We must keep in mind, however, that Śrīharṣa does not himself present the Advaita view on the subject but, given the nature of his work, confines himself to a critical treatment of other schools. As far as we know, there exists no detailed treatment of time from the perspective of Advaita Vedānta in secondary literature. For general treatments, see: [11, 14]. Reference [6] has a short section on Advaita Vedānta (pp. 57–72) in which he also briefly addresses Śrīharṣa's rejection of the Nyāya position.

4.3.1 Preamble [A]

As[7] is common in Indian philosophical debates, Śrīharṣa's argumentation in the KKK is framed in terms of a dialectical confrontation between the proponent of the thesis to be defended (*siddhānta*), Śrīharṣa himself, and an opponent, here a Naiyāyika. Śrīharṣa first points out a problem in the very formulation of the definition of causality (*kāraṇalakṣaṇa*) proposed by Naiyāyikas, wherein they state that the cause must precede the effect in time (*prākkālina*). Clearly, the word *prāk* [= *prāñc*] in the definition can only refer to something occurring in the past and as such must exclude whatever is present and future. The problem with this analysis (*vivecana*), according to Śrīharṣa, is that it assumes that time is real and that it is inherently divided into three time-forms, namely past, present and future. But on which logical basis are we to distinguish what is past from what is present and future? What allows the Naiyāyika to assume that time has a fundamental reality and that its division into past, present and future is meaningful in the first place? Śrīharṣa's claim that a time-based definition of causality is impossible without a prior definition of time itself serves as the starting point for his refutation of time (*kālakhaṇḍana*).

The opponent begins by stating how we can establish that time exists in the form of past, present and future: the three time-forms exist because we do cognize time as past, present and future in our everyday experience. The existence of cognitions (*buddhi*) such as "it exists", "it existed" and "it will exist" is itself a proof (*pramāṇa*) that time really exists—independently of these cognitions, as it were—in the form of past, present and future. This is a realist argument: the very fact that we cognize various objects proves that the world is *really* constituted of such objects. For Śrīharṣa, however, this argument amounts to a logical fallacy. He asks: but what is the content (*viṣaya*) of these cognitions of time? The opponent's answer is simple: particular time-forms (*kālaviśeṣa*), i.e., past, present and future. In other words, when you cognize an external event, you cognize it along with time: time is an essential component of each and every cognition. The content of the cognition of something cognized *as being present* points to the ontological existence of a present time that is different from past and future times.

4.3.2 Time-Forms as Inherent (svābhāvika) to Time *[B]*

But what does "particular time-forms" (*kālaviśeṣa*) really mean from an ontological standpoint? In the Indian philosophical terminology common to both opponents, a particularity (*viśeṣa*) can either be inherent (*svābhāvika*) to an entity or result from limiting conditions (*upādhi*) imposed on this entity. From now on, Śrīharṣa

---

[7] The letters in brackets refer to the corresponding sections of the Sanskrit text in the Appendix. We have introduced these subsections for the sake of clarity.

proceeds to show that neither of these alternatives makes sense in the case of time. First of all, if time-forms such as past, present and future are integral to time, i.e., part of his very nature, then their mutual difference (*bheda*) must also be real. In this case, the unity (*ekatva*) of time—a fundamental property of time according to Naiyāyikas—cannot be maintained. If Naiyāyikas were to uphold the unitary nature of time and that its unitary form is, say, "present," then how could they explain the experienced fact that the time cognized now as "present" was cognized as "future" yesterday and will be cognized as "past" tomorrow?

If, in order to explain these different cognitions, the opponent were to claim that the different time-forms result from the threefold nature (*trividhasvabhāva*) of time —a claim meaning that time, although unitary, possesses three aspects internal to it as it were—then he faces two major problems. The first problem raised by Śrīharṣa is of an ontological nature: how do we explain the reality of this division (*bheda*) into three aspects? If this is inherent to the nature of time, how again is the unity of time maintained? The second problem is epistemological. If past, present and future are merely three aspects of the same unitary time, there is no strict criterion for us to distinguish the cognition of past from that of present and future. Hence the cognitions "it happened" and "it will happen" could occur at the same time as the cognition "this is happening." Technically speaking, this position leads to the undesirable consequence (*prasaṅga*) that there is no proper distinction—or *vyavasthā*, a term which recurs in Śrīharṣa's argumentation—between the three types of cognitions. Therefore, the difference between the time-forms cannot be accounted for if we hold that the latter are integral to the nature of time.

4.3.3 Time-Forms as a Result from Limiting Conditions (aupādhika) *[C]*

Since it is untenable that time-forms, or any temporal distinction for that matter, are inherent to time, Naiyāyikas try to account for these on the basis of external factors. Technically, these external factors are termed *upādhi*—properties or "limiting conditions" imposed on a particular entity that pass their own quality to this entity (e.g., a red flower placed behind a clear crystal which makes the crystal look red). In the present context, the *upādhi* is defined by the opponent to be the particular relation that stands between the cognized event and activities such as the motion of the sun, etc. Time is not inherently divided into past, present and future; it is the association of particular activities (*kriyāsambandha*) with time that make us cognize time as past, present and future. Śrīharṣa disagrees with this explanation and presents his counterargument as follows. While the cognition of a certain event with reference to the motion of the sun enables us to define that this event occurs in the present, the same solar motion can also be used to describe the same event as past or future at some other time. For instance, if we were to claim a particular day as "today" based on a particular displacement of the sun, then the same day could be referred to as “yesterday” (*vṛtta*; a day has elapsed) or “tomorrow” (*vartsyat*; a day that has to elapse). Since solar motion is the common criterion for distinguishing the three time-forms, it cannot serve as a proper basis for cognizing them separately (i.e., *vyavasthā* is not possible).

The opponent tries to avoid this problem by clarifying how the cognized event relates to the limiting condition. The cognition of an event as present occurs when the relation (*sambandha*) between the event and the solar motion is actually “present” (*svarūpeṇa avatiṣṭhamāna*); the cognition of an event as past occurs when that same relation is no more there (*vinaṣṭa*); and the cognition of an event as future occurs when that same relation is yet to be established (*anāgata*). The distinguishing criterion here is not merely solar motion but precisely the relation that stands between it and the cognized event. Śrīharṣa’s reply to this proposition is that this distinguishing criterion is not adequate either because the relation in question must itself be located in time. In his definition of present time, the opponent had used the present participle *avatiṣṭhamāna* (from the verbal root *ava* + $\sqrt{}$*sthā*)—meaning “is taking place” or “is present”—to describe the relation between the cognized event and the limiting condition, here solar motion. Śrīharṣa’s point is that this verbal form, insofar as it derives from the present tense (*laṭ*), semantically involves the notion of present time. Since the definition (*lakṣaṇa*) of the present time (say t_2) itself involves the notion of present time (the object to be defined, *lakṣya*; say t_1), we have a clear instance of self-dependent reasoning (*ātmāśraya*). This problem arises if we assume that t_1 is identical to t_2. In order to avoid this problem of self-dependency, if the opponent were to say that t_2, conveyed through the verbal form *avatiṣṭhamāna*, is different from t_1, whose definition is being attempted, then we land into another logical problem, namely that of infinite regress (*anavasthā*). That is to say: to understand t_1 we need to understand t_2; and to understand t_2 we need to come up with another definition for another “present time” t_3, etc. The same problem occurs with respect to past and future times.

4.3.4 Time-Forms as a Consequence of the Presence and Absence of an External Activity *[D]*

At this point, the opponent turns to another, more technical definition of time-forms: the present corresponds to the time that is delimited (*avacchinna*) by a certain activity; the past corresponds to the time that is delimited by the prior absence (*prāgabhāva*) of this activity; and the future corresponds to the time that is delimited by the posterior absence (*pradhvaṃsābhāva*) of this activity.[8] While defining the

---

[8] Prior absence means the non-existence (*abhāva*) of something that is yet to come to existence (e.g., the absence of a house about to be constructed); posterior absence means the non-existence arising from the annihilation of something that was existing previously (e.g., the absence resulting from the demolition of the house).

three time-forms in this way, the opponent has recourse to a more abstract terminology based on the presence (*bhāva*) and absence (*abhāva*) of a certain activity (*kriyā*). When a certain activity is occurring, our cognition of this activity comes along with the cognition of present time; when this activity has yet to take place, the cognition of the (prior) absence of this activity comes along with the cognition of past time; and when this activity has come to completion, the cognition of the (posterior) absence of this activity comes along with the cognition of future time.

Once again, Śrīharṣa must disagree. First of all, to strictly define the present as that which is delimited by an activity (*kriyāvacchinna*) is too general a definition—a fallacy referred to as *atiprasaṅga*, or over-extensive application of the definition—for it applies also to past and future times insofar as these time-forms are also, in a way or another, delimited by activities. In addition, there are obvious problems with defining past and future in terms of the prior and posterior absence of an activity. Firstly, the expression "prior absence" (*prāgabhāva*) is hard to comprehend given that it contains the word "prior," a time-based notion that is yet to be defined properly. Secondly, it is difficult to distinguish posterior from prior absence. Śrīharṣa incidentally challenges the very definition of prior and posterior absence, based on which the opponent attempts to define time-forms. According to the opponent, prior absence is an absence that eventually gets destroyed (*vināśin*) and posterior absence is an absence that is eventually generated (*utpattimat*). Now, for Naiyāyikas, the destruction (*dhvaṃsa*) of the prior absence (*prāgabhāva*) is in the form of a counter-positive (*pratiyogin*) and so also is the prior absence of destruction.[9] Thus Śrīharṣa argues that it is untenable that only prior absence be *vināśin* and not posterior absence too, for the counter-positive (e.g., pot, etc.) is the same in both cases. Śrīharṣa then proceeds to analyze the term *utpattimat*: (a) if *utpattimat* means that something non-existent (*asat*) comes to existence (*sat*), and if this *sat* is taken in the technical sense of generality (*sāmānya*, a technical term in Nyāya), then it is inappropriate to assign it to absence (of any type) since generality exists only in substances (*dravya*), qualities (*guṇa*) and activities (*karman*), not in absence; (b) if, on the other hand, *sat* is taken in the more general sense of mere existence, then "prior absence" too becomes a candidate for this definition, for it also has existence before the pot comes to existence. Thus no counterargument can hold since the definitions of the two absences suffer from the problem of over-applicability of the definition.

---

[9] In Nyāya literature, the term *pratiyogin* (counter-positive or adjunct) is abundantly used. In the sentence, "that pot was broken," the counter-positive of the absence of the pot (as a result of being broken) is the pot itself. Such type of counter-positive is generally referred to as *abhāvapratiyogin* (absential counter-positive) in contradistinction to *bhedapratiyogin* (differential counter-positive) and *sambandhapratiyogin* (relational counter-positive). It is understood here that both the destruction of prior absence and the prior absence of destruction are essentially counter-positives themselves. As the saying goes: *prāgabhāvadhvaṃsaḥ pratiyogirūpaḥ, dhvaṃsaprāgabhāvo 'pi pratiyogirūpaḥ.*

4.3.5 Time-Forms Are Based on the Relation (apekṣā) to a Particular Activity *[E]*

Now that Śrīharṣa has successfully refuted the definition of time-forms based on the presence and absence of an activity, the opponent tries to defend his position by offering yet another definition of the time-form "present": the time delimited by a particular activity is cognized as "present" in relation (*apekṣayā*) to that particular activity and not some other activity (*anyāpekṣayā*). The term *apekṣā*—which literally means "with reference to" or "in relation to"—is the key term of this new definition, and this is the term that Śrīharṣa questions in the first place. What kind of "relation" exactly stands between the cognized activity (*kriyā*) and its "present-ness" (*vartamānatā*)? Is it a relation where the activity is a limiting condition (*upādhi*), a boundary-element (*avadhi*), a counter-positive (*pratiyogin*) or something that appears as a qualifier (*prakāra*) in the cognition? Śrīharṣa takes each of these separately and refutes them one by one:

- limiting condition: this alternative is quickly rejected on the basis that Śrīharṣa had already proven earlier that to invoke a limiting condition is problematic to the extent that the condition too is time-bound;
- boundary-element: an *avadhi* generally refers to a boundary, limit or reference point for a certain entity or process. In the cognition "this is long compared to that," the word "that" is the *avadhi*. Śrīharṣa first argues that time can be cognized without relying on such an *avadhi*. For instance, in the cognition "the tree exists" (*vṛkṣo vartate*), the time-form "present" is also cognized as the verbal form implies the cognition that the tree is "present," and yet the cognition does not involve any standard of comparison. Secondly, to postulate the existence of an *avadhi* with respect to which we cognize something as present, leads to the same problem raised earlier in connection with the *upādhi*: the impossibility to distinguish the cognitions of past, present and future (*pratyayāvyavasthā*). If we assume that a certain reference point in the sky allows us to define the day as "today," the same reference point can be used the next day to define "yesterday";
- counter-positive: Śrīharṣa does not elaborate here and sees the rejection of this position as a logical consequence of the rejection of the *avadhi* position;
- qualifier in cognition: if the opponent has in mind that the cognized activity relates to its "present-ness" through presenting itself as the qualifier (*prakāra*) in the cognition,[10] then it is tantamount to saying that even cognitions that are meant to convey past and present by appropriate verbal forms, end up conveying present-ness, since the criterion used to define present-ness is common to all the three time-forms.

---

[10] In the cognition of a cow, what is cognized is an entity that is endowed with the general property of cowness; this cowness is technically referred to as the qualifier (*prakāra*) in the cognition.

4.3.6 Time as a Simultaneous Object and Locus of Cognition *[F]*

The Naiyāyika opponent has up to now been forced to contend with Śrīharṣa's arguments. He now attempts to provide a last definition of the three time-forms: present time is the object (*viṣaya*) as well as the locus (*āśraya*) of a cognition; past time is the time delimited (*avacchinna*) by the prior absence (*prāgabhāva*) of the limiting condition (*upādhi*) imposed on present time; and future time is the time delimited by the posterior absence (*pradhvaṃsābhāva*) of that same limiting condition. We will focus here solely on Śrīharṣa's rejection of the definition of present time. The opponent understands present time to consist of two aspects—object and locus. To be the object (*viṣaya*) of a cognition means that time is cognized along with the cognition of a certain entity or activity (e.g., when cognizing a book, both the book and the time at which cognition occurs are objects of the cognition). However, if present time were defined only in this way, the definition would be over-extensive for several things can be objectified in the process of cognition. Thus the opponent adds another qualification (*viśeṣaṇa*) to his definition: present time is also the locus (*āśraya*) of the said cognition, which means that it is the condition *sine qua non* of its taking place. Just as cognition, according to Naiyāyikas, cannot take place without a self (*ātman*) that supports it, in the same way cognition cannot take place without time as its support. In other words, every cognition takes place *in* time. Note that without the first qualifier (i.e., *viṣaya*) the definition would also be over-extensive for it would also apply to the self.

In order to refute this definition based on the twofold function of time in the process of cognition, Śrīharṣa looks at the specific way in which Nyāya understands cognition. According to Naiyāyikas, every cognition is followed by an "after-cognition"—a process technically referred to as *anuvyavasāya*, and otherwise translated as "reflective awareness" or "apperception"—in which the first cognition itself becomes an object of cognition. First, a certain entity or activity is cognized; then, the fact that it is cognized is itself cognized. In other words, cognition is not self-evident (*svaprakāśa*) but can only be the object (*viṣaya*) of another cognition. The opponent has just argued that time, more specifically "present-ness," is such an object of cognition. Śrīharṣa's contention is essentially that this process of "after-cognition" does not take place at the moment of the first cognition—namely at the time when something was cognized as "present"—but only *afterwards*. The implication is that present-ness cannot be cognized at the "present" moment.

4.4 The Advaita View of Time: Citsukha and Madhusūdana

Since the primary purpose of the KKK is to uphold the principle of indefinability (*anirvacanīyatā*) dear to the Advaita Vedānta tradition, Śrīharṣa does not explicitly define how time is conceived in Advaita Vedānta. As a matter of fact, prior to

Śrīharṣa, the Vedānta tradition does not have much to say about time and its ontological properties. While the various references to space (*ākāśa*) in the *Upaniṣad*s—mostly in the context of describing the manifestation of the physical world (*sṛṣṭiprakriyā*)—and the exegetical works of Śaṅkara and others, makes it clear that the early Vedānta tradition accepted space as an entity having empirical or transactional reality (*vyāvahārikasattā*), the same cannot be said with certainty about time, as there is hardly any discussion about it either in the sections of the *Upaniṣad*s concerned with *sṛṣṭiprakriyā* or in their commentaries. It is only much later, well into the medieval period, that Advaitins come to engage the concept of time from their own perspective. One of the most detailed discussion of time after Śrīharṣa is found in the *Tattvapradīpikā* of Citsukha.

4.4.1 Citsukha

Citsukha (13th century) was a prolific writer in the tradition of Advaita Vedānta. In his most celebrated work, the *Tattvapradīpikā*, Citsukha engages with the work of several prominent Naiyāyikas including Śrīdhara (10th century), Udayanācārya (10th century) and Vallabhācārya (c. 12th century). The Tattvapradīpikā is a substantial scholarly treatise in four chapters that is thematically akin to the *Brahmasūtra*s. In the first chapter, Citsukha lays down the theoretical tools required for approaching the teachings of Advaita Vedānta. In the second chapter, he primarily engages with the views presented by Naiyāyikas; it is in this chapter that he dwells at length on their view of time. Unlike Śrīharṣa, however, Citsukha presents his own position while refuting the Nyāya position.

Throughout the Tattvapradīpikā, Citsukha presents succinctly the theme of a discussion in the form of verses (*śloka*) on which he then elaborates at great length, a style generally referred to as *samāsavyāsapaddhati*. He begins his discussion on time with the following verse, where he beautifully summarizes the detailed discussion that he is going to present in prose later on:

> Since time is not an object of sense-perception, given that posteriority, etc. are not [proper] inferential marks (*liṅga*) [in order to establish its existence]; since it is not a cause on its own (*svarūpataḥ*); since [introducing] a limiting condition (*upādhi*) [in order to establish its existence] [also turns out] to be futile; and [finally,] since it is possible [to establish] the connection between a material entity (*piṇḍa*) and the motion of the sun [which is the basis for reckoning time] merely through the all-pervasive conscious [being] (*cetana*), how is it that [time, as an independent, ontologically real entity] gets established?[11]

Of particular importance for Citsukha is *how* Naiyāyikas establish the reality of time. Citsukha argues at length that neither perception nor inference are proper

---

[11] *pratyakṣāgocaratvena paratvāder alingataḥ* | *svarūpato 'nimittatvāt upādhau niṣphalatvataḥ* || *divākaraparispandapiṇḍasaṃgatisaṃbhavāt* | *vyāpinaś cetanād eva kathaṃ kālaḥ prasiddhyati* || (*Tattvapradīpikā*: 510).

means of knowledge (*pramāṇa*) to establish time as an ontologically real entity. Recalling Vallabhācārya's realistic treatment of time in the *Nyāyalīlāvatī*, Citsukha says that the ontological reality of a substance (*dravya*) can be established through perception (*pratyakṣa*) if and only if the substance in question has a visible form (*rūpa*) and/or has the quality of touch (*sparśa*). However, according to Naiyāyikas themselves, time is without form and is intangible, and thus cannot be an object of perception. Time cannot be perceived through the other sense-organs either since it is understood that the nose, ears and tongue can only sense qualities (*guṇa*) and not substances. Nor can time be perceived through internal or mental perception (*mānasapratyakṣa*) since internal perception cannot perceive things that are external to the body without the assistance of the five sense-organs.

Citsukha then shows that inference (*anumāṇa*) too is of no avail to prove the existence of time. We recall from our previous discussion (Sect. 2) that Naiyāyikas stress inference as the only means of knowing time. In their view, time is inferred on the basis of such cognitions as posteriority (*paratva*), anteriority (*aparatva*), simultaneity (*yaugapadya*), etc., all of which serve as inferential marks (*liṅga*) in the inferential process. According to them, such notions could not arise unless there is an entity called "time" which allows for their arising. Now, inference is possible only when a relation of invariable concomitance (*vyāpti*) is established between a probandum (*sādhya*, here time) and its inferential mark (*liṅga*). Citsukha refutes all arguments adduced by Naiyāyikas to ascertain the concomitance between the probandum and its marks, required for the inference to take place.

He further asks whether posteriority, anteriority, etc. are manifestations (*kārya*) of time or are associated with time through certain external factors (*upādhi*). If one holds that they are manifestations of time (time being their inherent cause, *nimitta*), there is the problem that time, being unitary, cannot be the cause of various temporal distinctions. If one admits instead that these various temporal distinctions result from factors such as the motion of the sun, then the latter can just as well be said to be sufficient in order to account for conventional temporal usages. What need is there, in this case, to invoke an independent entity called "time"? Challenging this explanation offered by Citsukha, a Naiyāyika may well argue: temporal distinctions such as young, old, etc. are well known to be associated with entities (*piṇḍa*) that are present here on earth such as buildings, humans, etc. How can these entities be explained away simply by invoking some sort of activity of the sun, unless we are able to bring in a connection between that activity that is supposed to be taking place somewhere in space, with something that is being present here on earth?

At this point of his argumentation, Citsukha introduces the Advaita view. According to him, the connection in question can be explained by the all-pervasive (*vyāpin*) and conscious being (*cetana*), which is, he argues, well known and accepted by proponents of all schools. Of course, Citsukha has in mind the universal self (*ātman*) of the Vedāntic tradition, identified with Brahman in Advaita Vedānta. If the inert, unitary and all-pervasive (*vibhu*) time, can bring in, with the help of limiting conditions (*upādhi*), the necessary connection between the motion of the sun, etc. and entities on earth, why not the conscious, unitary and all-pervasive (*vibhu*) *ātman*? Thus to an Advaitin like Citsukha, there is simply no convincing reason for accepting the ontological reality of time. Not only is there no proper means of knowledge that can establish its existence, but there is also no logical ground to postulate its existence as a separate ontological category given that its only function is fulfilled by another principle, namely *ātman*.

4.4.2 Madhusūdana Sarasvatī

It is quite clear from the previous that Advaitins do not accept time as a separate ontological entity. In his *Siddhāntabindu*, the great 16th-century scholar Madhusūdana Sarasvatī argues in the same line but provides a different viewpoint with regard to time, worth mentioning here:

> As for time, it is none other than ignorance (*avidyā*) itself as precisely [ignorance] is the locus of everything (*sarvādhāra*).[12]

Firstly, Madhusūdana identifies time with *avidyā*, the primeval ignorance, which in late Advaita is viewed as the metaphysical principle explaining the origin and manifestation of the world as well as our mistaken cognitions of plurality and change. Secondly, he explains this identification on the basis that ignorance is the locus of everything. For Madhusūdana too time has no separate existence but unlike in Citsukha's view, it is not because it has no role to play in the working of the world. On the contrary, Madhusūdana identifies time with *avidyā* and in this sense time plays a fundamental role in manifesting the world as plural, changing and limited. It must be noted that the characterization of time as *sarvādhāra*—a function which Madhusūdhana here assigns to ignorance—has its equivalent in the Nyāya tradition. In the *Kārikavalī*, Viśvanātha Pañcānana (17th century), reiterating a point already made earlier by Praśastapāda and others, describes time as the producer of all the entities that are produced and the locus (*āśraya* = *ādhāra*) of the entire universe.[13] Since in Advaita Vedānta, everything, both tangible and intangible, is considered to be an effect of ignorance (*avidyākārya*), it is only appropriate that ignorance should be described here as *sarvādhāra*. However, the way in which time and ignorance are considered to "contain" all things differ considerably in Nyāya and Advaita Vedānta respectively. While ignorance is the locus of all that exists inasmuch as it is the material cause of all changeable entities (*pariṇāmyupādāna*), time is a locus inasmuch as it is the cause for the arising of cognitions involving time-forms.

Hence time in Madhusūdana's definition has a rather different meaning than in Nyāya. Through being identified with *avidyā* itself, time is relegated to the domain of things that have a transactional reality (*vyāvahārikasattā*), in contradistinction to

---

[12] *kālas tv avidyaiva tasyā eva sarvādhāratvāt* (*Siddhāntabindu*: 66).

[13] *janyānāṃ janakaḥ kālo jagatāṃ āśrayo mataḥ* (*pratyakṣakhaṇḍa*, *kālanirūpaṇa*, v. 45) (*Kārikāvalī*: 195).

Brahman, which alone has absolute reality (*pāramārthikasattā).* While Brahman is the only real existent in Advaita, everything that is phenomenal, including time, falls into the category of what is only apparently real (*mithyā*), dependent upon the higher reality of Brahman for its own existence.

4.5 Conclusion

Nothing is more commonplace than time, yet when it comes to conceptualising and describing it in words, nothing presents a greater predicament. The *experience* of time, whatever be its nature, makes us feel that it must be a real entity. Moreover, time presents itself as an indispensable tool for the description of natural phenomena, including our daily transactions in the world. In this line of thought, the realist school of Nyāya defends the view that time is an essential component of the world and classifies it among substances (*dravya*) in their scheme of categorizing the various entities that are ontologically real. In contrast to this, Advaitins hold that time is a phenomenal product and a human construct (*puruṣabuddhiprabhava*) that has no ontological reality. For them, what is "real" (*sat*) is something that cannot be denied through some other cognition in all three times (*trikālābādhya*)—past, present and future. In other words, what "really is" cannot be described as "it is not, it will not be, it was not": it is freed from any trace of time, it is pure timeless-ness. Since the "time" that is being experienced by us can be described as past, present and future, it cannot be *trikālābādhya* and hence cannot be given the status of something "real."

It is one thing to have a feel for something and claim that it must be real, but another to define it and help the other make sense of the definition. Though the Naiyāyikas try hard to define time in various ways, Śrīharṣa does not find these arguments convincing. Realizing the paradoxical nature of time, Madhusūdana identified time with ignorance (*avidyā*) whose delusive and indescribable nature has been vividly portrayed in the *Vivekacūḍāmaṇi*:

> It is neither existent nor non-existent, nor is it both; it is neither the same nor different, nor is it both; it is neither with parts nor partless, nor is it both; [ignorance] is a great wonder and cannot be described in words.[14]

Thus we find both logicians and philosophers struggling hard to come up with a coherent definition of time that is at once convincing and can capture what it is without beating around the bush. However useful the notion of time may turn out to be in our daily transactions and in the equations of physics, the more deeply we ponder over it to understand its nature, the more elusive it turns out to be. Hence, it is no surprise that Śrīharṣa, one of the most brilliant Advaitins of all times, upholds the view that time is utterly indefinable.

---

[14] *san nāpy asan nāpy ubhayātmikā no bhinnāpy abhinnāpy ubhayātmikā no* | *sāṅgāpy anaṅgā hy ubhayātmikā no mahādbhutā 'nirvacanīyarūpā* (v.109) (*Vivekacūḍāmaṇi*: 45).

Appendix

The Sanskrit text presented here is based on the printed edition of the KKK in the Chowkhamba Sanskrit Series no. 146, pp. 1233–1247 (see References).

[A] *evaṃ kāraṇalakṣaṇaṃ nirasya tadviśeṣībhūtakālakhaṇḍanam upakramate. niyame ca prākkālīnatayā ’bhidhīyamāne prāg ity asya vyavacchedyau vartamānabhaviṣyatkālau, prākkālaś ca vyavacchedako vivecanīyaḥ, na ca tadvivecanaṃ śakyam. vartamānādibuddhaya eva svaviṣayavaicitrye pramāṇam iti cen na. tathā hi vartamānādibuddher eva ko viṣayaḥ? kālaviśeṣa iti cen na.*

[B] *kālasya viśeṣaḥ svābhāvika aupādhiko vā. nādyaḥ kālasya bhavadbhir ekatvābhyupagamāt. ya eva kālo vartamānaḥ pratīyate sa eva pūrvaṃ bhāvīti paścāt bhūta iti ca na pratīyeta. trividhasvabhāva evāsāv iti cen na. bhedaprasaṅgāt vyavasthānupapattiprasaṅgāc ca. yadaiva vartata iti pratyayas tadaiva vṛtto vartiṣyata iti pratyayaprasaṅgāt.*

[C] *dvitīyaś ced upādhir abhidhīyatām. sūryādikriyāsambandhabhedaḥ sa iti cen na. bhūtabhaviṣyator api kriyāsambandhapratyayasyāvaśyaṃ vaktavyatvāt. ya eva divasaḥ sūryagativiśeṣāvacchinno vartata iti pratītaḥ, sa eva hi tadavacchinno vṛtta ity avagamyate vartsyann iti ca.*

[…]

nanu satyam etat, paraṃ yadā tadupādhisambandhas tasya svarūpeṇāvatiṣṭhamānas tadā vartamānapratyayaḥ, yadā sa eva vinaṣṭo bhavati tadā bhūtapratyayo, yadā ’nāgatas tadā bhaviṣyatpratyaya iti naitad asti.
yady atra laṭo vivakṣito ’rthas tadā tajjñānasyaiva tajjñānopāyatvam ity ātmāśrayānavasthayor anyataraprasaṅgaḥ […]

[D] *kriyāvacchinnaḥ kālo vartamānaḥ tatprāgabhāvāvacchinno bhūtas tatpradhvaṃsāvacchinno bhaviṣyann iti cen na. atītānāgatapratītikālo ’pi kriyāvacchinnaḥ pratīyata iti vartamānapratyayaprasaṅgasya tādavasthyāt kriyānavacchinnasya tatprāgabhāvapradhvaṃsabhāvāvacchedānupapatteḥ.*

prāgabhāvaś ca prāgarthāniruktau kathaṃ na duradhigamaḥ? pradhvaṃsasyāpi prāgabhāvāt kathaṃ viśeṣo vaktavyaḥ? abhāvo vināśī prāgabhāva, utpattimān pradhvaṃsa ity anayor viśeṣa iti cen na. ko hi prāgabhāvasya vināśo yena vināśīty ucyate? yadi pratiyogibhūto ghaṭādiḥ pradhvaṃsasyāpi prāgabhāvavat pratiyogīti so ’pi vināśī prāptaḥ. utpattimāṃś ca pradhvaṃsa ity utpattipadārtho vivecanīyaḥ. yady asāv asataḥ sattvaṃ tac ca sāmānyaṃ, tadā ’bhāve ’sambhava eva. atha svarūpasattvaṃ tadā prāgabhāve ’pi prasaṅgaḥ, tasyāpi kadācid asattvāt. pūrvam asataḥ paścāt sattvaṃ vivakṣitam iti cen na. pūrvedānīṃpaścādarthasyaivānirūpaṇāt. etena kāraṇāvacchinnaṃ sattvam utpattir ity api nirastam. pūrvāparanirvacanam antareṇa kāraṇārthānirvacanāt.

[E] *astu tāvad atītānāgatayor yathātathā niruktiḥ. yatkriyāvacchinno yaḥ kālaḥ sa tatkriyāpekṣayā vartamāno, na tv anyāpekṣayeti cet tadapekṣayeti ko ’rthaḥ. kiṃ* *tadupadhānena, uta tadavadhikatayā, uta tatpratiyogikatayā, uta tena prakāreṇety eva?*

nādyaḥ, upādhyavacchinnasyātītānāgatapratipattiviṣayatvam api tasyety asakṛd uktatvāt.

nāpi dvitīyaḥ, asmād ayaṃ dīrgha itivad asmād ayaṃ vartata ity avadhyapekṣām antareṇa pratīyamānatvāt, sarvadaiva ca trividhāvadhyapekṣayā āsīd asti bhaviṣyatīti pratyayāvyavasthāprasaṅgāt.

ata eva na tṛtīyaḥ.

nāpi caturthaḥ, atītānāgatapratītikāle kriyāvacchedaprakāreṇa vartamānapraty ayaviṣayatvaprasaṅgāt. nāsau kriyāvacchedalakṣaṇaḥ prakāro 'tītānāgatakāle vartata iti cen na. vartamānatāyā adyāpy anirūpaṇena vartata ity uktvā viśeṣasya darśayitum aśakyatvāt.

[…]

[F] *syād etat grāhakavijñānaviṣayo grāhakavijñānāśrayaś ca kālo vartamānaḥ, vartamānopādhiprāgabhāvāvacchinnaś ca pūrvas tatpradhvaṃsābhāvāvacchinnaś cānāgataḥ, prāgabhāvapradhvaṃsayoś ca svābhāvikam eva bhedam ādāya vyavasthā.* […] *iti maivam.*

jñānāsvaprakāśatāpakṣe svopahitasya svayaṃ grahaṇānupapatteḥ kathaṃ vartamānatāgrahaḥ, jñānāntareṇa ca tathāgrahe vartamānatāvabhāsāṅgīkāre tadā 'sau dṛṣṭo mayeti pratyayasya tadā 'sau mayā dṛśyata ity ākāratāpattiḥ.

References

Primary Sources

1. *Kārikāvalī* of Viśvanātha Pañcānana. Ā. N. Jere (Ed.). (1982). *Kārikāvalī of Viswanātha Nyāyapanchānana Bhattāchārya with the commentaries* Muktavalī, Dinakarī, Rāmarudri. Varanasi: Krishnadas Academy.
2. *Khaṇḍanakhaṇḍakhādya* of Śrīharṣa. Pandit L. S. Dravida (Ed.). (1904–1914). *Śrī Harsha's* Khaṇḍanakhaṇḍakhādya *with the commentary* Khaṇḍanaphakkikāvibhajana *of Ānandapūrṇa with extracts from the commentaries of Chitsukha, Śaṅkara Miśra and Raghunātha* (Vol. 2). Benares: Chowkhamba Sanskrit Series, no. 146.
3. *Tattvapradīpikā* of Citsukha. (1956). *Tattvapradīpikā* with the commentary *Nyāyaprasādinī* and Hindi translation. Varanasi: Udaseena Samskrita Vidyalaya.
4. Siddhāntabindu of Madhusūdana Sarasvatī, K. N. Subramanian (Ed.). (1989). *Siddhāntabindu: Madhusudana Sarasvati's Commentary on Sri Sankara's Dasasloki*. Varanasi: Rishi Publications.
5. *Vivekacūḍāmaṇi* (attributed to Śaṅkara). Swami Madhavananda (Ed., tr.) (1921). *Vivekachudamani of Sri Sankaracharya*. Almora: The Advaita Ashrama.

Secondary Sources

6. Balslev, A. N. (1983). *A study of time in Indian philosophy*. Wiesbaden: Otto Harrassowitz.
7. Bhaduri, S. (1947). *Studies in Nyāya-Vaiśeṣika metaphysics*. Poona: Bhandarkar Oriental Research Institute.
8. Faddegon, B. (1992). Space and time (in the Vaiśeṣika System). In H. S. Prasad (Ed.), *Time in Indian philosophy: A collection of essays* (pp. 465–478). Delhi: Sri Satguru Publications.
9. Granoff, P. (1978). *Philosophy and argument in late Vedānta: Śrī Harṣa's Khaṇḍanakhaṇḍakhādya*. Dordrecht/Boston/London: D. Reidel Publishing Company.
10. Halbfass, W. (1992). *On being and what there is: Classical Vaiśeṣika and the history of Indian ontology*. Albany: State University of New York Press.
11. Hiriyanna, M. (1992). Advaitic conception of time. In H. S. Prasad (Ed.), *Time in Indian Philosophy: A collection of essays* (pp. 541–542). Delhi: Sri Satguru Publications.
12. Keith, A. B. (1992). Time and space (in Nyāya-Vaiśeṣika). In H. S. Prasad (Ed.), *Time in Indian philosophy: a collection of essays* (pp. 457–463). Delhi: Sri Satguru Publications.
13. Lysenko, V. (2007). The Vaiśeṣika notions of *ākāśa* and *diś* from the perspective of Indian ideas of space. In E. Franco & K. Preisendanz (Eds.), *Beyond orientalism: The Work of Wilhelm Halbfass and its impact on Indian and cross-cultural studies* (pp. 417–447). Delhi: Motilal Banarsidass Publishers.
14. Mahadevan, T. M. P. (1992). The advaita view of time. In H. S. Prasad (Ed.), *Time in Indian philosophy: A collection of essays* (pp. 543–546). Delhi: Sri Satguru Publications.
15. Phillips, S. H. (1995). *Classical Indian metaphysics*. Chicago: Open Court.

Chapter 5
From Time to Time

Nathan Salmon

5.1 I

The apparent verdict of current theoretical physics is that the prospect of time travel does not violate general relativity. The philosopher interprets this as a judgment that time travel is, at a minimum, logically consistent with general relativity. It directly follows from this judgment that general relativity and time travel are each themselves consistent, hence that time travel is logically possible. On the other side of the coin, philosophers have argued that time travel of the sort depicted in the classic H.G. Wells' novella, *The Time Machine*, is inconsistent with common sense—for example the notion that one who travels to a past time when his paternal grandfather is alive but has not yet sired his father, could murder his own grandfather at that earlier time. This is the famous *grandfather paradox*. If the reasoning is sound, then Wellsian time travel—as depicted in *The Time Machine*, in numerous other science-fiction stories, in thought experiments, and the like—is metaphysically impossible and perhaps conceptually incoherent.

Time travel essentially involves someone or something changing its temporal location from one time to another in a manner analogous to motion in space, i.e., change in spatial location over time. The first of these temporal locations is the *time of origin*, t_o, the second is the *time of destination*, t_d. Time travel essentially

---

The thoughts expressed here are a result of ruminations on the first (and much better) film adaption of H. G. Wells' novella The Time Machine and on trenchant observations made by Hilary Putnam in his paper "It Ain't Necessarily So". The essay benefitted from discussions with C. Anthony Anderson, Mark Fiocco, Stephen Humphrey, Teresa Robertson. Heather Salazar, and the late Anthony Brueckner.

N. Salmon (✉)
Philosophy Department, University of California, Santa Barbara, South Hall 3431/3432,
Santa Barbara, CA 93106-3090, USA
e-mail: nsalmon@philosophy.ucsb.edu

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_5

involves something embarking on a purely temporal voyage, “moving” through time in a manner analogous to spatial motion. In one sense, mere temporal persistence is a kind of time travel; it is motion from present to immediate future at a rate of one (second per second), the constant rate of the passage of time. Insofar as ordinary persistence is a kind of temporal motion, it is arguably impossible *not* to travel in time. For what would it be to remain stationary in time?

Wellsian time travel is continuous temporal motion through time that deviates from standard temporal motion. It is traversing a span of time during some *other* interval of time. It essentially involves changing temporal location other than at the rate by which time passes from t_o to t_d—deviating from the timeline segment that begins with t_o (without also remaining in that timeline), traveling temporally other than forward and at the same rate as the passage of time, and finally “arriving” at the timeline segment that begins with t_d. Henceforth I shall normally use the phrase ‘time travel’ and its cognates to mean Wellsian time travel.

I here undertake a philosophical investigation into the concept of Wellsian time travel. I shall also consider the prospect of backward (or retro-) causation as a genuine possibility, in the weak sense of mere logical consistency. That is, I shall assume that there is no formal contradiction in the very idea of *A* causing *B* though *B* precedes *A*. Backward time travel, at least taken in conjunction with common sense, immediately entails backward causation. (The grandfather’s future progeny causes his own death.) The champion of Wellsian time travel is committed to this assumption. Gainsayers may object, but they would be mistaken. Experience with the law demonstrates that the assumption is correct.[1]

Objections to Wellsian time travel along the lines of the grandfather paradox and similar bilking arguments typically should be regarded as modal arguments to the effect that backward time travel has the consequence that someone can do or be something that it is impossible to do or be (e.g., be sired by a father who never exists). When the argument’s structure is laid bare, a significant weakness is immediately exposed. The premises do not support the conclusion that Wellsian time travel is impossible. At most they yield only that backward time travel is impossible.[2] The prospect of forward time travel is left unscathed.

---

[1] Backward causation by judicial decree is causation between institutional “events”, not between brute facts, but backward causation it is. It is not analytic that if *A* causes *B*, then *B* does not precede *A*.

[2] The argument is valid only in modal logics stronger than ***T***. I believe ***T*** is the correct logic of metaphysical modality, so that the argument is fallacious. The argument can be alternatively formulated so that it is valid even in ***T***.

5.2 II

As David Lewis noted, the concept of time travel presupposes a distinction between two kinds of time.[3] The time traveler is a two-timer (in at least one sense). First, there is *real time*, t_r, through which the time traveler travels. This is also called 'common time', 'objective', 'external', 'coordinate', 'global', 'local' (yes, both!), and more. It is also called simply 'time'. Real time is to time travel as physical space is to spatial motion: It is the "space" through which the time traveler travels. In addition there is the time traveler's *proper time* ('rest', 'home', 'clock', 'process', 'personal', 'time-traveler', 'intrinsic', 'subjective'), t_p. A time traveler's proper time is what the traveler's wristwatch tracks—as with the inattentive globe trotter who does not diligently readjust a personal timepiece when travelling across time zones. Proper time is to time travel as real time is to spatial motion. It is the dimension over or during which the change in location occurs. The real time of origin t_o coincides with the proper time of "departure", immediately after which proper time begins deviating from real time. The proper time at the real destination time t_d is the time traveler's time of arrival.

Trivially, real time and proper time each flows from past to future at a fixed rate relative to itself of one (second per second), irrespective of the time traveler's circumstances. To that extent, both kinds of time are objective. There is the mundane circumstance of trivial temporal motion if, and only if, the would-be time traveler moves forward through real time at the same rate as the passage of real time, and thus neither gains nor loses any real time over proper time. In it the two kinds of time coincide exactly; where there is only trivial temporal motion, $t_p = t_r$. There is Wellsian time travel if and only if the intrinsic arrival time is not the real destination time. The proper time of the Wellsian time traveler gains on real time, lags behind real time, reverses with respect to real time—or maybe freezes with respect to real time.

Imagine an amusement-park conveyance ride called *Travels through Time*. The rider is seated on a bicycle labeled 'TIME MACHINE'. The bicycle sits upon the conveyer belt at an illuminated position labeled 'HOME TIME'. The conveyance and illumination each move forward, carrying the rider and bicycle along with them, at a constant rate of one inch per second. Second by second, inch by inch, the rider and bicycle move their way through the long, darkened corridor labeled 'HALL OF TIME', their changing position constantly illuminated. Through the mist, to the right and left of the conveyer belt are signs pointing the way ahead and marked 'FUTURE' and other signs pointing backward and marked 'PAST', amidst scenes and photographs from history, all arranged in chronological sequence. The rider is free simply to maintain HOME TIME position on the conveyer belt, sit back, allow the conveyance to do all the transporting labor, and enjoy the show. Conveyer and rider then move forward through the HALL OF TIME in unison. But

---

[3] "The Paradoxes of Time Travel", *American Philosophical Quarterly*, 13 (April 1976), pp. 145–152, at p. 145.

the rider is encouraged to change position on the conveyer belt by cycling either forward or backward atop the moving belt. In so doing, the rider moves not merely relative to the HALL, but also relative to the HOME TIME conveyer position, which is itself in motion relative to the HALL.

It may be taken as a definition of ‘time travel’ that there is a process (the *temporal voyage*) that something x (the *time traveler*) undergoes immediately following upon a particular real time t_o (the *time of origin*), that has a duration d_p (a positive quantity) in x’s proper time, and as a result of which when x’s proper time is $t_o + d_p$ (the arrival time), x is at a real t_d other than $t_o + d_p$. A *time machine* is any device that produces Wellsian time travel. In Wellsian time travel there are no restrictions concerning the origin time, the proper-time duration, or the real destination time. These definitions are satisfactory only to the extent that the presupposed notion of proper time is independently and antecedently understood. In particular, the notion thus defined cannot be invoked to define proper time. The latter might be independently understood operationally, in terms of the time traveler’s “clocks,” the temporal order of the time traveler’s intrinsic periodic processes —circadian rhythm, digestion, growth, sleep, reasoning, learning, other mental processing, etc., and especially entropy (aging, deterioration, degradation, or decay).[4]

Some facts of time travel follow immediately from the concept. Among these truisms are the following facts about any temporal voyage:

TT1 The time traveler exists at the real time of origin, t_o.
TT2 The time traveler exists at the real time of destination, t_d.

Both *TT1* and *TT2* follow immediately from the fact that the statement ‘At time t, x is located at ℓ’ analytically entails ‘x exists at t’.[5]

Assume that the temporal voyage proper-time duration d_p is not zero. During the temporal voyage the rate of the passage of proper time deviates from that of real time. A single unit of proper time (e.g., 1 min) is of different duration from the same unit of real time. Without this deviation between intrinsic and real time there is only trivial temporal motion. There is a temporal exchange rate, which may vary during the temporal voyage. If 1 min of proper time uniformly buys the time traveler 1 h of real time, then, assuming there is conservation of time variance,

---

[4] The proposed definition is essentially Lewis’s characterization of time travel (*op. cit.*, p. 145), which is often taken as a definition. Since the temporal voyage duration d is a measure of intrinsic (“personal”) time, not of real time, Lewis’s characterization also presupposes the notion of proper time, and therefore cannot be used to define the latter. Lewis proposes that proper time be defined functionally rather than operationally. (He does not provide an actual definition.).

[5] If it is assumed that ‘x moves from ℓ_1 at t_1 to ℓ_2 at t_2’ analytically entails ‘Between t_1 and t_2, x traverses a path from ℓ_1 to ℓ_2’, then we might have the further result that the time traveler exists also during its temporal voyage, if the temporal-voyage duration $d > 0$. If it is assumed furthermore that motion requires traversing a continuous path, then we might have the further result that the time traveler exists continuously during its temporal voyage, if the temporal-voyage duration $d > 0$.

1 min of real time lasts only one second of proper time. Proper time is therewith contracted relative to real time. In science-fiction stories, proper time is typically contracted during a temporal voyage rather than dilated, but the reverse is conceptually possible. (Arguably it is even actual.)

TT3 Barring multiple temporal voyages, the time traveler's proper time, t_p, coincides with real time, t_r, at the time of origin, t_o, and deviates from real time beginning immediately thereafter.

An object's *(average) rate of time travel* over a proper-time interval i is the ratio of real time spanned during i (a positive or negative quantity) to elapsed proper time (the positive quantity of time in i), $\Delta t_r/\Delta t_p$. In Wellsian time travel, the time-travel rate may be any positive or negative ratio without restriction. Rather than express $\Delta t_r/\Delta t_p$ as a ratio between interval lengths, I shall convert it to a scalar, e.g., for the particular time-travel rate *spanning backward 4 s of real time for every 3 s of elapsed proper time* I write '$ttr = -1.33$'. If the time traveler is transported to another time instantaneously, then the ratio of spanned real time to elapsed proper time is division by zero, hence ttr is undefined. The greater its range of time-travel rates, the more powerful the time machine.

The average time-travel rate over an entire temporal-voyage v is $(t_d - t_o)/d_p$. A simple relationship thus obtains among the origin time (t_o), the real destination time (t_d), the temporal-voyage proper-time duration (d_p), and the temporal-voyage time-travel rate (ttr_v):

$$t_d = t_o + ttr_v(d_p).$$

In the special case where $ttr_v = 1$, $t_d = t_o + d_p$, i.e., the intrinsic arrival time is the real destination time. There is Wellsian time travel over the course of the temporal voyage v if and only if $ttr_v \neq 1$. This is the case if and only if there is non-zero acceleration or deceleration both immediately after the temporal-voyage origin time t_o (immediately after which proper time begins to deviate from real time) and at the proper time of arrival $t_o + d_p$ (when proper time resumes flowing at the constant real-time rate). A time traveler may also accelerate or decelerate during a single temporal voyage. An object's *instantaneous time-travel rate* during a sub-voyage is the limit of the average time-travel rate $\Delta t_r/\Delta t_p$ as Δt_p approaches 0. The real time reached during a temporal voyage v is a function of proper time, $t_r = \boldsymbol{\rho}_v(t_p)$. The instantaneous time-travel rate is the first time derivative, dt_r/dt_p. To simplify discussion we shall assume that the time traveler travels at a uniform time-travel rate throughout the temporal voyage—no acceleration or deceleration between the origin and arrival times—so that dt_r/dt_p remains constant throughout the temporal voyage (e.g., 60 = 1 h of real time to 1 min of proper time).

The *gain* of real time over proper time during a proper-time interval i is the total quantity of excess real time purchased during i, $\Delta t_r - \Delta t_p$. An object's *(average) rate of gain*, gr, of real time over proper time during a proper-time interval i is the ratio of gained real time to elapsed proper time, $(\Delta t_r - \Delta t_p)/\Delta t_p$. This figure is the difference between the time-travel rate and the constant no-time-travel rate of one unit of real time per unit of proper time, i.e., $gr = \Delta t_r/\Delta t_p - 1$. The gain over the entire temporal voyage is the difference between the quantity of spanned real time and the proper-time duration, $t_d - t_o - d_p$. This is also the difference between the real destination time and the intrinsic arrival time, $t_d - (t_o + d_p)$. If the proper-time duration is 5 min and the average time-travel rate is 12, so that the time traveler buys one real hour in just 5 intrinsic minutes (arriving 1 h into the future of the origin time), then there is a gain of 55 min of real time over proper time. By contrast, if the duration is 5 min and the average time-travel rate is −12, so that the time traveler arrives one hour into the past after only 5 min, then there is a loss of 65 min (a gain of −65 min) of real time over proper time. Gain rates reflect the asymmetry of time.

The analog of time-travel rate in the amusement-park ride is the rider's linear velocity relative to the real HALL OF TIME—the speed at which the rider moves through the HALL, forward (positive) or backward (negative). The analog of the gain rate is the rider's linear velocity relative to the HOME TIME position on the conveyance—the speed at which the rider moves ahead (positive) or behind (negative) the HOME TIME position, which is itself in motion relative to the HALL. If the rider simply maintains position on the conveyance, the former velocity is a ratio of 1 (inches:seconds), i.e., one inch (representing one second) of progress through the corridor for every second of rider time, while the latter velocity is 0. If the rider cycles backward atop the conveyance at two inches per second, the rider therewith travels relative to the HOME TIME position at a rate of −2. However, since the HOME TIME position is forging ahead relative to the HALL at 1, the rider therewith travels relative to the HALL at a ratio of only −1. Backward time travel is harder work than forward, because of relentless time pressure.

Time-travel rates fall into several categories, each category corresponding to a unique time-travel orientation. The most distinctive categories are: time-travel rates greater than 1; 1 itself; negative time-travel rates; those between 0 and 1; and 0 itself. These are collectively exhaustive of all time-travel rates. Among negative time-travel rates are three distinctive sub-categories: −1; those less than −1; and those between −1 and 0. Interest is typically focused on time-travel rates that are either greater than 1 or less than 0. There is *forward time travel*, whereby the time traveler gains on real time, if and only if the time-travel rate is greater than 1. If the time-travel rate is uniformly 1.33, then the time traveler travels forward 1.33 s through real time for every second of proper time. There is *backward time travel*, whereby the time traveler travels in reverse through real time, if and only if the time-travel rate is negative.

There are three *limit rates*, at which the orientation of time-travel takes on a very distinctive character: 1, 0, and −1. As we have seen, the time-travel rate is exactly 1 if and only if there is only trivial time travel. The time-travel rate is exactly 0 if and only if there is *freeze-frame time travel*, in which the time traveler remains stationary with respect to real time. This is the bleak and lonely circumstance in which time literally *stands still* for the time traveler, who is stalled at some real time. (Notice, however, that there is no real-time interval during which the time traveler is

Table 5.1 For most readers, your current time-travel rate = 1

Time-travel rate (*ttr*)	$ttr < -1$	-1	$-1 < ttr < 0$	0	$0 < ttr < 1$	1	$ttr > 1$
Time-travel orientation	Fast backward	Reverse real time	Slow backward	Freeze frame	Lag-behind	No time travel	Forward
Gain rate (*gr*)	$gr < -2$	-2	$-2 < gr < -1$	-1	$-1 < gr < 0$	0	$gr > 0$

in this peculiar circumstance.) Backward time travel in *reverse real time* is traveling to the past at a rate of −1. At this time-travel rate, ten minutes into the temporal voyage the time traveler travels exactly ten minutes into the past.

The time-travel rate is a positive fraction, more than 0 but less than 1, if and only if there is *lag-behind time travel*, in which time traveler moves forward in time but at a slower rate than the passage of real time. This is a circumstance in which the real world appears to the time traveler to be moving in slow motion, because a single second of proper time is of shorter duration that one second of real time.[6] Table 5.1 provides various time-travel orientations and their corresponding rates.

At the end of the temporal voyage, the proper and real times are distinct but the flow of the former re-synchronizes with the latter. Proper time returns to moving forward at a rate of exactly one real second per time-traveler second. The time-travel rate returns to exactly 1, the gain-rate to 0.

5.3 III

What real time is it *during* the temporal voyage? The narrative of a time-travel story typically follows proper time. In telling a time-travel story in time-traveler chronological sequential order, the storyteller implicitly relies on an important fact: There is a function, $t_r = \boldsymbol{\rho}(t_p)$, that specifies what real time it is at any particular proper time. If there is only one temporal voyage v, and the real time reached during v is given as a function of proper time, $t_r = \boldsymbol{\rho}_v(t_p)$, then $\boldsymbol{\rho}$ may be given as follows:

If t_p is earlier than the time of origin ($t_p \leq t_o$), then $\boldsymbol{\rho}(t_p) = t_p$;
if t_p is within the temporal voyage ($t_o < t_p < t_o + d_p$), then $\boldsymbol{\rho}(t_p) = t_o + \boldsymbol{\rho}_v(t_p - t_o) = t_o + ttr_v(t_p - t_o)$; and

---

[6] Special relativity has the confirmed consequence that a dilation of time relative to a frame of reference is achieved simply by moving about relative to that reference frame. Arguably, anything that is in motion relative to a reference-frame *ipso facto* has a time-travel rate less than 1 with respect to that reference-frame. It is highly relevant, however, that this phenomenon is not typically thought of or described as a form of time travel. Instead it is thought of and described as a "slowing" of the traveler's time relative to a perspective regarded as stationary. The present discussion envisions that a single frame of reference is held fixed throughout.

if t_p is the same as or later than the proper time of arrival ($t_p \geq t_o + d_p$), then $\boldsymbol{\rho}(t_p) = t_p + t_d - (t_o + d_p)$.

Upon arrival and afterward, the real time is the sum of the proper time and the gained real time. In particular, $\boldsymbol{\rho}(t_o + d_p) = t_d$, i.e., when proper time is the time of arrival, real time is the destination time.

In forward time travel ($ttr > 0$) there is a 1–1 function, $t_p = \boldsymbol{\tau}(t_r)$, that specifies what proper time it is at any particular real time. This function is simply the converse of $\boldsymbol{\rho}$. It may be given as follows:

If $t_r \leq t_o$, then $\boldsymbol{\tau}(t_r) = t_r$;
if $t_o < t_r < t_d$, then $\boldsymbol{\tau}(t_r) = t_o + (t_r - t_o)/ttr_v$;
and if $t_r \geq t_d$, then $\boldsymbol{\tau}(t_r) = (t_o + d_p) + t_r - t_d$.

When real time t_r is later than or the same as the destination time t_d, proper time t_p is later than the arrival time $t_o + d_p$ by exactly the same interval that t_r is later than the destination time t_d.

If the temporal voyage is backward, the correspondence of proper time to real time is one-many, hence not a function. This points to a significant asymmetry between forward time travel and backward. The converse of the one-many correlation is $\boldsymbol{\rho}$. For times within the negatively spanned real time, there are at least two different proper times: one before the temporal voyage and one during as seen from the inside. If the spanned real time is short enough, corresponding to a single real time t_r earlier than t_o, there can be three different proper times: one before, one during, and one after the temporal voyage. This correlation between real times and proper times is also highly systematic.

5.4 IV

The concept of proper time is integral to the concept of time travel. If the primary philosophical question concerning time travel is whether it is a metaphysical possibility, then a secondary but important philosophical question concerning time travel is: W*hat exactly is proper time?* What also is the nature of the correspondence between proper and real time? Lewis writes, "I [distinguish] time itself … from the *personal time* of a particular time traveler … the time-traveler's personal time … isn't really time, but it plays the role in his life that time plays in the life of a common person" (*op. cit.*, p. 146). Lewis's concession that proper time is not genuinely time is more serious than he recognizes. It would, if correct, exclude the very possibility of Wellsian time travel. Proper time plays the functional role of time in the analogy with spatial motion. The time traveler moves from one real time to another during an interval spanning from one proper time to a later one. Time-travel rate is defined in terms of elapsed proper time, not elapsed real time; the instantaneous time-travel rate is the time derivative of $\boldsymbol{\rho}_v$ with respect to proper time, not with respect to real time. If time travel, defined as the temporal analog of spatial motion, is genuine—if continuous nonstandard re-location in real time over proper time is a possible phenomenon—then proper time must be genuine time of some kind, and yet not real time. Since real time is just time (*simpliciter*), the very notion of time travel requires that proper time be a sub-phenomenon of real time.

Real time is made up of real times, the values of 't_r'. Proper time is made up of proper times, the values of 't_p'. If proper time is time of a certain kind—as it must be for time travel to be the temporal analog of spatial motion—then proper times are times of a certain kind. Are the values of 't_p' of a different character from those of 't_r'? For example, is the intrinsic arrival time, $t_o + d_p$, metaphysically different in nature from the real destination time, t_d? Is proper time a subjective phenomenon of some sort, a shadow of real time? Is the intrinsic arrival time, or is proper time while *en route*, perhaps an alien time, outside and orthogonal to the real timeline?

Proper time is none of these things. It is not subjective in any significant sense. Nor is it an alien time, nor metaphysically peculiar in any way.[7] If a time traveler from t_o were to arrive at t_d at an alien time, it would make no sense to attempt to place the intrinsic arrival time as earlier than, later than, or identical with, any particular real time. Proper times are times like any other. For each proper time t_p, there is a particular real time t_r such that t_p just *is* t_r. In particular, the proper time of arrival is identical with a particular real time, one that is not the real destination time. It is just the particular real time $t_o + d_p$ that is later than (subsequent to, downstream on the real timeline from) t_o by exactly the duration d_p of the temporal voyage. Contrary to Lewis, *the proper-time dimension is simply ordinary time*.

Even the proper times that elapse during the temporal voyage itself are identical with particular real times. In fact, despite *TT3*, the duration of an *en route* proper time is exactly the same as that of the real time with which it is identical. They are the very same thing. Proper times are just the real times from the departed timeframe—each one shining temporarily in sequence, then replaced, one following upon another, with real presentness flowing forward in step as if nothing remarkable has just taken place. Indeed, nothing remarkable happens intrinsically to the times themselves; it is the time traveler that undergoes a remarkable process. The time traveler's proper timeframe deviates from the timeframe of origin. Despite the time traveler's deviation from the original timeframe, that timeframe itself is unscathed; time continues to flow even if un-graced by the time traveler's presence. Both during and after the temporal voyage, the departed timeframe is identical with the traveler's proper time.

Proper times are also real times. Whether a time is designated 'real' or alternatively as 'proper' depends on whether it is treated as the value or the argument of

---

[7] Lewis, *op. cit.*, argues that intrinsic ("personal") time is not a temporal dimension orthogonal to real ("external") time (p. 145). His instinct is correct but his argument is fallacious. Contrary to his major premise, one who travels backward in time to visit with childhood friends is at the same real time as his/her friends but his/her proper time is then indeed different from theirs. Lewis's claim that proper time "isn't really time" is also importantly incorrect. What are different are the time traveler's relations to time.

ρ. Although every proper time is a real time, if Wellsian time travel is genuine, then the correspondence **ρ** between a proper time and a real time is not strict identity.

If proper times are just real times, and yet *TT3* is a fundamental truism about time travel, is time travel then metaphysically impossible? On the other side of the coin, if time travel is possible and governed by *TT3*, how can proper times be real times?

As a prelude to answering these questions, we must first engage in a different line of inquiry. What timeline segment does the time traveler persist in immediately following the temporal voyage? Is it a proper timeline, to be treated as an argument to **ρ**? Is it a real timeline, to be treated as a value of **ρ**? Or is the time traveler at two different timelines, both proper and real time simultaneously, i.e., both at a single real time?

In absolute time—alternatively, within a single frame of reference—the answer must be that the post-voyage backward time traveler persists in real time and not in proper time. Otherwise the time traveler would remain in the original timeframe; there would be no genuine time-travel *departure*, as such. Proper time is a logical construct generated by the temporal order and rhythm of the time traveler's "clocks" (periodic processes of deterioration and the like)—a genuine timeframe, to be sure, but one through which the backward time traveler does not persist at all (unless the time traveler does so independently of the temporal voyage, e.g., by engaging in a separate temporal voyage). The time showing on the time traveler's personal timepiece upon arrival is a reading of his/her proper time.

What timeline segment does the backward time traveler persist in *during* the temporal voyage? Here again, the answer has to be that the backward time traveler is at real time, even if proper time is dilated or contracted relative to real time. Indeed, each time during the voyage may be seen as a layover arrival/destination time. The layover is of the best kind: very short but not too short. (It is in fact instantaneous, but that is long enough.) If the time traveler passes through the Roaring 20's on the way to 1,000,000 B.C., then the traveler existed during the Roaring 20's, even though (as the traveler might say) "it went by very fast … and backward".

This is more or less how time travel is depicted in *The Time Machine*. The Time Traveler witnesses the world change (or return to its pre-change state) before his very eyes. But there is a flaw in Wells' reasoning about Wellsian time travel. The Time Traveler is there—or rather, he is *then*—at each real time through which he passes. If he is indeed "then", why is it that the people around him at that very time do not see him there? He sees them; they should see him as well. Unless something very strange is occurring, they would indeed see him—and he would appear to be behaving peculiarly, moving and talking very slowly … and backward.

From t_o forward, proper time is just the real time of the time-traveler's original timeframe. Proper time corresponds to the position marked 'HOME TIME' on the amusement-park ride. Proper time *is* home time, the very timeframe that the time traveler left behind. It is analogous to the international traveler who upon arriving at LAX says "I'm on Asian time". Proper time is the temporal analog of jetlag. It is *time-machine lag*. Proper time is evidently a construct of a certain kind, based entirely on the progressively changing states of the time traveler's intrinsic temporal processes.

In labeling a time 'proper' rather than 'real', we are not distinguishing it in its nature from time as we know it. In particular, we are not positing a subjective, shadowy, parallel temporal dimension orthogonal to real time. Rather, we are marking time off in terms of the time traveler's relations to it. The rationale for calling a time 'proper' rather than 'real' concerns the relation that the time traveler bears toward it. The (nontrivial) time traveler is at a particular time, and therewith bears a different relation to a different time, the latter time being labeled the time traveler's 'proper time'. Let us say that the time traveler is *at* real time and *on* proper time. The time traveler is *on* one time when *at* a different time; the time the traveler is at is called 'real time'; the time the traveler is on is called 'proper time'. Everything is on its proper time. Most of us are also on real time. The backward time traveler is not at proper time when on it. Asian time is a real time, but the jetlagged traveler is not at it. Analogously, *proper time is time, but the post-voyage backward time traveler is not at it.*

If something is travelling through time at a rate of 60, then one second of proper time buys 1 min of real time. A second of time lasts exactly 1 s, a minute exactly 1 min, an hour exactly 1 h. One must resist the temptation to think that the time traveler is at an hour for only a minute. Rather, the time traveler remains *on* a single minute while *at* an entire hour. The time traveler is at a time even if the intrinsic clocks are on a different timeframe.

5.5 V

Someone who is about to travel back to 1,000,000 B.C. *was there already*—or rather, the time traveler was *then* already. The soon-to-be time traveler already existed in prehistoric times. The pre-voyage time traveler visited the prehistoric world in pre-history, when it was present (in the non-indexical sense). The time traveler is about to cease to be. In general, if a soon-to-be time traveler is "about," in proper time, to travel to a past time t_d, then the traveler was *already* at t_d before the temporal voyage. Moreover, since the soon-to-be backward time traveler is really about to *depart* from the current real timeframe, the time traveler is not about to arrive at t_d; the traveler is about to cease to exist. Rejecting either of these truisms leads to a serious misunderstanding of what time travel is supposed to be. More generally, we have the following additional fact about any backward temporal voyage:

TT4 The backward time traveler does not exist at the particular time $t_o + d_p$—unless the traveler either exists at an earlier time and simply persists to $t_o + d_p$, or embarks on a separate temporal voyage and travels to $t_o + d_p$ (or both).

We do not make the claim about forward time travel analogous to *TT4*. Real-time $t_o + d_p$ is the time that the time traveler is on upon arrival at t_d. Where $ttr > 1$ and $d_p > 0$, $t_o + d_p < t_d$ and therefore,

$$\boldsymbol{\tau}\left(t_o + d_p\right) = t_o + d_p/ttr < t_o + d_p.$$

That is, in forward time travel, if the intrinsic duration d_p is non-zero, then $t_o + d_p$ precedes the destination time t_d, and therefore the proper time corresponding to real-time $t_o + d_p$ precedes $t_o + d_p$ itself. The proper time corresponding to $t_o + d_p$ is still within the temporal voyage.

TT5 The forward time traveler whose time-travel rate is greater than 1 is *en route* to the destination time t_d at the particular time $t_o + d_p$.

This is in keeping with *TT3*. The contrast between *TT4* and *TT5* points to another asymmetry between forward time travel and backward.

Putting *TT2* together with *TT4* and *TT5*, we obtain a puzzling result: At the proper arrival time $t_o + d_p$, the fast-forward time traveler is not yet at the real destination time t_d; worse yet, the backward time traveler does not even *exist* at $t_o + d_p$ (except in special circumstances and for independent reasons). Either way, *the time traveler makes it to the real destination time, but does not arrive there at the proper arrival time*. How is this possible?

The seemingly bizarre fact is made possible by the distinction between being at a time and being on it. When the time traveler is on t, the traveler is at $\boldsymbol{\rho}(t)$. Whether the traveler is at a time that the traveler is on is another matter. The time traveler is on the proper arrival time at the real destination time, but is not then at the proper arrival time. At the proper arrival time the time traveler is not on it. The backward time traveler is not at the proper arrival time at all (again, except in special circumstances).

The phenomenon generally referred to as 'time travel' is so-called only insofar as it is viewed from the perspective of the agent undergoing the process. From this perspective proper time is time *simpliciter*. But from the perspective of real time—the default perspective—it emerges that the phenomenon in question is not objectively temporal re-location. It is only temporal re-location when seen from the inside. As seen from the objective observor's contrasting vantage point, the object undergoing the process is undergoing a change, but not a change of temporal location (other than ordinary persistence). Rather, *the change is a de-synchronization of the object's intrinsic temporal processes with the passage of real time*. Viewed from the perspective of spectator, the object's intrinsic "clocks" (aging, deterioration, etc.) are running slow, or fast, or backward, or jumping instantaneously. Importantly, it is a phenomenon intrinsic to the object undergoing the process. What looks from the inside like temporal re-location is, as seen by the spectator, a diachronic recalibration of the object's temporal processes.

An object's *(average) rate of passage of proper time* over a real-time interval i is the ratio of proper time spanned during i (a positive or negative quantity) to elapsed

Table 5.2 For most readers, your current proper-time passage rate = 1

Proper-time passage rate (*ptr*)	$ptr < -1$	-1	$-1 < ptr < 0$	0	$0 < ptr < 1$	1	$ptr > 1$
Proper-time orientation	Fast backward	Benjamin button	Slow backward	Suspended	Slow	Perfect time	Fast
“Time-travel” orientation	Slow backward	Reverse real time	fast backward	Instantaneous	Forward	No time travel	Lag-behind

real time (the positive quantity of time in *i*), $\Delta t_p/\Delta t_r$. Likewise, an object's *instantaneous rate of passage of proper time* is the limit of $\Delta t_p/\Delta t_r$ as Δt_r approaches 0. The proper-time passage rate is thus the reciprocal of the (so-called) time-travel rate, the instantaneous proper-time passage rate the reciprocal of the instantaneous time-travel rate. An object is on real time if, and only if, its proper-time passage rate is 1. The spectator's perspective yields an inversion of the time-traveler's perspective. Table 5.2 provides various proper-time orientations with respect to real time, their corresponding passage rates, and corresponding “time-travel” orientations.

From the passive spectator's perspective, and from the perspective of the non-participant, so-called time travel is ordinary persistence, or discontinuous persistence, coupled with very peculiar phenomena—walking, talking, processing, aging, and deteriorating too slowly, or too quickly, or too abruptly, or too backward. In particular, from the spectator's vantage point the phenomenon of so-called backward time travel is persistence together with both a reversal of the subject's intrinsic processes and a philosophical profanity: backward causation. A backward time traveller's presence in 1912 was caused by events that would transpire years later, well after their effect. The causation is forward in proper time but backward in real time.

On this re-conceptualization of time travel, the notion of *freeze-frame time travel* corresponds to an instantaneous jump in proper time. The corresponding proper-time passage rate is division by zero, hence undefined. For the freeze-frame time traveler it is *as if* time were passing, yet this does not occur during any interval of real time. In effect, while *en route* to t_d the freeze-frame time traveler is ejected from the flow of time into a temporal limbo.

One possible phenomenon often referred to as ‘time travel’ is perfectly coherent, including backward time travel. There is no logical inconsistency in the idea of an individual's intrinsic temporal processes running backward. However, so-called backward time travel, construed as taking place over real time, involves an individual popping into existence *ex nihilo*. This conflicts with accepted natural laws. Worse, in some cases, the popping into existence occurs at a time t_d prior to the individual's birth. Even if this is logically coherent, it is arguably quite impossible metaphysically.

Philosophers debate whether a possible phenomenon should be regarded as time travel, as normally understood.[8] Yet it appears from the foregoing that there is a single metaphysically possible phenomenon that may be equally legitimately regarded as time travel or, alternatively, as the de-synchronization of an object's intrinsic clocks with real time, depending only on perspective. Whether one describes the phenomenon in question as a change in real time with respect to proper time, or instead as a change in proper time with respect to real time, in some sense the same facts are captured. The two descriptions are equivalent—two sides of the same coin, six of one and a half-dozen of the other, two inversions of the same chord, two *Sinne* converging on the same *Bedeutung*.

In many contexts, and especially in science fiction, the description in terms of time travel is favored over the other, although the other is generally the less misleading description. With spatial motion—continuous change in position from one spatial location to another over time—the spatial traveler is at the former location at one time and at the latter location at a later time, having traversed a path between the two locations in the interim. This appears to be integral to the very concept of *motion* (*re-location*; *change in position*). If time travel is the temporal analog of spatial relocation, then by analogy the time traveler is at one temporal location t_o at one time t, and at a different temporal location t_d at a time t' subsequent to t. But that much is true of everything that persists from t_o to t_d (or the other way as the case may be), time traveler or no, since t_o and t_d are themselves such a pair of times $<t, t'>$. If something x exists at a time t, then at that very time t, x is at t. There is no more relocation-to-a-different-time occurring with the intrepid time traveler than occurs with the guy flipping burgers at the local eatery. A so-called time traveler may be on proper time, but the traveler persists in real time.

As defined, time travel is not the temporal analog of spatial motion. Talk of 'time travel' relies heavily on shifting from real time to proper time at crucial junctures. We tend to think of proper time simply as time and of the destination time as a different place. These modes of thinking are confused. The destination time is not a place; it is a real time. Proper time is also time, but the so-called time traveler is merely on it. Traveling from a timeline segment beginning with t_o, to another beginning with t_d, the time traveler is at the former timeline segment when on t_o, and at the latter timeline segment when on $t_o + d_p$. But the forward time traveler is still *en route* to t_d when at $t_o + d_p$ and the backward time traveler does not even exist at $t_o + d_p$. In either case the time traveler exists at the corresponding real destination time, t_d, but in either case the time traveler also persists all the while from t_o to t_d, or from t_d to t_o.

Authentic Wellsian time travel requires a more full-blooded, non-metaphorical notion of two-timing. Insofar as spatial motion is relocation in space over time, authentic time travel is relocation in time over time at a rate other than 1. What is

---

[8] Hilary Putnam in "It Ain't Necessarily So," *The Journal of Philosophy*, 59, 22 (October 25, 1962), at Sect. 4; Robert Weingard, "On Traveling Backward in Time," *Synthese*, 24 (1972), pp. 117–132.

commonly referred to as ‘time travel’, both in science fiction and in science proper, is *simulated* time travel. The simulation is accomplished by treating real time as space while treating proper time as real time, pretending that the protagonist is at a time that in fact, or within the story, the traveler is merely on, not at. The storyteller or theorist completes the charade by letting the narrative follow proper time instead of (or in addition to) real time, while relegating real time to a kind of spatial dimension rather than temporal. The subject is depicted as changing “temporal location over time.” This is not the same thing as the subject’s proper time merely being out of sync with real time. In authentic Wellsian time travel there must be two non-overlapping times such that when the time traveler is at one of them, the traveler is not only on but also somehow *at the other*. At real time t_r, the traveler is not only on t_p, by virtue of the traveler’s intrinsic clocks being set to t_p, but also somehow *at* t_p *in addition to being at* t_r. At one time the traveler is at two times. But it is unclear what this could be. A new notion of *being at a time* would be required, one that allows for something to be at two non-overlapping times at one of them, and so at one time at a non-overlapping time. But this seems precluded by the very concept and logic of *being at a time*. This fact challenges the very conceptual possibility of Wellsian time travel.

Chapter 6
Why Spacetime Has a Life of Its Own

James Robert Brown

Thinking about the nature of space and time goes back to antiquity, but I will start with Isaac Newton (1642–1727). He famously wrote in his *Principia*, "Absolute, true, and mathematical time, in and of itself and of its own nature, without reference to anything external, flows uniformly and by another name is called duration.... Absolute space, of its own nature without reference to anything external, always remains homogeneous and immovable.... Absolute motion is the change of position of a body from one absolute place to another..." (Newton, *Principia* Scolium to definition 8)

The bucket thought experiment was offered as evidence for this view. In stage I in the following Fig. 6.1 the water and bucket are at rest with respect to one another and the water surface is flat. In stage II the water and bucket are rotating with respect to one another. In stage III the water and bucket are back to being at rest with respect to one another, but the surface is now concave. How do we explain the difference between I and III? The difference, says Newton, is that the water-bucket system is at rest with respect to space in I and is rotating with respect to space in III. Thus, he concludes, space exists. This is a standard form of explanation, used widely in the sciences. If the germ theory of disease is the best explanation for various forms of illness, then germs exist. If space is the best explanation for inertial motion, then space exists.

There are several ingredients involved in Newton's conception of absolute space and time.

1. Space (as well as time) is a substance, an entity in its own right. Consider a basket of apples. The basket is just as real as the apples in it. Space is taken to be a container, like a basket, and every bit as real. Of course, like any analogy, it can't be pushed too far. A basket is a material object like an apple, and one basket could

---

J.R. Brown (✉)
Department of Philosophy, University of Toronto, Toronto, Canada
e-mail: jrbrown@chass.utoronto.ca

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_6

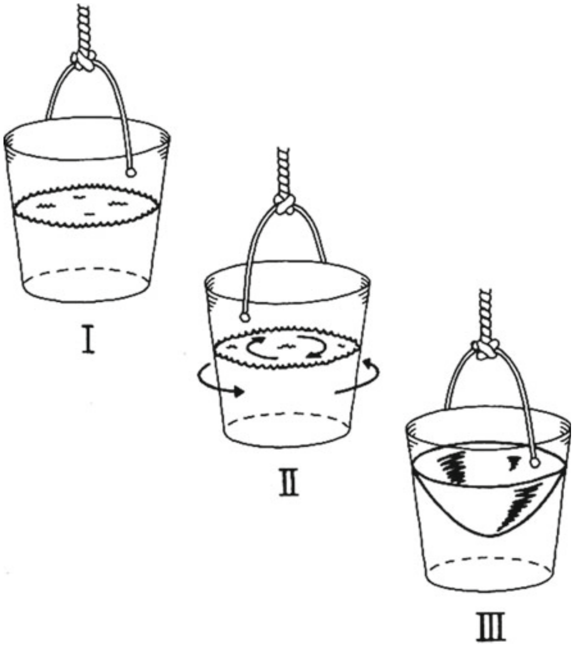


Fig. 6.1 Newton's bucket thought experiment

be put in another. Space is not a material object and is not itself in space. Space consists of points which are possible locations of bodies, but these points exist whether they are occupied or not.

2. Every point of space and every point of time is exactly like every other point. As we move west, for instance, the points of space are identical to those we occupied earlier—except for their location.
3. Location in space (and in time) is absolute, as is velocity. Thanks to Galilean relativity, however, the laws of nature are the same in any frame, as long as the frame is moving inertially.
4. Absolute location and velocity are not detectable. However, acceleration (non-inertial motion) is observable. Newton's famous bucket thought experiment is tied to this and is presented as evidence for the existence of absolute space.
5. Space is the source of inertia. Bodies, according to Newton's first law, tend to move in straight lines or remain at rest unless acted upon by a force. Space is causally responsible for inertial motion.

The correspondence between Samuel Clarke (1675–1729) and Gottfried Wilhelm von Leibniz (1646–1716) is one of the great scientific exchanges of all time. Clarke was a close associate of Newton, and it is widely believed Newton had a big hand in Clarke's side of the exchange. Leibniz and Clarke argued about many things including the principles of dynamics and the nature of God, but the chief topic of interest was the nature of space and time. Clarke upheld Newton's absolutism while Leibniz attacked it and offered his relational account in its place. He put his relationalism succinctly, saying, "I hold space to be something merely relative, as time is; that I hold it to be an order of coexistences, as time is an order of successions." (Leibniz, *Leibniz-Clarke Correspondence*, letter III, Sect. 4, p. 25f)

Leibniz attempted to refute any sort of absolutist view with a remarkable argument based on symmetry considerations. First he posits the Principle of Sufficient Reason: Nothing happens without a reason why it should be so and not otherwise. He then points out that there can be no reason why the universe is located where it is in absolute space and not at some other place, say a meter to the right, since there could be no reason for one location rather than the other. He takes this to be a *reductio ad absurdum* of Newton's account. Imagine a change of location for the whole universe, which is sometimes called a "Leibniz shift."

> ...if space were an absolute being, there would something happen for which it would be impossible there should be a sufficient reason. Which is against my axiom. And I prove it thus. Space [according to Newton] is something absolutely uniform; and without the things placed in it, one point of space does not absolutely differ in any respect whatsoever from another point of space. Now from hence it follows, (supposing space to be something in itself, besides the order of bodies among themselves,) that 'tis impossible there should be a reason, why God, preserving the same situations among bodies among themselves, should have placed them in space after one certain particular manner, and not otherwise; why every thing was not placed the quite contrary way, for instance, by changing East into West. (*Leibniz-Clarke Correspondence*, III-5, p. 26)

The two universes would be completely symmetrical. There could be no reason to favour the creation of one over the other, so God would have no reason to choose. But God only acts for a reason and the universe does exist, so Newton's claim can't be a correct account of how things are. The existence of absolute space leads to this absurdity. Note that this argument can be run considering absolute location (no difference between being located here or there) or considering absolute velocity (no difference between moving west with a velocity v or $2v$, $3v$, etc.) Whether this sort of argument also works for accelerations is questionable. Leibniz, unfortunately, did not address this case in any detail before he died.

The problem with Newton's absolute space, as Leibniz sees it, is easily solved, if we simply hold that the only thing that exists (and the only thing that matters) is the set of relations among the bodies.

> But if space is nothing else, but that order of relation; and is nothing at all without bodies, but the possibility of placing them; then those two states, the one such as it now is, the other supposed to be the quite contrary way, would not at all differ from one another. Their difference therefore is only to be found in our chimerical supposition of the reality of space in itself. But in truth the one would exactly be the same thing as the other, they being absolutely indiscernible... (*Leibniz-Clarke Correspondence*, III-5.)

Leibniz's Principle of Sufficient Reason may seem inappropriate since it's based on religious considerations, which today should play no role in these matters. It could readily be recast, it should be noted, in the form of a symmetry principle, devoid of any religious overtones.

Like Leibniz, Ernst Mach (1838–1916) is a relationalist, but he comes at it a different way. He holds that all legitimate concepts are acquired empirically and demands that scientists stick to the "actual facts," by which he means the facts of observable experience. Newton's absolute time, for example, is independent of all change, so Mach is completely dismissive. "This absolute time can be measured with no motion; it has therefore neither a practical nor a scientific value; and no one is justified in saying that he knows aught about it. It is an idle metaphysical conception." (Mach, *The Science of Mechanics*, 273).

As for absolute space, it is both philosophically illegitimate and quite unnecessary for science: "No one is competent to predicate things about absolute space and absolute motion; they are pure things of thought, pure mental constructs, that cannot be produced in experience. All our principles of mechanics are... experimental knowledge concerning the relative positions and motions of bodies." (Mach, *The Science of Mechanics*, 280) Mach claims that all observable motion is relative. This is the only type of motion that is comprehensible, and hence, the only type that is legitimate in science.

Before Mach's attack on Newton's absolute space and time, the empiricist George Berkeley (1685–1753) formulated a similar view. I include mention of Berkeley here, since his historical importance in these matters is under appreciated.

> ...it is clear that we ought not to define the true place of the body as the part of absolute space which the body occupies, and true or absolute motion as the change of true or absolute place; for all place is relative just as all motion is relative. But to make this appear more clearly we must point out that no motion can be understood unless besides the body in motion our own body also, or some other body, be understood to exist at the same time. For up, down, left, and right and all places and regions are founded in some relation, and necessarily connote and suppose a body different from the body moved. (Berkeley, "*De Motu*" §58.)

In replying to Newton's bucket argument, Mach formulated what is now known as "Mach's Principle." One version of this asserts that only relative rotations are relevant in accounting for this type of phenomena. The bucket rotating with respect to the "fixed stars" is equivalent to the fixed stars rotating with respect to the bucket. His view comes out in the following long passage.

> ...if we take our stand on the basis of facts, we shall find we have knowledge only of relative spaces and motions. Relatively, not considering the unknown and neglected medium of space, the motions of the universe are the same whether we adopt the Ptolemaic or the Copernican mode of view. Both views are, indeed, equally correct; only the latter is more simple and more practical. The universe is not twice given, with an earth at rest and an earth in motion; but only once with its relative motions, alone determinable. It is, accordingly, not permitted us to say how things would be if the earth did not rotate. We may interpret the one case that is given us, in different ways....
>
> Newton's experiment with the rotating vessel of water simply informs us, that the relative rotation of the water with respect to the sides of the vessel produces no noticeable centrifugal forces, but that such forces are produced by its relative rotation with respect to the mass of the earth and the other celestial bodies. No one is competent to say how the experiment would turn out if the sides of the vessel increased in thickness and mass till they were ultimately several leagues thick. (Mach, *The Science of Mechanics*, 283f.)

Back to Berkeley for a moment. In summing up his view, he remarks that we ought to "...consider motion as something sensible, or at least imaginable; and to be content with relative measures. If we do so, all the famous theorems of the mechanical philosophy by which the secrets of nature are unlocked, and by which the system of the world is reduced to human calculation, will remain untouched..." (Berkeley, "*De Motu*," §66.)

In short, we can have all the glories of Newton's physics without absolute space. Everything we want and need in physics is available to us, according to both Berkeley and Mach, without (as they see it) the metaphysically murky medium of absolute space and time in which objects exist and events happen. But is this true?

Albert Einstein (1879–1955) certainly thought we could do physics without it. In the opening pages of his paper on General Relativity (1916) he clearly endorses the Berkeley-Mach point of view. He begins with the remark that in classical mechanics there is an "epistemological defect...pointed out by Ernst Mach." (Einstein, "The Foundation of the General Principle of Relativity," Einstein, et. al., *The Principle of Relativity*, 1916, p. 112) Einstein then describes a thought experiment with two globes that are in observable rotation with respect to one another. One is a sphere, the other an ellipsoid of revolution (Fig. 6.2).

Einstein asks, "What is the reason for the difference in the two bodies?" He then sets empiricist—indeed, verificationist—conditions on any acceptable answer. His empiricism and verificationism leads directly to Mach's principle and the principle of general covariance.

> No answer can be admitted as epistemologically satisfactory, unless the reason given is an observable fact of experience. The law of causality has not the significance of a statement as to the world of experience, except when observable facts ultimately appear as causes and effects.

Einstein then declares that classical physics is not up to proper epistemological standards. That is, it does not comply with a strict Machine empiricism.

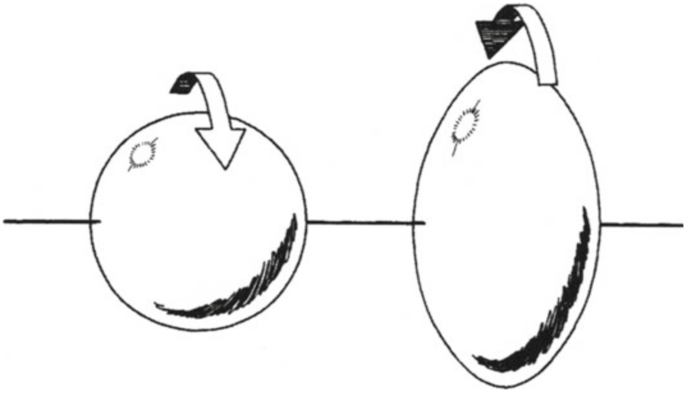


Fig. 6.2 Einstein's rotating spheres thought experiment

> Newtonian mechanics does not give a satisfactory answer to this question. It pronounces as follows:—The laws of mechanics apply to the space $R1$, in respect to which the body $S1$ is at rest, but not to the space $R2$, in respect to which the body $S2$ is at rest. But the privileged space $R1$ of Galileo, thus introduced, is a merely factitious cause, and not a thing that can be observed. It is therefore clear that Newton's mechanics does not really satisfy the requirement of causality in the case under consideration, but only apparently does so, since it makes the factitious cause $R1$ responsible for the observable difference in the bodies $S1$ and $S2$.

By "the privileged space $R1$ of Galileo," Einstein means an inertial frame, which plays a central role in Newtonian physics and in special relativity. But he no longer considers it satisfactory. Einstein then goes on to say how things should be properly viewed, introducing both Mach's principle and the principle of general co-variance.

> The only satisfactory answer must be that the physical system consisting of $S1$ and $S2$ reveals within itself no imaginable cause to which the differing behaviour of $S1$ and $S2$ can be referred. The cause must therefore lie *outside* this system. We have to take it that the general laws of motion, which in particular determine the shapes of $S1$ and $S2$, must be such that the mechanical behaviour of $S1$ and $S2$ is partly conditioned, in quite essential respects, by distant masses which we have not included in the system under consideration. These distant masses and their motions relative to $S1$ and $S2$ must then be regarded as the seat of the causes (which must be susceptible to observation) of the different behaviour of our two bodies $S1$ and $S2$. They take over the role of the factitious cause $R1$. Of all imaginable spaces $R1$, $R2$, etc., in any kind of motion relatively to one another, there is none which we may look upon as privileged a priori without reviving the above-mentioned epistemological objection. *The laws of physics must be of such a nature that they apply to systems of reference in any kind of motion.* (Einstein, ibid., 112f. Einstein's italics throughout.)

In spite of Einstein's authority, the later years of the 20th century saw a shift to Newton's outlook. Not in all aspects, but substantivalism became the dominant outlook. A strong consideration in favour of substantivalism stems from inertia. It is related to Newton's bucket argument in that acceleration and inertial motion are the key to absolute space. It is general enough to apply to general relativity (GR) as well. The argument, (here adapted from [1]), runs something like this:

1. Current physical theories employ the notion of inertia, which is of central importance to them. Thus, both classical mechanics and GR have as a fundamental principle that bodies move inertially unless acted upon by a force.
2. Inertial motion involves the idea of a straight line (or geodesic). Of course, "straight line" may have different properties in Euclidean and in non- Euclidean geometry. But in any case, the notion of a straight line is needed by various contemporary theories, including both Newton's and Einstein's. Similarly, the notion of a curved trajectory, one that deviates from straightness or from being a geodesic, is needed for cases when a force is present.
3. To say there are straight lines or geodesics or that there is an inertial structure, is to say some thing has these properties. This thing is obviously spacetime itself.
4. Thus, spacetime exists and is represented well in GR by a pseudo-Riemannian manifold.

Substantivalism asserts the independent existence of this manifold. Arguments such as the one just sketched carried the day for quite a while, but they were dealt a body blow by the *hole argument* [2], which I will now sketch.

Assume that we have a model $(M, g_{\mu\nu}, T_{\mu\nu})$, that is, a solution of Einstein's field equations. Then a new solution can be generated by performing an active diffeomorphism to generate the solution $(M, g'_{\mu\nu}, T'_{\mu\nu})$. For a point p in M, it is not the case that that $g_{\mu\nu}(p) = g'_{\mu\nu}(p)$. This is another model of the universe that is empirically indistinguishable from the original. And yet they are objectively different, according to many substantivalists.

Now imagine a region of spacetime, the hole, that is part of the whole universe. Next imagine a material object—an atom or a galaxy, it doesn't matter—moving through the hole. This will happen in our first model, $(M, g_{\mu\nu}, T_{\mu\nu})$. The second model is just like the first outside the hole, but inside it is $(M, g'_{\mu\nu}, T'_{\mu\nu})$. There are two distinct trajectories through the hole. Which one did the galaxy follow? There is no empirical test. It gets worse. GR is taken to be a deterministic theory, but it won't be if we adopt a substantivalist outlook. Determinism is characterized as follows: Two models governed by the same laws and having the same initial conditions, will have the same history. Otherwise the laws and the models are nondeterministic.

Our two models are governed by the same laws, GR, and have the same initial conditions up to the hole, but they diverge after that. The trajectory through the hole is not fixed by the laws and initial conditions. According to Earman and Norton, the price we pay for substantivalism is too high. We must embrace a ridiculous form of indeterminism. So, they conclude, substantivalism must be abandoned. The two models of the world, $(M, g_{\mu\nu}, T_{\mu\nu})$ and its diffeomorphic image $(M, g'_{\mu\nu}, T'_{\mu\nu})$, should be taken to be physically the same world, with merely different mathematical representations. In other words, they uphold so-called Leibniz equivalence and reject substantivalism.

There have been a number or responses to the hole argument. I will not attempt to survey them but will mention one. Earman and Norton take the manifold M to be the obvious counterpart to Newton's absolute space. If substantivalism is true, then M must exist in its own right and not be some system of relations. But what about the metric tensor, $g_{\mu\nu}$? Shouldn't spacetime be the manifold *and* the metric tensor, $(M, g_{\mu\nu})$, not just M alone? After all, Newton's absolute space included the standard Euclidean metric, so why not include the metric in the updated GR version? This is hugely important, since the hole argument will fail without a different metric tensor $g'_{\mu\nu}$. The diffeomorphic copies would be merely different mathematical representations of the same physical situation. Earman and Norton insist on the bare manifold M alone, arguing that the metric tensor is in an important sense more like a physical entity, possessing physical properties that make it quite unlike the familiar distance function of Euclidean space.

They are quite right, the metric is a very different thing in the space and time of Newtonian from what it is in the spacetime of GR. The debate has changed, though there is still considerable overlap with the pre-GR debates.

It is important to be completely clear on what is meant by substantivalism, especially after the hole argument. It was described when discussing the history of the problem, but now I need to provide more detail, beginning with the common term "absolutism." Absolutism about space could in principle include any or all of the following ingredients, some of which have been overturned by GR.

- *Metric*: Distance is independent of the amount and distribution of matter. (Overturned by GR; the presence of mass-energy affects distance.)
- *Geometry*: Space is Euclidean or some sort of non-Euclidean, but whatever geometry it has is independent of matter. (Overturned by GR; the presence of mass-energy affects curvature.)
- *Dimension*: The number of dimensions of spacetime (four) is independent of matter. (Possibly correct; it seems independent of mass-energy distribution.)
- *Topology*: The universe could be like a four-dimensional sphere, or like a four-dimensional doughnut, etc., but whatever it is, it is independent of matter. (Possibly correct; it seems independent of mass-energy distribution.)
- *Substance*: Space (or spacetime) is a substance, a thing in its own right. The other items in this list, metric, geometry, and so on are properties of this substance, space. These spatial properties might arise due to the amount and distribution of matter (mass-energy), but space (spacetime) itself has a life of its own, independent from everything else.

We are here concerned only with the last listed sense of absolute: Spacetime is a substance. Is this true? The argument above based on considerations of inertial structure says yes. The hole argument says no. We seem to be at a standoff. I am now going to turn to a new consideration, the idea of *metric expansion*, to see what answer it might suggest.

It is a generic feature of current cosmological models that the universe we inhabit is expanding. This is clear already in the traditional FLRW (Friedman-Lemaître-Robertson-Walker) models of GR. Nowadays it is also commonly believed that at the very early stages of the universe an even more extreme expansion occurred, known as cosmic inflation [3]. According to cosmic inflation, the very early universe expanded with extraordinary speed. A region around the Planck-scale expanded to something macro-sized. In a mere 10^{-36} to 10^{-32} s the universe grew by a factor of about 10^{26}. Cosmic inflation may be the most dramatic example of the general phenomenon known as metric expansion. Inflation is rejected by some critics, but the more general phenomenon of cosmic expansion is widely accepted; this is the ongoing expansion of the universe. Not only is the universe expanding, but it is expanding at an accelerating rate. Note that this is the expansion of space, not spacetime.

Since metric expansion is a significant feature of GR, associated with dark energy as well as cosmic inflation of the very early universe and even so-called eternal inflation, questions of its consequence for long-standing philosophical issues should be of interest to everyone.

If we take the claim literally that "space itself is expanding," then it suggests that space is a thing in its own right. Physicists speaking in a popular context offer remarks such as these:

> Brian Greene: What Einstein actually theorized is that no object can move through space faster than the speed of light. But nothing in his ideas prevents objects from riding the swell of space itself, and in that way, moving apart faster than the speed of light.
>
> Leonard Susskind: If the entirety of space is being swept away from you like the tide going out, then ripples on space very, very far away from you can appear to be going faster than the speed of light because all of space is being swept along with the tide. (Both quoted in a radio broadcast, [4].)

Of course, these are remarks in a popular context; as such they can be no more than suggestive. But what they do suggest is that: metric expansion implies substantivalism. That is, expanding space must be a substance, an entity in its own right. The reason is that two particles that are close at some time early in expansion might turn out to be far apart at a later time, and the time it took to separate them was such that the speed of separation is much greater than the speed of light. An observer at particle *a* watches particle *b* receded faster and faster until it finally disappears. The disappearance is not because the light is too dim, but because the light cones of *a* and *b* will not overlap in the future.

How is this possible, given the correctness of GR with its limiting speed *c*? The standard reply is simple: The particles are not moving apart in space; rather it is space itself that is expanding, so there is no conflict with relativity. Once we have accepted this claim, the conclusion seems inevitable: Space exists independent from matter.

It should be stressed that not everything expands along with space. There is a popular pedagogical model of the big bang expanding universe in which we start with a deflated balloon, use a marker to draw figures of various sizes and shapes on it. As the balloon is inflated, the figures on it grow at the same rate as the ballon expands. This is the wrong model of metric expansion. A better model to capture metric expansion is to glue small objects, say coins, onto the balloon. Now blow it up. The balloon grows, but the coins remain the same size ([5] suggest this analogy). This is what happens with metric expansion; space expands, but not the objects in space. The internal forces holding the objects together are more important than the expansion of spacetime. Local physics is unaffected by expansion. Even for rather large cosmological structures such as galaxies the gravitational forces holding them together are more important than the expansion. Hence, structures as large as galaxies can maintain their size despite cosmic expansion.[1]

The argument to follow is adapted from Earman's argument above. His case for substantivalism is based on considerations of inertia. The case for a similar substantivalist conclusion is based on metric expansion.

1. Current cosmology employs the notion of metric expansion. This notion is of central importance in explaining various features of the observable universe within the framework of GR (e.g., inflation, dark energy).
2. Metric expansion involves the idea of the universe changing size without material objects themselves changing size. That is, the universe could double while an atom or a galaxy remains the same size.

---

[1] At least, this is the standard view, but it is controversial. See [6] for a brief survey of rival views.

3. If the universe expands but there is no material object expanding and there is no rearrangement of material objects relative to one another, then something non-material expands. This something is obviously space.
4. Thus, space exists as an entity in its own right. It is not a system of relations among material objects, nor is it a mere mathematical representation or some sort of useful fiction. In short, substantivalism is correct.

This is a new argument for substantivalism. It is not based on metaphysical principles but rather on detailed aspects of prevailing physics. Current cosmological thinking is carried out in the framework of GR, which includes metric expansion as an essential part. This is currently our best bet at how things are. If the universe were different, then the augment could fail.

I should perhaps soften this argument to a disjunction: You can take it as an argument for substantivalism OR you can take it as a new challenge for relationalists to overcome. I'm betting on substantivalism. Space is a thing in its own right; it has a life of its own.

Acknowledgements The parts of this essay concerning metric expansion stem from joint work with Keizo Matsubara and Lee Smolin. I am grateful to the editors for their invitation and for comments on an earlier draft.

References

1. Earman, J. *World enough and spacetime*. Cambridge, MA: MIT Press.
2. Earman, J., & Norton, J. (1987). What price spacetime substantivalism: The hole argument. *British Journal for the Philosophy of Science*, *38*, 515–525.
3. Guth, A. (1981). The inflationary universe: a possible solution to the horizon and flatness problems. *Physical Review D*, *23*, 347. OCLC 4433735058.
4. Krasny, M. (2006). Retrieved February 1, 2016 from http://www.kqed.org/a/forum/R201403190900.
5. Misner, C., Thorne, K., & Wheeler, J. (1974). *Gravitation*. New York: Freeman.
6. Bonnor, W. B. (2000). Local Dynamics and the Expansion of the Universe. *General Relativity and Gravitation*, *32*(6), 1005–1007.

Chapter 7
The Phenomenology of Space and Time: Husserl, Sartre, Derrida

Hans Herlof Grelland

The notions of space and time have been a subject of philosophical discussion since antiquity. In modern philosophy, Newton's concepts of absolute space and absolute time was subject to criticism by empiricists like Berkeley and Hume, who prepared the ground for Einstein and the ideas behind his theories of relativity. Simultaneously with Einstein's revolutionary works in physics, the development of new philosophical methods by Edmund Husserl and others under the heading of *phenomenology* opened for renewed investigations of a broad range of scientific ideas, among them the concepts of space and time. The phenomenology of space and time is the subject of this article. I will emphasize that the concept of space to be discussed is the concept of empty space. Empty space is what people traditionally have thought that spatial bodies are "in", and which shows itself where there are no such bodies "filling" the space. One of the themes of this article is how to define a corresponding concept of "empty" time to obtain a symmetric description of space and time.

One of the lesser known but original phenomenological studies of space is found in Jean-Paul Sartre's monumental work, *Being and Nothingness* (BN) from 1943 [6]. Sartre does not discuss the notion of space extensively, but he brings in some novel and clarifying ideas, in particular by associating space with the phenomenological concept of absence or nothingness (*néant*). In the first part of this article I will give a presentation of Sartre's contribution to the phenomenology of space.

I will then consider the notion of time. Sartre's treatments of space and time are quite different. However, to see space and time together is important. When Hermann Minkowski suggested the unified concept of space-time in his famous 1908 paper [4], he made the prediction that "space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union between the two will preserve an independent reality" (p. 75). Now, the concept of space-time is formulated in a mathematical language, and is often considered as an abstract

---

H.H. Grelland (✉)
University of Agder, Jon Lilletuns vei 9, 4879 Grimstad, Norway
e-mail: Hans.grelland@uia.no

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_7

formalism far removed from our immediate experience of the world. But taken at face value, Minkowski's statement implies something more. If only space and time in combination is to be preserved as independent realities, the ontology and space and time must be closely related. I will outline how Sartre's analysis of space can be transferred to that of time, thus showing the symmetry of space and time and laying a phenomenological foundation for the notion of space-time.

Then I will compare Sartre's notion of space to Einstein's. In *The Meaning of Relativity* (MR) [2] Einstein redefines the concepts of space and time in a way appropriate for the theories of relativity. He starts with our experience of space in a way that can be seen as a simple phenomenological analysis, and I will show how a close reading of Einstein's text reveals a fundamental similarity to Sartre's ideas.

When we consider phenomenological description and abstract geometry, one may question how these combine. For instance, does a geometrical description, with its idealized concepts, imply an abstraction leading to a loss of meaning compared to the concrete experience of real bodies in space? Or does the mathematical language reveal some additional meaning which cannot be seen in the immediate experience, such that we are lead to understand the world better? This worried Husserl and it was the subject of many of his late works, among them the article *The Origin of Geometry*, written in 1936 and published posthumously in 1939. When Jacques Derrida translated this article into French in 1962, he wrote a introduction which was almost six times longer than the original article, in which he through a critical reading of Husserl's text developed his own ideas on mathematics, language, ideality and meaning. In the last part of this article I will give an outline of this one-way discussion between Husserl and Derrida. I will show how Sartre's study provides an illuminating case for this discussion.

Sartre's treatment of space is found in various places in BN, and none of the passages are very long. The notions presented are integrated as a part of a comprehensive ontological system. A terse statement of the essential idea can be found on p. 491: "… geometrical space … is a pure nothingness". To understand this, we need a brief explanation of Sartre's notion of nothingness. This is one of the most important concepts on BN, as seen from its title, and Sartre's use of the concept represents one of his novel contributions to phenomenology. BN is, according to its subtitle, a phenomenological ontology, in which the concept of nothingness plays an important role.

The core of phenomenology is the idea of intentionality. To be a conscious mind is always to be conscious of something, to see is to see something, to think is to think about something, to imagine is to imagine something, to know is to know something, to understand is to understand something, etc. This "directedness" is called Intentionality and is, according to phenomenology, the defining property of consciousness. To see, think, imagine, etc., are called intentional acts. In an intentional act, we are aware of something, and this something, the object of our awareness, is called the intentional object. Among intentional acts, the act of intuition will be particularly important. Intuition in phenomenology is a translation of German *Anschauung*, meaning sense perception. To be an intentional object is to appear to consciousness in some way. It can be something that we see, something that we think about, something that we imagine, etc. This is the way we relate to the world and the world available for our knowledge and understanding is a world of intentional objects. It is important to notice that intentional acts also include indirect observations, abstractions, and symbolic representations, but somehow even our most abstract theories of the world must be rooted in our experience of the world as appearance.

Now, Sartre points out that we do not only see what is there, but also what is not there, for instance because we expected to see something which for some reason does not appear. The absence which we in this way "see" is for Sartre an example of a general concept of nothingness. Very often the word nothingness in Sartre can be replaced by the word absence. It can be a concrete thing (I believed I had put my keys on the table, but when I look for it I only see no-keys, their absence) or a person (I had an appointment with Pierre in the café, but when I look for him, I see only no-Pierre, Pierre's absence), or a state of affairs (I believed that the sum of the angels of a triangle was 120°, but found out that it is not the case, that "sum-equals-120" is not an existing fact, it is absent from all reality). Nothingness is also a condition for logical negation. When I say that it is not the case that the Pierre is in the café, it is a statement of Pierre's absence, and hence of nothingness. Nothingness is also a condition for the constitution of objects, which appear in analogy with a paper cut, where what is not the figure is cut away, and the figure itself stands out against this nothingness. This brief sketch should indicate the wide range of situations where nothingness appears as an element of an intentional act. Sartre points out that nothingness is both subjective and objective. It is objectively true that Pierre is not in the café, but his absence "is there" because I, the subject, am looking for him. Thus, nothingness is a phenomenological concept, it concerns how the world appears, it is more than just logical negation, for which it is a precondition. And such is Sartre's phenomenological concept of space, that it is pure nothingness, pure absence, in this case absence of things, of bodies.

The concept of space as nothingness is mainly elaborated in BN on p. 20. This passage is a part of a section describing the kinds of being that includes non-being as a part of their way of being. Possible examples would be an empty box, where the empty space within is an essential part of making it a box, or, again, a paper cut where what is the figure is there only on the background of what is not there because it is cut away. In this context, space is introduced. Sartre considers space as derived from the notion of a spatial distance, exemplified by two points at a distance. He describes how this situation can be an intentional object in two ways, or, more precisely, how it can be two different intentional objects or appearances. There is one in which the distance itself is made into an intentional object, in the form of being thought of as a line segment connecting the two points and having a certain length. The other intentional object consists of the two points, where the distance is just an absence because the two points are not connected. The last alternative is the intention which gives us the notion of empty space as nothingness. You cannot see, think of, or imagine both these intentional objects at the same time, to see one dissolves the other, according to Sartre. The concept of a distance as absence is an fragment of a full concept of space as a combination of all distances of all spatial objects in a given situation. This means that, properly speaking, one cannot say that something is "in" space like a fish is in the water or the wine is in the bottle. Space is not a being of its own which something can be "in", but refers to a relation between the spatial objects.

Now we turn to the question of time. When Hermann Minkowski suggested the unified concept of space-time, one was faced with the question: How can the separate concepts of time and space, each being deeply rooted in our ordinary experience of the world, possibly "fade away", to be replaced by space-time, which is usually regarded as a pure mathematical abstraction? To solve this problem, we would need a phenomenology of time displaying the symmetry between time and space. In BN, time is treated quite differently from space. In his treatment of time, Sartre focuses on the experienced temporality of the subject. The question is if it is possible to see time as something analogous to space as nothingness. This is what we will try.

First, we remind ourselves that when considering space, we were talking about empty space. Empty space is that which shows itself in the nothingness of a distance. And it is also the notion of empty space which lies behind the traditional (and Newtonian) notion of a pre-existing space which spatial bodies are "in". The space as nothingness in Sartre is also an interpretation of empty space, a different one from the absolute space of Newton. So, we are looking for a corresponding concept of "empty" or pure time, not for the concept of the temporality of a process. This is a precondition for interpreting time as nothingness on par with space.

A distance, as Sartre describes it, is between two points in space. If we take a step further, we can say that a point in space is a projection into space of a space-time event. In physics, an event is not something with a content taken place at a moment in time and at place, it is just the moment and the place as such, but seen together, disregarding what actually might be happening there and then. In Minkowskis formulation of the special theory of relativity, the event is the primary thing, its division into a time and a space component depends on a frame of reference. In a phenomenological perspective, the frame of reference is given by the consciousness for which the described phenomena appear. So we can take a space-time split as given. As the two points p and q at a distance are the space projections of two events $u = (p, t)$ and $v = (q, s)$, the two moments s and t are at a distance in time. And as the distance in space creates space, so the distance in time creates time, in the meaning of "empty time" or "pure time", time *as such*. Pauses, breaks, intervals of silence, are examples of such pure time, pure nothingness in time in analogy with the pure nothingness of space between points at a distance.

An apparent problem could be that while to a consciousness two points at a distance can be experienced in one single experience, two moments do not happen simultaneously. Thus one would think that the two moments belong to two different intuitions. Here we need to bring in Husserl's study of time consciousness ([5], pp. 127–143). Husserl observes that our experience of temporal objects is not limited to the factual "now". His favorite example is a melody. We do not only hear one tone, and then another tone, the tones just passed is retained in consciousness in such a way that we actually hear a melody, not only single tones. Gradually the retained tones fade away and disappear from consciousness, but because of the retention we actually experience also time intervals, like we experience space intervals. Thus a "Sartrean" phenomenology of time becomes possible, and together with the phenomenology of space it represents an experiential foundation for the concept of space-time.

It is interesting to compare Sartre's concept of space to Einstein's, which is expressed in *The Meaning of Relativity* (MR). This book introduces space-time step by step, going from the immediate experienced space and time to the mathematically formulated theories of relativity. Einstein makes a point in going back to the immediate experience to get free from the heritage of Newtonian mechanics with its concepts of "absolute space" and "absolute time" given independently of observation. It is easy to discover the influence of Hume and Mach in Einstein's text, but we have no real indication that Einstein was familiar with Husserl's philosophy. He could have been: He was a friend of the mathematician Hermann Weyl, who knew Husserl and his philosophy very well. Weyl was in Göttingen at the same time as Husserl, went to some of his lectures and married one of his students. Later Weyl became a professor in Zürich while Einstein still was there. Moreover, Weyl taught Einstein's theory of general relativity in Zürich and published in 1918 the first textbook on this subject, *Space Time Matter* [7]. This book begins with a presentation of Husserl's phenomenology, which is used as a philosophical introduction to the concepts of space and time. It is highly probable that Einstein read the book, but although MR is based on a general phenomenological perspective, we do not find any specific phenomenological concepts in the book.

While Sartre starts derives the notion of space from the distance between points, Einstein starts with the distance between three-dimensional bodies. Thus, like Sartre, Einstein develops the concept of space from the concept of a distance. Where there is a distance between two bodies, there is space. To explain the concept of a distance, Einstein follows an elaborate procedure, beginning with the concepts of replacement (change of position) of a body and of what he calls a continuation. If we have given two three-dimensional bodies, we can think of one being moved relative to another. We can for instance bring one body up to another so that they are touching each other. By bringing other bodies up to a body A, A is said to be continued, i.e. extended. If we have another body, X, we can imagine A to be continued until it is in contact with X, i.e. such that there is no longer a distance between X and A. Such a continuation of A is something that is imagined, which has not taken place, but which we assume could be done. Thus, in the terminology of Sartre (which Einstein, of course, does not use), a continuation is nothingness. Now, Einstein defines the space of A as all imaginable (not realized) continuations of A. Thus, space is in fact an absence, precisely the absence of the continuations of A. Thus, we see that Einstein's notion is basically the same as Sartre's, although expressed in a slightly different way.

How would space appear to us in ordinary life, in my immediate visual experience, according to Einstein? Einstein's suggestion is, as we have seen, that space is nothing but the lack or absence of a physical body somewhere. Where there is space, there could have been a body, but it isn't. As bodies may have shapes and dimensions, pieces of space can be said to have the same, inherited from the absent body which would remove the space by its presence. We say that the body would replace the (empty) space. If we have a piece of space in front of us, and this piece of space could be replaced with a cube of length 20 cm in all three directions, then the space can be said to be cubic with the length 20 cm in the same directions. Einstein imagines this as taking place in relation to a given body (the body of reference) which we can imagine being extended, and which defines space (a piece of space) as the absence of this extension. Thus, he thinks that this concept of space always assumes that some body is originally given, to which the space, the absence, is defined. In ordinary life, he thinks that the earth functions as such a body of reference, which we are so used to that we never think about it. That is why we think of space as something that exists by itself, the everyday notion of absolute space.

Given the Sartre-Einstein phenomenological notion of space as absence or nothingness, we can now turn to the question of the meaning of the mathematical concepts of geometry in relation to the phenomenological description of space. Usually the introduction of the mathematics is considered as adding a "formalism" representing a new level of abstraction which removes us from the world of physical meaning, that is, the world appearing in intuition. This point of view is roughly what we will call intuitionism. It is highly misleading, and misses some interesting philosophical insights. The main objection to intuitionism is that "abstraction" is already implied in any given experience, and what mathematics does is rather to provide us with a language for unifying an otherwise fragmentary relation to the intentional object in question. A language displays what Wittgenstein called the logical structure of the phenomenon, which leads to better understanding, and hence—in fact—*more* meaning. And mathematics is nothing but a particularly sophisticated language.

This was the subject of the historical discussion between Edmund Husserl and Jacques Derrida taking place in the latter's translation of *The Origin of Geometry* (OG) into French. Derrida wrote an introduction [1] to his translation which was almost six times longer than Husserl's original article, and Derrida's criticism or "deconstruction" of Husserl's text and its philosophical assumptions is of profound interest for understanding what we do when we use mathematics in modern physics, including in the geometry of space, time and space-time. Husserl was concerned and worried about the development of science since Galileo into higher and higher degrees of abstraction and mathematisation. He thought that it was the immediate experiences and practical applications from which the scientific theory originated that gave a theory its fullness of meaning, and that the high degree of abstraction and the extensive mathematisation of science lead to a gradual loss of meaning. For Husserl, the development of geometry from a practical tool for design and building into a purely abstract mathematical science of ideal entities is an example of such a development. Husserl was not making an empirical historical investigation, what he presented was rather a reconstruction of what must have been the main steps in this development.

In his long and detailed introduction to Husserl's article [1], Derrida pays attention to the fine points in Husserl's text, pointing out the unresolved tension between admitting the necessity of the ideal, as we find it in the concepts of a language, and at the same time criticizing the alienation and degradation following from the idealization of mathematics in geometry. Derrida points out that what Husserl says about the ideal concepts of geometry, is true of language in general. To Husserl, geometry is both a means of expressing timeless truths and an abstraction hiding "sedimented" and forgotten layers of historical meaning. Thus, for Husserl, the establishment of ideality trough mathematical concepts and mathematical notation leads to forgetfulness of the meaning-giving origin in intuition. And at the same time this process is necessary for establishing the timeless validity of the mathematical theorems. Derrida finds in Husserl's text also statements to the effect that language in itself works in a similar fashion as mathematics. But the ideality of language is liberating meaning just as much as hiding or forgetting it. To Husserl, writing, notation, and hence language, is a system for representing a meaning originating in the immediate intuition, or, for the "pre-geometer", in his practical dealing with specific spatial problems. In Derrida's view, in any knowledge and understanding there are already implied concepts which are ideal entities established through the sign system of a language. All concepts are ideal in the sense of being repeatable and recognizable, which is in fact a condition for being applicable as language. Only through ideal concepts an experience makes sense, can be understood, repeated in memory and communicated to others. Thus, the immediacy of intuition imagined by Husserl as being the sole source of meaning of the ideal concepts is in itself an illusion. Derrida calls it the myth of the presence. There is no such thing as an immediate presence, untainted by language and thus ideality. The forgetfulness implied by developing geometry from a set of practical applications to an abstract theory is an unavoidable part of language in general and is not a particular problem for the mathematical concepts of geometry or other abstract sciences. It is just one side of the process which also leads to the establishment of meaning through idealization.

The points and straight lines of geometry are ideal entities, and Derrida's point is that so are the concepts of language. Any potential loss of meaning following from introducing such geometrical objects into a world of immediate experiences is similar to the potential loss of meaning in any language use.

The position that the meaning of language as well as mathematical ideas has its source an original intuition is called by Derrida *intuitionism.* Later writers, including myself, have argued that is possible to distinguish Husserl's intuitionism from phenomenology (see e.g. [3]).

We can use Sartre's analysis of space as an example. As mentioned above, Sartre considers the situation of having two points at a distance. He identifies two possible intentional objects which can be associated with this situation. One corresponds to space as nothingness, in which the end points are the foreground objects being at a distance. This being-at-a-distance is a lack or absence of material fullness, and this absence is space. The second intentional object is the distance itself as a line segment connecting the points, and the points are background objects being the end points of the line segment.

Here we observe that here is no original unmediated presence to relate to which represents the immediate intuition in Husserl's sense. In Husserl's own wording, the situation as it appears is always constituted, and in this case constituted in one of two possible ways. It is one of Derrida's points that there is an opposition between Husserl's theory of constitution and his theory of the original unmediated presence appearing in intuition. This is Derrida's method "deconstruction", to reveal the inner tension in the theory under consideration instead of criticizing from the outside. In this case the tension between Husserl's theory of constitution and his theory of presence.

The two intentional objects outlined above, corresponding to two different acts of constitution, are relevant here, because the duality of description differentiates between the notion of empty space, which induces anxiety of heights or open places and the mathematical space, which is based on the notion of the line segment between the points. Both descriptions are ideal descriptions, and the last one includes a concept which is a part of the mathematics of geometry, namely a line segment. Thus, we have here a simple example of a form of intentionality which can be developed into something much more complex by applying an elaborate mathematical structure instead of just a piece of a straight line.

The two ways of constituting the originally given situation are revealing. The geometrical description of space will never explain the phenomena of spatial anxieties, while the nothingness description cannot reveal the kind of structured reality represented by the elaborate geometrical mathematics needed for physics. We see that we can use a phenomenological analysis against Husserl; more precisely: against his intuitionism. This is necessary to rescue both the adequacy of scientific knowledge in modern physics and the fundamental philosophical and psychological insights of phenomenology.

References

1. Derrida, J. (1962). *Edmund husserl's origin of geometry*: *An introduction*. London: University of Nebraska Press. (The book includes Husserl, E. (1962). *The Origin of Geometry*).
2. Einstein, A. (1974). *The meaning of relativity*. Princeton: Princeton University Press.
3. Mensch, J.R. (2001). *Postfoundational phenomenology*. Pennsylvania State University.
4. Minkowski, H. (1952). Space and time. In H.A. Lorwnz, A. Einstein, H. Minkowski & H. Weyl (Eds.), *The Principle of Relativity*. New York: Dover.
5. Russell, M. (2006). *Husserl. A guide for the perplexed*. London: Continuum.
6. Sartre, J.-P. (1996). *Being and nothingness. An essay on phenomenological ontology*. London: Routledge.
7. Weyl, H. (1952). *Time space matter*. New York: Dover.

Chapter 8
Space, Time, and (How They) Matter

A Discussion of Some Metaphysical Insights about the Nature of Space and Time Provided by Our Best Fundamental Physical Theories

Valia Allori

Abstract This paper is a brief (and hopelessly incomplete) non-standard introduction to the philosophy of space and time. It is an introduction because I plan to give an overview of what I consider some of the main questions about space and time: Is space a substance over and above matter? How many dimensions does it have? Is space-time fundamental or emergent? Does time have a direction? Does time even exist? Nonetheless, this introduction is not standard because I conclude the discussion by presenting the material with an original spin, guided by a particular understanding of fundamental physical theories, the so-called primitive ontology approach.

8.1 Introduction

Scientific realists believe that our best scientific theories could be used as reliable guides to understand what the world is like. First, they tell us about the nature of matter: for instance, matter can be made of particles, or two-dimensional strings, or continuous fields. In addition, whatever matter fundamentally is, it seems there is matter *in* space, which moves *in* time. The topic of this paper is what our best physical theories tell us about the nature of space and time.

I will start in the next section with the traditional debate concerning the question of whether space is itself a substance or not. Section 8.3 is focused instead on the question of the dimensionality of space in light of string theory and quantum mechanics. I will then continue in Sect. 8.4 explaining how quantum gravity may suggest that space and time are not fundamental but emergent. In Sect. 8.5, I discuss whether physical theories successfully suggest the passage of time is

V. Allori (✉)
Department of Philosophy, Northern Illinois University,
Zulauf Hall 915, Dekalb, IL 60115, USA
e-mail: vallori@niu.edu

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_8

illusory, and whether change does not exist. I conclude the discussion illustrating how the primitive ontology approach to fundamental physical theory can shed some new light on to the issues discussed.

8.2 Is Space a Substance?

Material things seems to be arranged in space, but is that true? There are two positions: either there is space, and objects are arranged within it; or there is no space over and above the spatial relations between the objects. The first position, known as substantivalism, was first suggested by Newton and was heavily criticized by Leibniz. According to Leibniz, there is no absolute 'here' or 'there,' 'now' or 'then;' rather there are just spatial and temporal relations between material objects. This view is known as relationism: all motion is relative motion.

8.2.1 Arguments for Substantivalism

Newton's suggestion is that space and time are absolute, immutable quantities which provide the fundamental arena in which matter can exist and evolve. This view has been developed in the framework of classical (or Newtonian) mechanics but can also be generalized to other theories. Newton provided an argument, the bucket argument, to show that only substantivalism is able to account for certain observations. Consider a bucket of water hanged with a winded up rope. When the rope is let go and the bucket starts to rotate, the water's surface becomes concave. Why? The substantivalist's response is that it is accelerating with respect to absolute space. Instead, on the relationist view there has to be a body with respect to which the water will rotate, and the problem is that there is no such body: it cannot be the bucket itself (the water and the bucket are at rest with one another at the end when the water is concave, but also at the beginning when the water was not concave; the water was in motion relative to the bucket in the intermediate state, yet the water was not concave); it cannot be motion with respect of the walls of the room ("put the walls on wheels and spin them around a bucket of water, and the water will not go concave" [14]).

8.2.2 Arguments Against Substantivalism

Leibniz argued against substantivalism, suggesting that it is fundamentally problematic. Leibniz observes that a 'shifted' world, a world just like ours except that all matter is shifted in another place in absolute space without any change in the relations of one object to another, will count as a different world for the substantivalist, even if the two worlds do not differ in their fundamental properties. The same is true for a ‘boosted’ world, a world just like ours except that all matter is drifting through space at uniform velocity. Leibniz concludes that this is absurd because it violates either the principle of sufficient reason or the principle of identity of indiscernibles, namely that God had a reason to create the world exactly as it is, and that indiscernible possibilities are identical.

One can easily resist these arguments denying such principles are true, but there is a stronger argument against substantivalism based on symmetries that generalizes to theories other than Newtonian mechanics, as we will see. Translations and rotations are symmetry of shape: they are transformations that leave the shape of an object the same. Similarly, it is argued that symmetries of a law are transformations that preserve that law. In classical mechanics, shifts and boosts are symmetries: if w_1 is a classical world, then a world w_2 shifted 5 feet to the left will also be classical; likewise, a boosted world w_3. By definition, F is an invariant feature of a law if any two worlds related by a symmetry of the law agree on the values of F [14]. In classical mechanics, only relative (as opposed to absolute) distances and velocities are invariant: w_1, w_2 and w_3 are governed by the same laws, regardless of their absolute position and absolute velocity. This suggests that (all things being equal) invariant features are the only ones we should think of as real: if F is not preserved in symmetries, then there are systematic ways to alter its values and yet preserve the law. Tus, they ‘do not make any difference’ in the law, they are redundant or irrelevant [9, 37]. In addition (or alternatively), non-invariant features are by definition undetectable, and thus we do not have any reason to believe they exist [14]. In both cases, substantivalism is in trouble since according to her absolute position and velocity are real, even if they are not invariant.

There is consensus among substantivalists that the symmetry argument against absolute velocity is compelling. Thus they endorse Galilean substantivalism, in which absolute acceleration is defined without reference to absolute velocity. Like in Newtonian substantivalism, bodies traveling in straight lines have no acceleration, but differently from it, there is no fact of the matter about absolute velocities: a vertical straight trajectory in space-time or a tilted one will not be different. A body is accelerating only if it has a curved trajectory in space-time. The majority of substantivalists think that the symmetry argument against absolute position instead does not work. ‘Sophisticated’ substantivalists (see e.g. [12]) argue that they are not committed to maintain that the shifted and the actual worlds are different, as claimed by the relationists (see [14]).

The situation does not change much when we consider relativistic physics, in which we move from space and time to space-time, combined into a unique manifold. Therefore, space substantivalism arguably transforms into space-time (or manifold) substantivalism, the doctrine that the manifold of events in space-time is a substance. The basic idea of general relativity is that matter changes the geometry of space-time, curving it. Therefore, a crucial element is the metric, which captures all the geometric structure of space-time. The most famous (symmetry) argument against substantivalism in this framework is the ‘hole’ argument [18, 44]. Since there is no privileged coordinate system (the so-called general covariance of general relativity), we can 'spread' metric and matter in space-time in different ways without changing invariant properties. For instance, we can have a 'regular' translation of matter and metric, or we could have a 'hole' transformation: smoothly joined, we leave matter and metric unchanged outside the hole, and we spread them differently inside. Since the two distributions agree on all invariant features (i.e. on coordinate-independent properties such as the distance along spatial curves), they arguably describe the same physical situation. The problem is that according to manifold substantivalism they instead depict two distinct physical situations, characterized by undetectable non-invariant properties.

The most popular response to the 'hole' argument is again sophisticated substantivalism: one can regard these distributions as representing the same physical possibility (see [40]). Another response is 'metric essentialism:' contrarily to manifold substantivalism, points in space-time possess their relations essentially. That is, the metric is, so to speak, part of the container [27].

8.3 How Many Dimensions Does Space Have?

Another interesting question is what dimensions space (or space-time) has. In Newtonian mechanics and relativity theory, matter is represented respectively by three-dimensional entities evolving in time, or by four-dimensional 'worms' in space-time so it seems obvious that space has three dimensions, and space-time four. The situation changes in the framework of string theory and non-relativistic quantum mechanics.

8.3.1 Quantum Mechanics and Wave Function Realism

Classical mechanics dominated physics until the 20th century, when quantum mechanics and relativity were proposed as more successful alternatives. To get a clear metaphysical picture of the world out of quantum mechanics is notoriously difficult and controversial. The problem is that if, as quantum mechanics states, the complete description of any material object is provided by the so-called wave function, and the wave function evolves in time according to the equation developed by Schrödinger, then objects may have contradictory properties, like 'being here' and 'being not here' at the same time, which is extremely problematic. Nonetheless, in the last century few better quantum theories have been developed, most famously Bohmian mechanics [11], Everettian mechanics [19], and the GRW theory [20]. Bohmian mechanics avoids the inconsistencies of 'orthodox' quantum mechanics denying that the wave function provides the complete description, Everettian mechanics that it's problematic for objects to have contradictory properties as long as they are instantiated in different words, and the GRW theory denies that the wave function evolves according to the Schrödinger equation.

8.3.2 An Argument for Wave Function Realism

Nevertheless, it is controversial what matter is made of. One possibility is 'wave function realism:' the wave function represents matter in all the three theories above, even if they differ by either adding something to the wave function (like Bohmian or Everettian mechanics, which add particles and worlds), or modifying its dynamics. This view is motivated by focusing on the dynamics [39]: since in Newtonian mechanics the fundamental equation described the temporal evolution of three-dimensional points, then matter is made of point-particles in three-dimensional space. Similarly, since the fundamental equation of quantum theory, Schrödinger's equation, describes the evolution of the wave function, whatever object the wave function mathematically represents it is the fundamental constituent of matter, and space is whatever space the wave function lives in [3, 26, 33]. This space, introduced in the classical framework, is called 'configuration space:' the space of the configurations of all the particles in the world (given that matter is made of wave function, there are fundamentally particles, so the name 'configuration' should not be taken literally). Thus, if there are N particles in the universe (estimated to be 10^{80}), the dimension of configuration space is 3 N. If so, contrarily to what our everyday experiences suggest, space is a very high dimensional space.

8.3.3 Arguments Against Wave Function Realism

One problem for this view is that it cannot account for our experience of three-dimensional objects [5, 30, 31]. One could argue that they exist 'functionally' rather than fundamentally [2]. The preliminary problem here is that there are infinitely many functions from configuration space to three-dimensional space, and to select a privileged one amounts to add an ontology, which the wave function realist denies. Other proposals use symmetries [38], or grounding [39]. These approaches are all work-in-progress, and the debate over which is more promising is still open.

In addition, wave function realism may not be viable in the framework of quantum field theories [32, 46]. A first problem is that the definition of configuration space requires that the number of particles does not change in time, contrarily to what happens in quantum field theories, in which particles are created and annihilated. See most notably [34] for a strategy to address this problem.

Furthermore, some question the motivation for the approach: wave function realism prescribes that the world is very different from what we perceive it to be. Before accepting such a revisionary metaphysic, one should rule out the existence of viable, less counterintuitive alternatives. Since they exist (see Sect. 8.6), it is difficult to see the appeal of wave function realism [6, 31]. This is connected with the so-called problem of empirical incoherence [30]. Loosely speaking, a theory is supported by observations when it predicts that objects have certain features, and these features are actually observed. Since our observations are all observations of three-dimensional objects (pointer pointing in certain directions in three-dimensional space), a theory should make predictions about them. Wave function realism predicts instead that there are no three-dimensional object, so it cannot be supported by observations. Wave function realists (see e.g. [35]) respond that they do not deny three-dimensional object exists, rather they deny that they are fundamental, and accordingly they attempt to provide an account of how they emerge from a deeper reality.

8.3.4 String Theory and Extra Dimensions

Quantum field theory, the first proposed extension of quantum mechanics in a relativistic framework, is mathematically ill-defined. In order to overcome such difficulties, string theory, among other theories, has been proposed. In this theory, matter is described by one-dimensional objects called strings. On distances larger than the string scale, a string looks just like an ordinary particle, with properties determined by the vibrational state of the string. Since one of them corresponds to the graviton, a particle connected to gravity, string theory promises to be a unified description of all the fundamental forces. There are several versions of string theory, but for their mathematical consistency, they all require extra dimensions of space-time: for instance, in 'superstring' theory space-time is ten-dimensional. These extra dimensions are assumed to close up on themselves to form little circles, so that they are not macroscopically observable, similarly to what happens when we observe a garden hose from a distance and it appears to have only one dimension instead of two (this is the so-called 'compactification'). So, just like in the quantum framework the dimensionality of space is not as it seems, but contrarily to the quantum case in which the mappings from the high dimensional fundamental space to the perceived three-dimensional are arbitrary, here the compactification mechanism is part of the definition of the theory.

8.4 Is Space-Time Fundamental?

Many additionally have suggested that recent developments in quantum gravity, namely the theories that attempt to unify general relativity and quantum mechanics, imply that space-time is not fundamental but rather emergent.

8.4.1 Arguments for Emergence

As we saw, string theory was originally developed assuming a space-time background (as inert container) but the so-called ‘dualities,’ suggested to some otherwise. As symmetries relate possible physical description a given theory to one another, dualities connect different types of strung theories. Two theories are said to be dual, roughly, whenever they provide the same physics. There are various dualities. T-duality is a kind of scale invariance. As we saw, the extra dimensions are compactified but different theories have different compactification mechanisms. Suppose in a theory T_1 a dimension is wrapped around a circle of radius R. It turns out that, schematically, a theory T_2 in which the dimension is wrapped around a circle of radius 1/R is dual to T_1. That is, the transformation $R \rightarrow 1/R$ leaves the physics invariant. There is no difference between the physics generated by T_1, in which the space is ‘large,’ and by T_2, in which the space is ‘small.’ Mirror symmetry is the generalization of T-duality: the extra dimensions can be compactified so that they form a particular manifold (the Calabi-Yau manifold), which turns out to be dual to a manifold with a different topology. Then we have S-duality, which connects theories with different coupling constants (that is, the strength of the interaction is different): a theory T_1 with coupling constant g is dual to a theory T_2 with coupling constant 1/g. Another duality is the AdS/CFT (Anti-deSitter/Conformal Field Theory) duality, which connects a string theory, which includes gravity and is defined in ten dimensions on an Anti-deSitter space, with a quantum field theory, which does not include gravity, and is defined in three dimensions on the boundary of the AdS space. Some have argued that the metaphysical lesson we should draw from dualities is that space-time is emergent. The idea is very similar to the symmetry arguments we discussed previously, now applied at dualities: if T_1 and T_2 are dual, they are empirically indistinguishable, and we cannot choose between them. Only invariant properties describe something real: in the case of T-duality, for instance, there is no fact of the matter about whether the space is ‘small’ or ‘large:’ space is not fundamental [15, 24, 42].

The ‘rival’ of string theory is quantum gravity, in which general relativity is quantized. A particular type of quantization leads to canonical quantum gravity, newer approaches include loop quantum gravity, in which, arguably, space can be viewed as an extremely fine fabric or network of finite loops, called spin networks. The evolution of a spin network over time is called a spin foam. The spin network can either persist, fuse or split into several nodes, and “the resulting structure is taken to be the quantum analogue of a four-dimensional space-time and is called ‘spin foam’” [24]. The theory has not been completely developed but the idea is that the spin foam represents what is fundamental, rather than space-time, and that the perceived three-dimensionality thus have to suitably emerge from such structure.

8.4.2 Arguments Against Emergence of Space-Time

The view according to which space-time suitably emerges from a deeper physics faces very similar problems as wave function realism: first of all, how to account for the appearance of three-dimensional objects evolving in time? Why should we believe the theory? [23] take on the challenges and sketch possible solutions. The problem that seems to remain is about the motivation: given that there are alternatives, why would one commit to such a radical metaphysical picture?

8.5 What About Time?

So far, we have focused our attention on space, leaving aside many issues regarding the nature of time. I wish to outline just two connected questions, leaving aside many other interesting debates, namely whether time passes and whether it has a direction.

8.5.1 Arguments Against the Passage of Time

The ongoing debate in metaphysics about the nature of time is between those who believe that the passage of time is objective, and those who believe that this is just an illusion. Some have argued that in the framework of relativity, in which we go from space and time to space-time, we should think of time just as another dimension of a bigger fundamental space, and that the passage of time is just an illusion (see, e.g. [17]). Others instead argue that it is perfectly coherent to believe that time passes in a relativistic framework [28, 45].

In addition, there is a tension between microscopic laws and macroscopic behavior [1, 38]. In fact, on the one hand all macroscopic behavior has a natural temporal order: eggs break and do not un-break. On the other hand, the microscopic laws that govern the macroscopic behavior (whether classical, relativistic or quantum) are time-symmetric. That is, if a process is possible, then so is the process run backwards. So, why is it possible for the molecules that constitute an egg to follow both the trajectories corresponding to ‘the egg breaks’ and to ‘the egg un-breaks,’ while on the macroscopic level we only see eggs that break? The problem is to explain where the law that describes these macroscopic processes, the second law of thermodynamics according to which entropy never decreases, is coming from. Arguably, the puzzle has been solved in the framework of statistical mechanics by Boltzmann, in which it is overwhelmingly likely for a process to develop towards maximal entropy, but it is possible that entropy decreases. As such, eggs can un-break, it is just overwhelmingly unlikely to happen. In order for the derivation to go through, many, including Boltzmann, believe that it is necessary to assume the so-called ‘past hypothesis,’ the assumption that the universe started out with an extremely low entropy. Critics of this strategy complain that this condition calls for an explanation [41], since the probability of the universe starting in the requisite state is astronomically small (see [13] for a defense).

8.5.2 An Arguments for the Unreality of Time

Similarly to the argument for the emergence of space-time, some have argued that canonical quantum gravity, one of the contenders to unify general relativity and quantum mechanics, suggests time does not exist. Canonical quantum gravity gives rise to the Wheeler-de Witt equation for a universal wave function, the interpretation of which seems to describe a static universe. How can this theory describe a world like ours in which there is change? This is the so-called ‘problem of time’ (for a review, see [24]). The possible reactions to this problem can be either endorse timelessness or to attempt to quantize gravity in a different way. For the latter approach, see [25]. The former path has been taken most notably by [10, 16, 43, 47]. In particular, Rovelli’s basic idea is that we can describe change without time relating physical systems directly to one another.

8.5.3 Arguments Against the Unreality of Time

Objections to this suggestion are of two sorts: some suggests that the lack of change in the Wheeler-de Witt equation should not be taken metaphysically seriously, since it is an artifact of framing the theory in terms of canonical variables [21, 29]. Others have stressed that it is difficult to see how one can come to believe in a theory in which time does not exist [22]. As one can see, this is a variety of the problem of empirical incoherence mentioned above.

8.6 A New Look into the Debates: Primitive Ontology

If the reader has remained with me up to this point, if we set aside the issues connected with the direction of time, I hope she will be able to see a trend: on the one hand we have the relationists, the wave function realists, the space-time emergentists, the antirealist about time, which essentially resort to the intuition that if something is unobservable then we have no reason to believe it is real; on the other hand we have the substantivalists, the fundamentalists about space-time (i.e. the critics of wave function realism and of space-time emergentism), and the realist about time that instead seem to appeal to the idea that if something has an explanatory role then we have reason to believe that it exists, even if it is not detectable. In fact, many prominent arguments for the former positions are based on symmetries and invariant features: in classical mechanics, position and velocities are not invariant, and therefore theory are not real; general relativity is covariant, therefore space-time points are not real; since there are dualities connecting different string theories, we have no reason to believe one theory over the other. These, fundamentally, are all varieties of underdetermination arguments, and the opponents of these views essentially reply that observability is not the only virtue a theory may possess: explanatory power, in particular, should be taken into account.

The other arguments are slightly different: in quantum mechanics one needs nothing more than the wave function in order to account for the experimental results; general relativity suggests that space and time are part of the same continuum, so space and time are not fundamentally different and time does not pass; in canonical quantum gravity we have an equation that suggests that noting evolves in time, so time does not exist. Nonetheless, I think there is still something in common with the previous arguments: the wave function realist and the antirealist about time start off the bare formalism of the theory (respectively quantum mechanics, general relativity and canonical quantum gravity) in order to interpret it. In contrast, their opponents emphasize that the 'interpretation' comes first, and the theory should follow: we make a hypothesis about what the world is made of, and then we construct a mathematical theory to describe it. In particular, the problems of the macro-object and of empirical incoherence stem from this reflection: one theory should be able to account for what we experience, they should be able to explain empirical data, and three-dimensional space (or four-dimensional space-time) seems to be essential for that.

In this last section, I wish to briefly describe a unifying account of fundamental physical theory that essentially captures the ideas just outlined: the primitive ontology approach (for an updated version, see [4]). The main idea is that fundamental physical theories are about three-dimensional entities which evolve in time (the primitive ontology). The prototype of a theory with primitive ontology is classical mechanics, according to which macroscopic objects are composed of microscopic three-dimensional point-particles, and the temporal evolution of such objects is determined by Newton's equation. The proponents of this view point out the macro object problem and the problem of empirical coherence in quantum mechanics and quantum gravity as a motivation for their view: that is, theories with a primitive ontology are explanatory successful and empirically coherent, in contrast with their rival views. Because of these reasons, they propose that not only Bohmian mechanics, but also GRW and Everettian mechanics are actually theories about three-dimensional entities, may that be particles, or continuous three-dimensional fields, or space-time events ('flashes') [7, 8]. The idea is that the wave function does not describe matter, in contrast to what the wave function realist believe, but rather should be seen more like a nomological entity, needed to implement the law of evolution for the primitive ontology. Similarly, in the context of quantum gravity [21] dissolve the problem of time by arguing that the metric is the primitive ontology of the theory, whose evolution is governed by the wave function, which obeys to the Wheeler-de Witt equation. When confronted by dualities in string theory, the primitive ontologist will similarly argue that empirical adequacy and observability are not the only virtues a theory should have, and that space-time and objects in it are essential to explain theory experiences. Finally, in the context of the substantivalism-relationism debate, what does this approach have to say? At least one thing: in classical mechanics position is fundamental, being the primitive ontology, and symmetry arguments are completely ineffective in this context.

References

1. Albert, D.Z. (2000). *Time and chance*. Harvard University Press.
2. Albert, D.Z. (2013). Wave function realism. In D. Albert, A. Ney (Eds.), *The wave function: Essays in the metaphysics of quantum mechanics* (pp. 52–57). Oxford University Press.
3. Albert, D.Z. (1996). Elementary quantum metaphysics. In J. Cushing, A. Fine, & S. Goldstein (Eds.), *Bohmian mechanics and quantum theory: an appraisal* (pp. 277–284). Kluwer (1996).
4. Allori, V. (2015). Primitive ontology in a nutshell. *International Journal of Quantum Foundations, 1*(3), 107–122.
5. Allori, V. (2013). On the metaphysics of quantum mechanics. In S. Lebihan (Ed.), *Precis de la Philosophie de la Physique*. Vuibert.
6. Allori, V. (2013). Primitive ontology and the structure of fundamental physical theories. In D. Albert, A. Ney (Eds.), *The wave function: essays in the metaphysics of quantum mechanics* (pp. 58–75). Oxford University Press.
7. Allori, V., Goldstein, S., Tumulka, R., & Nino Zanghi, N. (2008). On the common structure of bohmian mechanics and the ghirardi-rimini-weber theory. *The British Journal for the Philosophy of Science, 59*(3), 353–389.
8. Allori, V., Goldstein, S., Tumulka, R., & Nino Zanghi, N. (2011). Many-Worlds and schrödinger's first quantum theory. *The British Journal for the Philosophy of Science, 62*(1), 1–27.
9. Baker, D. (2010). Symmetry and the metaphysics of physics. *Philosophy Compass, 5*, 1157–1166.
10. Barbour, J. (2001). *The end of time: the next revolution in physics*. Oxfrord University Press.
11. Bohm, D. (1952). A suggested interpretation of the quantum theory in terms of hidden variables II. *Physical Review, 85*, 180.
12. Butterfield, J. (1989). The hole truth. *The British Journal for the Philosophy of Science, 40*, 1–28.
13. Callender, C. (2004). There is no puzzle about the low entropy past. In C. Hitchcock (Ed.), *Contemporary debates in the philosophy of science*, Ch 12. Blackwell.
14. Dasgupta, S. (2015). Substantivalism versus relationalism about space in classical physics. *Philosophy Compass, 10*(9), 601–624.
15. Dawid, R. (2013). *String theory and the scientific method*. Cambridge University Press.
16. Earman, J. (2002). Thoroughly modern McTaggart: or what McTaggart would have said if he had learned general relativity. *Philosophers' Imprint, 2*(3), 1–28.
17. Godel, K. (1949). A remark about the relationship between relativity theory and the idealistic philosophy. In P. Schilpp (Ed.), *Albert Einstein. Philosopher-Scientist* (pp. 557–562). Open Court.
18. Earman, J., & Norton, J. (1987). What price substantivalism? The hole story. *The British Journal for the Philosophy of Science, 38*, 515–525.

19. Everett, H. (1957). Relative state formulation of quantum mechanics. *Review of Modern Physics, 29*, 454–462.
20. Ghirardi, G.C., Alberto, R., & Tulio, W. (1986). Unified dynamics for microscopic and macroscopic systems. *Physical Review D 34,* 91–470.
21. Goldstein, S., & S. Teufel (2001). Quantum spacetime without observers: ontological clarity and the conceptual foundations of quantum gravity. In C. Callender & N. Huggett (Eds.), *Physics meets philosophy at the planck scale* (pp. 275–289). Cambridge University Press.
22. Healey, R. (2002). Can physics coherently deny the reality of time?. In C. Callender (Ed.), *Time, reality and experience* (pp. 293–316). Cambridge University Press.
23. Huggett, N., Tiziana, V., & Christian W. (2012). Time in quantum gravity. In A. Bardon, & H. Dyke (Eds.), *The Blackwell companion to the philosophy of time* (pp. 242–261). Blackwell.
24. Huggett, N., & Wuthrich, C. (2013). Emergent spacetime and empirical (In)coherence. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics, 44*(3), 276–285.25.
25. Kuchar, K. (1992). Time and interpretations of quantum gravity. *Proceedings of the 4th Canadian Conference on General Relativity and Relativistic Astrophysics*. World Scientific, 211–314.
26. Lewis, P. (2004). Life in configuration space. *The British Journal for the Philosophy of Science, 55*(4), 713–729.
27. Maudlin, T. (1990). Substances and space-time: what aristotle would have said to Einstein. *Studies in History and Philosophy of Science Part A, 21*, 531–561.
28. Maudlin, T. (2002). Remarks on the passing of time. *Proceedings of the Aristotelian Society New Series, 102*, 259–274.
29. Maudlin, T. (2002). Thoroughly muddled McTaggart: or how to abuse gauge freedom to generate metaphysical monstrosities. *Philosophers' Imprint, 2*(4), 1–13.
30. Maudlin, T. (2007). Completeness, supervenience and ontology. *Journal of Physics A: Mathematical and Theoretical, 40*, 3151–3171.
31. Monton, B. (2002). Wave function ontology. *Synthese, 130*, 265–277.
32. Myrvold, W. (2015). What is a wave function? *Synthese, 192*(10), 3247–3274.
33. Ney, A. (2012). The status of our ordinary three dimensions in a quantum universe. *Nous, 46* (3), 525–560.
34. Ney, A. (2013). Introduction. In D. Albert, A. Ney (Eds.), *The wave function: essays in the metaphysics of quantum mechanics* (pp. 51–51). Oxford University Press.
35. Ney, A. (2015). Fundamental physical ontologies and the constraint of empirical coherence: a defense of wave function realism. *Synthese, 192*(10), 3105–3124.
36. Ney, A. Finding the world in the wave function: some strategies for solving the macro-problem. *Synthese* (forthcoming).
37. North, J. (2009). The structure of physics: a case study. *The Journal of Philosophy, 106*, 57–88.
38. North, J. (2011). Time in thermodynamics. In C. Callender (Ed.), *The oxford handbook of philosophy of time* (pp. 312–350). Oxford University Press (2011).
39. North, J (2013). The structure of a quantum world. In D. Albert, & A. Ney (Eds.), *The wave function: essays on the metaphysics of quantum mechanics* (pp. 184–202). Oxford University Press.
40. Pooley, O. (2013). Substantivalist and relationalist approaches to spacetime. In R. Batterman (Ed.), *the oxford handbook of philosophy of physics* (pp. 522–586). Oxford University Press (2013).
41. Price, H. (2004). On the origins of the arrow of time: why there is still a puzzle about the low entropy past. In C. Hitchcock (ed.), *Contemporary debates in the philosophy of science*, Ch 12. Blackwell.
42. Rickles, D. (2013). Mirror symmetry and other miracles in superstring theory. *Foundations of Physics, 43*(1), 54–80.

43. Rovelli, C. (2011). Forget time. *Foundations of Physics, 41*, 1475–1490.
44. Stachel, J. (1989). Einstein's search for general covariance. In D. Howard, & J. Stachel (eds.), *Einstein and the history of general relativity.* (pp. 63–100). Birkhäuser.
45. Zimmerman, D. (2007). In T. Sider, J. Hawthorn, & D. Zimmerman (Eds.) *Contemporary debates in metaphysics* (pp. 211–225). Blackwell.
46. Wallace, D., & Timpson, C. (2010). Quantum mechanics on spacetime I: Spacetime state realism. *The British Journal for the Philosophy of Science, 61*(4), 697–727.
47. Wheeler, J. A.: Time today. In J. Halliwell, J. Perez-Mercader & W. H. Zurek (Eds.), *Physical Origins of Time Asymmetry* (pp. 1–29). Cambridge University Press (1994).

Chapter 9
Relativity Theory May not Have the Last Word on the Nature of Time: Quantum Theory and Probabilism

Nicholas Maxwell

Abstract Two radically different views about time are possible. According to the first, the universe is three dimensional. It has a past and a future, but that does not mean it is spread out in time as it is spread out in the three dimensions of space. This view requires that there is an unambiguous, absolute, cosmic-wide “now” at each instant. According to the second view about time, the universe is four dimensional. It is spread out in both space and time—in space-time in short. Special and general relativity rule out the first view. There is, according to relativity theory, no such thing as an unambiguous, absolute cosmic-wide “now” at each instant. However, we have every reason to hold that both special and general relativity are false. Not only does the historical record tell us that physics advances from one false theory to another. Furthermore, elsewhere I have shown that we must interpret physics as having established physicalism—in so far as physics can ever establish anything theoretical. Physicalism, here, is to be interpreted as the thesis that the universe is such that some unified “theory of everything” is true. Granted physicalism, it follows immediately that any physical theory that is about a restricted range of phenomena only, cannot be true, whatever its empirical success may be. It follows that both special and general relativity are false. This does not mean of course that the *implication* of these two theories that there is no unambiguous cosmic-wide “now” at each instant is false. It still may be the case that the first view of time, indicated at the outset, is false. Are there grounds for holding that an unambiguous cosmic-wide “now” does exist, despite special and general relativity, both of which imply that it does not exist? There are such grounds. Elsewhere I have argued that, in order to solve the quantum wave/particle problem and make sense of the quantum domain we need to interpret quantum theory as a fundamentally *probabilistic* theory, a theory which specifies how quantum entities—electrons, photons, atoms—interact with one another probabilistically. It is conceivable that this is correct, and the ultimate laws of the universe are probabilistic in character. If so, probabilistic transitions could define unambiguous, absolute

---

N. Maxwell (✉)
Science and Technology Studies, University College London,
13 Tavistock Terrace, London N19-4BZ, UK
e-mail: nicholas.maxwell@ucl.ac.uk

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_9

cosmic-wide “nows” at each instant. It is entirely unsurprising that special and general relativity have nothing to say about the matter. Both theories are pre-quantum mechanical, classical theories, and general relativity in particular is deterministic. The universe may indeed be three dimensional, with a past and a future, but not spread out in four dimensional space-time, despite the fact that relativity theories appear to rule this out. These considerations, finally, have implications for views about the arrow of time and free will.

9.1 Two Views About the World

Ordinarily we think of the world as a three-dimensional place, with a past and a future. Things change. Time passes. Future events, after a time, occur now, and then become a part of the past, but none of this means that we ordinarily think of the universe as spread out in both space and time, objects having temporal parts and temporal extension in the same way that they have spatial parts and spatial extension.

Only when we think of the distant past—and perhaps the distant future—is there a temptation to think of time in spatial terms, so that the distant past is, as it were, “another place”. Science fiction about time travel exploits this tendency. It exploits our tendency to think of the distant past, and distant future, as distant “places” which we might travel to, for example in H.G. Well’s Time Machine.

Why are we prone to think of distant times on analogy with distant places? This, I think, is the reason. When we travel through space, we encounter new objects—new houses, cities, landscapes. But this is ordinarily hardly the case if we stay in the same place, and allow minutes or an hour or two to pass. The distant past—or the distant future—is, however, likely to contain objects quite different from those around us now (unless, perhaps, we are in some very ancient building). Hence, there is the temptation to think of distant times as “distant places” which we might visit if only we could discover how to create a viable time machine.

Much of the apparently baffling nature of time arises, I think, from a common sense tendency (to which some philosophers seem prone) to try to combine these two ways of thinking about time—or rather, two ways of thinking about the nature of the world, the three-dimensional view,[1] and the four-dimensional view.[2] We entertain the broadest possible perspective, and consider the entire history of the

---

[1] If string theory is correct, there may be 9 or 10 spatial dimensions. The “three dimensional view” becomes the nine or ten dimensional view. The crucial tenet of this view is, not the number of spatial dimensions, but rather that the world is not spread out in time in the way in which it is spread out in space.

[2] See [7], where I referred to the three and four dimensional views as C_2 and C_1 respectively, and briefly made the point that much confusion about time stems from attempting to combine these two incompatible views. The main point of the article was to argue that necessary connections between successive states of affairs *are* possible, despite Hume’s arguments to the contrary, but only if the three-dimensional, C_2, view is true.

universe, from its beginning in the big bang to its end in the big crunch, or from eternity to eternity (if eternal it should be). There is all of reality stretched out before the mind's eye, from the most distant past to the most distant future, and encompassing our present existence and time, now. But then we become aware of something missing: the present. If the universe really is stretched out in time, as it is stretched out in space, the vital *now* of present existence seems somehow to be missing. And so we attempt to add it, by adding a moving "now" along the time line, a brilliant light of existence, moving steadily from past to future.[3]

At once a host of baffling questions arise. Why is the dimension of time so profoundly different from the other dimensions of space? What is it that marks out "the present moment" from all other moments, past and future? What is it that causes "the present" to be so different from all other times? Why does "the present" move as it does, from past to future? What causes it so to move? How fast does it move? Do we need another dimension of time to mark the movement of the present along the time line? How long does "the present" endure? Does "the present" exist objectively, or is it no more than a subjective experience? Is time travel, back into the past, or forward into the future ahead of the passage of "the present", possible? What is it about time that enables it steadily to digest the future, turning it first into the present, and then into the past? What exactly do these mysterious transformations consist in? What keeps the present remorselessly travelling into the future and away from the past? There is, in the constitution of things, it seems, a mysterious process of *absolute becoming,* when events in the future enter the present and become *actualized*; for an instant they blaze into existence before disappearing into the shadows of the past. But what is this process of *absolute becoming*? How is it to be understood?

All these baffling questions only arise because, probably without even being aware of it, we have tried to combine two views about time, or rather two views about the nature of the universe, that are in flat contradiction with one another, and cannot be combined. The two views are the three-dimensional view of the universe, and the four dimensional view. Granted the four-dimensional view, "now" is like "here": it is just where you happen to be. There is no such thing as "the objective now"—or rather, all moments are equally entitled to be called "the objective now". Granted the three-dimensional view, there is no such thing as "the objective now" either. There is just the three-dimensional universe as it is *now*. That is indeed all there is. The past and the future are not "places" separated from us by temporal "distance". In speaking of events "that lie in the past", or "in the future" we are speaking metaphorically about what has been and what will be: no more. Objects are spread out in space, but not in time; it is facts about objects, their histories, that can be represented as being spread out in time. Such space-time representations depict,

---

[3] This incoherent view would seem to be implicit in McTaggart's A-series, according to which events are initially future, then present, then past, future events being converted into past ones by the passage of the present along time: see McTaggart (1908). It is explicit in all those attempts to rectify the perceived inadequacy of the spacetime view (or McTaggart's B-series) by adding "the present" to it, or "objective becoming", which is supposed to move steadily from past to future.

not objects, not the world, but facts about objects, facts about the world. And whereas the four-dimensional view needs to appeal to distinctive phenomena—irreversible processes, the second law of thermodynamics—in order to account for the direction of time, the distinction between the future and the past, the three-dimensional view does not need to make any such appeal. That time goes by in one direction, and the future becomes the past, is built into the nature of time, according to this view. Even if the universe one day degenerates into a state such that no irreversible processes occur, still time would pass from the future to the past.

9.2 The Impact of Relativity Theory

Up until 1905, it seemed as if science was indifferent between the two views, three and four dimensional. Then along came Einstein's special and general relativity, in 1905 and 1915, and it seemed that theoretical physics had declared unequivocally in favour of the four-dimensional view.[4] The three-dimensional view, sometimes called *presentism*, was, it seemed, decisively refuted.

Let E_1 and E_2 be any two events separated in space and time in such a way that light cannot travel from one to the other. E_1 might be the event of me typing "E_1" into my computer, while E_2 might be a short-lived flare occurring abruptly on the surface of the sun at a time that is simultaneous, for me, with me typing "E_1". Einstein's theory of special relativity now tells us the following. In some reference frames, E_1 occurs *before* E_2; in other reference frames, all moving with respect to the first set, E_1 occurs *after* E_2; and in a third set of reference frames, all stationary with respect to me, E_1 and E_2 occur simultaneously, at the same moment. Furthermore, according to special relativity, all these reference frames are equally viable. No one reference frame can be picked out as the proper, unique, objectively correct one.

The consequences for the three-dimensional view—presentism—are devastating. For the three-dimensional view only makes sense if there is, at every instant, a unique, cosmic-wide "now"—the three-dimensional universe as it is at this instant. If this does not exist—if there is no such thing as the unique, cosmic-wide "now" at each instant—as special relativity tells us, then the three-dimensional view becomes incoherent. What is required to exist for it to be viable does not obtain. The three-dimensional view is refuted by special relativity. And general relativity just confirms the point. Modern theoretical physics obliges us, it seems, to accept the four-dimensional view, with all its disturbing implications. The universe just *is*. The passage of time is an illusion. There is no such thing as the objective "now". All instants, everywhere, at all times and places, are equally entitled to be regarded as "now", just as all places are equally entitled to be regarded as "here". Free will becomes no more than an illusion

---

[4] It is really Hermann Minkowski who first interpreted special relativity in terms of the four-dimensional, space-time view in 1908: see [25]. Einstein's 1905 paper implicitly takes the three-dimensional view for granted.

since, if everything, past, present and future, just (tenselessly) *is*, what we do now cannot affect in any way what will come to be, in the future. Even more seriously, perhaps, our whole world—or at least the way we ordinarily conceive of the world in terms of the three-dimensional view—turns out to be an illusion. The four-dimensional "block universe" (as it is sometimes called) is very different from the three-dimensional world of human actions, persisting and changing things.

9.3 Special and General Relativity Are Both False

If general relativity is true, then special relativity is false. Special relativity requires space-time to be flat. General relativity, however, tells us that space-time becomes curved in the presence of matter, or energy-density more generally—energy associated with radiation as well as matter. Wherever you go in the universe, however far from stars and galaxies, the presence of matter, or energy-density, even if distant, still curves space-time very slightly, according to general relativity. Special relativity is thus false everywhere; and in some places, near black holes for example, where space-time is subjected to marked curvature, special relativity will be quite badly false.

But what of general relativity? Is it false? There are three grounds for holding that it is.

First, general relativity seems to be incompatible with quantum theory. Attempts to unify the two theories so far have failed. Many theoretical physicists believe string theory will unify relativity theory and quantum theory.[5] Other theoretical physicists strongly disagree.[6] So far, string theory has not been confirmed experimentally.

But even if string theory, or some other theory, does succeed in unifying general relativity and quantum theory—or the standard model, the quantum field theory of fundamental particles and the forces between them—in all likelihood, the unifying theory will reveal that general relativity is false. Almost always when two theories, T_1 and T_2 are unified by a new theory, T_3, it emerges that T_3 is incompatible with the predecessor theories, T_1 and T_2, and reveals these theories to be false.[7] Here, then, are grounds for holding that general relativity will eventually turn out to be false.

---

[5] For a popular exposition and defence of string theory see [5].

[6] For criticisms of string theory see [29, 27].

[7] The first great unifying theory in physics was Newtonian theory. This theory unifies Kepler's laws of planetary motion, and Galileo's laws of terrestrial motion. But in doing so, it reveals that both Kepler's and Galileo's laws are, strictly speaking, false. Granted Newtonian theory, planets deviate from precise Keplerian, elliptical motion because the planets attract each other gravitationally, and attract the sun, which leads to deviations. Again, granted Newtonian theory, a stone falling near the earth's surface does not fall with constant gravitation precisely because, as it falls, it gets closer to the centre of the earth, and thus the gravitational attraction on the stone increases very slightly, which means in turn that its acceleration increases very slightly. Newtonian theory explains why there are deviations from Kepler's and Galileo's laws. This almost always occurs whenever a new theory, T_3, unifies two predecessor theories, T_1 and T_2.

Second, almost all fundamental physical theories so far proposed that have been accepted because of their immense empirical success, have turned out subsequently to be, strictly speaking false. This is true of Kepler's laws of planetary motion, and Galileo's laws of terrestrial motion: see note 7. Special and general relativity reveal that Newtonian theory is false. James Clerk Maxwell's classical theory of the electromagnetic field reveals that predecessor laws of the electric force between charged particles, and the magnetic force between magnetic poles, are false. Quantum theory reveals that the whole of classical physics is false (Newtonian and Maxwellian theory). Relativistic quantum theory (quantum field theory) reveals that non-relativistic quantum theory is false. It is all too likely, then that current accepted fundamental theories of physics, general relativity and the standard model, will turn out to be false too, when a better theory emerges that unifies these current theories.

Third, there is an argument which provides strong grounds for holding that any dynamical theory of physics, which is about a restricted range of phenomena only, must be false. It goes like this. Theoretical physics only ever accepts *unified* fundamental dynamical theories, even though endlessly many empirically more successful disunified rivals are always available. This means that physics accepts a substantial metaphysical assumption about the nature of the universe: it is such that all (precise) disunified dynamical theories are false. The universe is such that some kind of unified pattern of physical law runs through all phenomena. Some such thesis as this is as secure an item of theoretical scientific knowledge as anything theoretical can be in physics. It is so secure, indeed, that physical theories that clash with it are rejected out of hand—not even considered—whatever their empirical success might be if they were considered.

Despite the validity of this argument, and its widespread dissemination in the literature since it was first published over 40 years ago in 1974,[8] it is still almost universally ignored. This is because it clashes with the orthodox conception of science firmly and almost unthinkingly taken for granted by scientists, philosophers of science and the public alike. This conception holds that, in science, evidence alone decides what theories are to be accepted and rejected. Considerations that have to do with simplicity, unity or explanatory power may influence choice of theory too, to a limited extent, but not in such a way that the world itself, or the phenomena, are permanently assumed to be simple, unified, or comprehensible. The decisive point is this: *no factual thesis about the nature of the universe can be accepted permanently, as a part of scientific knowledge independently of, and certainly not in violation of, the evidence*. This orthodox view, widely taken for granted, I call *standard empiricism*.

The above argument concerning persistent acceptance of unified theories demonstrates, however, that standard empiricism is false. Persistent acceptance of unified theories when endlessly many empirically more successful disunified rivals

---

[8] See, for example, Maxwell [9, 14, 16, 17, 22, 23].

are available means that science *does* make a big, persistent assumption about the universe independent of the evidence (even in a sense in violation of the evidence): the universe is such that some kind of unified pattern of physical law runs through all phenomena.

This thesis is, however, profoundly problematic. We have no reason to suppose that it is true. Even if it is true, the specific version of it that we hold at any stage in the development of physics is almost bound to be false. A glance at the history of physics reveals that we have changed our ideas a number of times as to what kind of unified pattern does run through all phenomena. I argue that we need a new conception of scientific method which facilitates the *improvement* of the big, problematic, metaphysical assumption physics is obliged to make. We need to represent this assumption in the form of a hierarchy of assumptions, these becoming less and less substantial as one goes up the hierarchy, and so more and more likely to be true, and more nearly such that the given assumption needs to be true for science, or the pursuit of knowledge, to be possible at all: see Fig. 9.1. As we descend the hierarchy, from level 7 to 3, assumptions become increasingly substantial, and increasingly likely to be false and in need of revision. At levels 6 to 3, we choose that assumption which (a) accords best with what is above in the hierarchy, and (b) has led to, or promises to lead to, the greatest empirical growth at levels 1 and 2, the levels of experimental results and testable theory. As physics advances, and we improve our knowledge and understanding of the universe, we improve assumptions and associated methods at levels 3 and 4; improving knowledge leads to improving assumptions and methods—improving knowledge about how to improve knowledge. There is something like positive feedback between improving scientific knowledge at levels 1 and 2, and improving knowledge about how to improve knowledge. Science adapts its nature to what it finds out about the nature of the universe—the key to scientific rationality which helps explain the astonishing, explosive growth in scientific knowledge. I call this hierarchical view of science *aim-oriented empiricism*.[9] One day, aim-oriented empiricism will replace the untenable but at present orthodox view of standard empiricism. In what follows, I take aim-oriented empiricism for granted.

At level 4 in the hierarchy of theses there is the thesis that the universe is such that a unified pattern of physical law runs through all phenomena. I call this thesis *physicalism*. According to aim-oriented empiricism, physicalism is a pretty secure item of theoretical scientific knowledge. So secure, indeed, that any physical theory that severely clashes with it is to be rejected on that account, whatever its empirical success may be.

Physicalism tells us that, whatever it is that determines precisely the way some specific kind of phenomenon evolves in space and time is precisely the same as that which determines the way *all* phenomena evolve, whether they be at the centre of

---

[9] For detailed expositions and defence of aim-oriented empiricism see works referred to in note 8, especially [23].

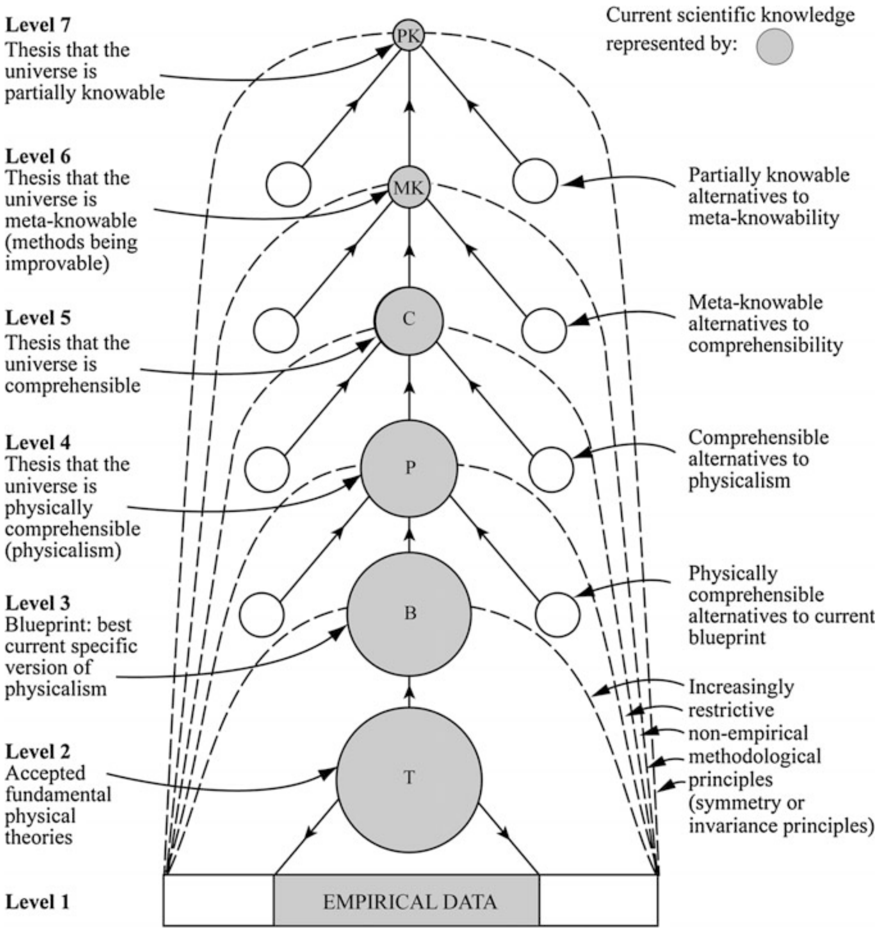


Fig. 9.1 Aim-Oriented Empiricism (AOE)

the earth, inside our heads, or inside the sun. This has the consequence that a theory that is precisely correct about the way any specific phenomenon, or any limited range of phenomena, evolve must thereby be correct about the way *all* phenomena evolve. This in turn means that a theory that is about a restricted range of phenomena only—so that it cannot be generalized to apply to all phenomena—must be only approximately valid about the phenomena to which it does apply. That is, it must be false. As far as fundamental physics is concerned, in order to be true about *anything* a theory must be true about *everything*. The only serious candidates for truth are so-called “theories of everything”—theories about all physical phenomena.

All physical theories so far developed (with the possible exception of string theory) are about restricted ranges of phenomena only. Hence, they are all false—even if some, such as quantum theory, are incredibly successful empirically, in making accurate predictions about a vast range of diverse phenomena. In particular, then, general relativity is false—despite its unity, and its empirical success.

Both special and general relativity are false.

9.4 Is the Three Dimensional View Viable Given that Relativity Theory Is False?

Does the falsity of relativity theory mean that the three dimensional view can be salvaged after all? That is by no means obvious. The falsity of relativity theory does not mean that all the empirical predictions of the theory are false as well. Indeed, we know that this is not the case. Special relativity is astonishingly successful empirically, and is integral to quantum electrodynamics, quantum electroweak theory, quantum chromodynamics, and the so-called *standard model* that puts these three theories together. And general relativity has met with predictive success as well. If special and general relativity imply that the 4 dimensional, spacetime view is true, then this implication may well be true even though the theories, from which it is derived, are ultimately false.

Everything depends on *how*, *in what way*, relativity theory is false.

Here, I explore the possibility that general relativity is false because it is a deterministic theory but the basic laws of nature are fundamentally probabilistic.[10]

9.5 Probabilism, Quantum Theory and the Three Dimensional View

Orthodox quantum theory (OQT) fudges two basic questions about the quantum domain:

1. What sort of entities are the objects of the quantum domain: electrons, photons, atoms?
2. Is the quantum domain fundamentally deterministic or probabilistic?

The first problem arises because electrons, atoms and other quantum entities seem, mysteriously to have both wave-like and particle-like properties. Niels Bohr, Werner Heisenberg, in developing quantum theory (QT), decided that the quantum wave/particle problem could not be solved and, as a result, developed QT as a

---

[10] For earlier discussion of this idea see [12, 20]. For a much more detailed discussion of the closely related issue of relativity and quantum non-locality see [6].

theory, not about quantum entities as such, but rather about the results of performing measurements on quantum entities. The version of quantum theory that emerged—OQT—avoids specifying what sort of entities electrons and atoms are when not being observed, but pays a very heavy price as a result. OQT is obliged to call upon some part of classical physics for a treatment of the measuring process. OQT cannot predict the outcome of measurement without classical physics, for that would require OQT to be able to specify, in a consistent way, what quantum entities are, and it is just this that OQT cannot do. (If OQT is applied to the measuring process, a further measurement must be made before OQT can issue in physical predictions.) Thus, the theory that makes physical predictions is quantum theory *plus some part of classical physics*. And this theory is unacceptably ad hoc and disunified, precisely because it is made up of two incompatible parts that are only rendered compatible by arbitrarily restricting their respective ranges of application (quantum postulates to quantum states, classical physics to macro measuring instruments).[11] As we saw in Sect. 3 above, a physical theory, in order to be acceptable, must be (a) sufficiently empirically successful and (b) sufficiently unified. OQT satisfies condition (a) magnificently, but fails dismally to satisfy condition (b).

Failure to solve the quantum wave/particle problem also leads to the failure of OQT to answer unambiguously whether the quantum domain is deterministic or probabilistic. According to OQT, quantum states evolve deterministically, in accordance with Schrödinger's time-dependent equation. Probabilism only enters in when measurements are made. But OQT is wholly ambiguous as to whether measurement really does involve the occurrence of objective probabilistic events. On the one hand, OQT in general makes probabilistic predictions about the outcome of measurements. But, on the other hand, OQT can hardly be interpreted as asserting that probabilistic events really do occur when, and only when, measurements are made. Is it conceivable that the universe would have had to wait for billions of years for physicists to make measurements, and thus provoke probabilistic events to occur for the first time? Secondly, if Schrödinger's equation is applied to the measuring process, everything proceeds *deterministically*!

In order to develop an acceptable version of quantum theory (QT), the above two questions must be clearly answered, and not just fudged. Granted that both determinism and probabilism are possible, are there any grounds for favouring one over the other?

There are. OQT, despite being unacceptably disunified, is nevertheless quite astonishingly successful empirically. OQT predicts a greater variety of phenomena with greater accuracy than any other physical theory. It has never been refuted. OQT has evidently got a great deal right about the quantum domain; it is just that it fails to solve the wave/particle problem, and thus must incorporate measurement into the theory, thus rendering it unacceptably disunified.

---

[11] This argument is spelled out in greater detail in Maxwell [8, 10]. See also [2, 13].

What all this suggests is that we should, initially, seek to keep as close to the structure of OQT as possible, and modify the theory just sufficiently to eliminate the devastating defects of the theory: the failure to solve the wave/particle problem, the resulting need to call upon measurement in an essential way, and the failure to answer the question: Is nature deterministic or probabilistic?

In order to do this, the decisive step we need to take is to eliminate "measurement" from the basic postulates of the theory.[12] That in turn requires that we solve the wave/particle problem. Probabilism, dramatically, enables us to do both!

OQT says: quantum states evolve deterministically, in accordance with Schrödinger's equation, until a measurement is made, when (in general) something ostensibly probabilistic occurs. If we adopt the conjecture of probabilism, we can modify this very slightly (implementing the above plan), so that QT asserts: quantum states evolve deterministically, in accordance with Schrödinger's equation, until *specific quantum mechanical physical conditions* arise, and an *objectively probabilistic event* occurs.

At once we have a solution to the wave/particle problem! Given probabilism, "Are quantum entities waves or particles?" is the wrong problem. Waves (or fields) and particles come from *deterministic* classical physics. If probabilism holds, and quantum entities interact with one another *probabilistically*, they ought to be quite different from classical deterministic entities, such as classical waves or particles. Granted probabilism, the wrong, traditional question "Waves or particles?" needs to be replaced by the two correct questions: "What kind of unproblematic, fundamentally probabilistic entities are there, as possibilities?" "Can we see quantum entities as some variety of these possible, unproblematic, fundamentally probabilistic entities?" Fundamentally probabilistic entities may be called *propensitons*.

Propensitons could be such that probabilistic events occur continuously in time, or they could be such that their physical state evolves deterministically until specific physical states of affairs arise, when and only when a probabilistic transition occurs. We may call the latter *intermittent propensitons*. Our strategy is to stick as close to the structure of OQT as possible. This requires that we hold that electrons, atoms and the rest are intermittent propensitons.

The simplest, most elementary example of an intermittent propensiton that one can think of is the following. It is, initially, a tiny sphere, of a definite radius. This expands, at some fixed rate until it touches another such sphere. That is the condition for both spheres to shrink instantaneously into the initial tiny spheres, each to be located probabilistically somewhere within the volume of its big sphere state.

We can now make this intermittent propensiton a bit more interesting by postulating that each sphere has within it a "stuff", which varies in density in a wave-like way. The density of this "stuff" determines the probability of location of the tiny sphere. Already we have an entity which can begin to mimic the behaviour of quantum entities. We can imagine sending a stream of these "spheres" at a

---

[12] I first made this point in [8].

two-slitted screen: most of the sphere's "stuff" is reflected from the screen, but some passes through both slits. The waviness of the "stuff" then interferes, producing bands on a second screen, where conditions are such as to localize probabilistically the spheres. This elementary intermittent propensiton is able to mimic the results of the two-slit experiment, generally held to illustrate most strikingly the enigmatic "wave/particle" behaviour of quantum entities.

There is just one final modification that needs to be made to our elementary intermittent propensiton, and we have full quantum theory. We specify that the propensity state of the propensiton is specified by $\Psi(r, t)$, a complex function that attributes a complex number to each point in space, r, at a given time, t. We specify (a) that $\Psi(r, t)$ evolves in time in accordance with Schrödinger's equation, and (b) that $|\Psi(r, t)|^2 dV$ gives the probability of the propensiton being localized in the volume element dV. We have all but recovered quantum theory, but with this difference: quantum theory has become a fully realistic, fundamentally probabilistic theory about the evolution and probabilistic interactions of intermittent propensitons. This version of QT—*propensiton quantum theory* (PQT) as we may call it—solves the wave/particle problem by declaring that electrons, photons, atoms, etc., are *intermittent propensitons*.

But a key problem remains to be solved: What is the quantum mechanical condition for probabilistic transitions to occur? My proposal here, first put forward in 1982, is that probabilistic transitions occur if and only if "particles", or bound systems are created or destroyed as a result of inelastic interactions.[13] For example if an electron collides with a hydrogen atom so that the system goes into a superposition of (a) an electron and hydrogen atom and (b) two electrons and one proton (the hydrogen atom being dissociated), then, although this superposition exists, it does not persist. Entirely in the absence of measurement, or interaction with an external environment, when the interaction is very nearly at an end, the superposition jumps probabilistically into *either* (a) the electron plus hydrogen atom state, *or* (b) the two electron and one proton state, the probabilities being those predicted by OQT (if a measurement were made). As I have shown elsewhere, this version of PQT recovers all the empirical success of OQT but nevertheless differs from OQT for as yet unperformed experiments.

Some years after I first put forward this version of PQT, Roger Penrose proposed another version. Both versions hold that probabilistic transitions are associated with superpositions involving mass. But whereas I hold that what is required is a superposition of states, each with different particles or bound systems associated with them (of different mass), Penrose's idea is that what is required is that a sufficiently massive body evolves into a state that is a superposition of different *positions* in space.[14]

Just conceivably, my version of PQT, Penrose's version, or some other version, may be correct. Let us, in any case, conjecture that this is so. The quantum domain

---

[13] See [11, 15, 18, 19, 21].

[14] Reference [26]. For a quite different proposal for probabilistic collapse see [4].

is such that, when specific quantum mechanical conditions arise, quantum states undergo instantaneous probabilistic transitions.

At once PQT comes into sharp conflict with special relativity.[15] Probabilistic changes of quantum state, of the kind required by PQT, will be such that the quantum state changes *instantaneously* throughout a region of space. But such a change will only be instantaneous in one reference frame.[16] In all other reference frames in motion with respect to this one frame, the change of state will not be instantaneous: it will travel at a faster than light, but finite, velocity, from one spatial position to another. If the change of state is caused by an interaction occurring in a relatively small spatial region, dV, then in some reference frames, a part of the quantum state far away spatially from dV, will begin to change, and to travel towards dV, before the event has occurred which causes the probabilistic transition. In these frames we have the absurdity that changes of quantum states *anticipate* future events!

Instantaneous probabilistic changes of quantum state, required by PQT, conflict with special relativity because they pick out a privileged reference frame, in which the change of state *is* instantaneous. Special relativity, however, demands that all inertial reference frames are equivalent, no one frame being uniquely privileged in representing the laws of nature.

Here, then, is a way in which relativity theory might be false in such a manner that the three dimensional view of the universe can be salvaged. Instantaneous probabilistic changes of quantum state pick out the unique cosmic “now” required to exist by the three dimensional view.

9.6 Ontological Probabilism and Relativity Theory Reconciled

It may be, however, that probabilism, if true, does not just *falsify* relativity theory. There is the possibility that a kind of *partial reconciliation* can be brought about between probabilism and relativity theory. Both might, in a sense, be true. Perhaps we can have our cake and eat it.

Let us conjecture that ontological probabilism is true, and hence the three dimensional view is true as well. Some version of PQT is correct: probabilistic changes of quantum state occur instantaneously across vast regions of space. These

---

[15] Probabilism and special relativity do not inevitably contradict one another, a point I made in [12]. This point is borne out by the existence of a fundamentally probabilistic, Lorentz invariant version of quantum theory developed by Tumulka [28]. This is a version of the theory of [4]. Tumulka’s theory suffers, however, from severe limitations: its ontology is that of discrete spacetime points, there are no interactions between “particles” and, most serious of all, nothing corresponds physically, in reality, to the quantum state.

[16] Here, and in what follows, I take “one rest frame” to mean “one set of reference frames all at rest with respect to each other”.

spatial regions may, incidentally, be very big indeed. Consider a photon emitted by an early star some tens of millions of years after the big bang. This has travelled in opposite directions for some 13 billion years, to be absorbed by the eye of someone gazing at the night sky. Instantaneously, the photon is annihilated. One instant it fills a region of space over 26 billion light years across; the next instant it has ceased to exist. Such an instantaneous collapse of state picks out a privileged reference, and an instantaneous "now", with a vengeance.

But let us suppose that such probabilistic collapses of quantum state are *all that there is in the constitution of things* to pick out "the reference frame at rest", and thus "the cosmic now". There is nothing else in the physical constitution of things that picks out the privileged reference frame to be the frame at rest. In particular, *that which determines how physical states of affairs evolve deterministically in time in between probabilistic transitions has nothing associated with it which can determine which frames are at rest, which in motion*. As far as deterministic evolutions of physical states are concerned, any inertial reference frame or, more generally, any coordinate system, is as good as any other. To a first approximation, special and general relativity are correct. Only probabilistic transitions reveal the serious inadequacy in relativity theory, in that they pick out the "cosmic-wide now" at each instant, thus picking out one frame of reference to be the privileged rest frame.

Our supposition is, then, that this state of affairs will be a feature of the yet-to-be-discovered true, unified "theory of everything"—a unification of probabilistic quantum theory (or the standard model) and deterministic general relativity which we may call *probabilistic quantum gravity* (PQG). This theory comes in two parts. There is (1) the *deterministic* part which specifies how physical states of affairs, propensity states, evolve in time in between probabilistic transitions; and there is (2) the *probabilistic* part which specifies the physical conditions that must obtain if probabilistic transitions are to occur, with the actualization of propensites, just one of many possible states of affairs becoming actual with such and such probability.

As far as part (1) is concerned, special and general relativity both hold, to a first approximation; only part (2) reveals the serious inadequacy in both theories.

The viewpoint just outlined is to be contrasted with the neo-Lorentzian view defended by Craig [3], and very effectively criticized by Balashov and Janssen [1]. According to special relativity, a body, A, set in motion with respect to another one, B, will be such that lengths in the direction of motion are contracted, clocks go slow, and masses increase, as measured by B. And all this will be true as well of lengths, clocks and masses travelling with B, as measured by A. These results follow from the basic postulates of special relativity: all inertial reference frames are equivalent, light has the same velocity c in all inertial reference frames, and space is homogeneous and isotropic. The neo-Lorentzian view supported by Craig rejects the first of these postulates, even as far as the deterministic evolution of physical states of affairs is concerned. Such evolutions of physical states pick out a unique reference frame to be the one at rest. Neo-Lorentzianism, as a result, faces two problems: (1) What exists physically that provides a basis for picking out that unique reference frame that is, objectively, at rest? (2) How can the behaviour of rods, clocks and masses in motion, successfully predicted and explained by special relativity, be explained by Neo-Lorentzianism? Neo-Lorentzianism fails to answer these two questions—especially question (2). In connection with question (1), Craig suggests that, given general relativity, and given the homogeneity and isotropy of the cosmos, a uniquely natural foliation of spacetime into space and time emerges which provides a rest frame at each point in space, and a unique cosmic time.[17] Neo-Lorentzianism provides no satisfactory answer to (2), no explanation for the behaviour of rods clocks and masses in motion so beautifully explained by special relativity. It fails, especially, to explain, given two bodies, A and B, in relative motion, why A observes B's lengths to shrink, clocks to go slow, and masses to increase, and B observes precisely the same of A!

The *reconciliation* viewpoint, indicated above, faces none of these difficulties. According to this view, *only quantum probabilistic transitions* determine, physically, that frame uniquely at rest. Furthermore, as far as deterministic evolutions of physical states of affairs are concerned, nothing exists that can pick out the rest frame. All inertial reference frames are equivalent. The postulates of special relativity hold (to a first approximation), and so the consequences of those postulates hold as well. The problem of *explaining* why rods, clocks and masses in motion behave as predicted by special relativity, which poses an insuperable problem for Neo-Lorentzianism, poses no problem at all for the reconciliation view, indicated above.

If we are to pursue this "reconciliation" view with complete integrity, we must acknowledge, it seems, that deterministically evolving physical states, being inherently indifferent to whether they are moving or at rest (in an absolute sense), have no way of determining what space-like hyperplane or hypersurface the instantaneous probabilistic change of physical state occurs, when conditions for it to occur obtain. The equations of motion, and the conditions for probabilistic collapse to occur, leave open in what hyperplane the collapse occurs. This is determined by "the cosmic now" of the three dimensional universe.

Can the hyperplane (or hypersurface) of probabilistic collapse be determined experimentally? What answer is to be given to that question will depend, I take it, on the specific form *probabilistic quantum gravity* takes.

References

1. Balashov, Y., & Janssen, M. (2003). Presentism and relativity. *British Journal for the Philosophy of Science, 54*, 327–346.
2. Bell, J. (1973). Subject and object. In J. Mehra (Ed.), *The physicist's conception of nature* (pp. 687–690). Dordrecht: Reidel.
3. Craig, W. L. (2001). *Time and the metaphysics of relativity*. Dordrecht: Kluwer.

---

[17] (Reference [3], ch. 10).

4. Ghirardi, G. C., Rimini, A., & Weber, T. (1986). Unified dynamics for microscopic and macroscopic systems. *Physical Review D, 34*, 470–491.
5. Greene, B. (1999). *The elegant universe*. London, W.W: Norton.
6. Maudlin, T. (2011). *Quantum non-locality & relativity*. Chichester: Wiley-Blackwell.
7. Maxwell, N. (1968). Can there be necessary connections between successive events? *British Journal for the Philosophy of Science, 19*, 1–25.
8. Maxwell, N. (1972). A new look at the quantum mechanical problem of measurement. *American Journal of Physics, 40*, 1431–1435.
9. Maxwell, N. (1974). *The rationality of scientific discovery, philosophy of science* (vol. 41, pp. 123–153 and 247–295).
10. Maxwell, N. (1976). Towards a micro realistic version of quantum mechanics. *Foundations of Physics, 6*, 275–292 and 661–676.
11. Maxwell, N. (1982). Instead of Particles and Fields. *Foundations of Physics, 12*, 607–631.
12. Maxwell, N. (1985). Are Probabilism and Special Relativity Incompatible? *Philosophy of Science, 52*, 23–43.
13. Maxwell, N. (1992). Beyond FAPP: three approaches to improving quantum theory and an experimental test. In A. van der Merwe, et al. (Eds.), *Bell's theorem and the foundations of modern physics* (pp. 362–370). Singapore: World Scientific.
14. Maxwell, N. (1993). Induction and scientific realism: Einstein versus van Fraassen. *British Journal for the Philosophy of Science*, *44*, 61–79, 81–101 and 275–305.
15. Maxwell, N. (1994). Particle creation as the quantum condition for probabilistic events to occur. *Physics Letters A, 187*, 351–355.
16. Maxwell, N. (1998). *The comprehensibility of the universe*. Oxford: Oxford University Press.
17. Maxwell, N. (2004). *Is science neurotic?*. London: Imperial College Press. chs. 1 and 2, and appendix.
18. Maxwell, N. (2004). Does probabilism solve the great quantum mystery? *Theoria, 19/3*(51), 321–336.
19. Maxwell, N. (1988). Quantum propensiton theory: a testable resolution of the wave/particle dilemma. *British Journal for the Philosophy of Science, 39*, 1–50.
20. Maxwell, N. (2006). Special relativity, time, probabilism, and ultimate reality. In D. Dieks (Ed.), *The ontology of spacetime* (pp. 229–245). : Elsevier.
21. Maxwell, N. (2011). Is the quantum world composed of propensitons? In M. Suárez (Ed.), *Probabilities, causes and propensities in physics* (pp. 221–243). Dordrecht: Synthese Library, Springer.
22. Maxwell, N. (2013). Has science established that the cosmos is physically comprehensible?. In A. Travena & B. Soen (Eds.) *Recent Advances in Cosmology* (pp. 1–56). New York: Nova Publishers Inc. Chapter one, Retrieved, from http://discovery.ucl.ac.uk/view/people/ANMAX22.date.html.
23. Maxwell, N. (2017). *Understanding scientific progress*. Saint Paul, Minnesota: Paragon House.
24. McTaggart, J. M. E. (1908). The unreality of time. *Mind, 17*, 457–473.
25. Minkowski, H. (1952). Space and time. In A. Sommerfeld (Ed.), *Einstein: the principle of relativity* (pp. 75–91). New York: Dover.
26. Penrose, R. (1986). Gravity and state reduction. In R. Penrose & C. J. Isham (Eds.), *Quantum concepts of space and time* (pp. 129–146). Oxford: Oxford University Press.
27. Smolin, L. (2007). *The trouble with physics*. London: Allen Lane.
28. Tumulka, R. (2006). A relativistic version of the ghiraardi-rimini-weber model. *Journal of Statistical Physics, 125*, 821–840.
29. Woit, P. (2006). *Not even wrong*. London: Jonathan Cape.

Part II
Physics

Chapter 10
Nature's Book Keeping System

Gerard 't Hooft

Abstract Establishing how one should describe and study nature's fundamental degrees of freedom is a notoriously difficult problem. It is tempting to assume that the number of bits (or qubits) needed in a given Planckian 3-volume, or perhaps 2-volume, is a fixed finite number, but this ansatz does not make the problem much easier. We come not even close to solving this problem, but we propose various ingredients in phrasing the questions, possibilities and limitations that may serve as starting points.

10.1 Introduction

When formulating a theory concerning the physical world, one usually just assumes that the basic kinematical variables are known, and can be enumerated, and studied by some experimental technique. Think of stars and planets in some star system, a star or planet itself, or eventually some agglomeration of elementary particles, as the case may be. The more basic and interesting part of the theory is then the attempt to explain the behaviour of these variables, in terms of some mathematical scheme, often using symmetries to make the problem more manageable.

One clearly assumes then that we know what the main variables are, how their properties, such as coordinates and momenta, can be listed, and most importantly, how we can maintain a book keeping system such that the results of our calculations can be checked against experimental observations.

In many practical cases, such a book keeping system appears to be totally under control, so that it requires little further discussion. Assuming that natural phenomena

---

A contribution to the collection "Space, Time, and Frontiers of Human Understanding".

G. 't Hooft (✉)
Institute for Theoretical Physics, Utrecht University and Spinoza Institute,
Postbus 80.089, 3508 TB Utrecht, The Netherlands
e-mail: g.thooft@uu.nl
URL: http://www.staff.science.uu.nl/~hooft101/

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_10

have a deductive ordering, we are tempted first to focus on nature's book keeping system at the tiniest distance scales. Once that it understood, we can go to higher levels and arrange our schemes for those branches of physics as well.

In this treatise, we focus on the more primary question: what *are* nature's most fundamental kinematic variables, and how do we find out about them? In short: what is nature's book keeping system at the most fundamental level?

10.2 The Sub-atomic World

Today, the terrain of the tiniest distance scales of physics where things are still pretty clear, is the world of the sub-atomic particles. How to do the book keeping system there, and indeed also how to perform the next step, which is the detailed description of nature's evolution laws, was unravelled during most of the 20th century.

The way to list the dynamical degrees of freedom for the subatomic particles is dictated by nature's symmetry laws, notably special relativity, in combination with the laws of quantum mechanics. We are in the fortunate situation that this combination is so restrictive that only one scheme survived, called quantum field theory. The situation still seemed to be so complex that most of the 20th century was needed to handle all obstacles, but the outcome was crystal clear.

At first sight, there appear to be two distinct descriptions for the subatomic particles, to be referred to as *field* space and *Fock* space. When we choose the field space option, we register the functional distribution of all *fields* in the system, of which there are three types: the gauge connection fields $A^a_\mu(\vec{x},t)$, which are vectors in 4 dimensional Minkowski space-time, the spinorial Dirac fields $\psi^i_\alpha(\vec{x},t)$, $\overline{\psi}_i^{\,\alpha}(\vec{x},t)$, which form a Grassmann algebra, and some scalar fields $\phi^a(\vec{x},t)$. Now by far the most practical way to register the fields is to list the values of their Fourier components, so that the fields are expanded in terms of plane waves. Each of these waves behaves as a harmonic oscillator. One subsequently subjects these oscillators to the rules of quantum mechanics. In accordance with whatever is dictated by the (relativistic) field equations, the oscillators will be weakly or strongly coupled. The (multi-dimensional) Schrödinger equations resulting from this can now be handled by standard methods in mathematical physics.

The Fock space solution may seem to be something altogether different. Here, we do not talk of fields, but of states with arbitrary numbers of elementary particles present. These particles, and their anti-particles, may have spin 1, $1/2$, or 0, and we list which of them are present, and what their coordinates $\vec{x}$ are, at the time t. The particles may form quantum wave functions, and their motion is dictated by relativistic equations. The particles may interact weakly or strongly, and their energies can be calculated. During an interaction, particles can be created or destroyed.

Surprisingly, it was found that these two ways to register what is going on, are different sides of the same coin. The operator fields describing the creation or annihilation of particles, are to be identified with the fields mentioned previously, while the energy quanta of these fields are identified with the particles. One then discovers that both descriptions are totally equivalent. They lead to exactly the same predictions for the outcomes of an experiment. Analysing the situation further, one finds that both schemes are subject to the *energy conservation law*: every configuration has a characteristic energy, and the total energy is a positive, conserved quantity. This law dictates that simple field or particle configurations also evolve into simple configurations. More involved configurations require more energy. Thus, we should be aware of the fact that the field / Fock space method owes its success to the energy conservation law.

10.3 Gravitation

All this is not at all the final word on nature's book keeping system. What was kept out of the picture so-far, was the gravitational force. This force obeys Einstein's equations for gravity, but these are subject to a new symmetry: general relativity.

Of course, what physicists in the second half of the 20th century tried to do, is to adapt the field / Fock description of nature's degrees of freedom to these new field equations, discovering that also new particles are needed: *gravitons*, which carry spin 2, and perhaps *gravitinos*, with spin $3/2$. What went wrong?

Technically, at least two things went wrong. One was that, space and time, or the coordinates that ought to be used to register our particles and our waves, themselves become dynamical, due to space-time curvature, as if the books in which we want to write the whereabouts of our particles, began to lead lives of their own. Closely related to this was the problem that the new field equations, and the particle equations alike, became *non-renormalizable*. This means that the ultra-small distance behaviour runs out of control.

But something else went wrong: something wicked happens with the energy conservation law: gravity is an *attractive* force, and it becomes stronger if mass, or equivalently, energy, conglomerates. Energy increases if we concentrate on smaller distance scales. The energies of particles, regardless how much mass they have, can approach to zero and even, in a sense, become negative. The strength of the gravitational force then totally runs out of control.

An other way of saying this is as follows. Since gravity *attracts* equal sign masses, while electric and magnetic forces, as well as the weak force, act repulsively when charges carry equal signs, gravitational fields carry a negative field energy density, which may compensate for any positive amount of energy that other particles or fields might have. This implies that the energy conservation law can no longer be called upon to keep the laws of physical interactions under control. Nature's book keeping rules will here have to be *entirely* different from what we are used to in particle physics.

In fact, there seems to be no strong impediment against the formation of ultra-tiny black holes. If two particles approach one another too closely with too much energy, black hole formation is inevitable. In quantised general relativity, neither Fock space nor field space provide useful lists of all possible configurations. It should not be a surprise that we still have not come with credible, let alone complete, laws on how things evolve at such small distances. The distance scale at which things are sure to go haywire, is the so-called Planck length,

$$L_{\text{Planck}} \;=\; \sqrt{\hbar\, G/c^3} \;=\; 1.6162 \times 10^{-33}\,\text{cm}\,, \tag{10.3.1}$$

where G is Newton's constant. This is roughly a billion times a billion times smaller than the sizes of particles studied in the Large Hadron Collider. Now imagine a volume of space of just a few Planck units cubed. What can we suggest about nature's book keeping system to describe *anything* one *can* imagine sitting in such a tiny region?

10.4 Superstrings and Black Holes

There are numerous clues that may be considered. A very powerful proposal is that of *superstring theory*. Representing the tiniest objects in physics as being line-like instead of point-like, yielded equations that are much more restrictive than the older relativistic equations for point-like objects, and the renormalization difficulties seem to disappear. Strings refuse to be crammed into tiny volumes because of their symmetry constraints.

There is much support for this idea, but I do see various problems with it. One is that, just because strings cannot be folded into tiny volumes, the notion of locality seems to be in jeopardy. An other is the lack of a local energy law, which could destabilise these systems. Then, even if infinities may cancel out, there still seem to be no obvious bounds as to how much information can sit in a tiny volume. Finally, there is the issue with the microscopic black holes.

In short, strings seem to be far from orderly book keepers. We should search for something better. Of course, we do not exclude the possibility that, after finding a more satisfactory answer, string theory may return to become the most promising game in town.

A more systematic attack comes from the black holes themselves. Black holes as tiny as the Planck scale will be difficult to understand as yet, but when their typical length scale, the radius $r = 2GM/c^2$, is more than an order of magnitude bigger, then standard quantum field theory dictates their behaviour: such black holes emit all sorts of particles, with a an apparently perfectly thermal[1] spectrum, called *Hawking radiation* [2]. From this calculation, it is easy to derive the entropy of a black hole, and a little statistical physics then reveals a startling feature: *The total number of distinct quantum states a black hole can be in, is an exponential function of the horizon area* [3].

---

[1] This neat picture has recently been questioned; thermodynamics alone does not suffice for the book keeping of particles near a black hole horizon [1].

This is only a small number of states, and the fact that it depends on the area rather than the total volume, came as a surprise. It looks as if black holes require a *very simple* book keeping law: one boolean degree of freedom (something like a spin variable that can only take the values ± 1), for every fundamental little square on the horizon. The fundamental surface of such a square will be

$$\delta\Sigma = (4\ \log 2)\, L^2_{\text{Planck}}\ . \tag{10.4.1}$$

Since black holes seem to be the most compact form of matter that can exist in nature, having one boolean degree of freedom per unit of surface area seems to be an absolute maximum [4]. This is referred to as the *holographic principle*: nature's information content can be measured in terms of bits per cm^2, with the inverse Planck length squared as an absolute bound [5]. One gets the impression that nature might obey a beautiful book keeping system.

All we have to do now is insert the details, and figure out nature's laws about processing these data.

This turns out to be prohibitively difficult today. One difficulty is Lorentz invariance. Suppose we have a situation with the maximal amount of information on a large surface. Now, perform a Lorentz transformation that keeps the surface in place. We get an other configuration, which is the previous one, but Lorentz contracted. There seems to be no limit on how many Lorentz contractions one can perform, but there should be such a limit, otherwise our book keeping system would not be finite.

Black holes do suggest what to do here: according to the statistical interpretation of the Hawking radiation emitted by the black hole, particles going in transfer their information onto the particles going out. One can calculate how this happens, by studying the curvature of space-time as caused by the gravitational fields of the in- and out-going objects [6]. The longitudinal location of the out-particles turns out to depend on the momenta of the in-going ones, and this is an important clue: nature's book keeper throws in-going particles into the basket of out-going ones, if only we could understand in more detail how this happens. We *calculate* how the gravitational force between in-going particles and out-going particles brings about this behaviour. These investigations may pay off already at the domain of the Standard Model itself: all particles species will have to leave their unique imprints in the fabric of space and time. Yet, this calculation did not lead to a properly finite book keeping device—although we think we came close [1, 7].

In short, the difficulty seems to be that Einstein's gravity equations include fields that describe the curvature of space and time. This adds a twist to our story: nature has to write its book keeping records on crumbled material: curved space-time. How do we make sense of that?

A daring answer could be that, for putting down information regarding space, time and matter, one has to replace this folded space-time material by something that is almost flat. It would be against Einstein's general invariance principle, but we may have to prepare for prices of this kind that will have to be paid. Nature might be giving us clues concerning this point: our universe appears to be *almost* flat, too flat to make sense in an uncensored Einsteinian theory.

One guiding principle was strongly adhered to by practically all researchers: anything as small as molecules, atoms, subatomic particles and other energy quanta, or smaller, always required the language of *quantum mechanics* to describe their properties. Most researchers take this to mean that we will need to use this quantum mechanical language no matter which approach we try. Yet, here also, one can have doubts. To me, quantum mechanics seems to be a *tool* rather than a theory. This means that, the most commonly employed interpretation of quantum mechanics, a discipline that sometimes seems to challenge our sense of logic to the extreme, may not always be justified. One can speculate that nature's ultimate book keeping device at the tiniest possible scales, must be simpler than any quantum mechanical one [8].

One should immediately add that, giving up general relativity and/or quantum mechanics, will bestow on us the burden to explain why both principles work so well in all domains of physics that have been explored at present.

Numerous other schemes have been proposed, over time. As is often the case in situations such as these, there have been an abundance of completely wild concoctions that serious researchers have come up with. Lack of phantasy is not our problem [9–11]. The question is rather, how to find a direct chain of solid arguments, with many calculations on the way to corroborate intermediate conclusions, leading to an unavoidable result that explains how particles, fields, energy quanta or *whatever*, can be listed in any tiny domain of our physical world, such that, when we take into account the physical laws controlling their behaviour, we find out how things behave at larger scales, until we hit the domain of ordinary quantum field theories, with which we can calculate almost anything we wish to know.

Note our emphasis on scales and sizes of things. In ordinary physics there seems to be no limit on how small one can imagine fundamental units of matter to be. In an orderly controlled physical world, there must come an end to these scale transformations. There is one natural way in which one may imagine such an "end of scales" to arrive in our book keeping system: scaling symmetry or more precisely, the introduction of *exact* local scale- and conformal invariance. One can then employ the *same* book keeping variables to describe all scales to come. The use of such a symmetry, which must be a *local* symmetry, is quite well-known in theoretical particle physics, and is called local gauge invariance. Local gauge invariance can be 'spontaneously broken' (the well-known BEH mechanism). But the known gauge theories mostly refer to internal rotation groups, not scaling transformations. Again, here General Relativity may help us to explore new grounds: coordinate transformations do include scale transformations, and this means that a locally conformal invariant description of gravitational effects [12, 13] is possible, in principle. It is the present author's view that one must look at gravitation from this perspective. General relativity features an exact local conformal gauge symmetry, which is spontaneously broken just as in BEH.

Thus, we emphasise, the bookkeeping system that we are after, should be able to describe literally all states that can occur in the physical world. Without going into any details, let us briefly recapitulate what we may have arrived at today.

10.5 Nature's Ultimate Book Keeping System

When trying to describe nature's most general state, we have to ask what the ultimate state will look like when we squeeze as much matter as is possible into the tiniest possible volume. One usually expects one or more black holes to form. All presently known laws of fundamental physics make use of quantum mechanics, so, even if one is skeptical about this, it is natural first to search within the quantum formalism. A very important notion then is *unitarity*: two states that start out to be orthogonal and normalised, evolve into another pair of states that is also orthogonal and normalised.

Laws of gravity, as they are known today, then suggest that all forms of matter will be *geometric*: the way they affect the curvature of space and time is the only form of information that is conserved [1, 6, 7]. We think we drew the important conclusion that observations of the sort mentioned here, will be the only way to reconcile finiteness of the degrees of freedom with an unbounded group of local Lorentz transformations. This is extremely important, if true. It means that Fock space will not be the appropriate language; rather, we get something that resembles a bit more string theory, which is also basically geometric. String theories known today, however, seem not yet to be based on very sound book keeping.

Indeed, for this, and other reasons, string theory has been under attack [14]. According to Smolin, the internal logic in string theories is too hazy to serve as a sound physical prescription. Consequently, it is all but impossible to extract reliable predictions from string theories, other than the claim that everything will soon be understood. These attacks are understandable, but not totally justified. String theory is a collection of quite impressive mathematical constructions and theorems. The feeling, shared by many of its practitioners, is that these mathematical notions must mean something, and that our physical world is likely to require such ingredients. String theory tells us that there are not only point particles, described by fields that can be used for a Fock space formalism, but in addition, we have stringlike features, as well as two-dimensional membranes and higher dimensional structures. This is an important lesson, and it may well be that this idea will be important for our theoretical thinking in the future.

The demand from string theory that space and time themselves must feature either 10 or 26 dimensions, however, seems to be too restrictive. If indeed, as we suspect, physical degrees of freedom form discrete sets, then dimensionality of space and time may be less well-defined notions, so that such 'predictions' from string theory may well be due to some mathematical idealisation having little to do with reality. All in all, we are badly in need of a more orderly listing of all conceivable configurations of physical variables in a small region of space and time.

As the reader may notice, this demand by itself is also not formulated very sharply. This is exactly the point we wish to bring forward in this paper. Part of our problem is the precise formulation of our question or questions. Nature's book keeping system must be outlined together with the answers to the question how the variables will evolve in time. It is important to observe that the big revolutions in science often came with improvements in our way to phrase the physical degrees of freedom, together with new proposals as to how these evolve. The grand total is what we call physical law.

Smolin complains that science today is slowing down considerably, and blames this to the rise of string theory and its staggering promises, at the expense of other approaches in the basic sciences.

But Smolin forgets that the real revolutions in science are often only recognised in hindsight. To our judgement, it is quite conceivable that many or all of our questions concerning nature's book keeping system will be solved in the not so distant future. However, the road towards these solutions will consist of very small but mathematically precisely formulated steps in our way of thinking. String theory was an interesting guess, but may well have been a too wild one. We are guessing the mathematical structures that are likely to play a role in the future, but we fall short on grasping their internal physical coherence and meaning. For this, more patience is needed.

References

1. 't Hooft, G. Black hole unitarity and antipodal entanglement, *Foundations of Physics, 49* (9), 1185–1198. doi:10.1007/s10701-016-0014-y, arXiv:1601.03447v4. [gr-qc].
2. Hawking, S. W. (1976). Particle Creation by Black Holes, *Communication in Mathematical Physics*, *43* (1975) 199, Erratum ibid. 46 (1976) 206; Hartle, J. B. & Hawking, S. W. Path (1976) Integral derivation of black hole radiance. *Physical Review*, *D13*, 2188.
3. Susskind, L., Thorlacius, L., & Uglum, J. (1993). The stretched horizon and black hole complementarity. *Physical Review D*, *48*, 3743–3761. arXiv: hep-th/9306069.
4. Bekenstein, J. D. (1981). Universal upper bound on the entropy-to-energy ratio for bounded systems. *Physical Review D*, *23*, 287–298. doi:10.1103/PhysRevD.23.287.
5. Susskind, L. (1995). The world as a hologram. *Journal of Mathematical Physics*, *36*, 6377. hep-th/ 9409089.
6. 't Hooft, G. (1991). The scattering matrix approach for the quantum black hole: an overview. *Journal of Modern Physics*, *A11* (1996) pp. 4623–4688. arXiv:gr-qc/9607022; id. The Black hole horizon as a quantum surface, *Physica Scripta*, *T36* (1991) 247.
7. 't Hooft, G. (2015). *Diagonalizing the Black Hole Information Retrieval Process*, arXiv:1509.01695; id., *Singularities, horizons, firewalls, and local conformal symmetry*, presented at the 2nd Karl Schwarzschild Meeting on Gravitational Physics, Frankfurt, July 23, 2015, arXiv:1511.04427. [gr-qc].
8. 't Hooft, G. *The Cellular Automaton Interpretation of Quantum Mechanics*, http://arxiv.org/abs/1405.1548, Springer, to be published
9. Sorkin, R. (2002). *Causal Sets: Discrete Gravity*, Notes for the Valdivia Summer School, Jan. 2002, arXiv:gr-qc/0309009, and references therein.
10. Rovelli, C. (1998). Loop Quantum Gravity, *Living Reviews in Relativity*, *1*; Smolin, L. (2001). Three Roads to Quantum Gravity. ISBN:0-465-07835-4.
11. Ambjorn et al. J. *Quantum Gravity via Causal Dynamical Triangulations*, arXiv:1302.2173v1 [hep-th], to appear in *Handbook of Spacetime*, Springer Verlag.
12. Mannheim, P. D. (2007). Solution to the ghost problem in fourth order derivative theories. *Foundations of Physics*, *37*, 532–571. arXiv:hep-th/0608154.

13. 't Hooft, G. (2015). Local conformal symmetry: the missing symmetry component for space and time, ITP-UU-14/25; SPIN-14/19, *International Journal of Modern Physics D*, *24*(12) 1543001 (5 pages), arXiv:1410.6675 v2 [gr-qc].
14. Smolin, L. (2006). *The Trouble with Physics, The Rise of String Theory, the Fall of a Science, and What Comes Next*. Boston, New Yark: Houghton Mifflin Com.

Chapter 11
Spacetime and Reality: Facing the Ultimate Judge

Vesselin Petkov

The views of space and time which I want to present to you arose from the domain of experimental physics, and therein lies their strength.

Minkowski [1]

Abstract Over a 100 years ago in his paper *Space and Time* Hermann Minkowski revealed the profound physical meaning of the relativity postulate—the experimental fact that physical phenomena are the same in all inertial reference frames implies that every inertial frame has its own space and time, which in turn implies that the Universe is an absolute four-dimensional world in which all moments of time have equal existence due to their belonging to the fourth (time) dimension. Since then there has been no consensus on the reality of this absolute world, which we now call Minkowski spacetime or simply spacetime. One might be tempted to interpret this situation in a sense that the question of the dimensionality of the world is so deep that we seem unable to comprehend it fully, which might be a manifestation of the first hints that there might exist some limits of our understanding of the world. I will argue that human abilities to understand the physical world are much greater than what most think by examining the issue of the reality of spacetime and showing that none of the experiments which confirmed the kinematic relativistic effects would be possible if the world were *not* four-dimensional. Therefore, facing the ultimate judge—the experimental evidence—allows us (i) to realize fully that in 1908 Minkowski had a better (than the present) understanding of the profound physical meaning of Einstein's special relativity as a theory of an absolute four-dimensional world, and (ii) to settle the issue of the reality of spacetime once and for all.

V. Petkov (✉)
Institute for Foundational Studies Hermann Minkowski, Montreal, Quebec, Canada
e-mail: vpetkov@minkowskiinstitute.org
URL: http://minkowskiinstitute.org

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_11

11.1 Introduction

Over a century after the publication of Minkowski's paper *Raum und Zeit* in 1909 [2] the issue of the reality[1] of spacetime (or the absolute world as Minkowski called it)—whether spacetime is just a mathematical four-dimensional space or it represents a real four-dimensional world—is still unresolved. What I think is even worse is that what seems to be the most shared view among both physicists and philosophers appears to be taking for granted the existence of some kind of objective becoming (which denies the reality of *all* events of spacetime) and thus denying the reality of the four-dimensional world, envisioned by Minkowski when he proposed the unification of space and time into an inseparable four-dimensional entity.

Explicitly or implicitly many physicists and philosophers regard spacetime as nothing more than a mathematical space which does not represent a real four-dimensional world. This view was explicitly defended by N. David Mermin in a relatively recent article *What's bad about this habit* in the May 2009 issue of *Physics Today* where he argued that "It is a bad habit of physicists to take their most successful abstractions to be real properties of our world" [3]. He gave the issue of the reality of spacetime as an example—"spacetime is an abstract four-dimensional mathematical continuum" [3]—and insisted that it is "a bad habit to reify the spacetime continuum" [3]. Mermin appeared to be so certain that the notion of spacetime does not represent a real four-dimensional world that he was openly lecturing *Physics Today*'s readers: "I would urge you to consider that this continuum is nothing more than an extremely effective way to represent relations between distinct events" and "The device of spacetime has been so powerful that we often reify that abstract bookkeeping structure, saying that we inhabit a world that is such a four- (or, for some of us, ten-) dimensional continuum" [3].

A 100 years after the publication of Minkowski's paper on spacetime it is difficult to explain the continued existence of views so clearly stating that spacetime does not have an ontological counterpart. By regarding spacetime as an abstract mathematical construction such views effectively ignore the revolutionary contribution of Minkowski to the advancement of our views of space and time. And indeed, if spacetime were merely a mathematical space with no counterpart in the world, we would call the unification of space and time Poincaré spacetime, not Minkowski spacetime, because it was Poincaré who first noticed in his paper *Sur la dynamique de l'électron* (before July 23, 1905 when he submitted the paper) that the Lorentz transformations had a geometric interpretation as rotations in what he seemed to have regarded as an *abstract* four-dimensional space [4, p. 168]. Poincaré appeared to have seen nothing revolutionary in the idea of a mathematical four-dimensional space since he believed that our physical theories are only convenient descriptions of the world and therefore it is really a matter of *convenience* and *our choice* which theory we would use. Here is Poincaré's own explanation [5]:

---

[1] By "reality of spacetime" I mean, following Minkowski, a real four-dimensional world.

> It quite seems, indeed, that it would be possible to translate our physics into the language of geometry of four dimensions. Attempting such a translation would be giving oneself a great deal of trouble for little profit, and I will content myself with mentioning Hertz's mechanics, in which something of the kind may be seen. Yet, it seems that the translation would always be less simple than the text, and that it would never lose the appearance of a translation, for the language of three dimensions seems the best suited to the description of our world, even though that description may be made, in case of necessity, in another idiom.

That Minkowski did not regard the absolute world implied by the relativity principle, as he argued, as an abstract four-dimensional continuum is seen even from the following general argument. Had he believed, like Poincaré, that uniting space and time into a four-dimensional space was only a convenient mathematical abstraction, he would not have written a paper whose title and content were devoted to something the main idea of which had already been published by Poincaré 2 years (and written more than 3 years) before Minkowski's talk on space and time given on September 21, 1908, and would not have begun his paper with the now famous introduction, which unequivocally announced the revolution in our views on space and time: "From now onwards space by itself and time by itself will recede completely to become mere shadows and only a type of union of the two will still stand independently on its own" [1].

However, the most convincing argument that Minkowski regarded the absolute four-dimensional world as real was provided by himself—he stressed that the strength of the new views of space and time he proposed comes from experimental physics [1]. So a century ago Minkowski was the first human (and please note—a mathematician) who faced the ultimate judge—the experimental evidence—on the issue of the ontological status of space and time. He did that by discussing one of the most important pieces of experimental evidence in the twentieth century—the negative result of the "famous interference experiment of Michelson" [1] that was carried out to detect the motion of the Earth with respect to the aether. Minkowski reviewed how Lorentz tried to explain it through a hypothesis that the arm of the interferometer used contracts in the direction of the Earth's motion due the contraction of its constituents (Lorentz used electrons to demonstrate the contraction hypothesis). Here is exactly how Minkowski realized that the experimental evidence forced upon us the new concept of space and time [1]:

> Lorentz called t', which is a combination of x and t, *local time* of the uniformly moving electron, and associated a physical construction with this concept for a better understanding of the contraction hypothesis. However, it is to the credit of A. Einstein who first realized clearly that the time of one of the electrons is as good as that of the other, i.e. that t and t' should be treated equally. With this, time was deposed from its status as a concept unambiguously determined by the phenomena. The concept of space was shaken neither by Einstein nor by Lorentz …

It was Minkowski who first made that attack on the concept of space when he realized that the postulated by Einstein equal reality of the time t (of an object believed to be at rest with respect to the aether) and the time t' (of an object in uniform motion) was the profound physical meaning of the experimental impossibility to detect the motion of the Earth in the aether. Minkowski pondered over the implications of the fact that objects in relative motion have different (equally real) times and arrived at the inescapable conclusion—if there exist more than one time, there should exist more than one space as well. Minkowski explained [1] that in the case of two inertial reference frames in relative motion along their x-axes

> one can call t' time, but then must necessarily, in connection with this, define space by the manifold of three parameters x', y, z in which the laws of physics would then have exactly the same expressions by means of x', y, z, t' as by means of x, y, z, t. Hereafter we would then have in the world no more *the* space, but an infinite number of spaces analogously as there is an infinite number of planes in three-dimensional space. Three-dimensional geometry becomes a chapter in four-dimensional physics. You see why I said at the beginning that space and time will recede completely to become mere shadows and only a world in itself will exist.

There is some irony in Minkowski's discovery that the world is four-dimensional. The mathematician Minkowski wanted to *understand* why physical phenomena are the same in all inertial reference frames, whereas the physicist Einstein merely postulated that fact and called it the relativity postulate (or the relativity principle) without *explaining* it.

That is why, Minkowski first realized the important hidden message in the *experimental fact* that physical phenomena are the same in all inertial reference frames—the experimental fact implies that the Universe is an absolute four-dimensional world in which space and time are inseparably amalgamated; *only in such a world one can talk about many spaces and many times*. And indeed, physical phenomena are the same in all inertial reference frames because every inertial observer describes the phenomena in his reference frame (i.e. in his own space and time) in which he is at rest. For example, the Earth is at rest with respect to its space and therefore all experiments confirm this state of rest.

Minkowski noted that "I think the word *relativity postulate* used for the requirement of invariance under the group G_c is very feeble. Since the meaning of the postulate is that through the phenomena only the four-dimensional world in space and time is given, but the projection in space and in time can still be made with a certain freedom, I want to give this affirmation rather the name *the postulate of the absolute world*" [1].

Since Minkowski arrived at the new view of space and time—that we live in an absolute four-dimensional world—*when he faced the ultimate judge* (a single experiment at that time[2]), it is difficult to understand how what now appears to be a widespread view could still regard spacetime as nothing more than a mathematical continuum given the fact that we have at out disposal more *experiments* that confirmed

---

[2] This is the Michaelson-Morley experiment which Minkowski mentioned explicitly and which used electromagnetic signals (light) to try to detect the Earth's motion with respect to the aether. That experiment showed that not only mechanical experiments (discussed and used by Galileo) fail to discover motion with constant velocity, but experiments involving electromagnetic phenomena fail too. The expression "physical phenomena are the same in all inertial frames" simply means that motion with constant velocity cannot be discovered; otherwise, if physical phenomena were different in some inertial frames, that would mean that those frames were in a state of absolute motion.

the relativistic effects. A century after Minkowski's insight I think it will be fair to take our turn and also face the ultimate judge on the reality of spacetime. We owe this to Minkowski.

The main purpose of this paper is to demonstrate that the theory of relativity would be impossible if the world were three-dimensional. Section 11.2 shows how relativity of simultaneity is possible only in a four-dimensional world. Section 11.3 revisits Minkowski's explanation of length contraction and discusses a more visualized version of his explanation through a thought experiment involving a meter stick, which demonstrates that no length contraction would be possible if the meter stick existed as a three-dimensional body.

11.2 Relativity of Simultaneity

Although no specially designed experiments have been carried out to test relativity of simultaneity, this major consequence of special relativity can be regarded as an experimental fact for two reasons:

- The experimental fact captured in the relativity postulate—physical phenomena are the same in all inertial reference frames—implies, as Minkowski demonstrated, that physical objects in relative motion have their own spaces and times, which can be explained if it is assumed that what exists is a four-dimensional world. But as a space constitutes a class of *simultaneous* events, it follows that relativity of simultaneity is a consequence of the existence of many spaces and therefore ultimately a consequence of the experimental fact of the invariance of physical phenomena, which is reflected in the relativity postulate.
- As we will see in Sect. 11.3 length contraction, which have been experimentally confirmed, is a specific manifestation of relativity of simultaneity.

In order to understand fully why, unlike Poincaré, Minkowski appeared to have realized that special relativity would be *impossible* in a three-dimensional world, is to ask explicitly the questions that Minkowski appeared to have implicitly considered—"What is the world?" and "How does the new concept of space and time affect our view of what exists?"

Concerning the first question, it seems we do not have much choice—most of those who regard spacetime as nothing more than a mathematical device appear (explicitly or implicitly) to hold the presentist view according to which the Universe is a single three-dimensional world defined as everything that exists *simultaneously* at the constantly changing moment 'now'. Minkowski might have started with this view too. Then, the answer to the second question follows naturally—the fact that objects in relative motion have different spaces (i.e. different classes of simultaneous events) implies that they have different three-dimensional worlds. But this is only possible if these worlds are different three-dimensional "cross sections" of an absolute four-dimensional world; *it is impossible to have many spaces and times and* *relativity of simultaneity in a three-dimensional world*. Indeed, if the world were three-dimensional (i.e., if spacetime were not representing a real four-dimensional world and were just an abstract mathematical space), there would exist one space and one absolute class of simultaneous events (since a three-dimensional space constitutes a single class of absolutely simultaneous events) in contradiction with relativity.

It should be stressed that the above argument is irrefutable only if the existence of physical objects and the world itself is regarded as *absolute*—only then the relativistic fact that observers in relative motion have different spaces and therefore different classes of simultaneous events implies an absolute four-dimensional world. But if existence is relativized, it appears to follow that relativity is possible in a three dimensional world—every observer would acknowledge the existence of only his own space (and three-dimensional world) and would deny the existence of the spaces (and the three-dimensional worlds) of the other observers in relative motion. So it appears that relativity of simultaneity (and therefore length contraction as well) can be explained either (i) in terms of absolute existence (in this case relativity of simultaneity implies a four-dimensional world—Minkowski's explanation) or (ii) in terms of relative existence (in this case the three-dimensionality of the world would be preserved—what exists for each of the inertial observers in relative motion would be his relativized three-dimensional world).

It might be tempting to take the relativization of existence seriously since it preserves the pre-relativistic (presentist) view that what exists is a three-dimensional world; moreover, originally Einstein formulated special relativity in a three-dimensional language. But a careful analysis shows that that option contradicts the *experiments* that confirmed the twin paradox [6, Chap. 5]. In fact, it can be immediately shown that the idea of relativization of existence is not a serious alternative to the deep intuition that the very essence of existence makes it impossible to regard it as relative. If we assume that relativization of existence is the correct interpretation of relativity of simultaneity (which means that for every observer only his three-dimensional world would exist), we arrive at total nonsense when we ask what exists for an observer in general relativity.

Let us consider a single inertial observer and assume that what exists relative to the observer is his three-dimensional world, i.e. his present. In flat spacetime an inertial observer is represented by a straight worldtube and the presents corresponding to different moments of the proper time of the observer are parallel to one another and do not intersect. In curved spacetime, however, the worldtube of any observer is curved, which means that the presents at different moments of the observer's proper time *intersect* one another. As a result, some events which were past at a given moment would be future at a *later* moment. This nonsensical conclusion follows from the assumption that reality for an observer in general relativity is a three-dimensional world. Therefore considering even a single observer in general relativity rules out presentism and also its relativized version since the view that existence should be relativized regards the world of an observer as three-dimensional.[3]

After ruling out relativization of existence, the only way to interpret relativity of simultaneity (i.e. the fact that observers in relative motion have different spaces and times) is Minkowski's interpretation—that the classes of simultaneous events (i.e. the spaces) of observers in relative motion are "cross sections" of an absolute four-dimensional world. Let me stress that there is no alternative to this interpretation—if one assumed that spacetime were a mathematical device and that the Universe were a three-dimensional world, there would exist *one* (i.e. *absolute*) space and therefore *one* (i.e. *absolute*) class of simultaneous events in contradiction with the experimental evidence that physical laws are the same in all inertial reference frames.

In the sixties [7, 8] pointed out that relativity of simultaneity does imply a four-dimensional world. In fact, their relativity of simultaneity argument, as we have seen, follows from the more general analysis (the existence of *many* spaces) that led Minkowski to the idea of spacetime. Like Minkowski's analysis, their argument has not been fully appreciated. Many physicists and philosophers have been refusing to accept the reality of the four-dimensional world of relativity, but *have not explained* how relativity of simultaneity would be possible if the world were three-dimensional. An example is Stein's criticism of the Rietdijk-Putnam argument [9, 10]. He correctly pointed out that one could not talk about distant present events in relativity but seemed to believe that he refuted the Rietdijk-Putnam argument (which, in essence, is Minkowski's argument) and some philosophers agreed with him. What he certainly refuted, however, is the presentist view according to which what exists is a single class of distant present events. Unfortunately, Stein criticized Rietdijk and Putnam for arguing that relativity of simultaneity implies a four-dimensional world, but explained neither how relativity of simultaneity would be possible if the world were *not* four-dimensional nor what the dimensionality of the world according to relativity would be.

So, when one explicitly asks the two questions above—what our view of the world is and how that view is affected by relativity—it follows that relativity of simultaneity is a manifestation of the four-dimensionality of the world (i.e. of the reality of spacetime) as Minkowski pointed out. Due to the *equal* existence of all events of spacetime, observers in relative motion can regard different "cross sections" of it as their classes of simultaneous events (i.e. as their spaces). If there existed just one class of privileged (due to their existence) events all observers in relative motion would share that class and no relativity of simultaneity would be possible. When this is taken into account it becomes evident that the lack of an objectively privileged class of simultaneous events implies not only relativity of simultaneity, but also conventionality of simultaneity. This should be specifically emphasized since, I think, any claim that simultaneity is relative but not conventional would amount to a contradiction in terms—there exists no objectively privileged class of simultaneous

---

[3] This argument can be also explained in the case of accelerated observers in special relativity [6, pp. 150–152] since an accelerated observer is represented by a curved worldtube in flat spacetime and therefore the presents corresponding to different events of the observer's worldtube intersect.

events (due to relativity of simultaneity), but there exists an objectively privileged class of simultaneous events (due to the non-conventionality of simultaneity).

The relativity of simultaneity argument can be also used to rule out a theory proposed by Broad in 1923, which "accepts the reality of the present and the past, but holds that the future is simply nothing at all" [11]. Broad's growing (or evolving) block universe model of the world has been recently revived by several physicists [12–14] as what appears to be the last remaining alternative to the Minkowski absolute four-dimensional world. The recent versions of the growing block universe claim (excluding [13]) that they do not allow any form of a preferred structure. But if it is explicitly assumed that the existence of physical bodies is *absolute*, this claim cannot be supported. The hypersurface on the edge of the growing universe, on which the birthing (or coming into being) of events happens, constitutes an objectively privileged class of events (due to the absoluteness of existence), which contradicts relativity. Also, the growing block universe model leads to the same nonsense when used to explain what exists for an accelerated observer in flat spacetime or for an observer in curved spacetime.

11.3 Length Contraction

Length contraction was experimentally tested, along with time dilation, by the muon experiment in the muon reference frame (see for instance [15]). With this in mind, let us now ask whether this relativistic effect would be possible if a body, subjected to length contraction, were what we perceive—a three-dimensional object.

The essence of Minkowski's explanation of the deep physical meaning of length contraction of two bodies is that it is a manifestation of the reality of the bodies' worldtubes (Minkowski called them strips). This can be best understood from Fig. 1 of his paper (the right-hand part of which is reproduced in Fig. 11.1 here)—length contraction would be *impossible* if the worldtubes of the two bodies, represented by the vertical and the inclined strips in Fig. 11.1, did not exist and were nothing more than "abstract geometric constructions" [3]. To see this even more clearly consider only the body represented by the vertical worldtube. The three-dimensional

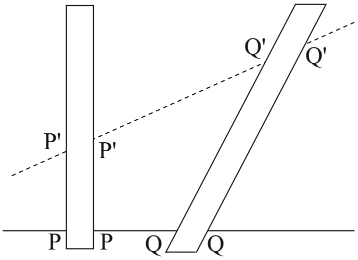


Fig. 11.1 In Minkowski's paper *Space and Time* Fig. 1, part of which is reproduced here, represents two bodies or Lorentzian electrons by their worldtubes or as Minkowski called them (world) strips

cross-section PP, resulting from the intersection of the body's worldtube and the space of an observer at rest with respect to the body, is the body's proper length. The three-dimensional cross-section $P'P'$, resulting from the intersection of the body's worldtube and the space of an observer moving with respect to the body, is the relativistically contracted length of the body measured by that observer. Minkowski stressed that "this is the meaning of the Lorentzian hypothesis of the contraction of electrons in motion" [1] and "that the Lorentzian hypothesis is completely equivalent to the new concept of space and time, which makes it much easier to understand" [1]. The worldtube of the body must be real in order that length contraction be possible because, while measuring the *same* body, the two observers in relative motion measure *two* three-dimensional bodies represented by the "cross-sections" PP and $P'P'$ in Fig. 11.1.

This is not so surprising when one takes into account relativity of simultaneity and the fact that a spatially extended three-dimensional object is defined in terms of *simultaneity*—all parts of a body taken *simultaneously* at a given moment (so length contraction is indeed a specific manifestation of relativity of simultaneity). If the worldtube of the body were an abstract geometric construction and what existed were a single three-dimensional body (a single class of simultaneous events) represented by the proper cross-section PP, both observers would measure the same three-dimensional body, i.e. the same class of simultaneous events, which means that simultaneity would be absolute.

That length contraction of a body would be impossible if the body existed as a three-dimensional object (not a worldtube) can be perhaps better demonstrated by the following thought experiment, which is a more visualized version of Minkowski explanation. An ordinary meter stick (Fig. 11.2) is at rest with respect to an observer A. What is shown in Fig. 11.2 is what we perceive and take for granted that it is what really exists. According to Minkowski, however, the meter stick exists equally at all moments of its history and what is ultimately real is the worldtube of the meter stick as shown in Fig. 11.3.

Assume that another meter stick at rest in another observer's (observer B's) reference frame moves relative to the first one at a distance 1 mm above it. Let us assume that at the event M the middle point of B's meter stick is instantaneously above the middle point of A's meter stick. Imagine also that lights are installed inside A's meter stick, which can simultaneously change their color at every instant in A's frame. At the event of the meeting M all lights are simultaneously white in A's frame. At all previous moments all lights were bright grey. At all moments after the meeting all lights will be dark grey. When A and B meet at event M this event is present for both of them. At that moment the present meter stick for A is white (that is, all parts of A's meter stick, which exist simultaneously for A, are white). All moments before M

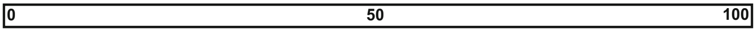


Fig. 11.2 An ordinary meter stick

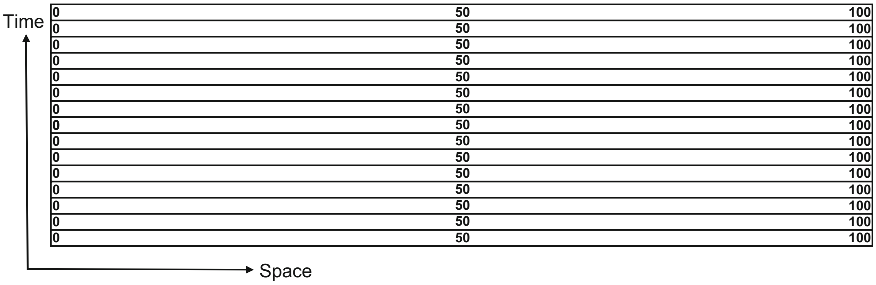


Fig. 11.3 The worldtube of the meter stick

Fig. 11.4 Relativistically contracted meter stick measured by observer *B*

when all lights of the meter were bright grey are past for *A*, whereas all moments when the meter stick will be dark grey are in *A*'s future.

Imagine that, instead of lights, *B*'s meter stick contains cameras at every point along its length. At the event of the meeting *M* all cameras take snapshots of the parts of *A*'s meter stick which the cameras face. At event *M* all snapshots are taken *simultaneously* in *B*'s reference frame.

Even without looking at the pictures taken by the cameras, it is clear that not all pictures will show a white part of *A*'s meter stick, because what is simultaneous for *A* is not simultaneous for *B*. When the picture of *A*'s meter stick is assembled from the pictures of all cameras it would show two things as depicted in Fig. 11.4—(i) *A*'s meter stick photographed by *B* is shorter, and (ii) only the middle part of the picture of *A*'s meter stick is white; half is bright grey and the other half is dark grey. So what is past (bright grey), present (white), and future (dark grey) for *A*, exists *simultaneously* as present for *B*. But this is only possible if the meter stick is the worldtube as shown in Fig. 11.5. The instantaneous space of *B* corresponding to the event *M* intersects the worldtube of the meter stick at an angle and the resulting three-color "cross section" is what is measured by *B*—a different three-dimensional meter stick, which is shorter[4] than the meter stick measured by *A*.

It should be stressed as strongly as possible that no length contraction would be possible if the meter stick's worldtube did not exist as a four-dimensional object. If the meter stick were a three-dimensional object, both observers would measure the *same* three-dimensional meter stick (the same set of *simultaneously* existing parts of the meter stick), which would mean that the observers would share the same (absolute) class of simultaneous events in a clear contradiction with relativity [16].

---

[4] In Fig. 11.5 the inclined "cross section," which represents the different three-dimensional meter stick measured by *B*, appears longer, not shorter, because a fact in the pseudo-Euclidean geometry of spacetime is represented on the Euclidean surface of the paper.

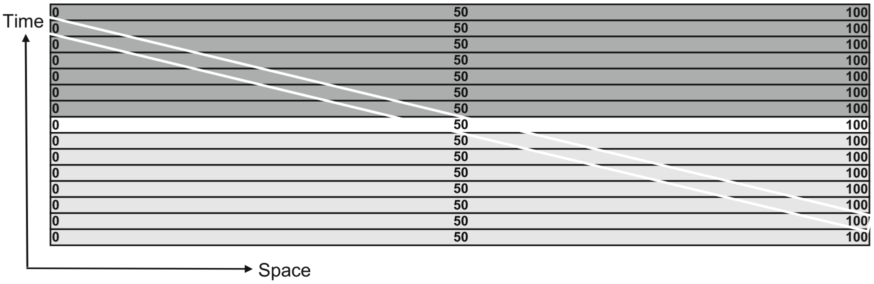


Fig. 11.5 The worldtube of the meter stick with different colors

11.4 Conclusion

Minkowski arrived at the new concept of space and time by decoding the hidden message in a single experiment—the impossibility to detect the motion of the Earth with respect to the aether implied that observers in relative motion have different spaces (not just different times as Einstein postulated), which is only possible in an absolute four-dimensional world. It was shown that a rigorous analysis, following Minkowski's line of thought, demonstrated two things:

(i) Minkowski had a deep understanding of the physical meaning of the experimental fact that physical phenomena are the same in all inertial reference frames—he did not postulate (as Einstein did) that that experimental fact should be simply accepted as a law of nature (the relativity postulate), but *explained* it: all inertial observers in relative motion have different spaces and times (only possible in a four-dimensional world) which explains why physical phenomena are the same for all inertial observers—each observer represents the phenomena in terms of his own space and time (for instance, the speed of light is the same for all inertial observer since for each observer light propagates in his space with respect to which he is at rest.)

(ii) The kinematic relativistic effects (here we discussed only relativity of simultaneity and length contraction) would be impossible if the world were *not* four-dimensional. This, in turn, provides *full explanation* of the physical meaning of these effects.

It was shown in the paper how the ultimate judge—the relativistic experimental evidence—settled the issue of the reality of spacetime once and for all. As the ruling of the ultimate judge cannot be appealed (that is, it is irrefutable), I believe it is clear that refusing to face the implications of Minkowski's view because of the huge challenges they pose, and trying to squeeze Nature into our pre-set and deceivingly comfortable views of the world should not be an option for anyone in the 21st century.

References

1. Minkowski, H. (2012). Space and time. In V. Petkov (Ed.), *Space and time: Minkowski's papers on relativity*. Montreal: Minkowski Institute Press.
2. Minkowski, H. (1909). Raum und zeit. *Physikalische Zeitschrift*, *10*, 104–111; *Jahresbericht der Deutschen Mathematiker-Vereinigung*, *18*, 75–88.
3. Mermin, N. D. (2009). What's bad about this habit? *Physics Today*, 8.
4. Poincaré, H. (1906). Sur la dynamique de l'électron. *Rendiconti del Circolo matematico Rendiconti del Circolo di Palermo*, *21*, 129–176.
5. Poincaré, H. (2001). Science and method. In *The value of science: Essential writings of henri poincaré* (p. 438). New York: Modern Library.
6. Petkov, V. (2009). *Relativity and the nature of spacetime* (2nd ed.). Heidelberg: Springer.
7. Rietdijk, C. W. (1966). A rigorous proof of determinism derived from the special theory of relativity. *Philosophy of Science*, *33*, 341–344.
8. Putnam, H. (1967). Time and physical geometry. *Journal of Philosophy*, *64*, 240–247.
9. Stein, H. (1968). On Einstein-Minkowski space-time. *Journal of Philosophy*, *65*, 5–23.
10. Stein, H. (1991). On relativity theory and the openness of the future. *Philosophy of Science*, *58*, 147–167.
11. Broad, C. D. (1959). *Scientific Thought*. Littlefield: Adams and Co.
12. Christian, J. (2007). Absolute being versus relative becoming. In V. Petkov (Ed.), *Relativity and the dimensionality of the world* (pp. 163–195). Heidelberg: Springer. arXiv:gr-qc/0610049.
13. Ellis, G. F. R. (2007). Physics in the real universe: Time and space-time. In V. Petkov (Ed.), *Relativity and the dimensionality of the world* (pp. 49–79). Heidelberg: Springer. arXiv:gr-qc/0605049.
14. Sorkin, R. D. (2007). Relativity theory does not imply that the future already exists. In V. Petkov (Ed.), *Relativity and the dimensionality of the world* (pp. 153–191). Heidelberg: Springer. arXiv:gr-qc/0703098.
15. Ellis, G. F. R., & Williams, R. M. (1988). *Flat and curved space times*. Oxford: Oxford University Press.
16. Petkov, V. (2007). On the reality of Minkowski space. *Foundations of Physics*, *37*, 1499–1502.

Chapter 12
The Future's Not Ours to See

Anthony Sudbery

> *Que sera sera*
> *Whatever will be will be*
> *The future's not ours to see*
> *Que sera sera.*

So sang Doris Day in 1956, expressing a near-universal belief of humankind. You can't know the future. In this chapter I will trace the different forms of this belief, both pre-scientific and scientific, and discuss some differing kinds of scientific justification for it in physics, culminating in the form of the statement provided by the best physical theory we have found to date, namely quantum mechanics. I will argue that quantum mechanics casts doubt on the second line of the song, which suggests that even if we can't know it, there *is* a definite future (and also that we can't do anything to change it—something I will not discuss). This denial is also an ancient belief: the future is *open*. If this line of the song expresses fatalism, the denial of it might be related to the existence of free will, but, again, I will not discuss this.

If it is not quite a universal belief of humankind that we *can't* know the future, the universal *experience* of humankind is that we *don't* know the future. We don't know it, that is, in the immediate way that we know parts of the present and the past. We see some things happening in the present, we remember some things in the past, but we don't see or remember the future. But perception can be deceptive, and memory can be unreliable; this kind of direct knowledge is not certain. And there are kinds of indirect knowledge of the future which can be as certain as anything we know by direct perception or memory. I reckon I know that the sun will rise tomorrow; if I throw a stone hard at my kitchen window, I know that it will break the window. On

---

A. Sudbery (✉)
Department of Mathematics, University of York,
Heslington, York YO10 5DD, UK
e-mail: as2@york.ac.uk

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_12

the other hand, I did not know on Christmas Eve last year that my home town of York was going to be be hit by heavy rain on Christmas Day and nearly isolated by floods on Boxing Day.

In the ancient world, and, I think, to our childhood selves, it is events like the York floods that make us believe that we cannot know the future. I may know some things about the future, but I cannot know everything; I am sure that some things will happen tomorrow that I have no inkling of, and that I could not possibly have known about, today. In the past such events might have been attributed to the unknowable will of the gods. York was flooded because the rain god was in a bad mood, or felt like playing with us. My insurance policy refers to such catastrophes as "acts of God"; when we feel that there is no knowing who will win an election, we say that the result is "in the lap of the gods".

Aristotle formulated the openness of the future in the language of logic. Living in Athens at a time when invasion from the sea was always a possibility, he considered the sentence "There will be a sea-battle tomorrow". One of the classical laws of logic is the "law of the excluded middle" which states that every sentence is either true or false: either the sentence is true or its negation is true. But Aristotle argued that neither "There will be a sea-battle tomorrow" nor "There will not be a sea-battle tomorrow" is definitely true, for both possibilities lead to fatalism; if the first statement was true, for example, there would be nothing anybody could do to avert the sea-battle. Therefore these statements belong to a third logical category, neither true nor false. In modern times this conclusion has been realised in the development of many-valued logic [1].

But some statements in the future tense do seem to be true; I have given the examples "The sun will rise tomorrow" and, after I have thrown the stone, "That window is going to break". Let's look at these more closely. In fact, none of these future statements are 100 % certain. The sun might not rise tomorrow; there might be a galactic star-trawler heading for the solar system, ready to scoop up the sun tonight and make off with it at nearly the speed of light. When I throw the stone at the window, my big brother, who is a responsible member of the family and a superb cricketer, might be coming round the corner of the house; he might see me throw the stone and catch it so as to save the window.

We did not know that the sun would fail to make its scheduled appearance tomorrow morning; I did not know that my naughtiness would be foiled. But this lack of knowledge is not a specific consequence of the fact that we are talking about the future. If the Spaceguard programme had had a wider remit we might have seen the star-trawler coming, and then we would have known that we had seen our last sunrise; if I had known my brother's whereabouts I could have predicted his window-saving catch. In both these scenarios the lack of knowledge of the future reduces to lack of knowledge about the present.

The success of modern science gave rise to the idea that this is always true: not knowing the future can always be traced back to not knowing something about the present. As more and more phenomena came to be explained in terms of the laws of physics, so that more and more events could be explained as being caused by previous events, so confidence grew that every future event could be predicted with certainty, given enough knowledge of the present. The most famous statement of this confidence was made by Laplace in 1814:

> We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes. ([2], p.4)

This idea goes back to Newton, who had a dream:

> I wish we could derive the rest of the phenomena of Nature by the same kind of reasoning from mechanical principles, for I am induced by many reasons to suspect that they may all depend upon certain forces by which the particles of bodies, by some causes hitherto unknown, are either mutually impelled towards one another, and cohere in regular figures, or are repelled and recede from one another. [3]

In this view, everything in the world is made up of point particles, and their behaviour is explained by the action of forces which make the particles move according to Newton's equations of motion. These completely determine the future motion of the particles if their positions and velocities are given at any one instant; the theory is *deterministic*. So if we fail to know the future, that is purely because we do not know enough about the present.

For a couple of centuries Newton's dream seemed to be coming true. More and more of the physical world came under the domain of physics, as matter was analysed into molecules and atoms, and the behaviour of matter, whether chemical, biological, geological or astronomical, was explained in terms of Newtonian forces. The particles of matter that Newton dreamed of had to be supplemented by electromagnetic fields to give the full picture of what the world was made of, but the basic idea remained that they all followed deterministic laws. Capricious events like storms and floods, formerly seen as unpredictable and attributed to the whims of the gods, became susceptible to weather forecasts; and if some such events, like earthquakes, remain unpredictable, we feel sure that advancing knowledge will make them also subject to being forecast.

This scientific programme has been so successful that we have forgotten there was ever any other way to think about the future. One author writes

> In ordinary life, and in science up until the advent of quantum mechanics, *all* the uncertainty that we encounter is presumed to be… uncertainty arising from ignorance. [4]

We have completely forgotten what an uncertain world was inhabited by the human race before the seventeenth century, and we take Newton's dream as a natural view of waking reality.

Well, it was a nice dream. But it didn't work out that way. In the early years of the twentieth century Ernest Rutherford, investigating the recently discovered phenomenon of radioactivity, realised that it showed random events happening at a fundamental level of matter, in the atom and its nucleus. This did not necessarily mean that

Newton's dream had to be abandoned—the nucleus is not the most fundamental level of matter, but is a complicated object made up of protons and neutrons, and maybe, if we knew exactly how these particles were situated and how they were moving, we would be able to predict when the radioactive decay of the nucleus would happen. But other, stranger, discoveries at around the same time led to the radical departure from Newtonian physics represented by quantum mechanics, which strongly reinforced the view that events at the smallest scale are indeed random, and there is no possibility of precisely knowing the future.

The discoveries that had to be confronted by the new physics of the 1920s were twofold. On the one hand, Planck's explanation of the distribution of wavelengths in the radiation emitted by hot matter, and Einstein's explanation of the photoelectric effect, showed that energy comes in discrete packets, instead of varying continuously as it must do in Newton's mechanics and Maxwell's electromagnetic theory. On the other hand, experiments on electrons by Thomson, Davisson and Germer showed that electrons, which had been firmly established to be particles, also sometimes behaved like waves. These puzzling facts found a systematic, coherent, unified mathematical description in the theory of quantum mechanics which emerged from the work of theorists after 1926. This theory was itself so puzzling that it is not clear that it should be described as an "explanation" of the puzzling facts it incorporated; but an essential feature of it, which seems inescapable, is that when applied to give predictions of physical effects, it yielded probabilities rather than precise numbers.

This is still not universally accepted. Some people believe that there are finer details to be discovered in the make-up of matter, which, if we knew them, would once again make it possible to predict their future behaviour precisely. This is indeed logically possible, but there would necessarily be aspects of such a theory which lead most physicists to think it highly unlikely.

The format of quantum mechanics is quite different from previous physical theories like Newtonian mechanics or electromagnetism (or both combined). These theories work with a mathematical description of the state of the world, or any part of the world; they have an equation of motion which takes such a mathematical description and tells you what it will change into after a given time. Quantum mechanics also works with a mathematical object which describes a state of the world; it is called a state vector (though it is not a vector in three dimensions like velocity), and is often denoted by the Greek letter ψ or some similar symbol. But this is a different kind of mathematical description from that in mechanics or electromagnetism, which consists of a set of numbers which measure physical quantities like the velocity of a specified particle, or the electric field at a specified point of space; the quantum state vector is a more abstract object whose relation to physical quantities is indirect. From the state vector you can obtain the values of physical quantities, but only some physical quantities—you can choose which quantities you would like to know, but you are not allowed to choose all of them. Moreover, once you have chosen which ones you would like to know, the state vector will not give you a definite answer; it will only give you probabilities for the different possible answers. This is where quantum mechanics departs from determinism. Strangely enough, in its treatment of change quantum mechanics looks like the old deterministic theories. Like them, it has an equation of motion which will tell you what a given state of the world will become after a given time; but because you can only get probabilities from this state vector, it cannot tell you what you will see after this time.

State vectors, in general, are puzzling things, and it is not at all clear how they describe physical objects. Some of them, however, do correspond to descriptions that we can understand. Among the state vectors of a cat, for example, is one describing a cat sitting and contentedly purring; there is another one describing it lying dead, having been poisoned in a diabolical contraption devised by the physicist Erwin Schrödinger. But there are others, obtained mathematically by "superposing" these two state vectors; such a superposed state vector could be made up of a part describing the cat as live and a part describing it as dead. These are not two cats; the point of Schrödinger's story was that one and the same cat seems to be described as both alive and dead, and we do not understand how such states could describe anything that could arise in the real world. How can we believe this theory, generations of physicists have asked, when we never see such alive-and-dead cats?

There is an answer to this puzzle. If I were to open the box in which Schrödinger has prepared this poor cat, then the ordinary laws of everyday physics would ensure that if the cat was alive, I would have the image of a living cat on my retina and in my visual cortex, and the system consisting of me and the cat would end up in a state in which the cat is alive and I see a living cat. If the cat was dead, I would have the image of a dead cat, and the system consisting of me and the cat would end up in a state in which the cat is dead and I see a dead cat. It now follows, according to the laws of quantum mechanics, that if the cat is in a superposition of being alive and being dead, then the system consisting of me and the cat ends up in a superposition of the two final states described above. In symbols,

$$|\,\ddot{-}\rangle_{\mathrm{me}}\big(|\mathrm{alive}\rangle_{\mathrm{cat}}+|\mathrm{dead}\rangle_{\mathrm{cat}}\big)\longrightarrow|\,\ddot{\smile}\rangle_{\mathrm{me}}|\mathrm{alive}\rangle_{\mathrm{cat}}+|\,\ddot{\frown}\rangle_{\mathrm{me}}|\mathrm{dead}\rangle_{\mathrm{cat}}. \qquad (12.1)$$

Here a symbol like $|\Psi\rangle_{\mathrm{cat}}$ denotes a state vector of the cat, with Ψ varying to describe different states (e.g. $\Psi =$ "alive"), and similarly $|\Phi\rangle_{\mathrm{me}}$ denotes a state vector of my body and brain, while $|\Psi\rangle_{\mathrm{cat}}|\Phi\rangle_{\mathrm{me}}$ denotes a state vector of the joint system of the cat and me. The plus sign between two state vectors denotes the operation of superposition (mathematically, it is very similar to the process of adding two vectors in space to give a third vector in between the first two).

Look hard at the Eq. (12.1). Nowhere in it is there a state of my brain seeing a peculiar alive-and-dead state of a cat; there are only the familiar states of seeing a live cat and seeing a dead cat. This is the answer to the question at the end of the paragraph before the last one; it follows from quantum mechanics itself that although cats have states in which they seem to be both alive and dead, we will never see a cat in such a state.

But now the combined system of me and the cat is in one of the strange superposition states introduced by quantum mechanics. It is called an *entangled* state of me and the cat. How are we to understand it? We can understand the states represented by $|\,\ddot{\smile}\rangle_{\mathrm{me}}|\mathrm{alive}\rangle_{\mathrm{cat}}$ and $|\,\ddot{\frown}\rangle_{\mathrm{me}}|\mathrm{dead}\rangle_{\mathrm{cat}}$ individually—the cat is alive in one of them, and dead in the other, and I have the corresponding visual experience—but what does this mathematical sum, this superposition, mean? Maybe the mathematical sign + just means “or”; that would make sense. But unfortunately this meaning, if applied to the states of an electron, is not compatible with the facts of interference observed in the experiments that show the electron behaving like a wave ([5], pp. 22–24, 207). Some people think that this + should be understood as “and”: when the cat and I are in the state (12.1), there is a world in which the cat has died and I see a dead cat, and another world in which the cat is still alive and I see a living cat. Others do not find this a helpful picture ([5, 6], p. 221). Let us just take it as a true description of the cat and me, whose meaning is problematic.

Now let us broaden our horizon and consider the whole universe, which contains each one of us considered as a sentient, observing physical system. According to quantum mechanics this has a description by a state vector like (12.1), with me replaced by any sentient observer and the cat replaced by the rest of the universe. A sentient system, like you or me, has a large number of possible experiences, each of which occurs physically in certain states of the sentient system. These are described by state vectors of the understandable kind like $|\ddot{\smile}\rangle_{\mathrm{me}}$ and $|\ddot{\frown}\rangle_{\mathrm{me}}$. The whole universe can then be described by a state vector generalising (12.1), in which the states of the cat are replaced by states of the rest of the universe which go with the possible experiences of the observer. For those who like equations, the state vector of the universe is of the form

$$|\Psi(t)\rangle = \sum_n |\eta_n\rangle|\Phi_n(t)\rangle \tag{12.2}$$

where $|\Psi(t)\rangle$ is the state vector of the whole universe at time t, $|\eta_n\rangle$ is a state vector of the observer in which they are having the experience η_n, and $|\Phi_n(t)\rangle$ is the corresponding state of the rest of the universe.

Saying that this is the truth about the universe seems to conflict with my knowledge of what I see. Let us suppose that the cat survived when I did Schrodinger’s experiment. Then I know that my state is $|\ddot{\smile}\rangle_{\mathrm{me}}$ and therefore the cat’s state is $|\mathrm{alive}\rangle_{\mathrm{cat}}$. The other part of (12.1) is not part of the truth; it describes something that might have happened but didn’t. In the general case of the whole universe, I know that I have just one of the experiences η_n and therefore that the state of me and the rest of the universe is just one of the terms $|\eta_n\rangle|\Phi_n(t)\rangle$ and not the whole of (12.2). But this contradicts what was asserted in the previous paragraph. Which of these is the truth?

This contradiction is of the same type as many familiar contradictions between objective and subjective statements. It can be resolved in the way put forward by Nagel [7, 8]: we must recognise that there are two positions from which we can make statements of fact or value, and statements made in these two contexts are not commensurable. In the *external* context (the God’s-eye view, or the “view from nowhere”) we step outside our own particular situation and talk about the whole universe. In the *internal* context (the view from *now here*), we make statements as physical objects inside the universe. Thus in the external view, the state vector $|\Psi(t)\rangle$ is the whole truth about the universe; the components describing my different possible experiences, and the corresponding states of the rest of the universe, are (unequal) parts of this truth. But in the internal view, from the perspective of some particular experience $|\eta_n\rangle$ which I know I am having, that experience, together with the corresponding state of the rest of the universe, is the actual truth. I may know what the other components are, because I can calculate $|\Psi(t)\rangle$ from the Schrödinger equation; but these other components, for me, represent things that *might* have happened but *didn't*.

We can now look at what quantum mechanics tells us about the future. As we should now expect, there are two answers, one for each of the two perspectives. From the external perspective, the universe at any one time is described by a universal state vector, and state vectors at different times are related by the Schrödinger equation. Given the state vector at the present time, the Schrödinger equation delivers a unique state vector at any future time: the theory is completely deterministic, in complete accord with Pascal's world-view (in a quantum version).

From the internal perspective, however, things are completely different. We now have to specify a particular observer (who has been me in the above discussion, but it could have been you or anyone else, or indeed the whole human race taken together), with respect to which we can carve up the universal state vector as in (12.2); and we have to specify a particular experience state of that observer, say $|\eta_0\rangle$. From the perspective of that experience, it is by definition true that the observer has the definite experience η_0, and it follows from (12.2) that the universe is in a product state $|\eta_0\rangle|\Phi_0(t)\rangle$. The Schrödinger equation applies, as before. But with this initial state the Schrödinger equation will yield a state of the universe at a future time which, in general, is not a product state; in particular, in situations like that of Schrödinger's cat in which quantum effects are amplified to a macroscopic level at which they can be seen by our observer, the initial product state will develop into a superposition like the right-hand side of (12.1). So at the future time there will be a universal state vector of the form (12.2), in which most terms will be zero but more than one experience state of the observer occurs with non-zero coefficient. (A "coefficient" here is a state vector of the rest of the universe, the magnitude being significant.) What is the present observer to make of this? The experience states in the non-zero components must describe possible experiences of the observer at the future time t. No one of them is singled out as being what the observer actually will experience at that time. Some of them, though, loom larger than others, according to the magnitudes of their coefficients in the universal state vector.

I found this rather startling when I first encountered it. I was used to thinking that there is something awaiting me in the future, even if I cannot know what it is, and even if there is no law of nature which determines what it is. Whatever will be will be, indeed. But Aristotle already saw that this is wrong. A statement about the future, in general, cannot be actually true, even when we are careful to distinguish being true from our knowing that it is true. But we can say more than that.

Aristotle pointed out that although no one statement about the future is actually true, some of them are more *likely* than others. Similarly, the universal state vector at the time t contains more information, for me, than my possible experiences at time t. These experience states, occurring as components of the universal state vector, contribute to it to different extents, measured by the magnitudes of their coefficients.

Such magnitudes are usually used in quantum mechanics to calculate *probabilities*. So we can understand the future universal state as giving information, not only about what experiences are possible for me at that future time, but also about how probable each experience is.

The nature of probability is a long-standing philosophical problem [9], to which scientists also need an answer. Many scientists take the view that the probability of an event only makes sense when there are many repetitions of the circumstances in which the event might occur, and we count up the number of times that it does occur; they hold that the probability of a single, unrepeated event does not make sense. But what we have just outlined does seem to be a calculation of a single event at a time which will only come once. In everyday life we often talk about the probability that something will happen on just one occasion—that it will rain tomorrow, or that a particular horse will win a race, or that there will be a sea-battle. A standard view of such single-event probability is that it refers to the strength of the belief of the person who is asserting the probability, and can be measured by the betting odds they are prepared to offer on the event happening. But the probability described in the previous paragraph is an objective fact about the universe. It has nothing to do with the beliefs of an individual, not even the individual whose experiences are in question; that individual is being told a fact about their future experiences, whether they believe it or not. What does this probability mean?

The probability that I will have a particular experience tomorrow—that I will see the cat alive, for example—arose in the context that there is no fact about what experience I will have tomorrow; the statement "I will see a living cat" is neither true nor false. In logical terms, its truth value is neither 1 nor 0. But if the probability of this experience is close to 1, the statement is nearly true; if it close to 0, the statement is nearly false. This suggests that the truth value of the statement should be identified with its probability, and that this tells us what the probability of a single event means. *The probability of a future event is the truth value of the future-tense proposition that that event will happen*. This view of probability, and the associated many-valued logic of tensed propositions, has been explored in [10].

It has now become clear that the universal state vector plays very different roles in the two perspectives, external and internal. From the external perspective, it is a full description of reality; it tells how the universe is constituted at a particular time. This complete reality can be analysed with respect to any sentient system as in (12.2), yielding a number of components, attached to different experiences of the chosen sentient system, which are all parts of the universal reality. From the internal perspective of this system, however, reality consists of just one of these experiences; the component attached to this experience is the complete truth about the universe for the sentient system. All the other non-zero components are things that *might* have happened, but *didn't*. The role of the universal state vector at a later time, in this perspective, is not to describe how the universe *will be* at that time, but to specify how the present state of the universe *might change* between now and then. It gives a list of possibilities at that later time, with a probability for each of them that it will become the truth.

It might seem that we can at least know these probabilities for the future, being able to calculate them from our certain knowledge of our present experience, using the Schrödinger equation. But even this is uncertain. Our present experience could well be only part of the universal state, and it is the whole universal state vector that must be put into the calculation of future probabilities. Those things that might have happened, but didn't, some of which we don't even know about, might still affect the future. This is a further limit to human knowledge. However, if those things are sufficiently different from our actual experience on a macroscopic scale, then quantum theory assures us that the effect they might have on the future is so small as to be utterly negligible. This consequence of the theory is known as *decoherence*.

Human knowledge of the future, therefore, is limited in a fundamental way. It is not that there are true facts but the knowledge of them is not accessible to us; there are no facts out there, and there is simply no certain knowledge to be had. Nevertheless, there are facts with partial degrees of truth, and our knowledge of them is itself partial. Our best knowledge of the future can only be probable.

References

1. Gottwald, S. (2001). *A treatise on many-valued logics*. Baldock, Hertfordshire: Research Studies Press.
2. Laplace, P.-S. (1951). *A philosophical essay on probabilities*. New York: Dover.
3. Newton, I. (1934). *Principia mathematica* (p. xviii). University of California Press. Preface to the first edition.
4. Alford, M. G. Ghostly action at a distance: A non-technical explanation of the Bell inequality. *American Journal of Physics*. To be published. arXiv:1506.02179.
5. Sudbery, A. (1986). *Quantum mechanics and the particles of nature*. Cambridge University Press.
6. Sauders, S., Barrett, J., Kent, A., & Wallace, D. (Eds.) (2010). *Many worlds?* Oxford University Press.
7. Nagel, T. (1979). Subjective and objective. In *Mortal Questions* (pp. 196–214). Cambridge University Press.
8. Nagel, T. (1986). *The view from nowhere*. Oxford University Press.
9. Gillies, D. (2000). *Philosophical theories of probability*. London: Routledge.
10. Sudbery, A. (2016). *The logic of the future in quantum theory*. arXiv:1409.0755v6.

Chapter 13
Hermann Weyl's Space-Time Geometry and Its Impact on Theories of Fundamental Interactions

Norbert Straumann

Abstract One of the major developments of twentieth century physics has been the gradual recognition that a common feature of the known fundamental interactions is their gauge structure. In this essay the early history of gauge theory is reviewed, emphasizing primarily Weyl's seminal contributions of 1918 and 1929, that originated in a new concept of space-time.

13.1 Introduction

In this contribution we first sketch in a non-technical manner Hermann Weyl's early attempt to unify gravitation and electromagnetism by extending the space-time structure of general relativity (GR). Einstein admired Weyl's theory as "*a coup of genius of the first rate…*", but immediately realized that it was physically untenable: "*Although your idea is so beautiful, I have to declare frankly that, in my opinion, it is impossible that the theory corresponds to nature.*" This led to an intense exchange of letters between Einstein (in Berlin) and Weyl (at the ETH in Zürich), which is now published in *The Collected Papers* of Einstein [1]. No agreement was reached, but Einstein's intuition proved to be right.

Although Weyl's attempt was a failure as a physical theory it paved the way for the correct understanding of what is called gauge invariance, a central symmetry principle of modern physics. Weyl himself re-interpreted his original theory after the advent of quantum theory in a seminal paper [2].

Before coming to a description of Wey's first paper early in 1918 [3], we have to indicate Einstein's great step from Special Relativity (SR) to GR, in which a completely novel geometrical understanding of gravity was reached.

N. Straumann (✉)
Physik-Institut, Universität Zürich, 8057 Zürich, Switzerland
e-mail: norbert.straumann@gmail.com

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_13

13.2 From Minkowski's Space-Time to the Dynamic Space-Time of General Relativity

The first four of his five papers from March to June of Einstein' annus mirabilis 1905 were announced in a letter to his friend and member of the *Olympia Academy* Conrad Habicht. About the fourth paper Einstein writes: *[This] "is only a rough draft at this point, and is an electrodynamics of moving bodies which employs a modification of the theory of space and time."* This work, soon called relativity theory, attracted also the great mathematician Hermann Minkowski, one of Einstein's teachers at ETH in Zürich. In September 1908, during the annual meeting of the German association of Scientists and Physicians in Cologne, Hermann Minkowski presented a transparent geometric interpretation of Einstein's modification of space and time. This far reaching conception began with the famous sentences: "*The views of space and time which I wish to lay before you have sprung from the soil of experimental physics, and therein lies their strength. They are radical. Henceforth space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality.*"

The structure of this union is close to that of a 4-dimensional Euclidean space, with one crucial difference: From school days we are all familiar with the Euclidean plane, and learned the theorem of Pythagoras, which implies that the distance between two points can be represented as a sum of two squares. In Minkowski's space-time the distance between two localized events is a sum of three squares *minus* one square, corresponding to a temporal direction. For this reason we say that this space-time geometry is pseudo-Euclidean. For sometime Einstein considered Minkowski's 4-dimensional geometry as "*superficial learnedness*" and once wrote to a colleague: "*Since the mathematicians have grabbed hold of the theory of relativity, I myself no longer understand it.*" We shall see that a few years later he changed this opinion.

With the completion of SR, Newton's theory of gravity had to be changed, because the Newtonian law of gravitational attraction is an action -at-a-distance law. We know from later recollections by Einstein that he was soon convinced that gravitation has no place in the framework of SR. In the course of a slow process he arrived at the hypothesis that the rigid pseudo-Euclidean metric of Minkowski had to be generalized to a *dynamical* field, subject to laws, analogous to what was known from electromagnetism. The physical metric field implies that space-time is curved and that it influences and is being influenced by all other physical processes, for instance the motion of stars. Einstein finally found the laws of this two-sided interactions, in particular field equations which determine the generation of gravitational fields by material sources of all kind.—This is not the place to say more about this greatest contribution of Einstein, called general relativity theory, that the world celebrated in November 2015—100 years after Einstein's completion of the theory.

This essay was begun just a few days after the announcement that gravitational waves (ripples of the metric field) had been recorded, that were created in the coalescence of two black holes of about 30 solar masses. In such processes the dynamical nature of space-time in GR is particularly impressive.

13.3 The Quest for Unification

After Einstein had reached his goal of a successful relativistic theory of gravity, he began to think about the remaining arbitrariness of the new theoretical framework. One of these was the separate existence of gravitation and electromagnetism. According to his views, they had to be unified. Furthermore, GR did not impose any restrictions on the properties of matter. The mass and charge of the electron and proton, and why there were (at the time) no other particle in nature, appeared to be arbitrary. A major goal of a unified theory was to explain the existence and properties of matter. This search was Einstein's main pursuit for more than half of his scientific life, without real success.

A first very interesting unification attempt was put forward by another great figure, namely Hermann Weyl. Before we come to this, it should be said, that before GR was born Weyl was exclusively occupied with central problems in pure mathematics. But with Einstein's new theory of gravity he became very interested in GR. He wrote the first systematic presentation of the theory with the title "Space-Time-Matter" (STM) [4], after his lectures on the subject in the Summer Term of 1917 at ETH in Zürich. In the preface of the first (now seven) edition he wrote: "*At the same time it was my wish to present this great subject as an illustration of the intermingling of philosophical, mathematical, and physical thought, a study which is dear to my heart.*" He continued with: "*But I have not been able to satisfy these self-imposing requirements: the mathematician predominates at the expense of the philosopher.*" Such books are not written any more.

STM fascinated me enormously during the first semesters of my studies at ETH. The preface begins with "*Einstein's Theory of Relativity has advanced our ideas of the structure of the cosmos a step further. It is as if a wall which separated us from Truth has collapsed. Wider expanses and greater depths are now exposed to the searching eye of knowledge, regions of which we had not even a presentiment.*"

Unfortunately, I never saw Hermann Weyl; he died in Zürich during my first semester in fall of 1955.

13.4 Weyl's Attempt to Unify Gravitation and Electromagnetism

On the 1st of March 1918 Weyl writes in a letter to Einstein: "*These days I succeeded, as I believe, to derive electricity and gravitation from a common source…*". Einstein's prompt reaction by postcard indicates already a physical objection which he explained in detail shortly afterwards. Before we come to this we indicate the main ideas of Weyl's theory of 1918 [3].

13.4.1 Weyl's Generalization of Riemannian Geometry

Weyl's starting point was purely mathematical. He felt a certain uneasiness about Riemannian geometry,[1] as is clearly expressed by the following sentences early in his paper:

> But in Riemannian geometry described above there is contained a last element of geometry "at a distance" (ferngeometrisches Element)—with no good reason, as far as I can see; it is due only to the accidental development of Riemannian geometry from Euclidean geometry. The metric allows the two magnitudes of two vectors to be compared, not only at the same point, but at any arbitrarily separated points. A true infinitesimal geometry should, however, recognize only a principle for transferring the magnitude of a vector to an infinitesimally close point and then, on transfer to an arbitrary distant point, the integrability of the magnitude of a vector is no more to be expected than the integrability of its direction.

After these remarks Weyl turns to physical speculation and continues as follows:

> On the removal of this inconsistency there appears a geometry that, surprisingly, when applied to the world, explains not only the gravitational phenomena but also the electrical. According to the resultant theory both spring from the same source, indeed in general one cannot separate gravitation and electromagnetism in a unique manner. In this theory all physical quantities have a world geometrical meaning; the action appears from the beginning as a pure number. It leads to an essentially unique universal law; it even allows us to understand in a certain sense why the world is four-dimensional.

For certain readers the following few technical explanations may be useful. (A detailed description can be found in [5].) In contrast to GR Weyl's geometry is equipped not with one, but a class $[g]$ of conformally equivalent metrics This corresponds to the requirement that it should only be possible to compare lengths at one and the same world point. In addition, the theory contains also a class of vector fields $[A]$. A crucial property is that substitutions of the form

$$g \mapsto e^{2\lambda} g, \quad A \mapsto A - d\lambda, \tag{13.1}$$

where λ is an arbitrary smooth space-time function, do not change the geometry. Pairs (g, A) related by (13.1) are considered to be equivalent. In Weyl's application to physics, they leave the physical laws unchanged. These transformations, called *gauge transformations*, play a central role. The first of the substitutions is interpreted by Weyl as a different choice of calibration (or gauge). This is accompanied by the substitution of the vector field A, a transformation physicists know since the 19th century from electrodynamics.

---

[1] This is the geometry based on the metric field of GR, that provides the mathematical tools for Einstein's theory of gravity. Einstein learned this (relatively new) mathematics from his friend Marcel Grossmann, who was professor for mathematics at ETH. The two colleagues had from summer 1912 to spring 1914 a very fruitful collaboration, in which they came close to the final theory.

13.4.2 Electromagnetism and Gravitation

Turning to physics, Weyl assumes that his "purely infinitesimal geometry" describes the structure of space-time and consequently he requires that physical laws should satisfy a double-invariance: 1. They must be invariant with respect to arbitrary smooth coordinate transformations. 2. They must be *gauge invariant*, i.e., invariant with respect to substitutions (13.1) for an arbitrary smooth function λ.

Nothing is more natural to Weyl, than identifying A with the vector potential and $F = dA$ with the field strength of electromagnetism.

Independent of the precise form of the action Weyl shows that in his theory gauge invariance implies the *conservation of electric charge* in much the same way as general coordinate invariance leads to the conservation of energy and momentum. This beautiful connection pleased him particularly: "… *[it] seems to me to be the strongest general argument in favour of the present theory—insofar as it is permissible to talk of justification in the context of pure speculation.*" Similar structural connections hold also in modern gauge theories.

13.4.3 Einstein's Objection

After this sketch of Weyl's theory we come to Einstein's striking counterargument which he first communicated to Weyl by postcard. The problem is that if the idea of a nonintegrable length connection (scale factor) is correct, then the behaviour of clocks would depend on their history. Consider two identical atomic clocks in adjacent world points and bring them along different world trajectories which meet again in adjacent world points. Then their frequencies would generally differ. This is in clear contradiction with empirical evidence, in particular with the existence of stable atomic spectra. Einstein therefore concludes:

> …(if) one drops the connection of the metric to the measurement of distance and time, then relativity looses all its empirical basis.

The author has described the intense and instructive subsequent correspondence between Weyl and Einstein elsewhere [6]. As an example, we quote from one of the last letters of Weyl to Einstein:

> This [insistence] irritates me of course, because experience has proven that one can rely on your intuition; so little convincing your counter arguments seem to me, as I have to admit…
>
> By the way, you should not believe that I was driven to introduce the linear differential form in addition to the quadratic one by physical reasons. I wanted, just to the contrary, to get rid of this 'methodological inconsistency (Inkonsequenz)' which has been a stone of contention to me already much earlier. And then, to my surprise, I realized that it looks as if it might explain electricity. You clap your hands above your head and shout: But physics is not made this way! (Weyl to Einstein 10.12.1918).

13.5 Weyl's 1929 Classic: "Electron and Gravitation"

Shortly before his death late in 1955, Weyl wrote for his *Selecta* [7] a postscript to his early attempt in 1918 to construct a 'unified field theory'. There he expressed his deep attachment to the gauge idea and adds (p. 192):

> Later the quantum-theory introduced the Schrödinger-Dirac potential ψ of the electron-positron field; it carried with it an experimentally-based principle of gauge-invariance which guaranteed the conservation of charge, and connected the ψ with the electromagnetic potentials ϕ_i in the same way that my speculative theory had connected the gravitational potentials g_{ik} with the ϕ_i, and measured the ϕ_i in known atomic, rather than unknown cosmological units. I have no doubt but that the correct context for the principle of gauge-invariance is here and not, as I believed in 1918, in the intertwining of electromagnetism and gravity.

This re-interpretation was developed by Weyl in one of the great papers of the twentieth century [8]. Weyl's classic does not only give a very clear formulation of the gauge principle, but contains, in addition, several other important concepts and results.

Much of Weyl's paper penetrated also into his classic book "The Theory of Groups and Quantum Mechanics" [8]. There he mentions also the transformation of his early gauge-theoretic ideas: "*This principle of gauge invariance is quite analogous to that previously set up by the author, on speculative grounds, in order to arrive at a unified theory of gravitation and electricity. But I now believe that this gauge invariance does not tie together electricity and gravitation, but rather electricity and matter.*"

Many years later, Weyl summarized this early tortuous history of gauge theory in an instructive letter [9] to the Swiss writer and Einstein biographer Seelig [9], which we reproduce in an English translation.

> The first attempt to develop a unified field theory of gravitation and electromagnetism dates to 1918, in which I added the principle of gauge-invariance to that of coordinate invariance. I myself have long since abandoned this theory in favour of its correct interpretation: gauge-invariance as a principle that connects electromagnetism not with gravitation but with the wave-field of the electron. —Einstein was against it [the original theory] from the beginning, and this led to many discussions. I thought that I could answer his concrete objections. In the end he said "Well, Weyl, let us leave it at that! In such a speculative manner, without any guiding physical principle, one cannot make Physics." Today one could say that in this respect we have exchanged our points of view. Einstein believes that in this field [Gravitation and Electromagnetism] the gap between ideas and experience is so wide that only the path of mathematical speculation, whose consequences must, of course, be developed and confronted with experiment, has a chance of success. Meanwhile my own confidence in pure speculation has diminished, and I see a need for a closer connection with quantum-physics experiments, since in my opinion it is not sufficient to unify Electromagnetism and Gravity. The wave-fields of the electron and whatever other irreducible elementary particles may appear must also be included.

References

1. *The collected papers of Albert Einstein*. (1987). (Vols. 1–9). Princeton University Press. http://www.einstein.caltech.edu/.
2. Weyl, H. (1929). Elektron und Gravitation. *I.Z. Physics*, *56*, 330.
3. Weyl, H. (1918). *Gravitation und elektrizität.* Sitzungsber (pp. 465–480). Berlin: Akademie der Wissenschaften.
4. Weyl, H. (1993). *Space · time · matter.* Translated from the 4th German edition. London: Methuen 1922. *Raum · Zeit · Materie*, 7. Auflage: Springer.
5. O'Raifeartaigh, L., & Straumann, N. (2000). Gauge Theory: Historical origins and some modern developments. *Review of Modern Physics*, *72*, 1–23.
6. Straumann, N. (1987). Zum Ursprung der Eichtheorien bei Hermann Weyl. *Physikalische Blätter*, *43*(11), 414–421.
7. Weyl, H. (1956). *Selecta.* Birkhäuser-Verlag.
8. Weyl, H. (1950). *Gruppentheorie und quantenmechanik.* Wissenschaftliche Buchgesellschaft, Darmstadt 1981 (Nachdruck der 2. Aufl., Leipzig 1931). English translation: Group Theory and Quantum Mechanics. New York: Dover.
9. Seelig, C. (1960). *Albert Einstein* (p. 274). Zürich: Europa Verlag.

Chapter 14
Matter, Space, Time, and Motion: A Unified Gravitational Perspective

C.S. Unnikrishnan

Abstract This article explores the links between space, time, matter and its motion in the essential context of the universe they constitute. Conventional and standard physics maintains a nearly dismissive separation between dynamics and the matter-filled universe and indeed between cosmic matter and a priori space-time while asserting their mutual influence in other physical situations through the Einstein equations of general relativity. This can be traced to the incompatibility between the special theory of relativity that rejects a preferred frame and the real situation of a single matter-filled universe in which observers in different states of motion have measurably different experiences. We will argue that the present position cannot be maintained consistent with empirical and physico-logical evidence. The historical fact that all our fundamental theories were formulated and completed before we acquired significant and crucial knowledge about the matter content of the universe and its gravity is contrasted with the unavoidable situation that all our empirical experience and dynamics happens in the presence of all other matter in the universe. A careful analysis leads to a new and necessary paradigm in which several of our notions and even fundamental theories involving space, time and matter need to be reformulated with cosmic gravity as the determining factor. This, when combined with some crucial experimental results, answers several open issues that have been debated for centuries in foundational physics.

14.1 Introduction

The central theme of discussion in this paper is the relation between matter and dynamics and its physical description. Hence we have to deal with the action of matter on matter and two essential elements in the backdrop, albeit their emergent nature, are space and time. They are emergent because they are part of our experience only in the ambience of sensible matter. There is no way to prove that they are a priori and independent of matter. There is not even a way to prove that they

---

C.S. Unnikrishnan (✉)
Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400005, India
e-mail: unni@tifr.res.in

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_14

exist except as relational entities. However, modern theoretical physics, especially the theories of relativity, treat them as a priori physical entities, ontologically prior and independent of matter, with an existence of their own. Their status in physics has evolved considerably, but this evolution is not in the direction that solves the fundamental problems that arose in the first place, in the context of the formal discussion of space and time in Newton's Principia and in its critiques. I hope to address these very issues and come to definite conclusions here. I will in fact argue that we are stuck in square one with regard to the question of the ontological status of space and time as physical entities due to the single important reason that we have ignored the matter-filled universe and its gravity in our 'completed' and 'well-tested' fundamental theories. The streak of inner wisdom that could have been integrated into modern physics starting from the Machian speculation on the origin of inertia and his critique on the physical reality of space have been ignored with serious implications to the very basis of modern physics, its consistency, general applicability, and its philosophy. The situation can be saved only by a paradigm change in which space and time are fully compatible with matter. The core issue is the incompatibility between the real gravitationally dominant and hence physically preferred cosmic frame, with its universal time providing an explicit reference for the state of 'rest', and the prevalent concept of a relativistic space and time in modern physics that rejects the physical prominence of any preferred frame. What is urgent then is to pay attention to logical consistency and empirical evidence with rigour, avoiding over-interpretations and focusing on the very postulates on which our fundamental theories are built. This results in a new theory in which dynamics and relativity and gravitational consequences of the matter-energy in the universe and the cosmic frame playing the role of an absolute frame for motion and time, all fully consistent with the totality of empirical evidence.

It is natural in physical vision that what is absolutely transparent remains invisible. But in the realm of ideas and insights the situation should be exactly the opposite. For the logical inner eye, a crystal clear evidence should form the basis a theoretical construction. Yet, it is a disturbing revelation that all of modern physics continue to base its fundamental postulates on empty space and associated abstract time, albeit modifiable, in spite of our mature knowledge that all physical processes and dynamics happens in the presence of vast amount of matter and its integrated gravity in a gigantic universe with no empty space. The interaction energy between the earth and us (or any body) is only a billionth of that between us and the rest of the matter in the universe. Yet, our experience misses out on cognition of this interaction, and this has reflected in how we formulate and accept our theories of physics.

All our fundamental theories were formulated and completed well before we had any significant knowledge about the matter and its extent in the universe, and hence with the nonexistent 'empty space' as their arena. The consistency of this paradigm is taken for granted, because of our physical insensitivity to enormous gravitational potentials of the vast amounts of matter in the Universe and also because of apparent agreement with empirical evidence. However, all experimental tests of any of our theories are always done in the presence of vast amounts of cosmic matter and its gravity, and if cosmic gravity had any effect on dynamics and time, we certainly have missed it in our fundamental theories. To stress this point, one may note that the totality of empirical experience unavoidably contains any effect of cosmic gravity, whereas the theories explicitly assume that such effects are negligible, because the theories were constructed before we knew about the universe. It is logically imperative that all physical theories dealing with space, time and motion, in particular for theories of dynamics and relativity, need to incorporate cosmic matter and gravity in the theory to the extent they have genuine and observable physical effects. That this is essential is not merely an issue of being logically consistent, but being empirically rigorous, as we shall prove. With this, our notions of space and time also will naturally need revision because the cosmic gravitational frame assumes the role of a preferred 'absolute' frame and hence compatible notions of space and time inherits the same features.

Perceptions of space and time are always in association with matter; otherwise space and time are unobservables. Dynamical space is abstracted out of separation or interval between discernible matter. Time, in all cases imaginable, is in fact just motion (evolution) of matter in a generalized sense and not a separate dimension. Motion in space is then naturally motion in the presence of matter, distributed as the universe itself. We will see that the conventional ontology of space and time in modern physics does not go beyond being metaphysical. However, this being at the very foundation of the theories of relativity and gravity, we need to carefully examine how a changed perspective modifies standard physics itself, which is considered robust. A new paradigm has to defend itself while being consistent with all empirical features of the old view. Rejection of unobservable space and time in favour of a universe defined by its matter content and gravity not only demands new physics, but also brings back the idea of absolute space and time in place of Einstein's relative space-time, all tied rigorously and consistently with the special reference frame determined by cosmic matter.

In such a paradigm, which I call cosmic relativity [1, 2], there is indeed a preferred gravitational frame of the universe, with its unmistakable and incontestable physical reality. An ontologically and logically robust framework and theory of space, time, matter and dynamics naturally emerges with the preferred cosmic frame as reference for motion, and a universally synchronizable time, and yet consistent with all known empirical results. All relativistic effects and the law of dynamics itself are derived as gravitational effects of matter in the universe and therefore I call the cosmic frame as the 'determinant frame', to indicate a deeper meaning than 'preferred frame'. Obviously, the relative velocity of light then has to be Galilean, and this is where the new theory goes well beyond anything that have been attempted in the Machian spirit (a factual experimental situation which confirms this is discussed in Appendix E). This will of course go beyond what we accept as a standard theory of motion and relativity, and key empirical tests where the conventional theory fails will be highlighted. The strength of cosmic relativity is its experimental support. Machian speculation cannot be taken forward and implemented without a preferred and privileged role for the cosmic gravitational frame and this is precisely why Einstein's general theory of relativity remained non-Machian, with no explanation for inertia or pseudo-forces. This paradigm then demands a modified Einstein's equation for gravity that includes cosmic matter as a permanent feature in the equation itself, with the advantage of being Machian and fully consistent with experimental tests, devoid of pseudo-forces. In such a framework, causality and locality are supreme and speculations of nonlocality, as in quantum entanglement, have no place. With inconsistencies and paradoxes that worried philosophers of science eliminated, the physical theory reclaims its rightful position in natural philosophy.

14.2 The Standard Scenario

The general theory of relativity (GTR) is the standard theory of gravity in which it deals with space, time and matter in an integrated way. It is built on just two foundational pillars, one is the special theory of relativity (STR) and the other the Einstein equivalence principle. STR rejects any notion of absolute rest and its fundamental mathematical structure involves boost transformations that preserves the isotropy and homogeneity of empty space and time. In fact, only empty space can remain isotropic in all inertial frames and therefore STR is a false start in a matter-filled universe, the only one known. One special feature of STR is that no physical reason (influence) is ascribable to relativistic changes, like time dilation of a relatively moving clock. The foundation for GTR is the Einstein equivalence principle that generalizes the equivalence of the inertial and gravitational mass and asserts the local equivalence of a uniform gravitational field $\vec{g}$ and an accelerated frame with acceleration $\vec{a} = -\vec{g}$. GTR does not seek the physical reason for the empirically observed validity of the equivalence principle in either form. It merely assumes it as a fundamental postulate consistent with precision experiments. The most interesting aspect, in the context of this article, of basing the theory on such a principle is that gravity is treated as equivalent to a pseudo-force. Therefore, the theory goes against its announced inspiration of Machian ideas on banishing pseudo-forces from physics. In some sense it is ironical that the GTR that gave rise to modern cosmology is firmly built on the idea of empty space and a matter filled cosmos in just one of the solutions of Einstein's equations. This indeed is the statement that GTR does not incorporate the Mach's conjecture concerning the Newtonian pseudo-forces [3]. The non-gravitational inertial forces retain their pseudo-status even within GTR.

The great improvement from Newton's absolute space to Einstein's space-time is that matter-energy modifies the spatio-temporal relations and thus acts on space and time and the behaviour of clocks and trajectories are in turn modified by the modified space and time. The GTR does not get rid of the Newtonian space with its own physical ontology and reality. It just makes it modifiable by the presence of matter. However, it does not make any difference if we say that space is modified and therefore all motion are affected, or that there are gravitational fields that affect length and time standards. It is in fact an over-interpretation to insist that space and time are modified by the presence of matter because that statement is untestable. What can be tested is whether relative distances between material particles change and whether material clocks at different positions have different rates. The only observables associated with space and time are spatial and temporal intervals of binary events marked by matter, including clocks. The two kinds of intervals (distance and duration) are to be measured individually and a space-time interval, if created by combining them, is a synthetic concept. Geometry is the set of relations on distances and durations. If matter in free motion is taken as reference, their trajectories are modified in the presence of other matter. So, are the rates of clocks. The modification of distances and durations in the presence of matter alter the geometric relations, and this is conventionally called the modification of the geometry of space-time, whereas in reality modifications refer to the geometrical relations between separations of material points and durations of material clocks. Because the modifications of intervals are independent of the properties of matter, as embodied in the equivalence principle, one may drop reference to specific matter and work legitimately with dynamics in the modified geometry of space-time.

However, it is important to understand that the relation between geometry and matter as specified by the Einstein's equations of GTR, equating components of 'curvature' of space and time to the components of energy-momentum of matter, merely relates how spatial and temporal intervals defined by matter vary in relation to other matter, in the presence of matter. Both the mathematical relation and a prescription of measurements to test the relations involve only material trajectories and clocks. The conventional interpretation of the 'left-hand side' of the equations involving the concept of a physical and malleable space-time independent of matter is just a multi-layer synthetic convenience. Instead of seeing gravity as action of matter on matter, sandwiching the metaphysical space and time as the dynamical entities, both of which have no gravitational charge whatsoever (mass or energy), is an over-interpretation, albeit being very convenient and useful mathematically, and consistent because of the equivalence principle.

All the modern developments from dynamics to cosmology relied on the theories of relativity. In fact, the construction of cosmology became possible only because of the structure of GTR that linked matter and space-time trajectories in a way that did not need dealing with forces and bounded mass distributions. One could explore how the 'metric' of space and time evolved, which in turn determined how matter distributions evolved and an entire consistent picture emerged. However, these very developments bring out deep questions of logical consistency that are powerful enough to question and shake the very foundations on which the theory is built. This may sound surprising, to say the least, but even a superficial examination of some of the issues may be enough to persuade anybody who insist on both logical and empirical rigour that this is indeed the case. The cosmological picture we have today from observations turns out to be that of a preferred frame, conceptually and practically, with its gravity determining both dynamics and relativity, and providing the physical reason for the validity of the equivalence principle, thereby rendering the very theory that initiated and facilitated all these studies incomplete. The universe is gravitationally self-dynamical and self-determinatory all the way down to the dynamics of its every elementary constituent. And the standard theory of gravity does not encompass any of these features.

That this is indeed the situation in GTR is to be stressed and I quote from Einstein's Nobel lecture [4], "Already Newton recognized that the law of inertia is unsatisfactory in a context so far unmentioned in this exposition, namely that it gives no real cause for the special physical position of the states of motion of the inertial frames relative to all other states of motion… For this reason E. Mach demanded a modification of the law of inertia in the sense that the inertia should be interpreted as an acceleration resistance of the bodies against one another and not against "space"… This interpretation is even more plausible according to general relativity which eliminates the distinction between inertial and gravitational effects. It amounts to stipulating that, apart from the arbitrariness governed by the free choice of coordinates, the $g_{\mu\nu}$-field shall be completely determined by the matter. Mach's stipulation is favoured in general relativity by the circumstance that acceleration induction in accordance with the gravitational field equations really exists, although of such slight intensity that direct detection by mechanical experiments is out of the question."

Einstein reaches so close to the real solution and misses it, proclaiming undetectability of the small GTR effects (as finally detected in experiments like the Gravity Probe-B), because he could not incorporate the vast amounts of matter in the universe into his equations as a consistent eternal factor to determine the large cosmic $g_{\mu\nu}$-field responsible for the readily observable inertial forces! In fact, he unwittingly sacrificed forever his original desired program that 'the $g_{\mu\nu}$-field shall be completely determined by the matter' by implying that space and time themselves could be gravitational sources, by introducing the cosmological constant.

14.3 The New Paradigm for Dynamics and Relativity

14.3.1 A Hypothetical Scenario in Electrodynamics

Imagine a situation where there is a slight charge asymmetry of less than 1 part in 10^{20} between the proton and electron and that atoms are slightly charged [5]. In the pre-cosmology era, while Maxwell was completing his theory, no knowledge about the matter content of the universe was known and such a tiny charge asymmetry in matter known then (solar system, Milky way, a few nebulae around us and distant stars) would not make any difference to experimentally observed features of electrodynamics, even if somebody cared to include the background charge in calculations. However, the true reality would have been the enormously charged universe with a huge electric potential generated by all the unseen charged matter, but undetectable in all inertial experiments. In fact, for uniform motion the equations of electrodynamics, including the Lorentz force law, would work perfectly with only local magnetic and electric fields included. That is because any uniform motion within the large charged sphere creates only constant four-vector potentials and they cannot be locally detected. However, in situations of accelerated motion of the charge, some extra forces would appear without any identifiable sources because the potentials becomes time dependent. An extra force will be required to accelerated a particle as if a pseudo-electric field is acting against the acceleration of all charges. Charges in rotation will experience pseudo-magnetic forces, and so on. A magnetic moment in a rotating frame would feel an anomalous torque, without any identifiable source. The situation would require postulating pseudo-forces in electrodynamics, proportional to the electric charge and one would be faced with an electromagnetic inertia, in addition to the usual inertia proportional to mass, forcing one to ascribe it as well to special motion relative to 'absolute' space, as Newton did to discuss his 'water in a rotating pail' experiment. A Machian critique would demand looking for an interaction involving the electric charge instead of folding one's tails between the warm comfort of pseudo-forces. We do not see such forces and the entire electrodynamics is accounted for with local sources only because the matter in the universe is essentially neutral. Yet, violent and irrational objections that go well beyond scientific method are raised when it is pointed out that the origin of the observed inertial (coupling to mass) pseudo-forces is indeed the gravitationally charged universe, with the huge amount of matter around us compensating for the weakness of the gravitational constant.

This immediately suggests that we are grossly wrong in ignoring the gravity of the matter around us in our fundamental theories of dynamics. The gravitational potential of just the visible matter in the universe adds up to about 10^8 times the gravitational potential of the earth in a terrestrial laboratory. It is also well-known that the Newtonian potential of a near critical-density universe is equal to the square of the velocity of light $\left(\Phi_U \simeq c^2\right)$, and should have been taken as the strong clue that all of relativistic effect are in fact gravitational. A phenomenological claim that all Lorentz factors are in fact $\left(1 - v^2/\Phi_U\right)^{1/2}$ and has nothing to do with velocity of light is empirically fully defendable [1]. Motion then generates velocity dependent potentials and acceleration transforms this into a time dependent potential which manifests as a force that resists motion. The electromagnetic analogy and our present knowledge about the gravitational active matter content in the universe should have been sufficient to see the truth of the claim that conventional pseudo-forces are indeed gravitational, as researchers like D. Sciama had already shown [6]. They are actually gravito-magnetic (arising from either the curl or the time dependence of the relative currents), but opening up to this wisdom is highly resisted without defendable reasons. However, stopping there gives an incomplete and inconsistent story, rejected several times in the past by the physics community, even though there is no empirical or logical objection to such a calculation and demonstration. The all-important consequence to realize is that such a framework also demands a preferred reference and notion of absolute rest, and hence compatible Galilean notions of space and time and propagation of light, which goes completely against what is accepted as the correct physics today.

14.3.2 Physics in the Gravitationally Charged Universe

Once we anchor a theory of relativity and dynamics on the matter and gravity of the real universe, as it is, there are no further freedom to tune any aspect of the theory of motion. All that we know and see about motion, and evolution and modifications of space-time intervals should follow as mere gravitational consequence of interactions with cosmic matter. There are no additional postulates or principles to depend on. Matter and its gravity becomes our physical priors, instead of the triad of space, time and matter and all else should follow from it. Do they? Remarkably so!

Empty space is the same everywhere (homogeneous) and in all directions (isotropic) for all inertial observers and the measures of spatial and temporal intervals or the metric remain homogenous and isotropic in every moving frame. The invariant isotropy is a property specific to the Lorentz transformation (LT) wherein the metric *diag* $\{-1, 1, 1, 1\}$ goes to *diag* $\{-1, 1, 1, 1\}$ under motion. This is the very basis of the special theory of relativity. The matter distribution in the universe is more or less homogenous and isotropic on very large scales, but only for an observer who is 'at rest' relative to the average matter distribution. It is clear that moving in matter-filled universe generates a relative flow of matter opposite to the motion of the observer and this current of gravitational charge is like any other physical current and naturally should generate physical effects akin to magnetic effects in electromagnetism. The current and magnetic potentials defines a direction in space and that makes space and its metric (measuring standards) anisotropic. However, in reality, motion results in anisotropy proportional to the velocity, inducing an anisotropic metric, which is impossible with Lorentz transformations. Therefore, the real universe is maximally Lorentz violating, contrary to general belief (the truth is not visible to two-way experiments, like the Michelson-Morley experiment or its variations, where this anisotropy cancels to first order). The Doppler dipole anisotropy of the temperature of the cosmic microwave background (CMB) enables fairly precise determination of one's motion relative to the cosmic frame. There is no more an excuse to claim that uniform motion is equivalent to rest. More serious is the fact that the monotonically decreasing temperature of the CMB provides us with a universal absolute clock and time, synchronized everywhere in the universe to a precision better than a part in 10^5. This demolishes the arguments for lack of simultaneity of spatially separated events.

Naturally, one might dismiss the deviation we are suggesting as pointless in spite of its logical strength because the (only) testable features of Lorentz transformations, physical effects like time dilation and velocity addition law, relativistic Doppler shift etc. all are experimentally confirmed. In contrast, absolute preferred frame requires Galilean transformations for coordinates and how can it reproduce fundamental and everyday effects like the time dilation?! Well, it turns out to be easier than one imagined, as explicitly shown in the Appendix T. Galilean transformation does indeed change the metrical measures of space and time and time coordinate picks up a correction $\sqrt{1 - v^2/c^2}$, which is the time dilation, now with v as the absolute velocity. All experiments are consistent with this and one can trace the well-documented east-west asymmetry of GPS clocks to exactly this, and also rest with a correct discussion of twin paradox etc. without raising vague and flimsy arguments about non-inertial motion etc. [7]. All other important relativistic results then follow, if we consistently include the fact that the relative velocity of light is also Galilean, as has been explicitly demonstrated in recent experiments (Appendix E).

14.3.3 The Complete Solution

How is the theory to be reconstructed in way that retains the agreement with all its impressive precision tests during a century and yet compatible and consistent with the prior background universe with so much matter and its gravity? This can be done only by modifying the Einstein equation itself, with cosmic relativity as the new basis and cosmic gravity incorporated into the equation as its permanent feature. Fortunately this is easy, maintaining the Bianchi identities and the consistency with all known experimental tests. The resulting Centenary Einstein's equation [8, 9] is

$$R_{ik} - \frac{1}{2}g_{ik}R - \frac{8\pi G}{c^4}T_{ik(U)} = \frac{8\pi G}{c^4}T_{ik} \tag{14.1}$$

The extra piece on the left is the energy momentum tensor of the universe, included as the non-removable integral part of the equation itself. Adding this term implies universe as a preferred frame. We have avoided the cosmological constant term deliberately.

Then, laws of motion emerge as a gravitational consequence of cosmic matter. Exactly the same physical reason then leads us to the principle of relativity, relativistic modification of duration and separation, and to foundations of the theory of gravity itself, forming one inseparable self-determined physical system. Dynamics acquires a new meaning and even content, purging pseudo-forces and causeless effects of motional relativity, from physical theory. This of course regains the “absolute” and there is indeed preferred space, defined by cosmic matter, and absolute universal time, synchronized all over the universe, and even practically available as a universal clock in the slowly decreasing temperature of the microwave background radiation. The price to pay is the abolition of untested beliefs, like the invariance of relative one-way velocity of light, held sacred for over a century. Yet, such a framework is remarkably consistent with all empirical evidence, contrary to what one might naively expect. After a century, we again have harmony between philosopher’s time and physicist’s time, and indeed modern common mans’ time.

14.3.4 Matter, Dynamics and Relativity

Here we summarize the relation between matter, dynamics and relativity, as supported by empirical evidence and logical consistency. In terms of a causal structure, we conclude that space is not a causal primary element for dynamics or relativity. Both dynamics and relativistic effects pertains to change in physical states of matter and requires a fundamental interaction as the cause. Any theoretical paradigm that does not identify and include the physical interaction responsible for physical changes has to be considered tentative and incomplete. It is in fact straightforward to derive the laws of dynamics, like the Newton's equation of motion, and relativistic modification of spatial and temporal measures (metric) in frames moving relative to the matter in the universe as due to direct gravitational interaction and since the matter content in the universe is known to good accuracy now, there is no ambiguity about these relations. Hence, all motional relativistic effects are gravitational effects, arising entirely because the gravitational potentials in a moving frame are different from those in a frame at rest in the cosmic frame.

14.3.5 Matter, Space and Time

Now we are in a position to reassess the notion of space and time that is consistent with gravitationally active matter in this universe. First conclusion is of course that there is indeed physical justification for notion of absolute rest, absolute velocity and universally synchronized time with the matter-frame of the universe as the preferred frame. Prevalent notions of space and time in modern physics is undefendable, logically and empirically. Space and time are not physical entities by themselves and are just convenient proxy elements that represent the features of the gravitational properties of the underlying matter distributions. The most important conclusion for physics is that the space-time is Galilean in the coordinate transformational sense and then spatial intervals and durations correctly show second order relativistic modifications in motion with the absolute velocity as the deciding factor in the Lorentz-Larmour formulae. This has been now demonstrated in a number of experiments (Appendix E).

14.3.6 Universe and Its Space and Time

The expansion dynamics of the universe itself, with a history of over 10 billion years, is not really visible in the four hundred years of dynamics and relativity in physics. Yet, we need to mention it because of the current thinking that it is the expansion of space that is manifesting as the observed recession of matter from matter in the universe, rather than the real motion of matter. The pictures are equivalent in all respects and insisting on one and denying the other as invalid is arbitrary. The picture of expansion of space also requires that compound matter like atoms do not expand with space, whereas ‘nearly free’ matter like galaxies expand with the space. So, the present picture is a mixture of notions in which an expandable space takes with it ‘embedded matter’, as if in a solid expanding framework, while allowing matter to move through it freely, as if in a frictionless fluid. Also, as we mentioned earlier, endowing space and time (metric) with properties that make them equivalent to a source of gravity by adding a cosmological constant term to Einstein’s equations is a retrograde step in Machian view. We believe that the current physical interpretation of the expansion dynamics of the universe will change eventually, in favour of motion of matter determined by matter, instead of expansion of the abstract physical space.

14.4 Philosophers’s Space and Time

A new agenda for dealing with the two fundamental encompassing entities for physics, space and time, was indicated by logical positivists in the 20th century. Reichenbach asserted that the philosophy of space and time is essentially the philosophy concerning relativity (theories). O. Neurath, a prominent philosopher of the Vienna circle wrote [10], “Although what is called philosophical speculation is undoubtedly on the decline, many of the practically minded have not yet freed themselves from a method of reasoning, which, in the last analysis, has its roots in theology and metaphysics. No science which pretends to be exact can accept an untested theory or doctrine; yet even in an exact science there is often an admixture of magic, theology and philosophy. It is one of the tasks of our time to aid scientific reasoning to attain its goal without hindrance…

Physics has been successfully purged of metaphysical formulas. For example, the conception of ‘absolute motion’ has been discarded, a conception which acquired meaning only if one thought of ‘absolute space’ as a gigantic glass case in which ‘coordinates’ were woven like spider webs so that it became possible to determine whether a body is at absolute rest or whether it is moving about within the case. The Mach-Einstein conception dispenses with this ‘absolute space’…”

However, it turns out that facts contradict what the logical positivists favoured. Our discussion (and the reasoning and empirical evidence mentioned in the appendix) clearly demand the return of the absolute into the core of physics and the philosophy of physics has to respond in a logically consistent manner. Given the cosmic markers, accurately represented by the cosmic microwave background, nobody can assert the equivalence of rest and motion. Also, the temperature is monotonically decreasing, as has been for billions of years, and provides a universal time that is everywhere same in the universe to high accuracy. A local clock of any nature can be synchronized to the locally measured averaged temperature of the background and will read the same time as any other clock anywhere else. The new paradigm with clear ontological priority for matter-energy and its gravity and an emergent and interpretational space and time is not just a philosophically complete and satisfactory world view; it provides a complete derivation and understanding of dynamics as entirely gravitational.

14.5 Summary

By recognizing and incorporating the gravitational influence of all the matter-energy in the universe in dynamics and relativity, we have completed the theory of dynamics and relativity in a way that integrates Galileo, Newton, Mach and Einstein, solving the outstanding problems related to inertia, pseudo-forces, equivalence principle, principle of relativity and propagation of light. The only consistent paradigm has the matter-filled universe as the preferred absolute frame and Einstein equations of general theory of relativity is modified to incorporate this feature. All relativistic effects and laws of dynamics are gravitational and the notions of space and time is linked and inseparable from the cosmic frame, providing an 'absolute' and universally synchronized reference for motion and time, which is the determinant gravitational frame for all of physics.

Appendix T

Here we summarize the essential theoretical results [1, 2, 8, 11], outlining the derivations that support the arguments in the paper. The conventional description of the properties of space and time under motion is contained in the Lorentz transformations (LT). This rejects Galilean relativity and replaces it with invariance of the homogeneity and isotropy of space. However this can be maintained only in empty space. Empty space is isotropic and homogenous and remains so in every moving frame. So, LT that keeps the isotropic and homogenous metric invariant is the valid description of motional transformations only in the nonexistent metaphysical empty space. The real space has matter—the charge of gravity—at an average density of 10^{-29} g/cc, which becomes a 'current' of the charge of gravity in every moving frame. Space becomes anisotropic with this flow, a physical vector potential proportional to velocity (which is in fact a part of the full 10-component symmetric tensor), $A_i/c = g_{0i} = v_i/c$ is generated, but LT cannot accommodate this anisotropy (see Fig. 14.1). Note that a similar electromagnetic vector potential would have been generated if the universe were slightly electrically charged. Not only the geometrical features, but also the 'transformational features' of space and time are to be determined by the matter around.

This is all we need to see the solution to the age-old problem of pseudo-forces, inertia etc., and the real source of relativistic effects. The fact that there are no locally measurable physical effects of a constant vector potential field is then the

Fig. 14.1 Motion converts a nearly homogenous and isotropic distribution of matter into a matter-current generating relativistic anisotropic gravitational potentials or metric components. Only those motional coordinate transformations that reflect this are physically relevant and legitimate

statement of the principle of relativity and it is strongly tied to the observed homogeneity of the matter-energy distribution. Surprisingly, the Galilean transformations (GT) correctly gives us, along with the observed anisotropy, the most important relativistic feature of motion—time dilation! To illustrate this use a limited version of the actual Robertson-Walker metric, ignoring the very slow time evolution. GT specifies the coordinate transformations $x' = x = vt$ and $t' = t$. Physical spatial and temporal intervals can be deduced by including the metric in the moving frame. Under GT, the metric coefficients transform from $\left\{g_{00} = -1, g_{0i} = g_{i0} = 0, g_{ii} = 1\right\}$ to $\left\{g_{00} = -(1 - v^2/c^2), g_{0i} = g_{i0} = v/c, g_{ii} = 1\right\}$. Nonzero g_{i0}, a gravito-magnetic potential, gives the observed anisotropy. The temporal duration in the moving frame is now $\sqrt{-g_{00}}dt' = (1 - v^2/c^2)^{1/2}dt$. GT indeed contains the physics of time dilation with v as the absolute velocity or velocity relative to the cosmic frame. Of course, the relative velocity of light is now Galilean, $c' = c \pm v$. Therefore, we also get the length contraction correctly.

This can also be treated in the language of gravitational potentials. It is well known that the Newtonian gravitational potential Φ_u at a point in this universe of size Hubble radius or so, evaluated using the observed matter-energy density, is numerically close to c^2. In moving frames, the relativistic potential will have velocity dependent 'vector potential' component

$$A_i = \frac{v_i \Phi_u}{c}\left(1 - v^2/c^2\right)^{-1/2} = \gamma \frac{v_i \Phi_u}{c} \tag{14.2}$$

leading to several large gravito-magnetic effects.

Though g_{i0} is homogeneous in a uniformly moving frame, which implies the principle of relativity, if there an acceleration g_{i0} becomes time dependent and the physical effect is a reactive force on the accelerated system,

$$F_i = -m_g \frac{dA_i}{cdt} = -\frac{m_g \Phi_u}{c^2}\left(\gamma a_i + \gamma^3 v_i (\vec{v} \cdot \vec{a})\right) \tag{14.3}$$

In other words, the vector potential modifies the momentum (enabling generalizing to quantum theory) as $p' = p - m_g A_i$. We see that accelerating a body requires overcoming this cosmic gravito-magnetic reaction and hence a force F_i which is the full relativistic form of Newton's law of dynamics and the conventional inertial mass is just $m_i = m_g \Phi_u / c^2$. The law of dynamics is indeed gravitational, operative only because the universe is gravitationally charged, a connection that physics never imagined so far. Newton's law is a relativistic gravito-magnetic consequence of cosmic matter and the analogue in electrodynamics is the Lenz's law. We see that not just the conventional pseudo-forces in rotation, but also the very fact that a force proportional to mass is required to accelerate an object is a direct consequence of cosmic gravity, going a step deeper than Mach's speculation on pseudo-forces in rotation.

Hence the ratio m_i / m_g is universal. The equivalence principle is a necessary implication of cosmic gravity. Therefore, Newton's law of dynamics and the equivalence principle have the same physical content and one implies the other through their cosmic gravity connection [11]. Needless to emphasize that both the centrifugal and the Coriolis forces follow as consequences of cosmic gravity, fully satisfying Machian speculations. The curl of g_{0i} or A_i is 2Ω,which is the cosmic gravito-magnetic field B_{cg} in any rotating frame and an object moving at velocity v will feel a gravito-magnetic force $F_{gm}(v) = m_g v \times 2\Omega$. Thus Coriolis force is just the gravitational Lorentz force. The particular pseudo-force that Mach addressed, the centrifugal force is from the time dependence of the direction of A_i; $F_c = m_g A_i \frac{d\hat{A}}{dt} = mv^2/r$.

The other important effect is the coupling of cosmic gravity to spin, both classical and quantum. Spin is the current of the charge of gravity and all spin dependent physical effects should be traced to the gravitational interaction. The coupling is $s \cdot B_{cg}/2$. Cosmic relativity addresses this completely and derives a whole lot of observed and observable physical effects, ranging from geometric phases and spin valve effects on particles in chiral motion to purely quantum effects like the fractional quantum effect and spin-statistics connection [12].

Appendix E

The Relative Velocity of Light

In the context of the debate between absolute and relative, there is in fact just one crucial foundational postulate we need to test and verify for deciding one way or other. That is the nature of propagation of light. A paradigm in which there is a preferred frame that determines the 'true' velocity of light, the one-way velocity of light relative to an inertial observer moving relative to the frame at 'absolute velocity' v is $c \pm v$. In contrast, the prevalent concept of space-time in physics derives from the postulate of the invariance of the relative velocity of light, which in turn rejects the idea of absolute rest, preferred frame and universal time. Then light takes the same duration to travel the relative distance L measured from the observer, independent of the velocity of the observer. If this could be directly checked, we can decide without ambiguity on the fundamental debate. The general premise is that this test is not possible because the measurement of duration requires two clocks separated by distance L, and their synchronization requires the very feature we are trying to ascertain. However, it turns out that this problem can be solved in an amazingly simple way [1, 13], once we break apart from the psychological shackles of classical two-way experiments like the Michelson-Morley experiment which are not decisive between the two alternatives. That the M-M experiment cannot decide the issue, contrary to common beliefs and statements is easy to prove. The expected signal in the M-M interferometer is proportional to the difference in two-way light travel time in the two arms;

$$\delta T = \frac{2L_1}{c(1 - v^2/c^2)} - \frac{2L_2}{c(1 - v^2/c^2)^{1/2}} \tag{14.4}$$

This was observed to be zero. One explanation is that relative velocity of light is Galilean, but the arm L_1 that is along the direction of the velocity has a physical contraction of $L_1(1 - v^2/c^2)^{1/2}$. The STR explanation is that there is no physical length contraction (and it is a coordinate effect that depends on relative velocities), but the relative velocity of light is an invariant constant and hence light takes time $2L/c$ in each arm, independent of the velocity of the interferometer. There is no way to resolve this ambiguity in any variation of the M-M experiment. One needs to find a way to compare one-way relative velocities.

The physical issue to determine by a measurement is the time taken to travel a distance L relative to the observer. Imagine a pre-defined light path with fixed length. One counts time when light starts from one end and it is clear that the time light takes to reach the other end does not depend on the shape of the light path. However, the second clock at the other end comes close to the first one when we loop around the path and one can dispense with the need for two clocks! If one chases any moving entity that is moving very fast on a looped one-dimensional path, one gets a crossing after a duration $\delta t = L/c'$ where L is the distance around the loop and c' is the relative velocity, all measured relative to the observer. This can be measured with a single clock and the observer can continue to be on a inertial trajectory (unlike in the controversial Sagnac rotating experiment). The crucial point is that correctness of the experimental method can be tested on a known Galilean wave, like sound and then it becomes simply a comparison between sound and light in identical configurations (Fig. 14.2). The experiment with sound returns the expected result; the duration depends linearly on the velocity of the observer as $\delta t = L/(c - v) \simeq L(1 + v/c)/c$. The experiment with light returns the unexpected result; the duration, measured interferometrically, depends linearly on the velocity of the observer as $\delta t \simeq L(1 + v/c)/c$. Light is indeed Galilean! Its true velocity is determined in the cosmic rest frame (and not the ether of the 19th century). Irrespective of what one believes, a direct empirical test falsifies the fundamental postulate of relativity theories and also supports the existence of a preferred frame. The result itself cannot pin down what this frame is, but supplementary evidence unambiguously picks the cosmic gravitational frame as the determinant frame.

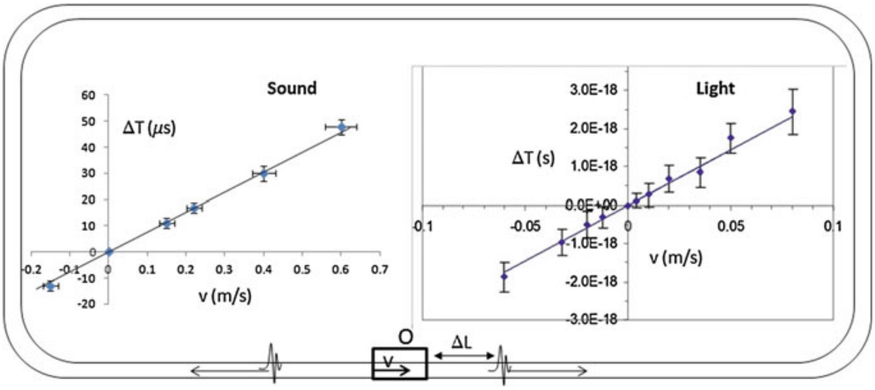


Fig. 14.2 When a wave-pulse is chased by an observer, the relative distance increases progressively, but the time taken to reach any point on the path can be determined with just one clock by looping the path around, while the observer is in inertial motion. The validity of the idea is tested by measuring the relative velocity of sound (*left panel*), which returns the expected Galilean dependence of the time delay on the velocity of the observer. The experiment is repeated with light in the same configuration and the result is similar, proving without ambiguity that light is indeed Galilean

It is important to mention that we have now completed several other experiments [12] involving currents and magnets in general relative motion and gyroscopes (gravito-magnets) in rotating frames, apart from analyzing several other experiments where particles with spin are transported along non-inertial trajectories, all of which unambiguously and firmly asserts that "phenomena of electrodynamics as well as of mechanics possess properties corresponding to the idea of absolute rest".

References

1. Unnikrishnan, C. S. Cosmic relativity: The fundamental theory of relativity, its implications and experimental tests. arXiv:gr-qc/0406023.
2. Unnikrishnan, C. S. (2011). Physics in the 'once-given' universe. In S. Ghosh & G. Kar (Eds.), *Recent developments in theoretical physics* (pp. 99–120). Singapore: World Scientific.
3. Mach, E. (1919). *The science of mechanics*, (4th ed., pp 222–245 and appendix XX). (London: The Open Court Publishing Co.).
4. Einstein, A. (1967). *Fundamental ideas and problems of the theory of relativity, nobel lectures, physics 1901–1921*. Amsterdam: Elsevier Publishing Company.
5. Unnikrishnan, C. S., & Gillies, G. T. (2004). The electrical neutrality of atoms and of bulk matter. *Metrologia*, *41*, S125–S135.
6. Sciama, D. (1953). On the origin of inertia. *MNRAS*, *113*, 34–42.
7. Unnikrishnan, C. S. (2005). On Einstein's resolution of the twin clock paradox. *Current Science*, *89*, 2009–2015.
8. Unnikrishnan, C. S. (2015). Dynamics, relativity and the equivalence Principle in the 'once-given' universe. In E. Augé, J. Dumarchez, & J. Trân Thanh Vân (Eds.), *Proceedings of the 50th Moriond meeting on gravitation* (pp. 177–180), La Thuil.

9. Unnikrishnan, C. S. (2015). General Theory of Relativity and the Universe between beauty and truth. Research Gate Open Access.
10. Neurath, O. (1983). Physicalism: The philosophy of the Vienna circle. In R.S. Cohen & M. Neurath (Eds.), *Philosophical papers 1913–1945*. Dordrecht: D. Reidel Publishing Company, 49.
11. Unnikrishnan, C. S. (2014). True dynamical tests of the equivalence principle. *International Journal of Modern Physics (Conference series)*, *30*, 1460267.
12. Unnikrishnan, C. S. Cosmic relativity: Dynamics and relativity in the 'once-given' universe, Monograph in preparation.
13. Unnikrishnan, C. S. (2011). Precision measurement of the one-way speed of light and implications to the theory of relativity. In M.C. Duffy, V.O. Gladyshev, A. Morozov, & P. Rowlands (Eds.), *Physical interpretations of relativity theory: Proceedings, London 2006* (pp. 505–511). Bauman Moscow State Technical University.

Chapter 15
An Anomaly in Space and Time and the Origin of Dynamics

Joan A. Vaccaro

Abstract The Hamiltonian defines the dynamical properties of the universe. Evidence from particle physics shows that there is a different version of the Hamiltonian for each direction of time. As there is no physical basis for the universe to be asymmetric in time, both versions must operate equally. However, conventional physical theories accommodate only one version of the Hamiltonian and one direction of time. This represents an unexplained anomaly in conventional physics and calls for a reworking of the concepts of time and space. Here I explain how the anomaly can be resolved by allowing dynamics to emerge phenomenologically. The resolution offers a picture of time and space that lies below our everyday experience, and one in which their differences are epiphenomenal rather than elemental.

15.1 Introduction

One of the earliest attempts to describe the nature of time and space comes from Parmenides (~500 BCE) [1]. He and his pupil Zeno argued for monism—that there was only a single reality—and so to them time was a complete whole without division. They argued that this gives less absurdities than the opposing pluralistic view where multiple realities catered for different modes of being. Zeno's well-known paradoxes were attempts to illustrate the absurdities that would follow from pluralism. However, a new way of looking at nature, based on empirical observations and mathematical calculations, emerged in the European Renaissance period. The perceived difficulties associated with Zeno's paradoxes were largely swept aside with the development of calculus. Building on the work of Copernicus and Galileo, Newton proposed that an absolute time flows uniformly throughout an absolute space [2]. Newton's framework represents a kind of pluralism where each moment in time

---

J.A. Vaccaro (✉)
School of Natural Sciences, Centre for Quantum Dynamics,
Griffith University, Brisbane, Australia
e-mail: J.A.Vaccaro@griffith.edu.au

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_15

represents a separate reality. Then, about a century ago, James [3] and McTaggart [4] resurrected a monist view of time in the form of the block universe, where time is seen to be one structure without a present, past or future. The block universe represents, for the most part, the orthodoxy among physicists in modern times [5, 6]. Nevertheless, a new kind of pluralism will reemerge later in this chapter.

The impetus for abandoning Newton's framework of space and time in physics came from its failure to account for the propagation of light in the Michelson-Morley experiments of 1887. This anomaly led Einstein in 1905 to propose a new framework for space and time in his special theory of relativity. How we think of time and space today in terms of a background geometry is moulded by Einstein's relativistic *spacetime*—an amalgamation of space and time into a single entity. An interval of either time or space for one reference frame can be an interval that extends over both time and space in another reference frame.[1] In this sense, one can say that space and time appear in special relativity on the same footing.

Yet time and space are quite different in other respects. For example, matter can be localised in a region of space but not in an interval of time. That is, a lump of matter—such as an atom, a coffee cup or even a galaxy—can exist in one region of space and no other, but conservation of mass[2] forbids matter from existing at one time interval and no other. To exist only at one time interval, for one second after midday say, would mean the matter not existing before midday, existing only during the second after midday and vanishing at the end of the second. We avoid this drastic violation of mass conservation in conventional physics by insisting that matter follows an *equation of motion* that translates it over all times. The upshot is that matter is presumed to undergo continuous translation over time (as time evolution) but there is no corresponding presumption about the matter undergoing translations over space.

Moreover, the presumed continuous translation over time occurs in a preferred direction which is described by various arrows of time. The first arrow to be named formally is the thermodynamic arrow [7] which points in the direction of increasing entropy. Other arrows include the cosmological arrow, which points away from the big bang, and the radiation arrow, which points in the direction of emission of waves [6]. In contrast, space is isotropic.

There is another, quite subtle, difference between time and space that has largely escaped attention until recently: translations in time and space have very different discrete symmetry properties [8–10]. The discrete symmetries here represent an invariance to the operations of charge conjugation (C), parity inversion (P) and time reversal (T). Although nature respects these symmetries in most situations, exceptions have been discovered in the last 60 years. The exceptions are observed as violations of particular combinations of the C, P and T symmetries in certain particle decays [11–16]. The violations are independent of position in space, and so they occur over *translations in time* (i.e. as a decay) and *not translations in space*.

---

[1] Appendix 1 discusses this in more detail.

[2] The terms mass and matter here can be taken to mean relativistic energy.

The fact that time and space have these differences does not, in itself, constitute a problem. On the surface, the differences don't appear to be pointing to a glaring anomaly that requires a reworking of the foundations of physics like the results of the Michelson-Morley experiment did. Yet there is an anomaly, one that has been around for so long that it risks being overlooked because of its familiarity. It is to do with the fact that there is no cause for the block universe to be anything other than symmetrical in time. In other words, there is no physical basis for one direction of time to be singled out [6]. This invites the question, so where is the other direction of time? It may be tempting to speculate that another part of the time axis may carry arrows pointing in the opposite direction. But this will not do, given the impact the discoveries of the violation of the discrete symmetries have for the Hamiltonian. The Hamiltonian is a mathematical object that defines the dynamics. The violation of time reversal symmetry, called T violation for short, implies that there is a different version of the Hamiltonian for each direction of time, yet we observe only one version in our universe and, not surprisingly, only the observed version of the Hamiltonian appears in conventional theories of physics. Where is the other direction of time and its concomitant version of the Hamiltonian? The fact that there is no answer in conventional physics constitutes a basic anomaly which calls for a fundamental shift in our thinking about time and space.

The purpose of this chapter is to expose the anomaly and then review my recent proposal [10] to resolve it through restructuring the way time and space appear in physical theory. The anomaly is articulated more precisely in Sect. 15.2 and then Sect. 15.3 prepares for the required restructuring in terms of a goal and three basic principles. Following that, the principles are applied to non-relativistic quantum mechanics in Sect. 15.4 and the chapter ends with a discussion in Sect. 15.5. Additional background material and specific details are left to the appendices: special relativity in Appendix 1, generators and translations in time and space in Appendix 2, and quantum virtual paths in Appendix 3. Full details of my proposed resolution can be found in Ref. [10].

15.2 An Anomaly: Missing Direction of Time and Its Hamiltonian

To expose the anomaly we must first lay to rest a common misconception that the arrows of time are melded in some way into the concept of time itself. In particular, if the only thing that distinguishes the two directions of time is an increase of entropy in one direction, then perhaps one could be forgiven for succumbing to a conceptual shorthand and regarding the entropy increase as somehow causing the direction of time. But, in truth, the arrows are only *evidence* that time has a direction and there is simply no basis for claiming them as the *cause* of that direction. An analogy will help make the distinction between evidence and cause clearer. Imagine that the leaves falling from a tree are blown by a steady wind to land preferentially on the downwind side of the tree. The pattern of leaves on the ground would then provide *evidence* of the direction that the wind is blowing, but there would be no basis for claiming that the leaves *cause* the wind to have any particular direction. The same situation occurs with the direction of time: the arrows are patterns that provide evidence of the direction of the translations over time, but those patterns do not cause the translations themselves nor do they cause the translations to be in a particular direction.

Having laid bare the evidential nature of the arrows, we now examine the thermodynamic arrow in particular. This arrow, like all the arrows, is phenomenological in origin. It arises because thermodynamics was developed to be in accord with nature and thus it was intentionally structured to have an increasing entropy in the direction of time we refer to as "forwards" or the "future". However, as Loschmidt pointed out long ago, thermodynamics is consistent with time-symmetric physical laws, such as Newton's laws of motion, and so any prediction of an increase in entropy in one direction of time is, necessarily, a prediction of an increase in the opposite direction of time. To ignore this and claim that the thermodynamic arrow, or any of the arrows, explains the direction of time, is to commit what Price calls a double standard fallacy [6]. Avoiding the fallacy leaves us with the problem of a missing direction of time.

Its resolution calls for a time-symmetric model of nature that accounts for both directions of time—a model in which there are reasons for arrows to point in both directions. There have been admirable attempts along these lines by Carroll, Barbour and their co-workers [17, 18], but there is something fundamental missing from their analyses because they only consider time-symmetric physical laws. The only fundamental law that is not time symmetric is usually dismissed as having little to do with large-scale effects [2, 5, 6, 19]. It is associated with the weak interaction, and its time asymmetry is observed as T violation in the decay of the K and B mesons [13–16]. However, despite being previously overlooked, I have shown that T violation is capable of producing large-scale physical effects [8–10]. Moreover, the experimentally observed T violation implies that the universe is described by two versions of the Hamiltonian, one for each direction of time. The double-headed arrows of Carroll, Barbour and co-workers do not account for this crucial fact.

The problem, then, is not only that there is a missing direction of time, but that the associated version of the Hamiltonian is missing along with it. The anomaly is the rather glaring absence of both directions of time and both versions of the Hamiltonian in conventional physical theories; it can be stated formally as follows.

Anomaly *There is no basis for nature to be asymmetric in time. Experiments in particle physics indicate that there are two versions of the Hamiltonian, one for each direction of time. A time-symmetric theory of nature must give an equal account of both directions of time and both versions of the Hamiltonian. Conventional theories fail in this regard because they can accommodate only one version of the Hamiltonian and one direction of time.*

The anomaly calls for a restructuring of the concepts of time and space in physics.

15.3 The Goal and Basic Principles

The goal of the restructuring might appear to be to simply find a time-symmetric description that includes both versions of the Hamiltonians. However, aiming the goal directly at the anomaly like this misses an opportunity for rebuilding from a deeper level. For example, if Einstein had been satisfied with a description of the propagation of light that was consistent with the Michelson-Morley experiment, he may have settled on some aether-dragging model. Instead, his search for an indirect, but deeper, solution led to his special theory of relativity, a natural consequence of which was the resolution of the light-propagation anomaly. In the same way, we need to take a step back from the anomaly itself. We have seen that the differences between time and space are related by the fact that they involve translations: conservation laws and the equation of motion represent translations over time, the direction of time describes an asymmetry in translations over time, and the violation of the discrete symmetries is observed for translations over time. Our understanding of the relationship between time and space would be advanced significantly if all differences could be shown to have a common origin. The least understood among the differences is the C, P and T symmetry violations. Although the violations are generally considered to represent profound properties of nature, they don't play any significant role in conventional physics. Indeed, they stand out as having been overlooked. To address this situation, we undertake the more ambitious goal as follows:

Goal *To treat time and space on an equal footing at a fundamental level, and to allow their familiar differences to emerge phenomenologically from the discrete symmetry violations.*

If the violations deliver the differences between space and time then we will have found a theory that incorporates both versions of the Hamiltonian in a way that gives rise to the familiar direction of time. The anomaly would then be resolved as a natural consequence of the goal.

Having settled on the goal, we now turn to the basic principles needed to achieve it. When the C, P and T symmetries are obeyed we want matter to be localisable both in time and space. This will require a formalism in which conservation laws do not apply and an equation of motion is not defined—this marks a serious departure from conventional physics. When the violation of the symmetries are introduced into the formalism, an effective equation of motion and conservation laws need to appear phenomenologically as a consequence—only then will it be in agreement with conventional physics. The symmetry violations clearly need to play a significant role in the formalism. The violations manifest as changes due to the C, P and T operations,[3] and so their impact would tend to be greater in a formalism in which the operations are more numerous. The P and T operations, in particular, are associated with reversing directions in space and time, respectively. It is clear from this that

---

[3] If the C, P and T operations do not change the system then the symmetries are obeyed. Violations represent the converse situation where changes result from the operations.

we need a formalism comprising paths in time and space which suffer innumerably-many reversals. A stochastic Wiener process involves paths of this kind in space. Feynman's path integral method [20] also involves similar kinds of paths over configuration space.

The important point about Feynman's method is that it underpins analytical mechanics in the limit that Planck's constant, $\hbar$, tends to zero. Indeed, his method shows that Hamilton's principle of least action arises as a consequence of destructive interference over all possible paths in configuration space between the initial and final points. But it stops short of considering paths that zigzag over time of the kind we need to consider here and, as a consequence, it stops short of considering the impact of the C, P and T symmetry violations that are the focus here. Nonetheless, it does demonstrate the importance of quantum path integrals for describing the universe on a large scale.

Although the paths need to comprise innumerably-many reversals, there are reasons to believe that there are physical limitations to the resolution of intervals in space and time [21]. For example, the position of an object can be determined by observing the photons it scatters, but the accuracy of the result cannot be better than the Planck length $L_P = 1.6 \times 10^{-35}$ m [22]. Correspondingly, the timing of the scattering events cannot be determined any better than the Planck time $L_T = L_P/c = 5.4 \times 10^{-44}$ s where c is the speed of light. We will assume that fundamental resolution limits of this kind exist without specifying their value. It would be physically impossible to resolve the structure of paths with step sizes smaller than the resolution limit, and so we need to treat such paths as having equal physical status.

With these ideas in mind we formulate three principles on which to base the development of the new formalism:

Principle 1 *A quantum state is represented as a superposition of paths, each containing many reversals. We call these "quantum virtual paths".*

Principle 2 *There is a lower limit to the resolution of intervals in space and time. Quantum virtual paths with step sizes smaller that this limit have an equal physical status.*

Principle 3 *States have the same construction in both time and space. Any differences between space and time, such as dynamics and conservation laws, emerge phenomenologically as a result of the violation of discrete symmetries C, P and T.*

15.4 Applying the Principles

We shall apply the three basic principles to represent the quantum state of an object.[4] The object represents the only matter in space and time and it could be an atom, planet or galaxy. Its details are not important. We will refer to it as the "galaxy" in

---

[4] We only use the static representation of a state from non-relativistic quantum mechanics. We do not apply an equation of motion nor do we impose conservation laws.

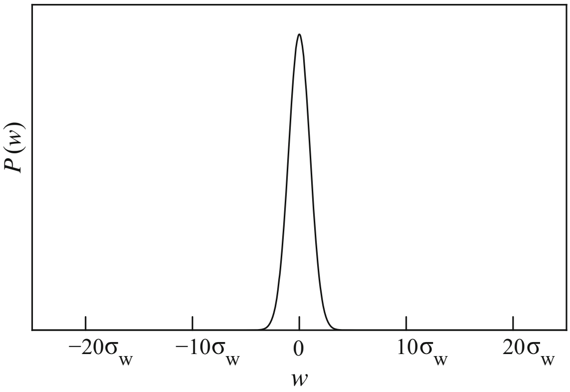


Fig. 15.1 *Bell-shaped* probability distribution $P(w)$ representing an object localised in the vicinity of the origin of the w coordinate. The standard deviation of the distribution is σ_{w}

the following. The first task is to develop the formalism in general terms without referring specifically to time or space. For that let w be a generic coordinate which will later be set to be either time or space. We want the galaxy to be localised with respect to w such that the spread in w is finite. The most general probability distribution with a finite spread has a bell-shape like $P(w)$ illustrated in Fig. 15.1.

15.4.1 Application of Principle 1

An equivalent representation is given by imagining that the galaxy takes a path that starts at the origin $w = 0$ and randomly steps back and forth along the w coordinate a number of times. Let there be N steps in the path and let the magnitude of each step be δw. For the final location of the galaxy to any value of w, the step size δw needs to be infinitesimally small and N needs to be correspondingly large. By setting

$$\delta w = \frac{\sqrt{2}\sigma_{\mathrm{w}}}{\sqrt{N}} \tag{15.1}$$

and choosing a suitably-large value of N we can make the step size, δw, as small as we like, and the maximum length of any path, $N\delta w$, correspondingly as large as we like, while keeping the standard deviation in the possible final locations fixed at σ_{w}. It needs to be emphasised that even though temporal references such as “starts”, “steps” and “final” are used here, the paths do not represent actual movement over a time interval. Rather they represent the galaxy executing a sequence of *virtual displacements* along w without any reference to time at all. That is, the galaxy is considered to be simply displaced from $w = 0$ to the point represented by the end of the random path. Virtual displacements arise in analytical mechanics when discussing constraints on motion [23]; here the accumulation of many random virtual displacements give the possible values of w.

For the location of the galaxy to be described by the smooth bell-shaped distribution $P(w)$ we need not just one path and its end point, but infinitely many. We don't know which end point describes the location of the galaxy and so we have to allow for the possibility that it could be the end point of any one of many paths. Technically, this means we represent the location of the galaxy by a *superposition* of the end points of all the paths. The superposition is called a "quantum virtual path", where quantum refers to the fact that it is a quantum superposition [10].

One can imagine a quantum virtual path for a specific value of N, say $N = 600$, as the sum of the end points of all the paths illustrated in Fig. 15.2. The step size δw for each zigzag path in the figure is given by Eq. (15.1) for some fixed value of the standard deviation σ_{w}. Another quantum virtual path can be constructed for $N = 601$ in a similar way for a correspondingly smaller step size δw. Imagine that this has been done for every positive integer value of N. As N increases in this imagined process, the step size δw reduces and the quantum virtual path represents an ever finer description of the state of the galaxy, eventually tending to the bell-shaped dashed curve shown in the figure. Each quantum virtual path so constructed represents a possible state of the galaxy in terms of its location along the w coordinate.

Each step of δw is produced using a particular operation called a "generator" of the translation. In particular, $\hat{W}_F$ is the generator for translations that increase the value of w and $\hat{W}_B$ is the generator for ones that decrease its value, as illustrated in the inset of Fig. 15.2. If the generators are invariant to reversals of direction then they are equivalent, i.e. $\hat{W}_F = \hat{W}_B$. More will be said about this later. A technical review of generators and translations is given in Appendix 2 and a brief discussion of how a quantum virtual path is related to the bell-shaped distribution $P(w)$ can be found in Appendix 3.

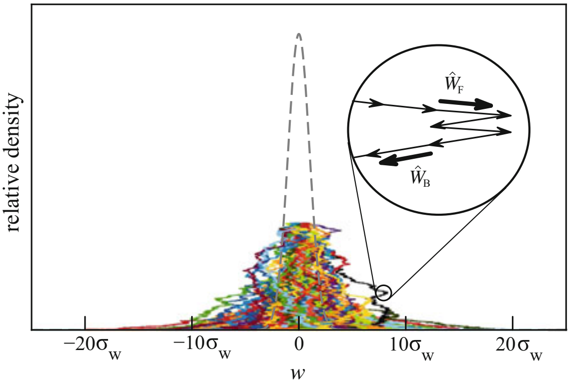


Fig. 15.2 Conceptual sketch of a quantum virtual path. Each curve represents a random path of N steps back and forth along the w coordinate starting at $w = 0$ and ending at a random value of w. The curves are displaced vertically to represent the relative density of paths. The *inset* illustrates the actions of the generators, $\hat{W}_F$ and $\hat{W}_B$, of translations in the $+w$ and $-w$ directions, respectively

15.4.2 Application of Principle 2

As the value of N increases, the step size δw from Eq. (15.1) becomes smaller. At some point δw will be smaller than the resolution limit δw_{min} for the w coordinate. All quantum virtual paths with a step size δw smaller than δw_{min} will give descriptions of equal status according to Principle 2. For convenience, we shall collect the equivalent quantum virtual paths in a set called $\mathbf{G}$. Each quantum virtual path in this set equally represents the state of the galaxy in terms of its location along the w coordinate. There are an infinite number of such quantum virtual paths in the set $\mathbf{G}$.

15.4.3 Application of Principle 3

We now discuss space and time explicitly. First consider the spatial case which, for brevity, we limit to just the x dimension. In this case the generic coordinate w is replaced with x and the generator of translations is replaced with $\hat{p}_x$, the component of momentum along the x axis. There is only one generator for translations in both directions of the x axis and so $\hat{W}_F = \hat{W}_B = \hat{p}_x$ here. Further technical details are given in Appendix 2. Figure 15.2 with w replaced by x illustrates a quantum virtual path over the x axis. Collecting the quantum virtual paths with a step size smaller than some minimum resolution limit yields the set of states of equal status which we will call $\boldsymbol{\Psi}$. All the quantum virtual paths in $\boldsymbol{\Psi}$ are physically indistinguishable from the bell-shaped distribution $P(x)$ represented in Fig. 15.1 with w replaced with x.

Next, we repeat the same exercise for time. In this case the coordinate is $w = t$ and, in general, there are two generators of translations given by the two versions of the Hamiltonian, i.e. $\hat{W}_F = \hat{H}_F$ and $\hat{W}_B = \hat{H}_B$ corresponding to the "forwards" and "backwards" directions of time, respectively. Technical details regarding these generators are given in Appendix 2. As with the spatial case, Fig. 15.2 with w replaced by t illustrates a quantum virtual path over the t axis, and collecting the quantum virtual paths which have a step size smaller than some minimum resolution limit yields the set of states of equal status which we will call $\boldsymbol{\Upsilon}$.

In a universe where the T symmetry holds, there is only one version of the Hamiltonian and so $\hat{H}_B = \hat{H}_F = \hat{H}$. In this case the galaxy is localised in time within a duration of the order of σ_{t} of the origin and all the states in $\boldsymbol{\Upsilon}$ are physically indistinguishable from the bell-shaped distribution $P(t)$ represented in Fig. 15.1 with w replaced with t. The galaxy only exists in time for a relatively short duration at the origin $t = 0$ and does not exist before or after this time. It can be imagined to come into existence momentarily and then promptly vanish. Clearly, in this case, the galaxy has the same representation in time as in space—it is localised in both—and the formalism places time and space on the same footing in this respect. This is far removed from conventional quantum mechanics as there is no equation of motion and the mass of the galaxy is not conserved.

The converse case, where T symmetry is violated, is defined by $\hat{H}_B \neq \hat{H}_F$. The key point here is that multiple paths that zigzag in different ways from the origin to the same end point can interfere. The interference can be compared to the way waves travelling on the surface of water behave; if the trough of one wave occurs at the same point as the crest of another, the two waves will tend to cancel each other in a process called destructive interference, whereas if two troughs or two crests meet they tend to reinforce each other as deeper troughs or higher crests, respectively, in a process called constructive interference. In a similar way, multiple paths that end at the same point on the time axis interfere either destructively or constructively. The result is that instead of the probability distribution having a maximum at the origin, like the bell-shaped curve in Fig. 15.1, destructive interference reduces the probability to zero in this region. This is compensated by constructive interference that yields two symmetrically positioned bell-shaped peaks further from the origin as illustrated in Fig. 15.3. In other words, each quantum virtual path is now composed of two bell-shaped peaks that represent the galaxy existing at two different times, $+t$ and $-t$, say. This situation is like Schrödinger's cat that exists in a superposition of being both dead and alive simultaneously, except that here the galaxy is at two different times. As the value of N increases, the two peaks become further separated as shown in Fig. 15.3, and the galaxy shifts in time accordingly. Each quantum virtual path represents one of the states in the set Υ, and according to Principle 2, has equal physical status. In terms of Fig. 15.3, this means that each double-peaked curve equally represent the position of the galaxy in time.

The presence of T violation clearly has a dramatic affect on the temporal description of the galaxy. For example, consider the question, where in time is the galaxy likely to be found? Without T violation, the unequivocal answer is only near the origin in accordance with Fig. 15.1, whereas with T violation, the answer implied by Fig. 15.3 would be at *any* time t.

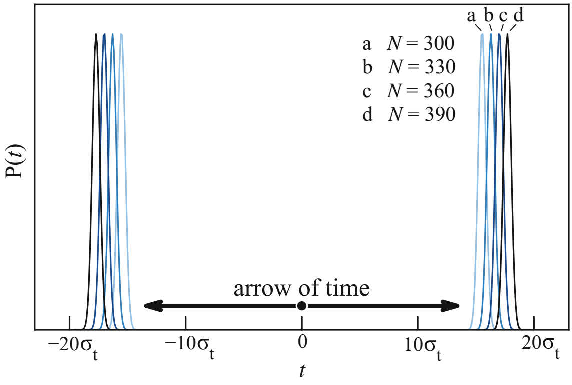


Fig. 15.3 The probability distribution $P(t)$ for various values of N in the case of T violation. Each curve has two *bell-shaped peaks* which move further apart as N increases

15.4.4 The Origin of Dynamics

We now focus on the T violation case. According to Principle 2, all states in the set Υ have an equal status in representing the galaxy in time. For any given value of time, t, there is a corresponding state in Υ that represents the galaxy being in a superposition of the times $+t$ and $-t$.[5] This implies that the galaxy exists at any time we wish to consider, and so its *mass is conserved*. This conservation law has not been imposed on the formalism, as it would need to be in conventional theories, but rather it is phenomenology arising from T violation.

The corresponding equation of motion is found as follows. The two peaks $+t$ and $-t$ in each curve in Fig. 15.3 represent time-reversed versions of the galaxy. An observer in the galaxy would not be able to distinguish between them and so we need only consider one, at $+t$ say. If the observer makes observations with a resolution in time that is broader than the width of the peak, the peak will appear to be instantaneous and a set of them will appear to form a continuous sequence. Under these circumstances, the observer would find evidence of an *equation of motion* that is consistent with the Schrödinger equation of conventional quantum mechanics. This equation has not been imposed on the formalism but rather it arises as phenomenology associated with T violation. This suggests that the origin of dynamics lies in T violation.

The remaining distinctive feature of time to consider is its direction, and the states in Υ have a time ordering in the following sense. According to the meaning of time evolution defined in Appendix 2, the peak labelled “d” Fig. 15.3 represents a state that has *evolved in time* from the state represented by the peak labelled “c”, and that state has evolved from the state represented by “b”, which has evolved from the state represented by “a”, but the converse is not true. This means that there is an arrow of time pointing in the direction of $+t$. The same argument applies to the time reversed states in regards to the $-t$ direction and so the *arrow is double headed*, like those of Carroll and Barbour and co-workers [17, 18]. The important point here is that both versions of the Hamiltonian, $\hat{H}_F$ and $\hat{H}_B$, are included in the formalism.

We have now achieved our goal: we treated time and space on an equal footing and found their familiar differences to emerge phenomenologically from T violation.

15.5 Discussion

We began by identifying a fundamental anomaly in physics, viz. conventional theories fail to give a time symmetric description that accounts equally for both versions of the Hamiltonian and both directions of time. We have presented a new formalism

---

[5] In principle, the time t could be chosen to be the current age of the universe, 13.8 billion years. There is a state in Υ that represents the galaxy being in a superposition of the times 13.8 and -13.8 billion years.

for quantum mechanics that resolves this anomaly. The new formalism is based on three principles that allow quantum states in time and space to be treated on an equal footing in terms of quantum virtual paths. The distinctive features associated with time, i.e. conservation laws, equation of motion and the direction of time, are not imposed on the formalism but rather emerge phenomenologically as a result of T violation. These key differences between time and space follow from the fact that the generators of translations in space and time, the momentum operator and the Hamiltonian, respectively, have different symmetry properties: the momentum operator is invariant to the C, P and T symmetry operations whereas the Hamiltonian is not. Accounting for these differences gives the *origin of dynamics*.

The new formalism also refines the meaning of time. In conventional theories, the word "time" refers to both a coordinate of a space-time *background* as well as the parameter describing *dynamical evolution*. Both concepts are firmly entwined by conservation laws. For example, the conservation of mass implies that a massive object will persist over all times and, accordingly, it is represented on a space-time background as existing at each time. The dynamical evolution of the object becomes the path of the object on the space-time background. Here, however, the two concepts of time as a background coordinate and as a dynamical parameter are distinct. Time and space have an equal footing as a background on which quantum states are represented. The states, as quantum virtual paths, represent objects that are localised in time and space: each state in the sets $\boldsymbol{\Psi}$ and $\boldsymbol{\Upsilon}$ represents a relatively-narrow bell-shaped distribution or a sum of two relatively-narrow bell-shaped distributions. In particular, mass is not conserved and there is no equation of motion for any *individual state* in $\boldsymbol{\Upsilon}$ (as illustrated by Fig. 15.3)—time appears only as a background coordinate. In contrast, mass conservation, the equation of motion and the direction of time, are properties of the *whole set* $\boldsymbol{\Upsilon}$ where time appears as a dynamical parameter. In other words, time as a background coordinate and as a dynamical parameter apply to distinct constructs in the formalism.

It might appear unusual that a quantum formalism is being proposed to explain large scale structure of nature given that quantum effects are typically seen only in relatively small systems under controlled conditions. However, Feynman's path integral method has already demonstrated how quantum phenomena underpins Hamilton's least action principle in analytical mechanics [20] and thus large scale structure. In this regard, the new formalism should be considered as an extension of Feynman's method to encompass paths over time and the C, P and T symmetry violations and, thus, to apply to nature on a large scale as well.

Finally, the set of states $\boldsymbol{\Upsilon}$ for T violation represents the galaxy at an infinite sequence of times. Each state in $\boldsymbol{\Upsilon}$ may be viewed as representing a different reality. In this sense, the formalism resurrects a kind of pluralism. The monism-pluralism cycle for time turns once more.

Appendix 1

We briefly review here the Lorentz transformation in special relativity. A point in spacetime is referred to as an event; it is specified by four coordinates x, y, z, t with respect to a reference frame. The Lorentz transformation gives the relationship between the coordinates of two different inertial reference frames. In particular, consider two events that occur a distance of Δx apart along the x axis and separated by a duration of Δt in time in the x, y, z, t reference frame. In the x', y', z', t' reference frame that is moving a constant speed v along the x axis of the first, the distance and duration along the x' and t' axes between the events are given by

$$\begin{aligned}\Delta x' &= \gamma(\Delta x - v\Delta t)\\ \Delta t' &= \gamma(\Delta t - v\Delta x/c^2),\end{aligned}$$

respectively, where $\gamma = \sqrt{1 - v^2/c^2}$ and c is the speed of light. The important point here is that for $\Delta t = 0$ what is considered to be solely a spatial interval, Δx, in one reference frame becomes part of a temporal interval $\Delta t'$ as well as being part of a spatial interval $\Delta x'$ in the other reference frame. That is, space and time are interchangeable.

Appendix 2

Here, we briefly review translations and their generators. Recall that the Taylor expansion of a function $f(x)$,

$$f(x + a) = f(x) + a\frac{\mathrm{d}}{\mathrm{d}x}f(x) + \frac{a^2}{2!}\frac{\mathrm{d}^2}{\mathrm{d}x^2}f(x) + \frac{a^3}{3!}\frac{\mathrm{d}^3}{\mathrm{d}x^3}f(x) + \dots,$$

can be written compactly in exponential form as

$$f(x + a) = \mathrm{e}^{-\mathrm{i}a(\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}x})}f(x).$$

When written in this form the differential operator $\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}x}$ is said to be the generator of translations in x. The generator of spatial translations along the x axis is $\hat{p}_x$, the operator representing the x component of momentum. We need only consider one dimension of space for our purposes here. Thus we write

$$|x + a\rangle_{\mathrm{x}} = \mathrm{e}^{-\mathrm{i}a\hat{p}_x}|x\rangle_{\mathrm{x}} \tag{15.2}$$

where $|x\rangle_{\mathrm{x}}$ represents a state vector for position x and, for convenience, we assume units in which $\hbar = 1$. Similarly, the generator of translations in time t is the Hamiltonian operator $\hat{H}$ and so

$$|\psi(t+a)\rangle_{\rm t} = {\rm e}^{-{\rm i}a\hat{H}}|\psi(t)\rangle_{\rm t} \tag{15.3}$$

where $|\psi(t)\rangle_{\rm t}$ represents a state at time t and evolving in the $+t$ time direction.

The symmetry operations relevant to these translations are the parity inversion $\hat{P}$ and the time reversal $\hat{T}$ operations[6] defined by Wigner [24]. Parity inversion interchanges x with $-x$, y with $-y$ and z with $-z$ and time reversal interchanges t with $-t$. For example, $\hat{P}|x\rangle_{\rm x} = |-x\rangle_{\rm x}$ and $\hat{T}|t\rangle_{\rm t} = |-t\rangle_{\rm t}$. The reverse of the translation in Eq. (15.2) can be written as

$$|x-a\rangle_{\rm x} = \hat{P}|-x+a\rangle_{\rm x} = \hat{P}{\rm e}^{-{\rm i}a\hat{p}_x}|-x\rangle_{\rm x} = \hat{P}{\rm e}^{-{\rm i}a\hat{p}_x}\hat{P}^{-1}|x\rangle_{\rm x}.$$

As $\hat{P}\hat{p}_x\hat{P}^{-1} = -\hat{p}_x$ we get

$$|x-a\rangle_{\rm x} = {\rm e}^{{\rm i}a\hat{p}_x}|x\rangle_{\rm x}$$

as expected directly from Eq. (15.2). This shows that the generator of translations in either direction of the x axis is the same. The reverse of the translation in Eq. (15.3) is somewhat different, however. Consider

$$\begin{aligned}|\phi(t-a)\rangle_{\rm t} &= \hat{T}|\phi(-t+a)\rangle_{\rm t} = \hat{T}{\rm e}^{-{\rm i}a\hat{H}}|\phi(-t)\rangle_{\rm t} = \hat{T}{\rm e}^{-{\rm i}a\hat{H}}\hat{T}^{-1}|\phi(t)\rangle_{\rm t}\\ &= {\rm e}^{{\rm i}a\hat{T}\hat{H}\hat{T}^{-1}}|\phi(t)\rangle_{\rm t}\end{aligned} \tag{15.4}$$

where $|\phi(t)\rangle_{\rm t}$ represents a state that evolves in the $-t$ direction and we have made use of the antiunitary nature of the time reversal operator, i.e. $\hat{T}{\rm i}\hat{T}^{-1} = -{\rm i}$, in the last line [24]. In general $\hat{T}\hat{H}\hat{T}^{-1} \neq \hat{H}$ and so we set, for convenience,

$$\begin{aligned}\hat{H}_B &= \hat{T}\hat{H}\hat{T}^{-1}\\ \hat{H}_F &= \hat{H}\end{aligned}$$

where the subscripts F and B refer to the "forwards" and "backwards" direction of time corresponding to the $+t$ and $-t$ time directions, respectively. If T symmetry is obeyed then

$$\hat{H}_B = \hat{H}_F = \hat{H}, \qquad \text{(T symmetry)}$$

and so there is a unique version of the Hamiltonian, whereas for T violation there is a different version of the Hamiltonian for each direction of time,

$$\hat{H}_B \neq \hat{H}_F. \qquad \text{(T violation)}$$

---

[6] We use the operator symbols $\hat{P}$ and $\hat{T}$ to represent the operations and the letters P and T to represent the corresponding symmetries. Thus, if the system is invariant to the $\hat{P}$ operation it obeys the P symmetry.

In general, we write Eqs. (15.3) and (15.4) as

$$|\psi(t+a)\rangle_{\mathrm{t}} = \mathrm{e}^{-\mathrm{i}a\hat{H}_F}|\psi(t)\rangle_{\mathrm{t}}$$
$$|\phi(t-a)\rangle_{\mathrm{t}} = \mathrm{e}^{\mathrm{i}a\hat{H}_B}|\phi(t)\rangle_{\mathrm{t}}.$$

The key point to be made here is that the generator of translations in space, $\hat{p}_x$, is invariant (up to a sign change) under any of the C, P and T operations. In contrast, the generator of translations in time, $\hat{H}$, is not invariant to the C, P and T operations, in general. This underlies the statement in the Introduction that the symmetry violations occur over translations in time and not translations in space.

In the case of T violation we need to take care with using the correct Hamiltonian associated with each direction of time. In particular, we need to apply the following principle:

Principle 4 *Physical time evolution is represented by the operators* $\mathrm{e}^{-\mathrm{i}a\hat{H}_F}$ *and* $\mathrm{e}^{\mathrm{i}a\hat{H}_B}$ *for the forward (+t) and backward (−t) directions of time, respectively. The operations* $\mathrm{e}^{\mathrm{i}a\hat{H}_F}$ *and* $\mathrm{e}^{-\mathrm{i}a\hat{H}_B}$ *represent the mathematical inverse operation of "unwinding" or "backtracking" the evolution produced by* $\mathrm{e}^{-\mathrm{i}a\hat{H}_F}$ *and* $\mathrm{e}^{\mathrm{i}a\hat{H}_B}$*, respectively.*

For example, $\mathrm{e}^{\mathrm{i}a\hat{H}_F}|\psi(t+a)\rangle_{\mathrm{t}} = |\psi(t)\rangle_{\mathrm{t}}$ represents unwinding the time evolution $\mathrm{e}^{-\mathrm{i}a\hat{H}_F}|\psi(t)\rangle_{\mathrm{t}} = |\psi(t+a)\rangle_{\mathrm{t}}$ whereas $\mathrm{e}^{\mathrm{i}a\hat{H}_B}|\psi(t+a)\rangle_{\mathrm{t}}$, which is not equal to $|\psi(t)\rangle_{\mathrm{t}}$ in general, represents time evolution of $|\psi(t+a)\rangle_{\mathrm{t}}$ in the $-t$ direction. More details are given in Ref. [10].

Appendix 3

In this Appendix we briefly discuss the mathematical construction of quantum virtual paths for the generic coordinate w. Let the generators of translations be given by $\hat{W}_F$ and $\hat{W}_B$ for the $+w$ and $-w$ directions, respectively. A quantum virtual path of the kind we want is given by [10]

$$|g\rangle_N \propto \frac{1}{2^N}\left(\mathrm{e}^{\mathrm{i}\hat{W}_B\delta w} + \mathrm{e}^{-\mathrm{i}\hat{W}_F\delta w}\right)^N |0\rangle_{\mathrm{w}} \tag{15.5}$$

where δw is given by Eq. (15.1) and represents an increment in w and $|w\rangle_{\mathrm{w}}$ represents a state for which w is well-defined.[7] Expanding the power on the right side gives 2^N terms each with N factors. Each term represents a path comprising N steps of δw over the w coordinate. For example, a term of the form

$$\ldots \mathrm{e}^{\mathrm{i}\hat{W}_B\delta w}\mathrm{e}^{-\mathrm{i}\hat{W}_F\delta w}\mathrm{e}^{\mathrm{i}\hat{W}_B\delta w}\mathrm{e}^{\mathrm{i}\hat{W}_B\delta w}\mathrm{e}^{-\mathrm{i}\hat{W}_F\delta w}|0\rangle_{\mathrm{w}}$$

---

[7] If w represents a spatial coordinate then $|w\rangle_{\mathrm{w}}$ would be a corresponding spatial eigenstate. For the case where w represents the time coordinate, however, we only need $|w\rangle_{\mathrm{w}}$ to represent a well-defined time. More details can be found in Ref. [10].

represents the object starting at the origin $w = 0$ and then undergoing virtual displacements to $w = \delta w$, $w = 0$, $w = -\delta w$, $w = 0$ $w = -\delta w$ and so on.

It is relatively straightforward to show that the state $|g\rangle_N$ in Eq. (15.5) approaches a Gaussian state in the limit of large N when the discrete symmetry holds. To see this set $\hat{W}_B = \hat{W}_F = \hat{W}$ and use

$$\exp(-A^2/2) = \lim_{N\to\infty} \cos^N(A/\sqrt{N})$$

to find

$$\lim_{N\to\infty} |g\rangle_N \propto \mathrm{e}^{-\hat{W}^2\sigma_\mathrm{w}^2}|0\rangle_\mathrm{w},$$

and then, assuming that $\hat{W}$ has a complete orthonormal basis, rewrite this as the Fourier integral

$$\begin{aligned}\lim_{N\to\infty} |g\rangle_N &\propto \int \mathrm{d}w\, \mathrm{e}^{-w^2/4\sigma_\mathrm{w}^2}\mathrm{e}^{-\mathrm{i}\hat{W}w}|0\rangle_\mathrm{w}\\ &= \int g(w)|w\rangle_\mathrm{w}\mathrm{d}w\end{aligned}$$

where $g(w)$ is given by

$$g(w) = \mathrm{e}^{-w^2/4\sigma_\mathrm{w}^2}.$$

The square of this, $g^2(w)$, is proportional to the bell-shaped probability distribution $P(w)$ represented in Fig. 15.1.

References

1. Jowett, B. (1892). *The dialogues of Plato*. Oxford: Clarendon Press.
2. Novikov, I. D. (1998). *The river of time*. Translated from the Russian by V: Kisin. Cambridge UK: Cambridge University Press.
3. James, W. (1884). The Dilemma of determinism. *Unitarian Review*, September (1884). Reprinted in James, W. (1912). *The will to believe, and other essays in popular philosophy*. New York: Longmans, Green and Co.
4. McTaggart, J. E. (1908). The unreality of time. *Mind*, *17*, 457–474.
5. Zeh, H. D. (2007). *The physical basis for the direction of time*. Berlin: Springer.
6. Price, H. (1996). *Time's arrow and Archimedes' point*. New York: Oxford University Press.
7. Eddington, A. S. (1928). *The nature of the physical world*. New York: Macmillan.
8. Vaccaro, J. A. (2011). T violation and the unidirectionality of time. *Foundations of Physics*, *41*, 1569–1596. doi:10.1007/s10701-011-9568-x.
9. Vaccaro, J. A. (2015). T violation and the unidirectionality of time: Further details of the interference. *Foundations of Physics*, *45*, 691–706. doi:10.1007/s10701-015-9896-3.
10. Vaccaro, J. A. (2016). Quantum asymmetry between time and space. *Proceedings of Royal Society A*, *472*, 20150670. doi:10.1098/rspa.2015.0670.
11. Lee, T. D., & Yang, C. N. (1956). Question of parity conservation in weak interactions. *Physical Review*, *104*, 254–258.
12. Christenson, J. H., Cronin, J. W., Fitch, V. L., & Turlay, R. (1964). Evidence for the 2π decay of the ${K_2}^0$ Meson. *Physical Review Letters*, *13*, 138–140. doi:10.1103/PhysRevLett.13.138.
13. Angelopoulos, A., et al. (1998). First direct observation of time-reversal non-invariance in the neutral-kaon system. *Physics Letter B*, *444*, 43–51. doi:10.1016/S0370-2693(98)01356-2.
14. Sehgal, L. M., & van Leusen, J. (1999). Violation of time reversal invariance in the decays $K_L \to \pi^+\pi^-\gamma$ and $K_L \to \pi^+\pi^-e^+e^-$. *Physical Review Letters*, *83*, 4933–4936. doi:10.1103/PhysRevLett.83.4933.
15. Alvarez-Gaume, L., Kounnas, C., Lola, S., & Pavlopoulos, P. (1999). Violation of time-reversal invariance and CPLEAR measurements. *Physics Letter A*, *458*, 347–354. doi:10.1016/S0370-2693(99)00520-1.
16. Lees, J. P., et al. (2012). Observation of time-reversal violation in the B^0 Meson system. *Physical Review Letter*, *109*, 211801. doi:10.1103/PhysRevLett.109.211801.
17. Carroll, S. M., Chen, J. (2004). Spontaneous inflation and the origin of the arrow of time. http://arxiv.org/abs/hep-th/0410270
18. Barbour, J., Koslowski, T., & Mercati, F. (2014). Identification of a gravitational arrow of time. *Physical Review Letters*, *113*, 181101. doi:10.1103/PhysRevLett.113.181101.
19. Carroll, S. M. (2010). *From eternity to here: the quest for the ultimate theory of time*. London: Dutton.
20. Feynman, R. P. (1948). Space-time approach to non-relativistic quantum mechanics. *Review of Modern Physics*, *20*, 367–387. doi:10.1103/RevModPhys.20.367.
21. Amelino-Camelia, G. (1999). Gravity-wave interferometers as quantum-gravity detectors. *Nature*, *398*, 216–218. doi:10.1038/18377.
22. Mead, C. A. (1964). Possible connection between gravitation and fundamental length. *Physical Review*, *135*, B849–B862. doi:10.1103/PhysRev.135.B849.
23. Goldstein, H., Poole, C., & Safko, J. (2002). *Classical mechanics* (3rd ed.). San Francisco: Addison-Wesley.
24. Wigner, E. P. (1959). *Group theory and its application to the quantum mechanics of atomic spectra*. New York: Academic Press.

Chapter 16
Space, Time, and Adynamical Explanation in the Relational Blockworld

W.M. Stuckey, Michael Silberstein and Timothy McDevitt

The Relational Blockworld (RBW) interpretation [1–3] of quantum mechanics (QM) was spawned in response to quantum nonlocality, which is summed up nicely in this Wikipedia quote [4], "the phenomenon by which measurements made at a microscopic level contradict a collection of notions known as local realism that are regarded as intuitively true in classical mechanics." So, what is "local realism" and what does it have to do with space and time? Let us start with a so-called "delayed choice" experiment by Zeilinger.

The Zeilinger experiment [5] is shown in Fig. 16.1. The mystery of this experimental outcome does not require any familiarity with the formalism of quantum mechanics (QM). A laser beamed into a crystal produces a pair of entangled photons. "Entangled" simply means that a measurement outcome on one of the photons (at detector D1 here) is inextricably related to the outcome of a measurement on the other photon (at detector D2 here), as we will soon see. One of the photons proceeds down through the double slit to detector D2 while its entangled partner travels to the right through a lens to detector D1. The experimentalist can decide whether to locate D1 at a distance of one focal length (*f*, the "momentum measurement") or two focal lengths (2*f*, the "position measurement") behind the lens (Fig. 16.1). The photon detected at D2 then produces two different corresponding patterns, i.e., the upper right pattern (labeled "D1: Position measurement") and the lower right pattern (labeled "D1: Momentum measurement").

W.M. Stuckey (✉)
Department of Physics, Elizabethtown College, Elizabethtown, PA 17022, USA
e-mail: stuckeym@etown.edu

M. Silberstein
Department of Philosophy and Foundations of Physics, Committee for Philosophy and the Sciences, University of Maryland, College Park, MD 20742, USA
e-mail: msilbers@umd.edu

T. McDevitt
Department of Mathematical Sciences, Elizabethtown College, Elizabethtown, PA 17022, USA
e-mail: mcdevittt@etown.edu

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_16

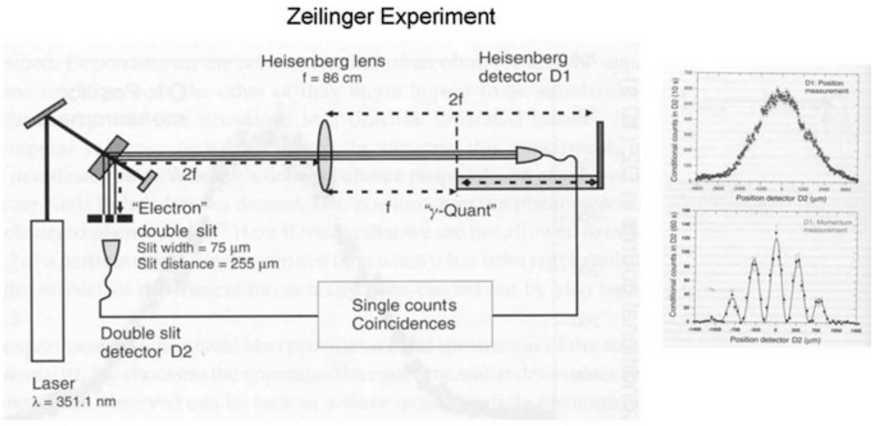


Fig. 16.1 Zeilinger's delayed choice experiment

That the photon at D2 acts like a particle (upper outcome) or a wave (lower outcome) depending on how its entangled partner is measured at D1 is interesting, but that's not the mystery. Notice that the distance from the photon source to D2 is much shorter than the distance from the photon source to D1. The two photons are emitted at the same time traveling at the same speed, so the first photon to reach a detector is at D2. But, the photon's behavior at D2 is supposedly determined by the experimentalist's choice of where to locate D1. How does the photon at D2 know where D1 will be placed? Doesn't the experimentalist have a real (delayed) choice that can be made after the photon has been detected at D2 and before the other photon has been detected at D1? If causal influences must proceed from past to future and the outcomes at D1 and D2 are causally related, then we have to conclude that the photon at D2 determines the experimentalist's choice of where to put D1? QM doesn't care about "choice," delayed or otherwise, it simply predicts the correlation in the outcomes without regard for which outcome occurs first [6].

We are dynamic creatures, our perceptions are formed in time-evolved fashion, so we're predisposed to think dynamically and, therefore, we want to understand/explain what we experience dynamically. However, physics may be telling us that despite our time-evolved perceptions, dynamical explanation is not fundamental. The Newtonian Schema [7] is the mathematical approach corresponding to "dynamical explanation." One inputs initial conditions (the initial state of the system in question) and a dynamical law tells you how that initial state is time-evolved to some final state. Sean Carroll sums it up nicely in this quote [8]:

> Let's talk about the actual way physics works, as we understand it. Ever since Newton, the paradigm for fundamental physics has been the same, and includes three pieces. First, there is the "space of states": basically, a list of all the possible configurations the universe could conceivably be in. Second, there is some particular state representing the universe at some time, typically taken to be the present. Third, there is some rule for saying how the universe evolves with time. You give me the universe now, the laws of physics say what it will

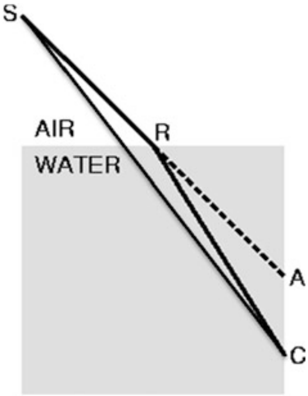


Fig. 16.2 A light ray emitted at S will be received at C after being refracted at R. This is the path of least time. The direct path S to C requires more time, so it is not the path taken. If the water is not present, the light ray would not be refracted at R and would instead proceed to A http://www-rohan.sdsu.edu/~aty/explain/optics/refr.html

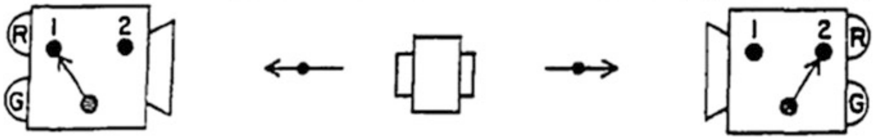
> become in the future. This way of thinking is just as true for quantum mechanics or general relativity or quantum field theory as it was for Newtonian mechanics or Maxwell's electrodynamics.

As Carroll points out, all the theories of physics can be written in the Newtonian Schema. In contrast, the Lagrangian Schema [9] says the behavior of objects isn't due to forces governing their behavior instant by instant, but rather their behavior from start to finish is the result of a minimal (extremal) action. The action S is the Lagrangian L integrated over the time interval in question.

$$S = \int L\,dt$$

where L is the difference between the system's kinetic and potential energies, so S is a spatiotemporally global quantity. The dynamical laws of the Newtonian Schema follow for each infinitesimal time dt in the Lagrangian Schema when you demand $\delta S = 0$, i.e., that S is extremal. If we are realists about the Lagrangian Schema approach, which people typically are not, we get the adynamical, spatiotemporally holistic view called the blockworld (BW). Let us imagine then that the Lagrangian Schema is not just a calculation device, but represents what we call an adynamical global constraint (AGC). It is this constraint that ultimately explains the strangeness of the quantum, as opposed to any dynamical explanation. Geroch sums up the BW perspective nicely with this quote [10]:

> There is no dynamics within space-time itself: nothing ever moves therein; nothing happens; nothing changes. In particular, one does not think of particles as moving through space-time, or as following along their world-lines. Rather, particles are just in space-time, once and for all, and the world-line represents, all at once, the complete life history of the particle.


The Hardy Experiment

Mermin, D.: Quantum mysteries refined. American Journal of Physics 62 (10), 880-887 (1994).

Fig. 16.3 Mermin's rendition of the Hardy experiment

An easy way to see the difference between the Newtonian Schema (dynamical view) and the Lagrangian Schema (BW view) is to consider the example of a light ray emitted from point S in air and arriving at point C in water (Fig. 16.2). The explanation per the Lagrangian Schema is that the path S to R to C is the path of least time. The direct path S to C (shown) takes longer, so it's not the path taken. The Newtonian Schema explanation is that the light ray emitted from S towards R proceeds without deviation, since nothing is interacting with it, until it hits the water at point R. The water then refracts the light towards point C. While both methods yield the same result, we tend to favor the Newtonian Schema explanation, since the Lagrangian Schema explanation sounds a bit like the light ray intended to go to C and calculated its path based on the presence of the water. According to the Newtonian Schema, the light ray that reached C was refracted at R, otherwise it would've gone to point A, i.e., it didn't adjust its emission direction with the intention of reaching C and it certainly doesn't 'know' it will encounter water when it's emitted. But the BW view doesn't attribute any intentionality or prescience to the light ray, rather it simply relegates dynamical explanation to secondary status a priori. If you want to know why something happens the way it does, the fundamental reason has to do with a spatiotemporally holistic characterization of the phenomenon under investigation. If the subsequent BW analysis permits it, a time-evolved story can be told in conjunction with the adynamical explanation. Otherwise, the fundamental BW perspective just does not permit a dynamical story, so the adynamical explanation just has to be accepted as the final answer. As Wharton says [9]:

> When examined critically, the Newtonian Schema Universe assumption is exactly the sort of anthropocentric argument that physicists usually shy away from. It's basically the assumption that the way we humans solve physics problems must be the way the universe actually operates.

No counterfactual definiteness is another such violation of our dynamical experience based in "local realism." Counterfactual definiteness is the claim that the particles being measured have a pre-existing value that is recorded faithfully, such that had you measured otherwise you would have gotten a different value. Let's look at a simple example of no counterfactual definiteness due to Hardy as explained by Mermin [11].

The Hardy experiment is shown and explained in Fig. 16.3. As with the Zeilinger experiment, you don't have to be familiar with the formalism of QM to appreciate this mysterious experimental outcome. In experimental trials (called "runs") where the detector settings are not the same, i.e., left detector set at 1 and right detector set at 2 ("12" for short) or the converse "21," you never see both detector 1 and detector 2 outcomes of green ("GG" for short). So, 12GG and 21GG never happen. The other two facts are that you occasionally see 22GG and you never see 11RR. Ok, so what's so mysterious about that? Well, let's try to explain the situation for the particles in a 22GG run. The particles don't 'know' how they will be measured in any given run, in fact the detector settings can be chosen an instant before the particles arrive at their respective detectors. And, assuming the measurements are space-like separated,[1] no information about the setting and outcome at one detector can reach the other detector before the particle is measured there, i.e., the measurements are said to be "local." Thus, the particles don't 'know' how they will be measured when they are emitted and they can't transmit that information to each other during the run. So, it seems the particles' possible outcomes with respect to each possible detector setting must be coordinated to prevent the proscribed outcomes 12GG, 21GG, and 11RR. Mermin calls the coordination of outcomes "instruction sets" and the fact that the particles have definite properties corresponding to possible measurements whether they are carried out or not is called "realism" or counterfactual definiteness. The combination of these assumptions is often called "local realism." Both assumptions seem reasonable, locality because the temporal order of space-like separated events is frame dependent according to special relativity, which would seem to rule out superluminal information transfer, and realism because the particles don't 'know' how they will be measured until the last moment and they have to avoid proscribed outcomes. So, let's continue with the analysis.

Since the result is 22GG we know the particles were both 1X2G at emission (X = R or G). We know particle 1 wasn't 1G2G, because had the detectors been set to 12 we would've gotten a 12GG outcome and that never happens. Likewise, particle 2 couldn't be 1G2G or a setting of 21 would've produced a 21GG outcome which never happens. Thus, both particles must've been 1R2G in order to give the 22GG outcome and rule out the possibility of a 12GG or a 21GG outcome. But, both particles being 1R2G can't be right either, because then a setting of 11 would've produced a 11RR outcome which never happens. So, what are the states of the particles when they are emitted in a run giving 22GG? We've exhausted all the logical possibilities and none of them work!

---

[1] Space-like separated events means information exchange between the events requires a faster-than-light (aka superluminal) signal. In special relativity, the order of space-like separated events A and B depends on your frame of reference. Suppose Bob sees event A happen before event B. Alice, who is in motion with respect to Bob, can then see event B happen before event A. This is different than time-like separated events (information exchange between the events can be done with a slower-than-light signal) where all observers agree on the temporal order, regardless of their relative motion.

There are only conjectural answers to that question in the foundations of physics community, but most seem to agree that there is no counterfactual definiteness, at least for some subset of properties. If there is no counterfactual definiteness, how does one account for the 22GG outcome of the Hardy experiment? Again, the mystery arises because we're looking for a Newtonian Schema (dynamical) explanation, i.e., a spatiotemporally local, time-evolved story in which the particles have definite properties prior to measurement and definite trajectories. In such a story, the state of each of the particles is time-evolved instant by instant from emission to detection, so we're simply asking for an instant-by-instant account of the particles in this experiment, as we tried unsuccessfully to provide above. But, if reality is fundamentally a BW whose observed patterns are governed by an AGC, e.g., extremal action, then the observed patterns in this experiment are just the outcomes and the Feynman path integral [12] of QM is the AGC. On this view, configuration space and the Schrödinger equation are just calculation devices. Furthermore, we can dispense with counterfactual definiteness or realism and jettison the very idea of particles as entities with autonomous existence and continuous trajectories from source to detector. We can't observe the particles between emission and detection, QM tells us and we find experimentally that additional detectors introduced between the source and detectors already in place will change the final outcomes. Thus, we're forbidden from acquiring such intermediate information without changing the experimental configuration, i.e., the BW pattern. The AGC (Feynman path integral) therefore doesn't have to say anything about counterfactual definiteness, it only has to account for the actual spatiotemporal experimental configuration and outcomes. Consequently, in the adynamical explanation, we're not concerned with counterfactual outcomes, i.e., BW patterns that aren't 'there'. The BW patterns are 4D involving the experimental configuration and process in its spatiotemporal entirety. Thus, in RBW the so-called quantum systems cannot be separated from the experimental set-up, but must include it (ontological contextuality). And, the quantum exchange of energy-momentum occurs via direct action, meaning it can get from source to detector without traversing the intervening space.

Price and Wharton may have said it best when they wrote [13]:

> In putting future and past on an equal footing, this kind of approach is different in spirit from (and quite possibly formally incompatible with) a more familiar style of physics: one in which the past continually generates the future, like a computer running through the steps in an algorithm. However, our usual preference for the computer-like model may simply reflect an anthropocentric bias. It is a good model for creatures like us who acquire knowledge sequentially, past to future, and hence find it useful to update their predictions in the same way. But there is no guarantee that the principles on which the universe is constructed are of the sort that happen to be useful to creatures in our particular situation.
>
> Physics has certainly overcome such biases before – the Earth isn't the center of the universe, our sun is just one of many, there is no preferred frame of reference. Now, perhaps there's one further anthropocentric attitude that needs to go: the idea that the universe is as "in the dark" about the future as we are ourselves.

The term "universe" in this quote is the model of objective reality per physics. And we would clarify Price and Wharton's claim that we need to abandon the notion that "the universe is as 'in the dark' about the future as we are ourselves," by saying "the most explanatorily powerful model of objective reality uses future boundary conditions that we, as dynamic creatures, experience conditionally as in the mechanical universe rather than absolutely as in the block universe." We now introduce our specific form of adynamical explanation called the Relational Blockworld (RBW).

RBW is a unique spatiotemporal form of ontological contextuality. Thus we reject the idea that reality is ultimately *composed of things*, i.e., self-subsisting entities, individuals or trans-temporal objects with intrinsic properties and "primitive thisness," haecceity, etc. Rather, contexts, constraints and relations are primary while relata are derivative, thus rejecting "building block" atomism or Lego-philosophy. Relata inherit their individuality and identity from the structure of relations such as the experimental set-up.

We do not claim there are relations without relata, just that the relata are not individuals (e.g., things with primitive thisness and intrinsic properties), but always

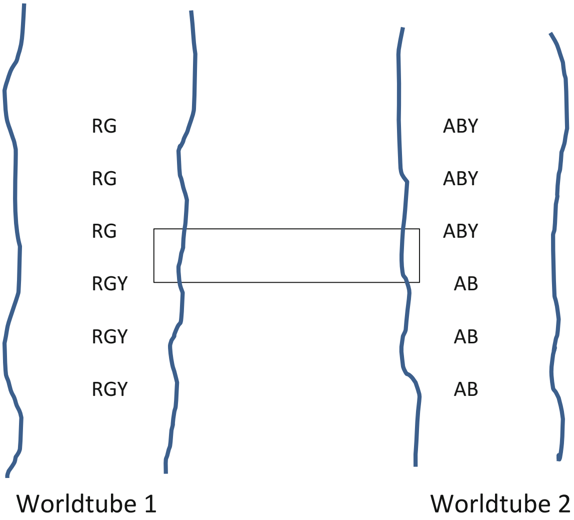


Fig. 16.4 Quantum exchange of energy-momentum—The property Y is associated with the source on the spacetimesource element (rectangle) shared by the worldtubes. As a result, property Y disappears from Worldtube 1 (Y source) and reappears later at Worldtube 2 (Y detector) without mediation. That is, there is no third worldtube/line needed to explain the exchange of energy-momentum associated with property Y between Worldtube 1 and Worldtube 2. While these properties are depicted as residing in the worldtubes, they don't represent something truly intrinsic to the worldtubes, but are ultimately contextual/relational, i.e., being the source of Y only makes sense in the context of (in relation to) a "Y detector", and vice versa. The A, B, R, and G properties shown might be established with respect to classical objects not shown in this Figure, for example

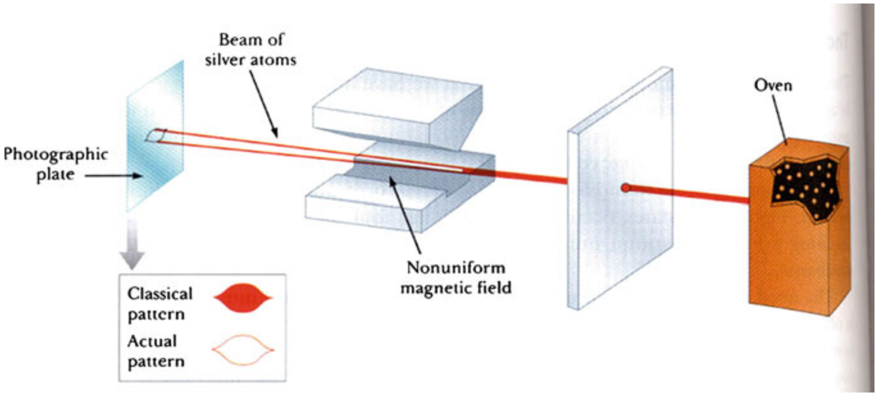


Fig. 16.5 Reproduced from Serway and Jewitt, Physics for Scientists and Engineers, Brooks/Cole CENGAGE Learning. The silver atoms deflected towards the north magnet pole of the Stern-Gerlach magnet are said to have spin "up." Those deflected towards the south magnet pole are said to have spin "down"

ultimately analyzable as relations as well (Fig. 16.4). RBW already somewhat violates the dynamical bias by rejecting *things* with intrinsic properties as fundamental *building blocks* of reality – the world isn't fundamentally *compositional*—the deepest conception of reality is not one in which we decompose things into other things at ever smaller length and time scales. The RBW 'beables', "spacetimesource elements,[2]" are certainly a violation of a compositional picture of realty, since their properties are inherited from their classical context.[3]

For example, let Hardy's experiment (Fig. 16.2) represent spin measurements (Fig. 16.5) on a pair of spin entangled particles (Fig. 16.6) with the two settings 1 and 2 of the Mermin device (Fig. 16.2) for the Hardy experiment corresponding to Stern-Gerlach magnet orientations X and Y (Figs. 16.5 and 16.6). Further, let the Mermin device outcomes R and G correspond to spin measurement outcomes "up" and "down," respectively. The spacetimesource element for any particular trial to include outcomes is constructed as in Fig. 16.7 (21RG there) and a corresponding probability amplitude is computed using the path integral without reference to counterfactuals. Each outcome represents a different spatiotemporal distribution of mass/energy per this particular quantum exchange, and the probability amplitude squared gives the relative probability of occurrence. The RBW rule for computing the action in the path integral for the probability amplitude is an AGC [14]. Thus, we go even further in rejecting dynamism, not merely because it is a BW, but

---

[2] The term "source" is used here in the quantum field theory context, i.e., it refers to sources *and* sinks of energy/momentum. We should note that these beables don't exist autonomously from measurement as Bell conceived of them.

[3] By "classical context" we mean the model of objective reality obtained using classical physics, e.g., general relativity.

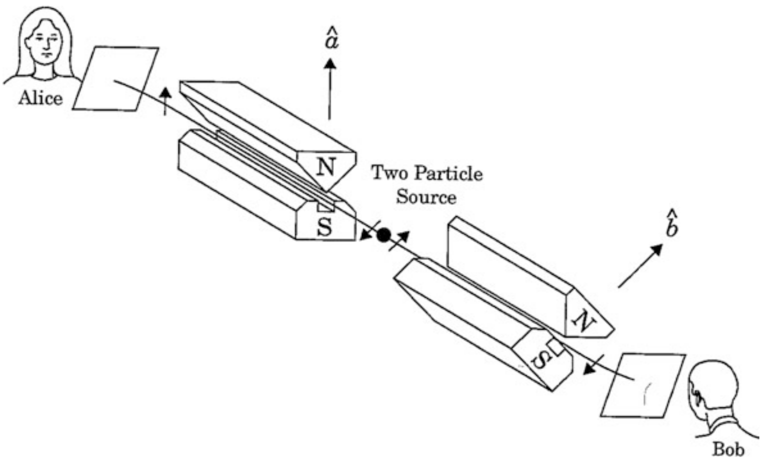


Fig. 16.6 *Alice* and *Bob* can orient their Stern-Gerlach magnets in one of two directions, X or Y, which correspond to the two settings 1 or 2 on Mermin's device for the Hardy experiment (Fig. 16.2). The two possible outcomes are "up" or "down" with respect to those orientations per Fig. 16.5, which correspond to the two outcomes R and G of the Hardy experiment

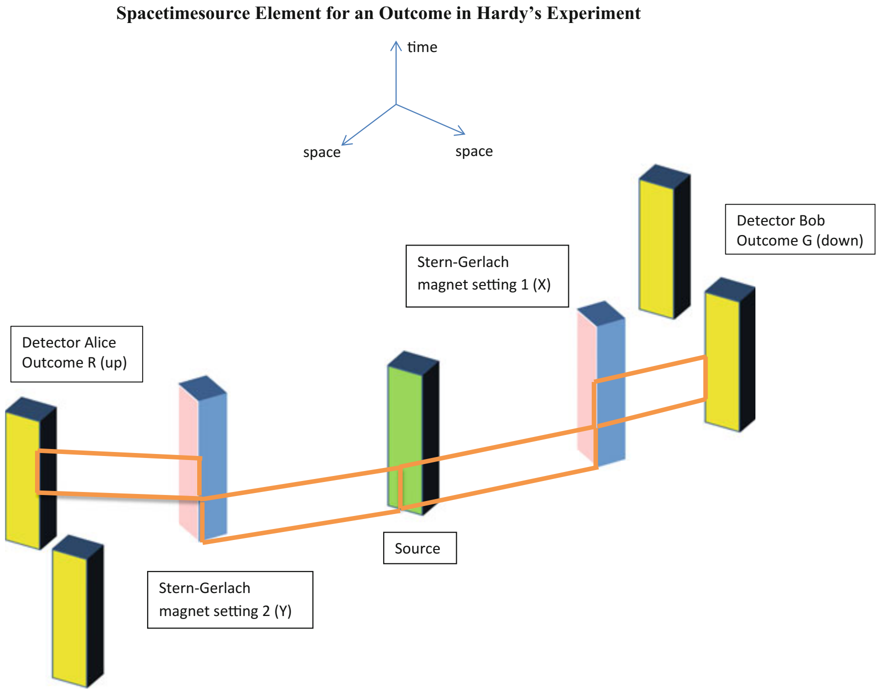


Fig. 16.7 The *orange blocks* (hollow rectangles) depict components of the spacetimesource element for a particular pair of Stern-Gerlach magnet settings and outcomes in Hardy's experiment (Color figure online)

because the fundamental AGC is not a dynamical law or even spacetime symmetries. The AGC constrains the probability amplitude for our beables, i.e., spacetimesource elements, which are spatiotemporal 4D ontological entities. For example in the twin-slit experiment, the spatiotemporal distribution of detector clicks is in accord with the distribution of spacetimesource elements per the probability amplitude obtained in accord with the AGC. To be clear, a spacetimesource element is not *in* spacetime, it is *of* spacetime, even while a distribution of detector clicks is viewed in the classical spacetime context of the experimental equipment and process from initiation to termination.

RBW gives Rickles and Bloom exactly what they are looking for in the following [15]:

> While the basic idea defended here (a fundamental ontology of brute relations) can be found elsewhere in the philosophical literature on 'structural realism', we have yet to see the idea used as an argument for advancing physics, nor have we seen a truly convincing argument, involving a real construction based in modern physics, that successfully evades the objection that there can be no relations without first (in logical order) having things so related.

As they say in the following passage, this idea has the potential to re-ground physics, dissolve current quagmires and lead to new physics [15]:

> Viewing the world as structurally constituted by primitive relations has the potential to lead to new kinds of research in physics, and knowledge of a more stable sort. Indeed, in the past those theories that have adopted a broadly similar approach (along the lines of what Einstein labeled 'principle theories') have led to just the kinds of advances that this essay competition seeks to capture: areas "where thinkers were 'stuck' and had to let go of some cherished assumptions to make progress." Principle theory approaches often look to general 'structural aspects' of physical behaviour over 'thing aspects' (what Einstein labeled 'constructive'), promoting invariances of world-structure to general principles.

Rickles and Bloom lament the fact that views such as ontological structural realism have yet to be so motivated and they further anticipate RBW almost perfectly when they say [15]:

> The position I have described involves the idea that physical systems (which I take to be characterized by the values for their observables) are exhausted by extrinsic or relational properties: they have no intrinsic, local properties at all! This is a curious consequence of background independence coupled with gauge invariance and leads to a rather odd picture in which objects and [spacetime] structure are deeply entangled. Inasmuch as there are objects at all, any properties they possess are structurally conferred: they have no reality outside some correlation. What this means is that the objects don't *ground* structure, they are nothing independently of the structure, which takes the form of a (gauge invariant) correlation between (non-gauge invariant) field values. With this view one can both evade the standard 'no relations without relata' objection and the problem of accounting for the appearance of time (in a timeless structure) in the same way.

For example, consider the particle tracks in a high energy physics detector. The tracks are worldlines, so they constitute what we mean by time-evolved "classical objects" and each worldline can be deduced one detection event (click) at a time in succession using $\psi(x, t)$, as shown by Mott for alpha particles in a cloud chamber [16].

Therefore, a probability amplitude could be computed for each worldline using spacetimesource elements detection event by detection event with each click providing empirical evidence of an otherwise unobservable, underlying spacetimesource element. However, as shown by Mott, after the first click the remaining clicks follow a classical trajectory with high probability, so the only real quantum computation needed is for the probability amplitude of the spacetimesource element of the set of first clicks, i.e., the first click for each worldline in the collection (again, a single spacetimesource element can have many components and represent many detection events and still be considered a unity, see Fig. 16.7). And, the properties (mass, charge, momentum, energy, etc.) for that spacetimesource element would simply be the properties of the subsequent worldlines (particles) defined relationally in the classical context of the accelerator source and particle detector. Basically, we are claiming that the worldtube of any particular classical object in space and time (defined relationally by its surrounding classical objects) can be decomposed into spacetimesource elements of space, time, and sources organized per an AGC using the context of those surrounding classical objects.

Accordingly, a particle physics detector event is one giant interference pattern (interference *a la* RBW), and the way to understand a particular pattern involving thousands of clicks can only realistically be accomplished by parsing an event into smaller subsets, and the choice of subsets is empirically obvious, i.e., spacetime trajectories. These trajectories are then characterized by mass, spin, and charge. Per RBW's adynamical explanation, the colliding beams in the accelerator and the detector surrounding the collision point form the graphical input that, in conjunction with the AGC, dictate the spacetime distribution of configurations of spacetimesource elements responsible for particle trajectories.

This severely undermines the dynamical picture of perturbations moving through a continuum medium (naïve field) between sources, i.e., it undermines the naïve notion of a particle as traditionally understood. In fact, the typical notion of a particle is associated with the global particle state of n-particle Fock space and per Colosi and Rovelli [17] "the notion of global particle state is ambiguous, ill-defined, or completely impossible to define." What we mean by "particle" is a collection of detector hits forming a spacetime trajectory resulting from a collection of adynamically constrained spacetimesource elements in the presence of colliding beams and a detector. And this doesn't entail the existence of an object with intrinsic properties, such as mass and charge, moving through the detector to cause the hits.

Our view of particles agrees with Colosi and Rovelli on two important counts. First, that particles are best modeled by local particle states rather than n-particle Fock states computed over infinite regions, squaring with the fact that particle detectors are finite in size and experiments are finite in time. The advantage to this approach is that one can unambiguously define the notion of particles in curved spacetime as excitations in a local (flat) region, which makes it amenable to Regge calculus (graphical version of general relativity). Second, this theory of particles is much more compatible with the quantum notion of complementary observables in that every detector has its own Hamiltonian (different sized graph with different properties), and therefore its own particle basis (unlike the unique basis of Fock space). Per Colosi & Rovelli [17], "In other words, we are in a genuine quantum mechanical situation in which distinct particle numbers are complementary observables. Different bases that diagonalize different HR [Hamiltonian] operators have equal footing. Whether a particle exists or not depends on what we decide to measure." Thus, in our view, particles simply describe how detectors and sources are relationally co-defined via RBW's adynamical global constraint.

In conclusion, RBW takes the Minkowskian fusion of space and time one step further by incorporating sources in its fundamental ontological entity, the 4D spacetimesource element. This results in a form of Wharton's Lagrangian Schema explanation governed by a fundamental adynamical global constraint, rather than Smolin's Newtonian Schema explanation of laws governing the behavior of time-evolved entities. In RBW, ontological contextuality is fundamental just as envisioned by Bohr and Wheeler. But unlike them, we have shown how to dispense with realism about the wavefunction along with all the problems it brings. In RBW, there is no measurement problem and no violation of locality. While it remains to be seen whether or not RBW is too radical a departure from the Newtonian Schema Universe, it seems certain that quantum nonlocality in general will entail a serious revision of our Newtonian mechanistic worldview. So, we close with this quote by Bertlmann [18]:

> I have a confession: I am not the realist one might expect after reading Bell's article "Bertlmann's socks and the nature of reality"; the world in its very foundations is much more abstract than we think with our "anschauliche" (intuitive) concepts, to borrow Werner Heisenberg's term. My personal feeling is that Bell's theorem, which reveals an apparent nonlocality in nature, points to a more radical conception whose onset we do not yet have.

References

1. Stuckey, W. M, Silberstein, M., & McDevitt, T. (2015). Relational Blockworld: Providing a realist psi-epistemic account of quantum mechanics. *International Journal of Quantum Foundations 1*, 123–170. http://www.ijqf.org/wps/wpcontent/uploads/2015/06/IJQF2015v1n3p2.pdf;
2. Stuckey, W.M., Silberstein, M., & Cifone, M.: Reconciling spacetime and the quantum: Relational Blockworld and the quantum liar paradox. *Foundations of Physics, 38*(4), 348–383 (2008). http://arxiv.org/abs/quant-ph/0510090;
3. Silberstein, M., Stuckey, W. M., & Cifone, M. (2008). Why quantum mechanics favors adynamical and acausal interpretations such as Relational Blockworld over backwardly causal and time-symmetric rivals. *Studies in History & Philosophy of Modern Physics, 39*(4), 736–751.
4. https://en.wikipedia.org/wiki/Quantum_nonlocality.
5. Zeilinger, A. (2004). Why the quantum? 'It' from 'bit'? A participatory universe? Three far-reaching challenges from John Archibald Wheeler and their relation to experiment. In J. D. Barrow, P. C. W. Davies, & C. Harper (Eds.), *Science and ultimate reality: Quantum theory, cosmology and complexity* (pp. 201–220). Cambridge: Cambridge University Press.
6. Aharonov, Y., & Zubairy, M.S. (2005).Time and the quantum: Erasing the past and impacting the future. *Science*, *307*, 875–879. Another example of this type of experiment

7. Smolin, L. (2009). *The unique universe: Against the timeless multiverse* (pp. 21–26). Physics World.
8. Carroll, S. (2012). A Universe from Nothing? Discover Magazine Online (28 Apr 2012) http://blogs.discovermagazine.com/cosmicvariance/2012/04/28/a-universe-from-nothing/.
9. Wharton, K. (2015). The Universe is not a Computer. In A. Aguirre, B. Foster, & Z. Merali (Eds.) *Questioning the foundations of physics* (pp. 177–190). Springer. http://arxiv.org/abs/1211.7081.
10. Geroch, R. (1978). *General Relativity from A to B* (pp. 20–21). Chicago: University of Chicago Press.
11. Mermin, D. (1994). Quantum mysteries refined. *American Journal of Physics, 62*(10), 880–887.
12. Feynman, R. P. (2005). Nobel Lecture. In L. M. Brown (Ed.) *Feynman's thesis: A new approach to quantum theory*. World Scientific Press.
13. Price, H., & Wharton, K. (2013). Dispelling the quantum spooks—a Clue that Einstein Missed? http://arxiv.org/abs/1307.7744.
14. Stuckey, et al. (2015). section 3.
15. Rickles, D., & Bloom, J. (2012).Things ain't what they used to be. FQXi Essay Contest (Spring 2012) http://www.fqxi.org/community/forum/topic/1537.
16. Mott, N. (1929). The wave mechanics of alpha-ray tracks. *Proceedings of the Royal Society, A126*, 79–84.
17. Colosi, D., & Rovelli, R. (2009). What is a particle? *Classical and Quantum Gravity 26*, 025002. http://arxiv.org/abs/gr-qc/0409054.
18. Bertlmann, R. (2015). Readers offer their own magic moments with John Bell. *Physics Today, 68*(12), 8–12.

Chapter 17
Spacetime Is Doomed

George Musser

17.1 Space as Order

Theoretical physicists have not wanted for imagination when it comes to developing a quantum theory of gravity. String theory, loop quantum gravity, causal-set theory, twistor theory: the approaches are diverse and the disagreements among their proponents are often vehement. And yet they have a common feature: that classical spacetime is not a fundamental ingredient of the world, but a construction consisting of more fundamental degrees of freedom. Those degrees of freedom become structured in very specific ways to give rise to the observed features of classical spacetime.

This is a radical shift in our conception of physics and its implications have yet to be fully assimilated. Nearly every physicist and philosopher from Democritus onward has assumed that space is the deepest level of physical reality. Just as the script of a play describes what actors do on stage, but presupposes the stage, the laws of physics have traditionally taken the existence of space as a given. (The same is true for time, but most of the approaches I will discuss treat space and time asymmetrically.) Space serves as the organizing principle of the natural world—the glue that binds the universe together, as the English physicist Julian Barbour has put it.[1] Physical objects do not interact willy-nilly; their behavior is dictated by how they are related to one another, which depends on where they lie in space at a given

---

Edited excerpt from "Spacetime Is Doomed" from SPOOKY ACTION AT A DISTANCE by George Musser. Copyright © 2015 by George Musser. Reprinted by permission of Farrar, Straus and Giroux, LLC.

---

[1] Barbour, *The End of Time: the Next Revolution in Physics*, 18.

G. Musser (✉)
Scientific American, 80 Glen Ridge Ave, Glen Ridge, NJ 07028, USA
e-mail: gmusser@nasw.org

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_17

time. This structuring role is easiest to see in the classical laws of mechanical motion, but also occurs in field theories. The value and rate of change of a field at different points in space fully determine what the field does, and points in the field interact only with their immediate neighbors.

What the candidate quantum theories of gravity do, in effect, is to invert this reasoning. Instead of saying that space brings order to the world, they say that the world is ordered and space is a convenient notion for describing that order. We perceive that things affect one another in a certain way and, from that, we assign them locations in space. The nice thing about defining space in terms of structure is that it sidesteps some of the long-running philosophical disputes over the nature of space. The ancient atomists (and, later, Newton) conceived of space as a thing in its own right, whereas Aristotle (and, later, Leibniz) deemed it an abstraction that describes how the contents of the universe are packed together. Either way, though, space reflects a structure that the natural world possesses. If the atomists were right and space has an independent existence, it must be highly ordered, like a neatly woven fabric, so that it can serve the functions that physics demands of it. If space is merely an abstraction, then the contents of the universe must fit together in just the right way to give meaning to the abstraction.

Consider, for example, causal-set theory. Theorems in general relativity theory show that you can map out the entire world from its web of cause-effect relations—you could put everything in the right place just by knowing what has to give rise to what.[2] (Technically, the web of relations does leave some ambiguity: its tells you that one event occurs earlier than another, but not how much earlier; it provides no scale. But causal-set theorists argue that space and time already have a natural scale if they are built up out of discrete units—“atoms” of space. Distance is then determined by counting the number of such atoms.)[3] The web of cause-effect relations must be highly ordered if it is to re-create space and time. A generic causal set does not correspond to a classical spacetime.[4]

17.2 Quantum Graphity

In many approaches, space is modeled as a network. The concept of space as a network goes back to the 1960s and the brainstorms of such theorists as John Wheeler, David Bohm, Roger Penrose, and David Finkelstein.[5] Wheeler, for one, imagined taking a bucket of “dust” or “rings”—primitive grains of matter that do

---

[2] Bombelli et al., “Space-Time as a Causal Set.”

[3] Riemann, “On the Hypotheses Which Lie at the Bases of Geometry,” 37.

[4] Dowker, “Causal Sets and the Deep Structure of Spacetime,” 454; Henson, “The Causal Set Approach to Quantum Gravity,” 405.

[5] Misner, Thorne, and Wheeler, *Gravitation*, 1203–1212; Bohm and Hiley, *The Undivided Universe: an Ontological Interpretation of Quantum Theory*, 374–378; Penrose, *The Road to Reality: a Complete Guide to the Laws of the Universe*, 946–950; Finkelstein, “Space-Time Code.”

not exist within space, but simply exist—and stringing them together to form space. More recently, Fotini Markopoulou and her colleagues developed a toy model of the stringing-together process, an approach they call "quantum graphity." Quantum graphity does not specify what the Wheelerian grains actually are—that is a job for a full-up theory of quantum gravity such as loop quantum gravity or string theory. Quantum graphity is a theory-in-miniature that focuses narrowly on what you might build with those grains.[6]

Indeed, Markopoulou and her colleagues' philosophy is that the detailed composition should not matter; the principles of organization should be universal. After all, physicists have found that similar rules govern a huge diversity of complex systems, from earthquakes to ecosystems to economies.[7] On the downside, quantum graphity is so bare-bones that it faces the problem of meshing with known physics.

The links between the elementary grains are as simple as can be. Two grains are either connected to each other or not, like Facebook users who can either be friends or not—just an on-or-off relationship. The resulting network looks like a string-art craft project in which you hammer nails (representing the grains) into a sheet of wood and stretch threads (the links) among some of them. Despite the simplicity of its construction, the network can take on a huge variety of shapes, ranging from skeletal outlines to elaborate mandalas.

To breathe life into the network—to give it the capacity to transform and evolve—Markopoulou and her colleagues suppose that the links switch on or off depending on the amount of available energy. This process is ad hoc, but again the goal is not to create a bulletproof theory, but to reconnoiter possibilities for how to construct space. Each link represents a certain amount of energy. Chains of links contain less energy than an equivalent number of isolated links, so the total energy of a network depends not only on the sheer number of links, but also on how they are put together. The more intricate the pattern is, the more energy it embodies.

The energy maxes out in a fully interconnected network, where every grain is linked to every other grain. In such a network, the principle of locality does not hold; you can go from any grain to any other grain in one hop, without passing through any intermediate points. The network lacks the hierarchy of relations—near versus far, small versus big—which is characteristic of space. You cannot subdivide it into separate chunks; it is an indivisible whole.

To see why the high-energy network is not spatial, try assigning locations to the grains. Every grain has to be equidistant (a single hop) from every other. For the first three grains, that is no problem: arrange them in an equilateral triangle. Four can be stacked in a pyramid. But where does a fifth go? There is nowhere equidistant to the first four points, at least not within ordinary, three-dimensional space. You need a four-dimensional pyramid. In fact, each additional grain requires a whole new dimension of space. Before long, you enter an ultra-higher-dimensional realm beyond our capacity to visualize. And most of that vast venue is wasted: the network is only

---

[6] Konopka, Markopoulou, and Severini, "Quantum Graphity: a Model of Emergent Locality."

[7] Stanley et al., "Scale Invariance and Universality: Organizing Principles in Complex Systems."

one hop wide in any direction and does a good impression of a balled-up spider's web. So although you might still talk of the network as existing within space, it is not the kind of space we want: three dimensions that extend as far as we can see in every direction and that provide an economical description of the relations among objects.

Lower-energy patterns are a different story. They're just what we want. Each grain connects to just a few others, forming a regular grid like a honeycomb or woven fabric. The notion of distance regains meaning: some grains are close together, the rest far apart. The network is nice and roomy. The principle of locality holds: for an influence to go from one place to another, it cannot hop straight there, but must work its way through the network. The passage of the signal takes time, which would explain why the speed of objects through space is limited (by the speed of light).[8]

In short, spacelessness and space are just two different phases of the same network of grains. One can metamorphose into the other; a crumpled wad can unfurl into a flat expanse. Theorists have proposed a couple of ways this might happen. The reshaping could be a process that occurs in time. The network starts off as sizzlingly hot—a highly interconnected pattern containing an enormous amount of energy. Then it cools off and crystallizes like a tray of water freezing to ice, as links dissolve and reorganize to create a tidy arrangement. The trick is to explain the cooling. Things do not just cool down on their own; something must drain them of heat. Markopoulou and her colleagues have speculated that the energy could go into the creation of matter. The primordial grains could coalesce into elementary particles, so that matter emerges hand in hand with space.

Alternatively, the transition may not be a process that unfolds in time, but a structuring that arises at the quantum level. The network can exist in multiple conditions at once, a limbo known as superposition. Although most of those conditions are nonspatial, they can fuse together into something that is spatial. The best-developed account of superposition of space goes by the somewhat unwieldy name of "causal dynamical triangulations." Its inventors have shown that nonspatial geometries neutralize one another, as long as events are highly ordered, with a distinction between cause and effect built in from the outset.[9]

17.3 Matrix Models

String theorists have explored ideas similar to quantum graphity. In the 1990s they pioneered "matrix models." A matrix in the mathematical sense has nothing to do with the virtual-reality "matrix" of the movie *The Matrix*, yet the premise is eerily similar: the world we experience is a kind of simulation generated by a deeper level

---

[8] Hamma and Markopoulou, "Background-Independent Condensed Matter Models for Quantum Gravity."

[9] Loll, Ambjørn, and Jurkiewicz, "The Universe From Scratch."

of reality. The best-known matrix model was developed by a quartet of theorists, Tom Banks, Willy Fischler, Steve Shenker, and Leonard Susskind.[10] Their model, like quantum graphity, supposes that the universe is a cat's cradle of interconnections among grains of primitive matter. Under the right conditions, extraneous connections rupture and the grains snap into a regular spatial grid.

String theory outgrew its name long ago. It postulates not just one-dimensional strings, but also two-dimensional membranes and higher-dimensional analogues—as theorists call them, 1-branes, 2-branes, 3-branes, 4-branes, and so on. Some branes, designated by D, can act as the endpoints of strings.[11] At the bottom of this pecking order is the humble D0-brane, a type of particle. Being a true geometric point lacking size or any other spatial attribute, the D0-brane is the perfect building block for space. Confirming this intuition, theorists calculate that the D0-brane has the right properties to serve as the graviton, the particle that has been hypothesized for decades to convey the force of gravity.

Matrix models take this particle as fundamental and construct the universe entirely from lots of them. Every particle can interact with every particle, and their interactions are not simply on or off, but can vary in strength and in quality. The more energy you inject into a pair of particles, the tighter their bond will become. The namesake matrix of numbers quantifies this web of interactions. For example, if you read down to the eighth row and then across to the twelfth column, the number there will tell you how strongly particle number eight interacts with particle number twelve. To express not just the raw strength but also the quality of the connection, you need several such matrices.

Each matrix is a square, and running diagonally from the top left corner to the bottom right is a special set of numbers—where the eighth row meets the eighth column, the twelfth row meets the twelfth column, and so on. These tell you how much each particle interacts with itself. Self-interactions are a core feature of matrix models. The particles are subatomic narcissists, the physics equivalent of Facebook users who always "Like" their own posts. Their self-interactions have a carefree, unrestrained quality; you can dial their strength up or down without having to pump in energy.

Whereas the workings of quantum graphity are somewhat ad hoc, the laws governing D0-branes are dictated by considerations of symmetry. The mathematical balance of the equations is the organizing principle of this model. Symmetry ensures that the off-diagonal values in the matrix are yoked to the diagonal values—in other words, that the branes' mutual interactions depend on their self-interactions. Particles that self-interact by comparable amounts forge a bond, whereas particles with differing levels of self-interaction remain aloof. Put simply, like attracts like. Consequently, the branes agglomerate into separate clusters like the social circles in your Facebook network. These clusters constitute the ordinary

---

[10] Banks et al., "M Theory as a Matrix Model: a Conjecture."

[11] Musser, *The Complete Idiot's Guide to String Theory*, 155.

subatomic particles of physics. Each cluster can be compactly described by a few numbers—namely, the strength and quality of its constituents' self-interactions.[12]

That is how space arises in matrix models. The D0-branes do not live or move within space. Mathematically they all sit on top of one another at a single point. But because they are so selective about their interactions, they produce our experience of living within space. What we call "position" is simply the set of numbers that uniquely identifies a given cluster. It's like pigeonholing your friends as "physics lovers," "Radiohead groupies," or "Cuban-style dancers."

That is just the start. You can take all our familiar spatial notions—movement, size, locality—and explain them in terms of brane dynamics. Movement: things shift their position because the D0-branes' self-interactions are varying. It's like saying the Cuban dancers suddenly all get interested in Dominican music. They "move" as a group to a new passion. Such movement may sound metaphorical, but in matrix models it is the origin of physical movement. Size: the self-interactions of the branes in an object are not exactly equal, but have a slight spread, so that the object spans a range of positions. Locality: clusters at separate locations are independent because their self-interactions differ, which suppresses their mutual interactions according to the logic of symmetry. This is like saying that Cuban dancers and Radiohead groupies never have much to say to each other.

Another departure from spatiality occurs inside clusters. The internal group dynamics are intense and every brane is interacting with every other. The branes scramble one another's self-interactions, and the matrix values representing those interactions lose the qualities of spatial coordinates. Ordinarily, coordinates are independent numbers: you can measure the latitude of a city separately from its longitude. But you cannot do that for branes within a cluster. If you measure the latitude of a brane first, then its longitude, you might get a different result than if you measured the longitude, then the latitude. This kind of ordering effect is known mathematically as "noncommutativity." In effect, the particle seems to be located in two different places, like Salt Lake City in my cities example. The degree of ambiguity is a measure of just how nonlocal and nonspatial the system is.

Indeed, the cluster does not really have an "inside"—there is no volume of space where the D0-branes bustle around. Arguably there are not even any D0-branes anymore, either, because they surrender their individuality and become assimilated into the collective. If you look at a cluster from the outside, what you see is not the outer surface of a material thing, but the end of space; and if you poke your hand into the cluster, you will not reach into its interior, for the cluster has no interior. Instead, your hand will become assimilated, too (which can't be good for it). If you wisely refrain from touching the cluster and instead throw particles into it, you will notice that the cluster's storage capacity depends on its area rather than on its interior volume—again, for the simple reason that it does not actually have an interior volume. Space has no meaning at this level.

---

[12] Banks, "The State of Matrix Theory," 342–343; Martinec, "Evolving Notions of Geometry in String Theory," 167–168.

17.4 Noncommutative Geometry

Some theorists such as Michael Heller have developed noncommutativity into a stand-alone theory called noncommutative geometry, which is a slightly ironic name: the aim is actually to do away with geometry and describe the universe as a big algebraic equation. The equation is formulated in terms of matrices, like those of matrix models, but given a new interpretation. The matrices are no longer arrays of numbers describing the connections among building blocks such as D0-branes, but individual entities in their own right, the primary ingredients from which everything else is made.

In concrete terms, those pursuing this approach take a top-down view of physics, in which global structures—ones that span the entire universe—are fundamental, and local geometric concepts such as "points" and "things" derive from those global structures, rather than the usual bottom-up view in which the universe is built from zillions of localized things. By analogy, imagine that instead of defining society as millions of individuals who assemble into sundry groups, we define it as millions of groups and identify each individual as a bundle of group memberships. (This isn't so far-fetched: philosophers and sociologists such as Hegel have argued that individual identity is constructed largely from social identity.[13] You could draw a big Venn diagram of groups and every person would be the unique intersection of some set of circles. For this top-down definition to work, groups must overlap in just the right way. If they don't, individuals lose their distinct identity; you'd get more than one person with the same set of group identities and be unable to tell them apart.

Something like this could happen with the global structures of noncommutative geometry. Under the right conditions, these structures interlock to produce a spatial universe governed by the usual laws of physics. Under other conditions, they do not—and their imperfect meshing would produce nonlocal phenomena such as entangled particles. If two distinct objects are merely different mixes of the same global entities, it stands to reason that they can retain a connection that transcends space. What happens in one place will be sensitive to what happens elsewhere, even without any communication in an ordinary sense.[14]

17.5 AdS/CFT Duality

Matrix models do have some peculiarities, but they establish a remarkable principle: a bunch of particles obeying quantum physics can organize themselves so that you would swear they live and move within space, even if space was not in the

---

[13] Pettit, *The Common Mind: an Essay on Psychology, Society, and Politics*, 166–173.

[14] Heller and Sasin, "Einstein-Podolski-Rosen Experiment From Noncommutative Quantum Gravity"; Heller and Sasin, "Nonlocal Phenomena From Noncommutative Pre-Planckian Regime."

original specification of the system. And it turns out that this principle is very general. Not just a swarm of D0-branes but almost any quantum system contains spatial dimensions folded inside it like a figure in a pop-up book. Most such systems do not bootstrap space from utter spacelessness, as matrix models do, but prime the pump with a low-dimensional space in order to generate a higher-dimensional one.

The AdS/CFT duality developed by Juan Maldacena is such a system. It starts with a three-dimensional space and generates a nine-dimensional one. One reason string theorists like this scenario so much is that it neatly explains the holographic principle, the idea that the universe can sustain much less complexity than the principle of locality would lead you to expect. The complexity is reduced by just the amount you would expect if one of the dimensions of space were illusory. In the AdS/CFT scenario, that is because the dimension in question *is* illusory. It can be collapsed down like an accordion because it was never really there. ("Illusory" is perhaps the wrong word. "Derived" or "constructed" would be better, if less poetic. The dimension may not exist at the lowest level, but it is still very real to anything larger than a brane.)

The disposable dimension reflects a particular aspect of order in the underlying quantum system. In fact, the requisite order is familiar to us from everyday life—specifically, the fact that big things and small things live as if in worlds apart. Our planet trundles around its orbit oblivious to human affairs, just as we spare little thought for the bacteria that lodge in our skin. Conversely, we have only a vague awareness of riding on a giant ball of rock, and bacteria know nothing of our daily struggles. Nature is stratified by scale.

Sound waves are an especially simple example of this stratification. Sounds of long and short wavelengths are oblivious to each other; if you sound a deep bass note and a high treble pitch simultaneously, each ripples through the room as though it were the only sound in the world. Their mutual independence is analogous to the autonomy of spatially separated objects. Suppose you play two piano keys, middle C and the adjoining D key. The C key creates a sound wave with a wavelength of 1 m 32 cm, and D produces one with a wavelength 14 cm shorter. These waves overlap in the three dimensions of space through which they propagate, yet they are independent of each other, as if they were located in different places. In a sense, you can think of the sound waves as residing 14 cm apart within a fourth spatial dimension.

The farther apart the keys are on a piano keyboard, the farther apart they are within this imaginary dimension; a given distance along the keyboard translates into a given distance within the dimension. You do not see this dimension as such; to you, it is an abstraction that captures the acoustical independence of sound waves. But it is a remarkably fitting abstraction. Musicians call the difference between pitches a musical "interval," which has connotations of distance, as if our brains really do think of the differences between pitches as spatial separation. AdS/CFT duality takes this abstraction literally and suggests that one of the dimensions of the space we occupy represents the energy or, equivalently, the size of waves within the underlying system[15].

Things of different sizes are not strictly independent; they interact with things of comparable size, and the effects can cascade from one scale to the next. Consider the proverb of the nail: for want of a nail, the shoe was lost; for want of a shoe, the horse was lost; then the knight, the battle, and the kingdom. A nail shortage in a single blacksmith shop didn't immediately cause the monarch's downfall; it exerted its influence indirectly, via systems of intermediate scales. Sound waves of different pitches can also behave like this. A Chinese gong begins rumbling at a low pitch and gradually vibrates at successively higher pitches.[16] The necessity of propagating through scale explains why spatial locality holds in the emergent dimension.[17] What happens in one place does not jump to another without passing through the points in between.

It is not automatic that the underlying quantum system would possess this kind of hierarchical order. Just as a painting must be composed in just the right way to produce the sense of depth, so must the system have a certain degree of internal coherence to give rise to space. What ensures this cohesion is entanglement among the system's particles or fields. To produce space as we know it, those particles or fields must be entangled by scale: each particle with its neighbor, each pair of particles with another pair, each group with another group. Other patterns lead to different geometries or systems that cannot be thought of as spatial at all. If the system is less than fully entangled, then the emergent space is disjointed, and an inhabitant of the universe would be trapped inside one region, unable to venture elsewhere. Furthermore, several recent studies suggest you can start with the constraints on entanglement and derive the general-relativistic equations for gravitation. On the face of it, entanglement seems to transcend space. Today, though, physicists think it might be what creates space.[18]

17.6 Scattering Amplitudes

In 2013 Nima Arkani-Hamed and Jaroslav Trnka unveiled a geometric technique for calculating probabilities, known in the jargon as "amplitudes," for particle processes. They gave their technique a suitably funky name: the "amplituhedron."[19] Based on the particles involved in a given process, you draw a polyhedron with one

---

[15] Susskind and Witten, "The Holographic Bound in Anti-De Sitter Space"; Balasubramanian and Kraus, "Spacetime and the Holographic Renormalization Group."

[16] Fletcher, "Nonlinear Dynamics and Chaos in Musical Instruments."

[17] Balasubramanian et al., "Holographic Probes of Anti–De Sitter Spacetimes"; Heemskerk et al., "Holography From Conformal Field Theory."

[18] Nishioka, Ryu, and Takayanagi, "Holographic Entanglement Entropy: an Overview"; Van Raamsdonk, "Building Up Spacetime with Quantum Entanglement"; Swingle, "Constructing Holographic Spacetimes Using Entanglement Renormalization."

[19] Arkani-Hamed and Trnka, "The Amplituhedron."

vertex for each particle. For instance, if you have two incoming particles that create four outgoing particles, you need a total of six vertices—a hexagon or one of its higher-dimensional counterparts. The momentum of a particle sets the size of its corresponding polyhedral face. Having formed this shape, you then calculate its interior volume, and that quantity, by the rules of the procedure, equals the desired amplitude.

The polyhedron is not a real object sitting in ordinary space but an abstract mathematical shape that captures the structure of particles' interactions. It subsumes all the previous calculation techniques that physicists have used to compute amplitudes, including Richard Feynman's diagrams and Zvi Bern and his colleagues' minimalist alternative[20]. These different techniques correspond to different ways to carve up the polyhedron for purposes of calculating its volume. The polyhedron also exhibits symmetries that nature possesses, but which theorists had never glimpsed before.

This procedure does not presume that the process plays out in spacetime. The locality we observe in daily life is a consequence of the way the faces fit together—specifically, that they form a closed shape, as opposed to disconnected planes. Those six vertices link into a hexagon rather than, say, an asterisk. In general, the faces will not fit together; locality is therefore a special case. The main lesson, as with the other approaches to emergent spacetime, is that space represents a type of order in the world, one that you might not expect a priori.

References

1. Arkani-Hamed, N., & Trnka, J. (2014). The Amplituhedron. *Journal of High Energy Physics* (10):30.
2. Balasubramanian, V., Kraus, P., Lawrence, A., & Trivedi, S. (1999). Holographic probes of anti–de sitter spacetimes. *Physical Review D, 59*(10), 104021.
3. Balasubramanian, V. Per Kraus Spacetime and the holographic renormalization group. *Physical Review Letters 83*(18).
4. Banks, T. (1998). The State of matrix theory. *Nuclear Physics B-Proceedings Supplements, 62*(1), 341–347.
5. Banks, T, Fischler, W, Shenker, S. H., & Susskind, L. (1997). M theory as a matrix model: a conjecture. *Physical Review D, 55*(8), .
6. Barbour, J. B. (2001). *The end of time: the next revolution in physics*. New York: Oxford University Press.
7. Bern, Z., Dixon, L. J., Dunbar, D. C., & Kosower, D. A. (1995). Fusing gauge theory tree amplitudes into loop amplitudes. *Nuclear Physics B, 435*(1), 59–101.
8. Bohm, D., & Hiley, B. J. (1995). *The undivided universe: an ontological interpretation of quantum theory*. Reprint. London: Routledge.
9. Bombelli, L., Lee, J., Meyer, D., & Sorkin, R. D. (1987). Space-Time as a causal set. *Physical Review Letters, 59*(5), 521–524.

---

[20] Bern et al., "Fusing Gauge Theory Tree Amplitudes Into Loop Amplitudes."

10. Dowker, F. (2005). Causal sets and the deep structure of spacetime. In A. Ashtekar (Ed.) *One hundred years of relativity* (pp. 445–64). Hackensack, N.J: World Scientific.
11. Finkelstein, D. R. (1969). Space-Time code. *Physical Review, 184*(5), 1261–1271.
12. Fletcher, N. H. (1993). Nonlinear dynamics and chaos in musical instruments. In Green, D. G., BOssomaier, T. (Eds.) Complex systems: from biology to computation (pp. 106–117). Amsterdam: IOS Press.
13. Hamma, A., & Markopoulou, F. (2011). Background-independent condensed matter models for quantum gravity. *New Journal of Physics, 13*(9), 095006.
14. Heemskerk, I., Penedones, J., Polchinski, J., & Sully, J. (2009). Holography from conformal field theory. *Journal of High Energy Physics, 2009*(10), 079.
15. Heller, M., & Sasin, W. (1998). Einstein-Podolski-Rosen experiment from noncommutative quantum gravity. *AIP Conference Proceedings, 453*, 234–241.
16. Heller, M., Sasin, W. (1999). Nonlocal phenomena from noncommutative pre-planckian regime. arXiv.org, 17 Jun 1999.
17. Henson, J. (2009). The causal set approach to quantum gravity. In D. Oriti (Ed.) *Approaches to quantum gravity: toward a new understanding of space, time and matter* (pp. 393–413). New York: Cambridge University Pres.
18. Konopka, T., Markopoulou, F., Severini, S. (2008). Quantum graphity: a model of emergent locality. Physical Review D *77*(10).
19. Loll, R., Ambjørn, J., Jurkiewicz, J. (2006). The universe from scratch. Contemporary Physics *47*(2).
20. Martinec, E. J. (2013). Evolving notions of geometry in string theory. *Foundations of Physics, 43*(1), 156–173.
21. Misner, C. W., Kip, S. (1973). Thorne, and John Archibald wheeler. In *Gravitation*, 1st edn. San Francisco: W. H. Freeman.
22. Musser, G. S., Jr. (2008). *The complete idiot's guide to string theory*. New York: Penguin Group.
23. Nishioka, T., Ryu, S., & Takayanagi, T. (2009). Holographic entanglement entropy: an overview. *Journal of Physics a: Mathematical and Theoretical, 42*(50), 504008.
24. Penrose, R. (2005). *The road to reality: A complete guide to the laws of the universe*. London: Jonathan Cape.
25. Pettit, P. (1996). *The Common mind: An essay on psychology, society, and politics*. New York: Oxford University Press.
26. Riemann, G. (1873). On the hypotheses which lie at the bases of geometry. *Nature, 8*(1), 36–37.
27. Stanley, H. E., Amaral, L. A. N., Gopikrishnan, P., Ivanov, P Ch., Keitt, T. H., & Plerou, V. (2000). Scale invariance and universality: organizing principles in complex systems. *Physica A: Statistical Mechanics and Its Applications, 281*(1), 60–68.
28. Susskind, L., & Witten, E. (1998). The holographic bound in anti-de sitter space.arXiv.org, May 19, 1998.
29. Swingle, B. (2012). Constructing holographic spacetimes using entanglement renormalization. arXiv.org, September 14, 2012.
30. Van Raamsdonk, M. (2010). Building up spacetime with quantum entanglement. In *General Relativity and Gravitation 42*(10):

Part III
Mathematics

Chapter 18
Geometry and Physical Space

Mary Leng

Geometry as a branch of mathematics studies the properties of points, lines, planes, solids, and their higher dimensional analogues. As applied to physical world, geometry is the study of figures in three-dimensional physical space. Until the mid-nineteenth century, ‘geometry’ meant Euclidean geometry, the axiomatic theory that forms the basis for the plane geometry will be familiar to most with a high school education, and the postulates on which this theory is based were considered to be indubitable truths about physical points and lines. But with the development of non-Euclidean geometries in the nineteenth century, mathematicians began to distinguish between geometry as a theory of physical space and geometries as theories of mathematical spaces. Doing so raises the question of the status of geometry considered as a theory of physical space.

Euclid presented his geometry in the *Elements* (c. 300 BCE), which gathered together and systematized the geometrical knowledge of the day. The presentation is still the paradigm of an axiomatic theory. The *Elements* starts with 23 definitions, five ‘common notions’ (essentially logical and arithmetical assumptions), and five postulates (now more commonly known as Euclid’s *axioms*), as follows:

1. It is possible to draw a straight line from any point to another point.
2. It is possible to produce a finite straight line continuously in a straight line.
3. It is possible to describe a circle with any center and radius.
4. All right angles are equal to one another.
5. If a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, the straight lines (if extended indefinitely) meet on the side on which the angles which are less than two right angles lie.

(Wolfram Mathworld *Elements*, http://mathworld.wolfram.com/Elements.html (accessed April 2016)

M. Leng (✉)
Department of Philosophy, University of York, Heslington,
N. Yorkshire Y010 5DD, UK
e-mail: mary.leng@york.ac.uk

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_18

Of these axioms, the first three can be thought of as idealized claims about what can be done with a ruler and pair of compasses, if we assume that these instruments could be arbitrary large. The fourth, though not about our constructing abilities, still has a very basic and intuitive character that makes it seem appropriate as an axiom rather than something in need of proof. The fifth postulate (which is equivalent to the claim that, given a straight line and a point not on that line, there exists one and only one parallel line through that point), however, was long thought to be different from the rest, looking much more like the *theorems* that can be proved on the basis of these axioms than something so basic as to be axiomatic. So while mathematicians had no doubt that all of these axioms were *true* of points and lines in physical space, there were questions about the appropriateness of the fifth axiom *as an axiom*, and attempts were made to prove it from the other four.

A standard method of mathematical proof is proof by contradiction, or *reductio ad absurdum*: we prove that a proposition P follows from a collection of assumptions by deriving a contradiction (or absurdity) from those assumptions together with the negation of P. So in trying to derive Euclid's parallel postulate, a reasonable approach to take would be to start by assuming the first four axioms and the negation of the parallel postulate, and show that this combination of assumptions leads to a contradiction. This approach was taken by many mathematicians including Gerolamo Saccheri (1667–1733), who derived sufficient bizarre seeming consequences from the assumption that the parallel postulate was false that he published his results under the title *Euclides ab Omni Naevo Vindicatus* ('Euclid Vindicated from All Faults'), declaring that parallel axiom as established [1].

However, Saccheri had not succeeded in proving a formal contradiction from the negation of the parallel postulate, and mathematicians began to suspect that the postulate was indeed independent of the other axioms. In the early 19th century Gauss, Bolyai, and Lobachevsky independently came to the conclusion that the assumption that, given a line and a point not on that line there is more than one parallel line running through it, was logically possible (if not true of physical points and lines). Later Gauss's student Riemann explored the hypothesis that there are no parallel lines, and again came to view this assumption as consistent. The independence of the parallel postulate was finally established in 1868 by Beltrami, whose 'Essay on the Interpretation of non-Euclidean Geometry' presented a model for a two dimensional non-Euclidean geometry on a three dimensional Euclidean surface (a pseudosphere). Before long, models had been given for both Bolyai-Lobachevsky (or hyperbolic) geometry and Riemannian (or elliptic) geometry, and it was clear that it was consistent to assume exactly one parallel (Euclid—surfaces of zero curvature); more than one (Bolyai-Lobachevsky—surfaces of negative curvature) or no parallels (Riemann—surfaces of positive curvature). Looking back on Saccheri's book, it could be recognised that his supposed absurdities derived when assuming the parallel postulate to be false were straightforward theorems of these new non-Euclidean geometries.

The development of non-Euclidean geometries was of ground breaking importance in mathematics in allowing the distinction for the first time between geometry as the theory of mathematically possible spaces, and geometry as the theory of physical space. As Michael Scanlan puts it, in a paper on the proof of the independence of the parallel postulate, “In the past, mathematical practice did not involve a distinction between theory and interpretation. In the eighteenth century, mathematics was seen as the ‘abstract’ study of certain aspects of nature” [2]. Geometry was a body of truths about physical points and straight lines, and it could reasonably be argued by Immanuel Kant that the truths of geometry are synthetic a priori, substantial (not merely definitional) truths about physical space that are knowable a priori through reflection on our intuition of the nature of our experience of space. The development of non-Euclidean geometries, and their acceptance as part of mathematics, brought to the fore the question of what the proper subject matter of mathematics is. While it could be thought that the points and lines of Euclidean geometry were just slightly idealized abstractions from the physically inscribed points and lines of the diagrams used to convince us of Euclid’s proofs, we now had new geometries, with their own terminology of ‘points’ and ‘lines’ but with different assumptions about parallels. At most one could be true of (idealized) points and lines in physical space, but all were equally good considered as mathematical theories. The distinction between mathematical spaces and physical space was thus drawn, raising the question of the status of mathematical objects as nonspatiotemporal abstracta.

What interests me here, though, is not so much the status of geometry as the theory of mathematical spaces, but the status of geometry as a theory of physical space. Even once the conceptual possibility of non-Euclidean geometries was recognized, it remained in theory acceptable still to think that Euclidean geometry was knowable a priori to be true as a theory of *physical* points and straight lines. Beltrami’s and Klein’s models of non-Euclidean geometries showed the consistency of these axiomatic theories by reinterpreting ‘point’ and particularly ‘straight line’ to apply to things that were not, by our own lights, *really* straight lines. It remained then possible to argue that, if by point we mean *point in physical space*, and if by straight line we mean *straight line in physical space*, then the mere consistency of alternative geometries should in no way shake our confidence in the truth of Euclid’s axioms when understood as a theory of physical points and lines. But in fact, Scanlan tells us, “the mathematicians who originally conceived of non-euclidean geometry, Bolyai, Lobachevsky, and to some extend Gauss, seem all to have conceived of the theory as one which is potentially applicable to physical space” [2], and while our experience locally is Euclidean, the question was raised as to whether on a large scale the Euclidean laws continue to hold. The story is told of Gauss measuring the angles of the triangle formed by three mountain peaks looking for evidence that on a large scale the angles did not add up to two right angles (which is equivalent to the falsity of the parallel postulate), though it is unclear that Gauss’s interest here was testing the possibility of curvature in space, as opposed to effects on measurements due to the curvature of the earth’s surface. Lobachevsky, however, explicitly conceived of a test of the geometry of space suggesting that one might measure a stellar triangle consisting of the distant star Sirius together with two different positions of the earth at six month intervals, to determine whether the angles were as predicted in a Euclidean or non-Euclidean geometry (see [3], p. 15).

For these mathematicians, then, the question of the correct geometry of physical space was now an empirical matter, to be determined by experiment.

We now know that our best physical theory of space and time is general relativity, according to which spacetime has *variable* curvature (neither Euclidean, nor elliptic, nor hyperbolic, all of which are geometric theories of constant curvatures). One major confirmation of this theory was via the measurement of distant stars. In 1919 the physicist Arthur Eddington travelled to the island of Principe, close to the equator and just off the coast of western Africa, to photograph the solar eclipse of 29 May (see Kennefick [4]). With the sun's light dimmed by the eclipse, it was possible to photograph positions of bright stars from the Hyades cluster beyond the sun, and to achieve measurement results that are generally taken to have confirmed Einstein's prediction of a space curved by the presence of the sun over Newton's assumed flat Euclidean space.

Should we conclude, then, that the question of whether physical space is Euclidean is an empirical one, answerable—and indeed answered in the negative—by experiment? Henri Poincaré argued forcefully against this conclusion, and in favour of the view that the question of the appropriate geometry of physical space is not a priori or *empirical* but rather a matter of convention. Thus at the turn of the twentieth century, Poincaré, well aware of the proposals for empirical tests of geometrical hypotheses, though prior to the development and testing of general relativity, could write:

> If Lobatschewsky's geometry is true, the parallax of a very distant star will be finite. If Riemann's is true, it will be negative. These are the results which seem within the reach of experiment, and it is hoped that astronomical observations may enable us to decide between the two geometries. But what we call a straight line in astronomy is simply the path of a ray of light. If, therefore, we were to discover negative parallaxes, or to prove that all parallaxes are higher than a certain limit, we should have a choice between two conclusions: we could give up Euclidean geometry, or modify the laws of optics, and suppose that light is not rigorously propagated in a straight line. It is needless to add that every one would look upon this solution as the more advantageous. Euclidean geometry, therefore, has nothing to fear from fresh experiments. [5]

According to Poincaré, then, the status of geometry as a theory of space is neither a priori nor empirical, but *conventional*, simply a matter of how we define our terms.

> What, then, are we to think of the question: Is Euclidean geometry true? It has no meaning. We might as well ask if the metric system is true, and if the old weights and measures are false; if Cartesian co-ordinates are true and polar coordinates false. One geometry cannot be more true than another; it can only be more convenient. Now, Euclidean geometry is, and will remain, the most convenient [5].

In stating that, in the light of apparent experimental refutation we could choose to alter our hypothesis that light propagates in straight lines rather than altering our geometry, Poincaré's discussion suggests that his conventionalism is simply an application of what has become known as the Quine-Duhem thesis, the claim that, given that no theoretical statement can be tested in isolation, but only against a backdrop of further theoretical assumptions, it is always possible to hold on to any statement in light of recalcitrant experience, simply by adjusting assumptions in our background theory: "Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system." (Quine [6], p. 43). But if Poincaré's claim is simply that we could redefine our terms so that 'straight line' doesn't mean 'path taken by a light ray', then this is not a terribly exciting form of conventionalism: there would still be the empirical question of what paths light rays take, and it would still be an empirical matter whether the paths taken by light rays are best described by a Euclidean or a non-Euclidean geometry.

In fact, Poincaré's conventionalism is about more than simply choosing our terms, as is seen in a thought experiment he presents of a world enclosed within a sphere with some rather peculiar properties. In this world, there is a property like temperature, which varies according to the distance from the centre. If R is the radius of the sphere and r the distance of a point in the sphere from the centre, the temperature at this point is proportional to $R^2 - r^2$. Bodies in this world expand and contract at a uniform rate according to changes in temperature, so that a rod that is a metre long by our standards at the centre will get smaller and smaller as it is moved away from the centre, approaching but never reaching zero (the world consists of all points inside the sphere, but not the sphere's boundary). Imagine a plane in this sphere consisting of a great circle of the sphere (i.e. cutting through the centre and with diameter R), and imagine an inhabitant starting at the centre of the plane with a surveyor's wheel with circumference 1 m. As they walk along a radius of the sphere towards the edge, both they and their wheel will contract uniformly. To them, they will feel as though their universe is unbounded of infinite extent as however many metres they travel they will be able to continue. From our perspective this is an error—the universe has finite bounds, but the strange behavior of their measuring instruments mean that the inhabitants are unable to realise this. If surveyors in Poincaré's sphere universe continue to take measurements they will conclude that they are living in a hyperbolic geometry of infinite extent, whereas from our perspective they are living in a Euclidean sphere with physical features that affect their ability to measure.

On one reading of this picture, the possibility of the sphere world presents an epistemic challenge to the claim that the question of the proper geometry of physical space is an empirical matter. On this view, two accounts are available that fit the observed phenomena for the inhabitants within the sphere world. One holds that they and their measuring instruments do not change size as they move around, and that the geometry of their world is hyperbolic. The other holds that they and their measuring instruments change size as they move around, and that the geometry of their world is Euclidean. The inhabitants can't choose between these two hypotheses, so for them the question of the 'true' geometry of physical space cannot be determined empirically (even though as a matter of fact the true geometry is Euclidean). Poincaré's own view, though, is that the lack of knowledge in this case is not because the truth is out there but beyond the inhabitants' grasp, but rather, that there is nothing 'out there' to be known. The epistemic view of the inhabitants' predicament holds that there is a fact of the matter about whether their measuring instruments shrink and grow as they move around or stay the same size, but that this fact is unavailable to the inhabitants. Poincaré, on the other hand, holds that the question of whether the inhabitants shrink or not is itself not an empirical matter, but dependent on a conventional choice for us to make about what we are going to take as counting as 'congruence'. There is no 'God's eye view' which determines what is *really* going on in this example. The two descriptions: the measuring instruments stay the same size and the geometry is hyperbolic, and the measuring instruments change in size and the geometry is Euclidean, are equally good ways of describing the same basic facts. We can choose to define 'congruence' in terms of the behaviour of measuring instruments (so that lengths that measure the same when measured by a meter stick that has been transferred from one to the other are counted as congruent), or we could choose to define it so that measuring whether lengths are congruent depends on knowledge of their distance from the centre. Each choice is a matter of conventional decision, and each leads to a different conclusion about 'the' geometry of the space, so the question of which geometry is correct turns out to be answered by conventional decision rather than empirical investigation.

Poincaré's picture can seem compelling once we consider that we too are in the position of the sphere dwellers. We assume by and large that our measuring instruments remain the same size as we move around in our universe, but an alternative picture according to which we grow and shrink according to location could also be made compatible with our observations. Should we, then, conclude that the question of the correct geometry of physical space can only be made sense of downstream of a conventional decision, and as such, is itself a matter of convention rather than empirical fact? Despite the conventionalist elements of the aforementioned 'Quine-Duhem' thesis, the empiricist response to this conventionalist claim is actually to be found in the work of W.V. Quine. Poincaré's conventionalism depends on holding that there are some elements of our theories that are purely conventional choices about how to set the meanings of terms, that before we can measure and build theories we have to define our terms, and these definitions are a matter of pure convention. Quine argues forcefully against this picture, holding that in the web of beliefs that makes up our best empirically tested theory of the world any element, including those that were originally introduced as conventional definitions to get theorizing going, can be amended in the light of recalcitrant experience. So even though decisions that may seem arbitrary or conventional may need to be made to get theorizing going, those decisions can be revised in the light of recalcitrant experience, making them as empirical as any other elements of our theories. Thus, Quine writes,

> The lore of our fathers is a fabric of sentences. In our hands it develops and changes, through more or less arbitrary and deliberate additions and revisions of our own, more or less directly occasioned by the continuing stimulation of our sense organs. It is a pale grey lore, black with fact and white with convention. But I have found no substantial reasons for concluding that there are any quite black threads in it, or any white ones [7].

In Quine's view, then, the fact that conventional choices about how to use our terms are made on the way to theorizing does not stop our theories—conventions included—from being empirically tested as a package. Indeed, in special and general relativity our previous assumptions about congruence and the behaviour of measuring instruments are challenged; we now adopt a theory according to which our measurements of length are relative to frame of reference (special relativity) and relative to our location with respect to the distribution of mass in the universe (general relativity). In Quine's view, the success of the theoretical package that includes these assumptions is confirmation of the package as a whole. All truths depend in part on the meaning of terms and in part on how the world is, but in this respect, the claim (supported by general relativity) that the spacetime we inhabit has a non-Euclidean geometry of variable curvature is as empirical as anything can be.

References

1. Kline, M. (1990). *Mathematical Thought from Ancient to Modern Times* (Vol. 3). New York: Oxford University Press.
2. Scanlan, M. J. (1988). Beltrami's model and the independence of the parallel postulate. *History and Philosophy of Logic, 9*, 13–34.
3. van Fraassen, 'B. (2006). Weyl's paradox: The distance between structure and perspective. In C. Suhm, A. Berg-Hildebrand (Eds.) *Bas C. van Fraassen: The Fortunes of Empiricism*. Frankfurt: Ontos Verlag.
4. Kennefick, D. (2009). Testing relativity from the 1919 eclipse—a question of bias *Physics Today* 37–42.
5. Poincaré, H. (1905). *Science and hypothesis*. London: Walter Scott Publishing Ltd.
6. Quine, W. V. (1961) The two Dogmas of empiricism. In *From a Logical Point of View* (2nd revised edn.). Cambridge, MA: Harvard University Pres. Reprinted in W.V. Quine.
7. Quine, W. V. (1976). Carnap and logical truth. Reprinted in *Ways of Paradox and Other Essays* (2nd edn.). Cambridge, MA: Harvard University Press.

Chapter 19
The Geometry of Manifolds and the Perception of Space

Raymond O. Wells Jr.

19.1 Introduction

In our contemporary world, we see ourselves immersed in a vast universe filled with galaxies, black holes, dark matter and other aspects of our cosmological surroundings. An important aspect to all of this is that we imagine ourselves to be in a variably curved four-dimensional space-time, as first put forward by Einstein in his theory of general relativity in 1916 [8]. In February of 2016, a century after Einstein's paper appeared, it was announced [1] by a large team of astronomers that gravitational waves were first detected by amazingly accurate sensors. This result produced headlines around the world, as did the confirmation of Einstein's theory by Dyson, Eddington, and Davidson in their celebrated eclipse paper [7] of 1920 showing that light was bent by the gravitational attraction of the sun, as was predicted by Einstein. The measurement of the bending of a specific star was $1.65''$, which compared favorably with Einstein's prediction of $1.75''$.

This four-dimensional space-time is a four-dimensional manifold with a Lorentzian metric which satisfies the Einstein field equations. The question we want to discuss in this essay is: what are abstract manifolds (four-dimensional or otherwise), and how did this mathematical concept and its associated geometric ideas arise, so that it could be used for modeling our universe as Einstein and so many others after him have done?

R.O. Wells Jr. (✉)
University of Colorado at Boulder, Jacobs University Bremen, Rice University,
Boulder, USA
e-mail: row@raw.com

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_19

19.2 The Origins of the Concept of a Manifold

The notion of a manifold is a relatively recent one, but the theory of curves and surfaces in Euclidean three-space $\mathbf{R}^3$ originated in Greek mathematical culture. For instance, the book on the study of conic sections by Apollonius [18] described mathematically the way we understand conic sections today. The intersection of a plane with a cone in $\mathbf{R}^3$ generated these curves, and Apollonius showed moreover that any plane intersecting a *skew cone* gave one of the three classical conic sections (ellipse, hyperbola, parabola), excepting the degenerate cases of a point or intersecting straight lines. This was a difficult and important theorem at the time. The Greek geometers studied intersections of other surfaces as well, generating additional curves useful for solving problems and they also introduced some of the first curves defined by transcendental functions (although that terminology was not used at the time), e.g., the quadratrix, which was used to solve the problem of squaring the circle (which they couldn't solve with straight edge and compass and which was shown many centuries later not to be possible). See for instance Kline's very fine book on the history of mathematics for a discussion of these issues [21].

Approximately 1000 years after the major works by the Greek geometers, Descartes published in 1637 [6] a revolutionary book which contained a fundamentally new way to look at geometry, namely as the solutions of algebraic equations. In particular the solutions of equations of second degree in two variables described precisely the conic sections that Apollonius had so carefully treated. This new bridge between algebra and geometry became known during the 18th and 19th centuries as *analytic geometry* to distinguish it from *synthetic geometry*, which was the treatment of geometry as in the book of Euclid [17], the methods of which were standard in classrooms and in research treatises for approximately two millenia. Towards the end of the 19th century and up until today *algebraic geometry* refers to the relationship between algebra and geometry, as initiated by Descartes, as other forms of geometry had arisen to take their place in modern mathematics, e.g., topology, differential geometry, complex manifolds and spaces, and many other types of geometries.

Newton and Leibniz made their major discoveries concerning differential and integral calculus in the latter half of the 17th century (Newton's version was only published later; see [23], a translation into English by John Colson from a Latin manuscript from 1671). Leibniz published his major work on calculus in 1684 [22]. As is well known, there was more than a century of controversy about priority issues in this discovery of calculus between the British and Continental scholars (see, e.g., [21]). In any event, the 18th century saw the growth of analysis as a major force in mathematics (differential equations, both ordinary and partial, calculus of variations, etc.). In addition the variety of curves and surfaces in $\mathbf{R}^3$ that could be represented by many new transcendental functions expanded greatly the families of curves and surfaces in $\mathbf{R}^3$ beyond those describable by solutions of algebraic equations.

Moreover, a major development that grew up at that time, and which concerns us in this paper, was the growing interaction between analysis and geometry. An important first step was the analytic description of the curvature of a curve in the plane at a given point by Newton. This was published in Newton's 1736 monograph [23] mentioned earlier. In the text of this reference (in a section entitled "To find the quantity of curvature in any curve") one finds the well-known formula for the curvature of a curve defined as the graph of a function $y = f(x)$

$$K_P = \pm \frac{f''(x)}{[1 + (f'(x)^2]^{\frac{3}{2}}}, \tag{19.1}$$

which one learns early on in the study of calculus. Curvature had been studied earlier quite extensively by Huygens [19], and he computed the curvature of many explicit curves (including the conic sections, cycloids, etc.). He did not use calculus, per se, but he used limiting processes of geometric approximations that are indeed the essence of calculus. Even earlier Apollonius had been able to compute the curvature of a conic section at a specific point (see [18] and Heath's discussion of this point).

The next development concerning analysis and geometry was the study of space curves in $\mathbf{R}^3$ with its notion of curvature and torsion as we understand it today. This work on space curves was initiated by a very young (16 years old) Clairaut [5] in 1731. Over the course of the next century there were numerous contributions to this subject by Euler, Cauchy, and many others, culminating in the definitive papers of Frenet [11] and Serret [32] in the mid-19th century giving us the formulas for space curves in $\mathbf{R}^3$ that we learn in textbooks today. Space curves were called "curves of double curvature" (French: *courbes à double courbure*) throughout the 18th and 19th centuries.

A major aspect of studying curves in two or three dimensions is the notion of the measurement of arc length, which has been studied since the time of Archimedes, where he gave the first substantive approximate value of π by approximating the arc length of an arc of a circle (see, e.g., [27]). Using calculus one was able to formalize the length of a segment of a curve by the well-known formula

$$\int_\Gamma ds = \int_a^b \sqrt{(x'(t))^2 + (y'(t))^2}dt,$$

if Γ is a curve in $\mathbf{R}^2$ parametrized by $(x(t), y(t))$ for $a \le t \le b$. The expression $ds^2 = dx^2 + dy^2$ in $\mathbf{R}^2$ or $ds^2 = dx^2 + dy^2 + dz^2$ in $\mathbf{R}^3$ became known as the *line element* or infinitesimal measurement of arc length for any parametrized curve. What this meant is that a given curve could be approximated by a set of secants, and measuring the lengths of the secants by the measurement of the length of a segment of a straight line in $\mathbf{R}^2$ or $\mathbf{R}^3$ (using Pythagoras!), and taking a limit, gave the required arc length. The line elements $ds^2 = dx^2 + dy^2$ and $d^2 = dx^2 + dy^2 + dz^2$ expressed *both* the use of measuring straight-line distance in the ambient Euclidean space as well as the limiting process of calculus. We will see later how generating such a line element to be *independent* of an ambient Euclidean space became one of the great creations of the 19th century.

The concept of curvature of a curve in $\mathbf{R}^3$ was well understood at the end of the 18th century, and the later work of Cauchy, Serret and Frenet completed this set of investigations begun by the young Clairaut a century earlier. The problem arose: how can one define the *curvature of a surface* defined either locally or globally in $\mathbf{R}^3$?

Euler recognized the difficulty of this problem, as we see from this quote of Euler's from his paper on curvature of surfaces from 1767, "Recherches sur la courbure des surfaces" [9]:

> In order to know the curvature of a curve, the determination of the radius of the osculating circle furnishes us the best measure, where for each point of the curve, we find a circle whose curvature is precisely the same. However, when one looks for the curvature of a surface, the question is very equivocal and not at all susceptible to an absolute response, as in the case above. There are only spherical surfaces where one would be able to measure the curvature, assuming the curvature of the sphere is the curvature of its great circles, and whose radius could be considered the appropriate measure. But for other surfaces, one doesn't know even how to compare a surface with a sphere, as when one can always compare the curvature of a curve with that of a circle. The reason is evident, since at each point of a surface there are an infinite number of different curvatures. One has to only consider a cylinder, where along the directions parallel to the axis, there is no curvature, whereas in the directions perpendicular to the axis, which are circles, the curvatures are all the same, and all other obliques sections to the axis give a particular curvature. It's the same for all other surfaces, where it can happen that in one direction the curvature is convex, and in another it is concave, as in those resembling a saddle.

In the quote above, we see that Euler recognized the difficulties in defining curvature for a surfaces at any given point. He does not resolve this issue in this paper, but he makes extensive calculations and several major contributions to the subject. In particular, he computes the curvatures at a point of the surface of the curves which are the intersections of planes passing though a normal vector at the point with the surface and finds what are now called the *principal curvatures* and the *principal directions* of the surface at the given point.

19.3 Gauss and Intrinsic Differential Geometry

A major development at the beginning of the 19th century was the publication in 1828 by Gauss of his historic landmark paper concerning differential geometry of surfaces, entitled *Disquisitiones circa superficies curvas* [15]. The year before, he published a very readable announcement and summary of his major results in [14], and we shall quote from this announcement paper somewhat later, letting Gauss tell us in his own words what he thinks the significance of his discoveries is. For the moment, we will simply say that this paper laid the foundation for doing intrinsic differential geometry on a surface and was an important step in the creation of a theory of abstract manifolds which was developed a century later.

Gauss defined *curvature* κ of a surface $S \subset \mathbf{R}^3$ to be the derivative of what is now called the Gauss mapping, which maps a point on the surface to its unit normal vector, which is a point on the unit sphere in $\mathbf{R}^3$. He showed first that the curvature (now called the *Gaussian curvature*) is the product of the Eulerian principal curvatures at that point. This definition of curvature (either via the Gauss mapping or the product of the principal curvatures) depends explicitly on the embedding of the surface in $\mathbf{R}^3$. They both use the notion of the normal vector at the given point, which depends very much on the embedding.

Suppose that the surface is described in terms of local coordinates (u, v), i.e.,

$$x(u, v), y(u, v), z(u, v) \tag{19.2}$$

are three smooth functions that represent the surface parametrically, then the line element

$$ds^2 = dx^2 + dy^2 + dz^2$$

in $\mathbf{R}^3$ induces a metric on the surface S of the form

$$ds^2 = Edu^2 + 2Fdudv + Gdv^2, \tag{19.3}$$

where E, F, and G are functions of (u, v) obtained by differentiating the parametrizing functions $x(u, v)$, $y(u, v)$, and $z(u, v)$. One can use the line element ds^2 on S to measure the length of a curve on the surface, to measure the angle formed by two intersecting curves on S, and to compute other geometric quantities.

The great achievement of Gauss in this paper was to show (his *Theorema Egregium*) that the Gaussian curvature could be defined in terms of the first and second derivatives of the coefficients E, F, and G of the line element in (19.3), and was thus independent of the embedding. A major motivation for Gauss's work on curvature in this paper were his parallel experimental geodesic measurements, which were trying to estimate the curvature of the earth near Gttingen.

Gauss understood full well the significance of his work and the fact that this was the beginning of the study of a new type of geometry (which later generations have called *intrinisic differential geometry*). We quote here from his announcement of his results published some months earlier from pp. 344–345 of [14]:

> These theorems lead us to consider the theory of curved surfaces from a new point of view, whereby the investigations open to a quite new undeveloped field. If one doesn't consider the surfaces as boundaries of domains, but as domains with one vanishing dimension, and at the same time as bendable but not as stretchable, then one understands that one needs to differentiate between two different types of relations, namely, those which assume the surface has a particular form in space and those that are independent of the different forms a surface might take. It is this latter type that we are talking about here. From what was remarked earlier, the curvature belongs to this type of concept; and, moreover, figures constructed on the surface, their angles, their surface area, their total curvature as well as the connecting of points by curves of shortest length, and similar concepts, all belong to this class.

The results described in this short announcement (and the details in the much longer Latin paper on the subject) formed the basis of most of what became modern differential geometry.

An important point that we should make here is that Gauss did significant experimental work on measuring the curvature of the earth in the area around Göttingen, where he spent his whole scientific career. This involved measurements over hundreds of miles, and involved communicating between signal towers from one point to another. He developed his theory of differential geometry as he was conducting the experiments, and at the end of his announcement [14], he summarizes one of his experiments to say that a geodesic triangle with the longest side 12 miles long has the sum of its three angles measuring $2''$ greater that 180°, indicating in a precise manner the positive curvature of the earth near that triangle. It is interesting to compare these experiments with those concerning the bending of light one century later and the measurement of gravitational waves two centuries later.

19.4 Riemann's Higher-Dimensional Geometry

In mathematics, we sometimes see striking examples of brilliant contributions or completely new ideas that change the ways mathematics develops in a significant fashion. A prime example of this is the work of Descartes [6], which completely changed how mathematicians looked at geometric problems. But it is rare that a single mathematician makes as many singular advances in his lifetime as did Riemann in the middle of the 19th century. In this section, we will discuss in some detail his fundamental creation of the theory of higher-dimensional manifolds and the additional creation of what is now called Riemannian or simply differential geometry. However, it is worth noting that he only published nine papers in his short lifetime (he lived to be only 40 years old), and several other important works, including those that concern us in this section, were published posthumously from the writings he left behind. His collected works (including in particular these posthumously published papers) were edited and published in 1876 and are still in print today [29].

Looking through the titles, one is struck by the wide diversity as well as the originality. Let us give a few examples here. In Paper I (his dissertation), he formulated and proved the Riemann mapping theorem and dramatically moved the theory of functions of one complex variable in new directions. In Paper VI, in order to study Abelian functions, he formulated what became known as Riemann surfaces and this led to the general theory of complex manifolds in the 20th century. In Paper VII, he proved the Prime Number Theorem and formulated the Riemann Hypothesis, which is surely the outstanding mathematical problem in the world today. In Paper XII, he formulated the first rigorous definition of a definite integral (the Riemann integral) and applied it to trigonometric series, setting the stage for Lebesgue and others in the early 20th century to develop many consequences of the powerful theory of Fourier analysis. In Papers XIII and XXII, he formulated the theory of higher-dimensional manifolds, including the important concepts of Riemannian metric, normal coordinates, and the Riemann curvature tensor, which we will visit very soon in the paragraphs below. Paper XVI contains correspondence with Enrico Betti leading to the first higher-dimensional topological invariants beyond those already known for two-dimensional manifolds (that had been developed by Riemann in Papers I and VI).

His paper [31] (Paper XIII above) is a posthumously published version of a public lecture Riemann gave as his *Habilitationsvortrag* in 1854. This was part of the process for obtaining his *Habilitation*, a German advanced degree beyond the doctorate necessary to qualify for a professorship in Germany at the time (such requirements are still in place at most German universities today as well as in other European countries, e.g., France and Russia; it is similar to the research requirements in the US to be qualified for tenure). This paper, being a public lecture, has very few formulas, is at times quite philosophical and is amazing in its depth of vision and clarity. On the other hand, it is quite a difficult paper to understand in detail, as we shall see.

Before this paper was written, manifolds were all one- or two-dimensional curves and surfaces in $\mathbf{R}^3$, including their extension to points at infinity (which was developed as a part of projective geometry in the first half of the 19th century). In fact, some mathematicians who had to study systems parametrized by more than three variables declined to call the parametrization space a manifold or give such a parametrization a geometric significance. In addition, these one- and two-dimensional manifolds always had a differential-geometric structure which was induced by the ambient Euclidean space (this was true for Gauss, as well).

In Riemann's paper [31], he discusses the distinction between discrete and continuous manifolds, where one can make comparisons of quantities by either counting or by measurement, and gives a hint, on p. 256, of the concepts of set theory, which was only developed later in a single-handed effort by Cantor. Riemann begins his discussion of manifolds by moving a one-dimensional manifold, which he intuitively describes, in a transverse direction (moving in some type of undescribed ambient "space"), and inductively, generating an n-dimensional manifold by moving an $(n-1)$-manifold transversally in the same manner. Conversely, he discusses having a nonconstant function on an n-dimensional manifold, and the set of points where the function is constant is (generically) a lower-dimensional manifold; and by varying the constant, one obtains a one-dimensional family of $(n-1)$-manifolds (similar to his construction above).

Riemann formulates local coordinate systems $(x^1, x^2, ..., x^n)$ on a manifold of n dimensions near some given point, taken here to be the origin. He formulates a curve in the manifold as being simply n functions $(x^1(t), x^2(t), ..., x^n(t))$ of a single variable t. The concepts of set theory and topological space were developed only later in the 19th century, and so the global nature of manifolds is not really touched on by Riemann (except in his later work on Riemann surfaces and his correspondence with Betti, mentioned above). It seems clear on reading his paper that he thought of n-dimensional manifolds as being extended beyond Euclidean space in some manner, but the language for this was not yet available.

At the beginning of this paper Riemann acknowledges the difficulty he faces in formulating his new results. Here is a quote from the second page of his paper (p. 255):

> In that my first task is to try to develop the concept of a multiply spread-out quantity [he uses the word manifold later], I believe even more in being allowed an indulgent evaluation, as in such works of a philosophical nature, where the difficulties are more in the concepts than in the construction, wherein I have little experience, and except for the paper by Mr. Privy Councilor Gauss in his second commentary on biquadratic residues in the Göttingen Gelehrte Anzeige [1831] and in his Jubiläumsschrift and some investigations by Hebart, I have no precedents I could use.

The paper of Gauss that he cites here [13] refers to Gauss's dealing with the philosophical issue of understanding the complex number plane after some thirty years of experience with its development. We will mention this paper again somewhat later in this paper. Hebart was a philosopher whose metaphysical investigations influenced Riemann's thinking. Riemann was very aware of the speculative nature of his theory, and he used this philosophical point of view, as the technical language he needed (set theory and topological spaces) was not yet available. This was very similar to Gauss's struggle with the complex plane, as we shall see later.

As mentioned earlier, measurement of the length of curves goes back to the Archimedian study of the length of a circle. The basic idea there and up to the work of Gauss was to approximate a given curve by a collection of straight line segments and take a limit. The *length* of each straight line segment was determined by the Euclidean ambient space, and the formula, using calculus for the limiting process, became, in the plane for instance,

$$\int_\Gamma ds = \int_a^b \sqrt{(x'(t))^2 + (y'(t))^2}dt,$$

where $ds^2 = dx^2 + dy^2$ is the line element of arc length in $\mathbf{R}^2$. As we saw in Sect. 19.3, Gauss formulated in [15] on a two-dimensional manifold with coordinates (p, q) the line element

$$ds^2 = Edp^2 + 2Fdpdq + Gdq^2 \tag{19.4}$$

where E, F, and G are induced from the ambient space. He didn't consider any examples of such a line element (19.4) that weren't induced from an ambient Euclidean space, but his remarks (see the quote above in Sect. 19.3) clearly indicate that this could be a ripe area for study, and this could well include allowing coefficients of the line element (19.4) being more general than induced from an ambient space.

Since Riemann formulated an abstract n-dimensional manifold (with a local coordinate system) with no ambient space, and since he wanted to be able to measure the length of a curve on his manifold, he formulated, or rather postulated, an independent measuring system which mimics Gauss's formula (19.4). Namely, he prescribes for a given coordinate system a metric (line element) of the form

$$ds^2 = \sum_{i,j=1}^{n} g_{ij}(x)dx^i dx^j, \tag{19.5}$$

where $g_{ij}(x)$ is, for each x, a symmetric positive definite matrix, and he postulates by the usual change of variables formula

$$ds^2 = \sum_{i,j=1}^{n} \tilde{g}_{ij}(\tilde{x}) d\tilde{x}^i d\tilde{x}^j,$$

where $\tilde{g}_{ij}(\tilde{x})$ is the transformed positive definite matrix in the new coordinate system $(\tilde{x}_1, ..., \tilde{x}^n)$.

Using the line element (19.4), the length of a curve is defined by

$$l(\Gamma) := \int_a^b \sqrt{\sum_{i,j=1}^{n} g_{ij}(x(t)) \frac{dx^i}{dt}(t) \frac{dx^j}{dt}(t) dt}. \tag{19.6}$$

The line element (19.5) is what is called a *Riemannian metric* today, and the 2-form ds^2 is considered as a positive definite bilinear form giving an inner product on the tangent plane $T_p(M)$ for p a point on the manifold M. This has become the basis for almost all of modern differential geometry (with the extension to Lorentzian-type spaces where $g_{ij}(x)$ is not positive definite, *à la* Minkoswki space). Riemann merely says on page 260 of his paper (no notation here at all),

> I restrict myself therefore to manifolds where the line element is expressed by the square root of a differential expression of second degree.

Earlier he had remarked that a line element should be homogeneous of degree one and one could also consider the fourth root of a differential expression of fourth degree, for instance. Hence his restriction in the quote above.

The next step in Riemann's paper is his formulation of curvature. This occurs on a single page (p. 261 of [31]), It is extremely dense and not at all easy to understand. Over time, however, it became understood by several generations of mathematicians, and this became the basis for the work of Einstein.

We summarize briefly what Riemann described on this page. He formulated the notion of geodesic coordinates, i.e., coordinates $(x_1, \ldots, s_n)$, where the coordinates are geodesics in the metric ds^2 from (19.6), and he expanded the functions g_{ij} in a Taylor series at the origin and considered the second-order terms as a biquadratic form

$$Q(x, dx) = \sum_{ijkl} c_{ijkl} x^k x^l dx^i dx^j, \tag{19.7}$$

and the coefficients c_{ijkl} had certain important symmetry properties and effectively defined the *Riemannian curvature tensor* written classically and still today as a tensor of the form

$$R^{\rho}_{\sigma\mu\nu}.$$

How does one define such a curvature tensor for n-dimensional manifolds with a Riemannian metric in a general coordinate system? In a quite short paper written in

Latin for a particular mathematical prize in Paris in 1861, (Paper No. XXII) Riemann provides the first glimpse of the general Riemann curvature tensor. The purpose of Riemann's paper was to answer a question in the Paris competition dealing with the flow of heat in a homogeneous solid body.

Riemann's ideas in these two papers were developed and expanded considerably in the following decades in the work of Christoffel, Levi-Cevita, Ricci, Beltrami, and many others. The main point of our discussion has been that Riemann created on these few pages the basic idea of an n-dimensional manifold not considered as a subset of Euclidean space *and* of the independent concept of a Riemannian metric and the Riemannian curvature tensor. What is missing at this point in time is the notion of a topological space on the basis of which one can formulate the contemporary concept of an abstract manifold or an abstract Riemannian manifold.

19.5 Hermann Weyl: Prelude to the 20th Century

Other major developments in geometry that led to the study of abstract manifolds in the 20th century, and which we won't consider in any detail in this paper, are the creations of complex geometry, transformations groups, set theory and topology, among others. Complex numbers were called *imaginary numbers* (among other appellations) in the centuries preceding the 19th century. This meant, in the eyes of the beholders that these were not concrete real (or realistic) numbers (to use a pun!), but were simply imaginary artifacts that had no real meaning, but were useful in the way they arose as the would-be solutions of algebraic equations. Gauss, among others, used these numbers extensively in his career, for instance in his ground-breaking work on number theory early in his life, *Disquisitiones Arithmeticae* [12]. Only many decades later did he make the case for a definitive geometric interpretation of what we call complex numbers today in a brief paper [13], which was a commentary on some of his earlier work on number theory and which Riemann cited earlier in this paper as a philosophical work which was a guide for him. Gauss called these numbers (for the first time) *complex numbers*, with their real and imaginary parts representing coordinates in a two-dimensional plane—the *complex plane*. Gauss pleads with his readers to consider complex numbers and the complex plane to be considered as a normal part of mathematics, not as something "imaginary" or "unreal." It's very enlightening to read this quite readable short paper. Later in the century, Klein and others started to consider complex solutions of homogeneous algebraic equations and thus began the study of complex algebraic manifolds and varieties. Klein along with Lie fostered the study of transformation groups, and the notion of manifolds as quotients of such groups became an important development as well. See, for instance, the very informative book by Klein [20] (first published in 1928, but based on lectures of Klein towards the end of the 19th century), which described most of the developments in geometry in the 19th century in a succinct fashion.

The creation of set theory in the latter half of the 19th century was a singular effort of Georg Cantor over several decades of work. The first article in this direction was published in 1874 [3]. He then published six major papers in the Mathematische Annalen between 1879 and 1884, which established the basic tenets of set theory, and laid the foundation for work in multiple directions for the next century and beyond, in particular in logic and foundations of mathematics, as well as point set topology, to mention a particular part of geometry related to our thesis in this paper (see his collected works [4] for these papers).

After Cantor's groundbreaking discovery (some aspects of which were very controversial for quite some time) the notion of topological space evolved to its contemporary state as a set satisfying certain axioms concerning either neighborhoods or open sets. This became the development of point set topology, which has its own very interesting history, and we mention only the fundamental contributions of Frchet in 1906 [10] (metric spaces, Frchet spaces, in particular in the infinite-dimensional setting) and Hausdorff in 1914 [16] (which, among other things, defined specifically Hausdorff spaces, one of many specialized types of topological spaces used often in geometry today). In 1895 Poincaré launched the theory of topological manifolds with a cornerstone paper called *Analysis Situs* [25], which was followed up by five supplements to this work over the next 17 years (see Vol. 6 of his collected works [26] as well as a very nice translation of all of these topology papers of Poincaré by John Stillwell [24]). This paper and its supplements became the foundation of what is now called *algebraic topology* (following the pioneering work of Riemann [30] on Riemann surfaces and Betti on homological invariants of higher-dimensional manifolds [2]).

The final step in our journey is the book by Hermann Weyl in 1913 [33] entitled *Die Idee der Riemannschen Fläche*, in which he gives the first very specific definition of an abstract manifold with more structure than a topological manifold (which Poincaré and others had already investigated quite thoroughly). His motivation was to give a better understanding of Riemann surfaces which transcended Riemann's original description as a multisheeted covering of the complex plane with branch points of various kinds. Specifically, he considered a topological manifold of two dimensions (locally homeomorphic to an open set in $\mathbf{R}^2$) which had a finite or countable triangulation and which had the additional property that there were coordinate charts mapping to an open disk in the complex plane $\mathbf{C}$ whose transition functions on overlapping coordinate charts were holomorphic. He showed how all of the previous work on Riemann surfaces fits into this new picture, and he proved a fundamental existence theorem of global holomorphic or meromorphic functions on such surfaces, utilizing the Dirichlet principle, that had been first used by Riemann in his conformal mapping theorem from his dissertation [28]. This definition of a topological manifold with such additional structure became the model for all the various kinds of manifolds studied in the following century up to the current time. This included, for instance, differentiable (C^∞) manifolds, complex manifolds of arbitrary dimension, Riemannian manifolds, symplectic manifolds, and real-analytic manifolds, among many others. Weyl knew that this was new territory, and like Gauss with intrinsic differential geometry, and Riemann with n-dimensional manifolds, he carefully explained to his readers that he was introducing a new way of thinking. Here, in Weyl's own words, are what he thought about this ([33], p. V):

> Such a rigorous presentation, which, namely by the establishing of the fundamental concepts and theorems in function theory and using theorems of the analysis situs which do not just depend on intuitive plausibility, but have set-theoretic exact proofs, does not exist. The scientific work that remains to be done in this regard may perhaps not be particularly highly valued. But, nevertheless, I believe I can maintain that I have tried in a serious and conscientious manner to find the simplest and most appropriate methods that lead to the asserted goal; and at many points, I have had to proceed in a different manner than that which has become traditional in the literature since the appearance of C. Neumann's classical book "Lectures on Riemann's theory of Abelian Integrals" (1865).

As we mentioned in the Introduction, Einstein's paper utilizing a Riemannian structure with a Lorentzian metric appeared in 1916. Almost immediately thereafter in the summer of 1917, Hermann Weyl gave lectures in Zurich on Einstein's theory, which appeared in a book published by Springer soon thereafter. The book was entitled, in German, *Raum, Zeit, Materie*, and its third edition appeared in 1919 [34] (an English edition, *Space, Time, Matter*, became available in 1950, Dover Publications, New York, and is still available today). Weyl gave more mathematical background for Einstein's theory and had a major influence on the propagation of Einstein's ideas. Weyl considers arbitrary n-dimensional Riemannian manifolds and the corresponding theory of tensors on such a manifold. The global structure of such a manifold is not discussed; he concentrates primarily on the local theory for varying coordinate systems near a specific point.

If we jump almost a century later, we see that modern string theory considers the four-dimensional space-time of Einstein, which we can denote by M^4, as a basis for a 10-dimensional string-theoretic manifold. This is a four-dimensional space-time M^4 with an additional compact three-dimensional complex manifold X^3 (a Calabi–Yau manifold) attached to M^4 at each point, and which is extremely small relative to our usual perception of space around us. See, for example, the very interesting exposition of Shing-Tung Yau and Steve Nadis [35] for a discussion of this and with many references to the string-theoretic literature. The notion of an abstract manifold, as first formulated by Hermann Weyl in 1913, is essential for this theory. It is also interesting to note that Weyl's book dealt with the important example of Riemann surfaces, which were one-dimensional complex manifolds (with local holomorphic coordinate systems) with specific global properties (for instance, connectivity, which Riemann had initiated). This is, in fact, part of the nature of Calabi–Yau manifolds, which are three-dimensional generalizations of the one complex-dimensional Riemann surfaces of Hermann Weyl, and which have specific global properties, which distinguishes them from from other three-dimensional complex manifolds.

References

1. Abbott, B. P., et al. (2016). Observation of gravitational waves from a binary black hole merger. *Physical Review Letters*, *116*, 061102–1–061102–15.
2. Betti, E. (1871). Sopra gli spazi di un numero qualunque di dimensioni. *Annali di Matematica Pura ed Applicata*, *4*, 140–158.

3. Cantor, G. (1874). Ueber eine Eigenschaft des Inbegriffes aller reelen algebraischen Zahlen. *Journal fur die reine und angewandte Mathematik*, *77*, 258–262.
4. Cantor, G. (1932). In Zermelo, E. (Ed.), *Gesammelte Abhandlungen*. Berlin: Springer.
5. de Clairaut, A. C. (1731). *Recherche sur les courbe à double courbure*. Nyons.
6. Descartes, R. (1886). *La Géométrie*. Paris: Hermann. Cornell University Library Digital Collections, first published in 1637.
7. Eddington, A. S., Dyson, F. W., & Davidson, C. (1920). A determination of the deflection of light by the Sun's gravitational field, from observations made at the solar eclipse of May 29, 1919. *Philosophical Transactions of the Royal Society of London*, *220*, 291–333.
8. Einstein, A. (1916). Die Grundlage der allgemein Relativitätstheorie. *Annlen der Physik*, *49*, 769–822.
9. Euler, L. (1767). Recherches sur la courbure des surfaces. In *Memoires de l'academie des sciences de Berlin*, *16*, 119–143.
10. Fréchet, M. (1906). Sur quelques points du calcul fonctionnel. *Rendiconti del Circolo Matematico di Palermo*, *22*, 1–72.
11. Frenet, F. (1852). Sur les courbes à double courbure. *Journal de Mathématiques Pures et Appliquées*, *17*, 437–447.
12. Gauss, C. F. (1863). Disquisitiones arithmeticae. In *Gauss Werke* (Vol. 1, pp. 1–478).
13. Gauss, C. F. (1863). Theoria residuorum biquadraticorum, commentatio secunda. In *Gauss Werke*, (Vol. 2, pp. 169–178). Königlichen Gesellschaft der Wissenschaften, Göttingen, in German (Latin title), originally published in Göttingsche Gelehrte Anzeige, 1831, V. 1. *Apr.*, *23*, 625–638.
14. Gauss, C. F. (1873). Anzeige: Disquisitiones generales circa superficies curvas. In *Gauss Werke*, (Vol. 4, pp. 341–352). Göttingen: Königlichen Gesellschaft der Wissenschaften.
15. Gauss, C. F. (1873). Disquisitiones generales circa superficies curvas. In *Gauss Werke* (Vol. 4, pp. 217–258). Göttingen: Königlichen Gesellschaft der Wissenschaften. Originally published *Commentationes Societatis Regiae Scientiaarum Gottingensis recentiores*, Vol. VI,1828, pp 99-146.
16. Hausdorff, F. (1914). *Grundzüge der Mengenlehre*. Leipzig: Veit. later editions: 1927, 1935, de Gruyter; English translation editions (Set Theory): 1957, 1962, Chlesea.
17. Heath, T. L. (1956). *The Thirteen Books of Euclid's Elements* (Vols. 3, 2nd edn.). New York: Dover.
18. Heath, T. L. (1896). *Apollonius*. Cambridge: Treatise on Conic Sections.
19. Huygens, C. (1673). *Horologium Oscillatorium*. Paris: Muget. Translation into English online by Ian Bruce at www.17centurymaths.com.
20. Klein, F. (1928). In Rosemann, W. (Ed.), *Vorlesungen über Nicht-Euklidische Geometrie*. Berlin: Julius Springer.
21. Kline, M. (1972). *Mathematical thought from ancient to modern times*. New York: Oxford University Press.
22. Leibniz, G. W. (1684). Nova methodus pro maximis et minimis, itemque tangentibus, quae nec fractas, nec irrationales quantitates moratur. In *Acta Eruditorum* (pp. 467–473).
23. Newton, I. (1736). *The method of fluxions and infinite series*. London: Henry Woodfall. Translated by John Colson from unpublished Latin ms of Newton (Methodus Fluxionum et Serierum Infinitarum) from 1671; download available from http://www.archive.org/details/methodoffluxions00newt.
24. Poincaré, H. (2010). Papers on topology; Analysis situs and its five supplements. American Mathematical Society Providence, RI. Translation by John Stillwell.
25. Poincaré, H. (1895). Analysis situs. *Journal de l'École Polytechnique*, *1*, 1–121.
26. Poincaré, H. (1953). *Oeuvres de Henri Poincaré*. Paris: Tome VI. Gauthiers-Villars.
27. Resnikoff, H. L., Wells, R. O. (2015). *Mathematics in Civilization* (3rd edn.) New York: Dover.
28. Riemann, B. (1876). Grundlagen für eine allgemeine Theorie der Funktionen einer veränderlichen complexen Grösse. In H. Weber (Ed.) *Gesammelte Mathematische Werke* (pp. 3–47). Leipzig: Teubner. Inaugural Dissertation, Göttingen, 1851.

29. Riemann, B. (1876). In H. Weber (Ed.) *Gesammelte Mathematische Werke*. Leipzig: B. G. Teubner. 1990 edition with addtional notes from numerous commentators on Riemann's work was published by Springer and edited by R. Narasimhan.
30. Riemann, B. (1876). Theorie der Abel'sche Functionen. In H. Weber, editor, *Gesammelte Mathematische Werke* (pp. 81–135). Leipzig: Teubner. This paper in Riemann's Werke combines four paper from the Journal für reine und angewandte Mathematik, Vol. 54, 1857.
31. Riemann, B. (1876). Ueber die Hypothesen, welche der Geometrie zu Grunde liegen. In *Gesammelte Mathematisch Werke* (Chapter XIII, pp. 254–269). Leipzig: B. G. Teubner.
32. Serret, J. A. (1851). Sur quelques formules relatives à la théorie des courbes à double courbure. *Journal de Mathématiques Pures et Appliquées*, *16*, 193–207.
33. Weyl, H. (1913). *Die Idee der Riemannschen Fläche*. Leipzig: B. G. Teubner.
34. Weyl, H. (1919). *Raum, Zeit, Materie: Vorlesungen über allgemeine Relatvitätstheorie*. Berlin: Springer-Verlag.
35. Yau, S.-T., & Nadis, S. (2011). Geometry of the universe's hidden dimensions notices. *American Mathematical Society*, *58*, 1067–1076.

Chapter 20
Paradox? the Mathematics of Space-Time and the Limits of Human Understanding

Paul Ernest

There is a paradox in exploring space-time and the limits of human understanding from the perspective of mathematics. For mathematics provides the language in which theories of space-time are formulated and yet is understood to be outside of space and time itself. Thus geometry is the science of space, and yet is itself beyond space.

A second paradox is that the humanly created science of mathematics is widely accepted to transcend the limits of human understanding and knowing. It expresses truths that apply beyond and across all space and time, imposed on all rational beings by logical necessity. But how can this be as mathematics is itself human knowledge, and the history of its creation is there for all to see? How can it exceed the knowledge capacities of its authors? In accounting for the relationships between mathematics, space-time and the limits of human understanding I suggest a resolution to these paradoxes. However, this requires the Promethean task of bringing mathematics from heaven back down to earth.

20.1 Mathematics and Space-Time

Mathematics as the science of structure studies those structures underpinning space and time. Inquiries into the space around us began with mensuration and geometry, based on our actions in, and ideas about, the physical world. Mensuration addresses practical problems concerning the measurement of objects in space, including lengths, surface areas, volumes, capacities, etc. It was developed in Mesopotamia and Ancient Egypt in the millennia before the Christian era. Deductive geometry emerged around 500–600 years BCE (Before the Christian Era) in Ancient Greece,

---

P. Ernest (✉)
Graduate School of Education, Exeter University, Exeter EX1 2LU, UK
e-mail: P.Ernest@exeter.ac.uk

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_20

reaching its high point in Euclid's elegant and systematic text *The Elements of Geometry*, compiled around 300 BCE. For nearly two millennia it was believed that Euclid's geometry described the true geometric properties of the physical world. However, doubts emerged about the logical necessity of one of its basic assumptions, the parallel postulate, which states that there is one and only one parallel line at a given distance from any other given line. Rejecting this postulate as a true axiom led to the emergence of non-Euclidean geometries almost two centuries ago. In these geometries the sum of the angles of a triangle is not 180° but greater (in spherical geometries) or lesser (in hyperbolic geometries), showing that many different geometries of space are possible. Further generalizations of geometry have been developed, including projective geometry, topology, and finally set theory in which all but the last vestiges of space are erased. Projective geometry studies what is left invariant when shapes are projected onto a surface. Topology, also called rubber sheet geometry, studies what is left unchanged when shapes are stretched and distorted but not punctured or cut. Set theory is so abstract that the only quasi-geometric ideas it models concern boundaries and their relations of containment, intersection and disjunction.

At the same time as geometry was emerging historically so too was astronomy, with its models of the heavens and the movements of stars and planets. A great deal of mathematics including trigonometry was developed to model heavenly motions and make astronomical predictions. Trigonometry studies the properties of triangles when one or two sides and angles are known. As well as in astronomy its application in navigation has been very important for travel, trade and empire. So mathematics has helped us conquer space including both the earth we live on and the skies above and around us.

The role of time in mathematics is less straight-forward. Although written number and calculation were developed for the purposes of trade and taxation, an important application is in the marking and measurement of time for calendrical uses. These include astronomic, ritual and social regulation functions. Some philosophers such as Kant and Brouwer have suggested that the human experience of the passage of time provides a basis for counting and numbers. The idea is that numerical succession (n followed $n + 1$) models the before-after ordering of time (tick followed by tock). So one might say that time is included in mathematics in metaphorical form, because any ordered sequence serves as an analogue for time. This includes the sequence of Natural numbers (1, 2, 3, …) as well as mathematical proofs represented as sequences of deductive steps $D_1, D_2, D_3, \ldots, D_n$, where the Ds stand for the n steps in a mathematical proof of D_n.[1] Proofs embody, in their progress towards the final result, directionality and development, an analogue of time.

A breakthrough in the treatment of space and time came with Descartes' invention of coordinate geometry. This combines a number of spatial dimensions

---

[1] Each proposition D_i, for $0 < i < n + 1$, is deduced from D_{i-1} or other preceding propositions in the proof, or is an axiom. Note that the same analogy applies to a calculation represented as $t_1, t_2, \ldots, t_n$, where the ts are the successive terms in a calculation and t_n is the final answer.

with numerical measures along axes enabling the unique specification of locations (points), as well as lines and planes in mathematical space. For example, the equation $x = y$ describes the diagonal line that goes through points (1, 1), (2, 2), (3, 3), etc. If time is chosen as one of the axes, then time is represented graphically and is explicitly related to other variables, such as distance from the origin (the central point (0, 0) in the Cartesian plane), by means of algebraic equations.

The invention of the calculus by Newton (and Leibniz) provided a model of dynamical functions representing speed, acceleration of bodies in space, etc., within the framework of the Cartesian model of space-time. This allowed Newton to formulate his important theories of motion and gravity. Einstein extended this model to four dimensions to include both space and time, and by combining it with the new non-Euclidean geometries was able to formulate his theories of relativity. These are still accepted as the best current cosmological theories. The extension to n-dimensional space enables the framework to represent *avant garde* developments such as String Theory with its 11 dimensions. Of course mathematics is not limited by finitude as the nearby physical world is, and the infinite-dimensioned spaces of mathematics are now part of the toolbox that science can employ to formulate new theories of space-time.

Thus since its inception, mathematics has provided the conceptual framework for physical theories of space-time. Indeed these cosmological theories from Ptolemy and Copernicus to Einstein and String Theory could not be conceived let alone formulated without the ideas and systems of mathematics. The underpinning mathematical theories have often been developed as internally driven exercises in pure mathematics, with no idea as to possible applications. For example, with the invention of non-Euclidean geometries the great mathematician Poincare was able to say that at last a branch of mathematics had been invented that could never be applied. Ten years later, in 1905, Einstein published his special theory of relativity, the greatest breakthrough in physics since the time of Newton. Relativity is based on non-Euclidean geometry, and could not exist without it.

Because of its dual role mathematics has been described as both the queen and servant of science [1]. The historical development of the mathematical ideas of space and time suggests that virtually any idea dreamed up by mathematicians can be adopted by science as an underlying framework for their theories. As I write this, perhaps some brilliant young and as yet little known physicist is developing a new cosmological theory of space and time based on set theory, infinite dimensional geometry, algebraic cohomology or some other abstruse pure mathematical theory. As queen of the sciences mathematics conjures up pure crystalline realms of intricately interwoven structures that stretch off to infinity like diamond spangled spiders' webs, the Web of Indra.[2] As the servant of science mathematics provides

---

[2] The Web of Indra is an image from Hindu scripture of a web that stretches to infinity in all directions. It has no centre and its strands intersect and at each intersection there is a thousand-facetted jewel that reflects the whole web.

the basic toolbox and language from which scientific theories are built, including the basic frameworks of space and time.

However, there is a paradox in exploring space-time from the perspective of mathematics. For although mathematics provides the language in which theories of space-time are formulated it is itself understood to be outside of space and time. Geometry is the science of space, and yet is beyond space. In this respect it differs from science, for scientific theories are understood to be part of the world that they describe. Whether located in the minds of scientists, shared culturally in the community of scientists, present in published texts, or some combinations of these three, scientific theories are understood to be part of the physical world. Admittedly, a few philosophers of science such as Popper [25] disagree with this perspective and locate scientific theories in some objective realm beyond space, time and the physical world. However, this is a minority view. In contrast, within the mathematics community the positing of mathematical objects and theories as existing in some non-physical world beyond space and time is the standard view. It is termed Platonism or mathematical realism and refers to existence in an ideal world beyond our physical world or reality. In this respect it differs from the use of the term 'realism' in the physical sciences which concerns the empirical world.

For the scientist the paradox of geometry is not a paradox at all. Geometry and other mathematical theories are merely tools available to hand for application, and are in this world just like any other scientific theories. However, for many philosophers of mathematics mathematical objects and theories are ideal objects not of this world, yet they give us knowledge of all the possible forms of space and time including those used in current cosmological theories. Thus the paradox belongs to the philosophy of mathematics, not to the philosophy of science. I shall suggest a solution to this paradox below by adopting a social constructivist philosophy of mathematics. This undertakes the Promethean task of bringing mathematical objects and theories back down to earth.

20.2 Mathematics and the Limits of Human Understanding

Because of our nature, human understanding is finite and limited. After all we are higher primates and our brain capacities although large are necessarily finite and limited. Consequently, so is our knowledge and understanding. We inhabit the Earth, which is but a tiny speck in an unbelievably vast universe. Beyond this vastness, modern science postulates infinitely many parallel universes existing alongside our own, which remains the only one we can know anything about. That part of the universe that we can observe is tiny, and the time span we experience is so very short, even acknowledging that the light that reaches us from distant galaxies was emitted millions of years ago.

From a theistic position, what we poor finite creatures can come to know in an infinite universe created by an infinitely complex god is necessarily miniscule and subject to all the errors of our limited and fallible being. Even if god has created a universe of regularities and patterns we can never be sure we have uncovered them with our bounded, finite knowing.

From an atheistic perspective, which is my own position, there is no conscious author of our world. So we have even less reason to believe that our universe rests on necessary laws and universally invariant and intelligible structures. Whatever their faith, there is a consensus among scientists and philosophers of science that our best scientific theories must always remain conjectures that can never be proved true. Confirmation of our theories enables us to keep on using them tentatively, and only falsification provides any certainty about our theories. Unfortunately this is the certainty that our theories are wrong and that we need to look again for new working theories of the world. As biological creatures produced by evolution and chance, we have amazing capacities for making meanings and tools.[3] But there is little reason to believe that we can uncover absolute truths about reality. To believe otherwise is commit the sin of hubris, to gravely overestimate our capacities as knowers. It would exaggerate our biologically limited powers.

A more humble view of human knowing capacities finds validation in modern scientific theories. Relativity theory tells us that there is no absolute frame of reference in space, and that space-time-mass work together to shape what currently appears to be the non-Euclidean space we inhabit. Quantum theory tells us that there is a built-in uncertainty in measures of distance and momentum, enshrined in Heisenberg's *uncertainly principle,* so that our knowledge of the physical world can never be exact. Nor can we predict the future as those who subscribed to a mechanical model of the universe in the past mistakenly believed. Paradoxes concerning the dual wave-particle nature of matter remain unresolved, and Einstein worked on these paradoxes without resolution throughout the second half of his life, so it seems unlikely that they will be resolved in the foreseeable future. Thus our current theories of the physical world have the limits of our possible knowledge and knowing built-in. Of course these theories are themselves only well confirmed conjectures, but so too must their unknown successors be, in the future.

Does mathematics offer a firm bedrock, an island fortress in this sea of uncertainty? Mathematics is believed by many to offer objective truths that apply universally, forced by logical necessity on any rational being, no matter how remote such beings might be from us in space and time. However, there is an outstanding controversy in mathematics and its philosophy concerning the certainty of mathematical knowledge and what it means. The traditional absolutist view, going back to Plato, contends that mathematics provides infallible certainty that is both objective and universal. According to this view, mathematical knowledge is absolutely and

---

[3] Of course Darwin's theory of evolution, based on the idea that we are biological creatures produced by evolution and chance, may also be jettisoned in favour of a better biological theory sometime in the future. But this does not weaken my claims about the fallibility of human knowing.

eternally true and infallible, independent of humanity, at all times and places in all possible universes. So when correctly formulated, mathematical knowledge would be forever beyond error and correction. Any possible errors in published results would be due to human error, carelessness, oversight or misformulation. From this perspective certainty, objectivity and universality are essential defining attributes of mathematics and mathematical knowledge.

In contrast, there is a more recent and alternative ‘maverick’ tradition in the philosophy of mathematics according to which mathematical knowledge is humanly constructed and fallible [17]. This tradition includes the perspectives known as fallibilism [11, 21] and my own position of social constructivism [6]. As Lakatos [20, p. 184] puts it: “Why not honestly admit mathematical fallibility, and try to defend the dignity of fallible knowledge from cynical scepticism, rather than delude ourselves that we shall be able to mend invisibly the latest tear in the fabric of our ‘ultimate’ intuitions.”

The maverick tradition in the philosophy of mathematics rejects the claim of the absolute and universal truth of mathematical knowledge [5, 6, 11, 27]. It argues that mathematical knowledge does not constitute objective truth that is valid for all possible knowers and all possible places and times. In this tradition, the certainty of mathematical knowledge is acknowledged, but the concepts of certainty and objectivity are circumscribed by the limits of human knowing.

This ambiguity in the term certainty is best understood in terms of the concept of objectivity. On the one hand, what may be termed *absolute objectivity* refers to knowledge that is validated in the physical world as a brute fact verifiable by the senses, or in the domain of non-empirical knowledge by dint of logical necessity. Furthermore, the logic underlying such necessity is itself guaranteed to be absolutely valid and above and beyond any conceivable doubt throughout all possible worlds and universes. On the other hand, what may be termed *cultural objectivity* refers to knowledge that has a warrant going beyond any individual knower’s beliefs and is thus objective in the sense opposite to subjective. Laws, money and language are culturally objective because their existence is independent of any particular person or small groups, but not of humankind as a whole. These two meanings are not the same because, for example, mathematical objects and truths might exist in the social and cultural realm beyond any individual beliefs, thus being culturally objective, without having independent physical existence or existence due to logical necessity, that is, being absolutely objective.

In the second, cultural sense, objectivity is redefined as social, as I argue in Ernest [6]. There I extend the social theory of objectivity proposed by Bloor [2], Harding [10], Fuller [7], and others.[4] This cultural sense of objectivity is how social constructivism views mathematical objects and truths. This perspective has a strong bearing on a discussion of the limits of human understanding and knowing because it posits that mathematics and mathematical knowledge are wholly located in the

---

[4] The social theory of objectivity is also paralleled by the theory of social reality put forward by Searle [26] and his followers in the philosophy of mathematics such as Cole [3].

cultural domain and are, at least in part, contingent on human history and culture. Social constructivism does not, however, accept that what any group of people accept as true is necessarily true. Establishing mathematical truths requires proofs using accepted forms of reasoning. The methods, rules and criteria used in proofs have been built up over time by a community that is open to criticism and is self correcting. We can have faith that accepted mathematical results have been established with certainty, or at least as close to certainty as humans can get. And where flaws are uncovered, as was done in the great mathematician Hilbert's criticisms of Euclid's geometry, modern mathematicians rectify them, as Hilbert did.

The controversy between the traditional absolutist philosophies of mathematics and the maverick philosophies can be largely captured in terms of these two concepts of objectivity. One consequence is that both of these schools can be said to acknowledge the certainty of mathematical knowledge although the meaning differs according to the interpretation of 'objectivity'. Mathematical knowledge consists of those mathematical propositions that are objectively warranted as true or logically valid, and hence can be claimed to be known with certainty.

20.3 The Limits of Mathematical Truth

The key issue concerning mathematical knowledge is that of its truth. What is the truth status of mathematical knowledge including the theorems of mathematics? To what extent is mathematical knowledge true with certainty? To answer this question it is necessary to examine the warrants for mathematical knowledge that establish their truth. The immediate warranting of mathematical truths by intuition will not do, for two reasons. First, intuition cannot be shared with others as a public warrant for truth. Second, intuition varies between individuals as well as by time, place, culture, and soon, so it is not a reliable warrant. So warrants for the truths of mathematics must be provided via reason or proof. If reason in the form of proof is used, then to establish the truth of mathematical knowledge with certainty the following conditions are needed:

1. A set of true axioms or postulates as the foundation for reasoning;
2. A set of rules (and practices) of proof with which to derive truths;
3. A guarantee that the rules of proof are adequate to establish all the truths (completeness); and
4. A guarantee that the rules of proof are safe in warranting only truths (consistency).

However, each of these conditions raises problems.

1. It is not possible to warrant a starting set of axioms or postulates as true directly, as this leads to an infinite regress [20]. Some assumed truths are required as a starting point in any proof, and intuition is not enough to guarantee their truth; it

has misled some of the greatest minds in history. So the axioms and postulates must be assumed, and mathematical proofs take on a hypothetico-deductive form. That is, theorem T is true provided assumptions A are true: A entails T. This is acceptable but it means that mathematical truths are not absolute but relative to the set of assumptions made. By saying that ‘A entails T’ is our truth rather than ‘T’, we circumvent this, but it means that we have offloaded the assumptions on which mathematical truth rests onto the system of proof itself, the rules and practices of mathematical deduction and proof. These latter, rather than axioms or postulates, are assumed to be the certain foundations for mathematical knowledge.

2. Current mathematical practices exhibit a variety of accepted reasoning and proof styles. Published proofs are accepted as valid by communities of mathematicians, based on professional expertise rather than explicit rules. In addition, different mathematical specialisms require different proof styles and levels of rigour [19]. None of the proofs published are fully explicit or fully rigorous, even with respect to the relevant sub-discipline standards [18]. My further claim is that they cannot be made so. Published proofs cannot be translated into fully rigorous formal proofs [6]. Even if they were so translatable, because of their sheer size rigorous formal proofs could not be checked for correctness with any guarantee of certainty [23]. Thus the rules of proof provide practical certainty, relative to the prevailing accepted proof practices, rather than any absolute level of certainty.
3. It is accepted following Gödel’s [9] first incompleteness theorem that in any but the simplest mathematical theories the rules of proof are inadequate for establishing all of the relevant mathematical truths, providing they are self-consistent. Gödel proved this by constructing a true but unprovable sentence, a formal version of “This sentence cannot be proven”. If it is true is unprovable. If it is false then it can be proven, leading to a contradiction. Since sentences like this can be defined in most mathematical theories, provided they are consistent there are unprovable truths in them. Thus mathematical provability does not capture mathematical truth and indeed falls some way short of it [24].
4. It is also well known following Gödel’s [9] second incompleteness theorem that no guarantee can be given that rules and procedures of proof are safe in warranting nothing but truths of mathematics. It is not possible to prove the consistency of any sufficiently complex formal theory using only its assumptions and axioms. Further assumptions are required for such a proof, and these are assumptions that exceed those of the formal theory to be safeguarded. For example, Gentzen [8] proved the consistency of Peano arithmetic but he needed to use transfinite (infinite) induction. This might satisfy a working mathematician, but philosophically it remains problematic. Using ‘intuitively obvious’ non-finitistic principles in addition to some of the axioms of the system and logic means that what has to be assumed exceeds that which is safeguarded. This cannot be described as an absolute guarantee of consistency and hence of safety.

What this shows is that the truth of mathematical knowledge cannot be shown with absolute certainty. It cannot be said, free from caveats, that mathematical knowledge is absolutely true. Proofs in general constitute the strongest evidence for the certainty of mathematical knowledge, and mathematical proof lies at the heart of claims of mathematical certainty. It is the strongest weapon in the armoury used to persuade others of the certainty of mathematical knowledge. But proof cannot guarantee the absolute truth of mathematical knowledge.[5]

The conclusion that mathematical knowledge cannot be claimed with absolute certainty is agreed by many leading mathematicians, logicians and philosophers, scores of whom are cited in Ernest [5, 6] and Lakatos [22]. Furthermore, in addition to the Gödel theorems mentioned above, there are a number of other results about the limits of what can be known within mathematics. For example, Tarski proved that truth itself is indefinable within any mathematical theory, on pain of contradiction. The Lowenheim-Skolem theorem about the uncontrollable size of models of theories offers the paradoxical result that even the largest uncountably infinite sets can be fully and consistently interpreted within countable sets.[6] Church's theorem reveals that there is no decision procedure to determine whether an arbitrary proposition in arithmetic is true or false. Some attempts to so determine the truth of a claim may continue endlessly without ever reaching a result either way. Craig's Interpolation Theorem shows the impossibility of expressing proofs in any ultimate, final canonical form, because a further relevant step can always be inserted between any two steps in a proof.

Rather than signs of weakness these results should be celebrated as advances in knowledge that come from an understanding of the limitations of human knowing. Modern physics has well known analogues including general relativity and quantum theory mentioned above, showing the relative and uncertain nature of our knowledge. By showing the limits of our knowing such results and theories represent advances in our knowledge.

In mathematics, as our knowledge has become better founded and we learn more about its basis, we come to realize that the absolutist view of perfect knowledge is an idealization, a myth. This new understanding represents an advance in knowledge, not a retreat from what was viewed in the past as a stronger position of absolute certainty.

---

[5] It might be argued that I am assuming the truth of Gödel's incompleteness theorems to argue for the fallibility of mathematical knowledge and truth. But virtually all mathematicians of all philosophical camps accept these results as certain, in the cultural sense. If the objective certainty of mathematical results like Gödel's theorem are challenged on general grounds, rather than because an error is discovered in the proof, then my case is made for me.

[6] Uncountable sets have more elements in them than can be counted using the usual counting numbers, whatever order we try to count them. For example, the real numbers (representable as infinitely extended endless decimals) are uncountable. Cantor's diagonal theorem shows that any attempt to enumerate an uncountably infinite set demonstrably leaves out many elements. So the Lowenheim-Skolem theorem results in the paradox that uncountably infinite sets, no matter how large, can be fully represented and accommodated in a countably finite model.

The fact that mathematical certainty and objectivity are socially and culturally defined, does not mean that the fact '1 + 1 = 2' has the same status and objectivity as 'Desdemona is the wife of Othello' or that 'Paris is the capital of France'. The fact that 2 denotes two has the same status as these two contingent truths, because names are arbitrary. But the fact '1 + 1 = 2' is not arbitrary, because in arithmetic 2 is defined as the successor (1') of 1, and 1 + 1 = 1' by definition of '+'. In other words, it follows by logic. Similar but more extended arguments can be used for '5 + 7 = 12', Pythagoras' Theorem, or Gödel's first incompleteness theorem. Namely, that there is a logically deductive chain that proves the statement. Such mathematical truths can be known with certainty, cultural certainty that is, unlike above the contingent facts from Shakespeare and geography.

Relinquishing the hope of absolute certainties in mathematics does not represent a loss of knowledge [18]. It was Vico's (1710) great insight that we only know with certainty that which we have made. Mathematics is one of the greatest of human inventions, constructed and reconstructed over millennia. Our certainty in mathematics resides in knowing what we have constructed. Overall, what this means is that we have to be circumspect in what we claim in terms of mathematical certainty and objectivity. Mathematical knowledge is known with certainty within the bounds of what can be humanly known. Mathematical knowledge and mathematical objects are objective, but in the circumscribed sense of cultural objectivity. We do not have to relinquish the certainty and objectivity of mathematics. We merely have to be more circumspect in what we claim by ascribing certainty and objectivity to mathematics. Thus mathematics is a central case in establishing the limits of human knowing.

20.4 Knowledge Versus Understanding

In discussing the limits of human understanding I have focussed on knowledge, both scientific knowledge made up of theories and observational facts, and more centrally, mathematical knowledge, which is made up of theorems and theories.[7] Although there is an overlap of meaning between knowledge and understanding, there are also differences. The emphasis with knowledge is epistemological, concerning not only the knowledge claim itself but also the warrant for the claim. Knowledge is justified true belief. Understanding, however, concerns meaning, sense making and explanation. Thus as well as epistemology it concerns rhetoric, how explanations are formulated, as well as communication and psychology, what sense is made of propositions and theories and how well they relate to or explain other phenomena. Thus one might say that human understanding encompasses more than knowledge, understood in the epistemological sense. Within

---

[7] Mathematical theories are deductively structured collections of axioms and theorems, such as, for example, Euclidean geometry.

knowledge-related representations we can distinguish between information, belief and knowledge itself. Information is made up of representations, with signs designating state of affairs. The designations may be the outcomes of individual or group intentions, and information per se has no epistemological status beyond being a representation. Beyond this, beliefs not only designate some state of affairs but also imply a commitment to their validity by their holders or proponents. Lastly, knowledge representations are beliefs that are warranted as true. Thus there is an increasing epistemological strength gradient from information through belief and then on to knowledge. Likewise, the pool of representations diminishes as we pass from information through beliefs to knowledge, because all beliefs might be said to be information representations plus commitment, and all knowledge representations are beliefs plus warrants and truth status.

Understanding encompasses all of these types, including information, belief and knowledge. What we are acquainted with, believe or know comprises a wide variety of suspicions, guesses, rumours, conjectures, desired and inferred connections, as well as more robust facts, generalizations and theories. Understanding represents connected and better or worse buttressed claims, ideas, concepts and knowledge. Thus human understanding is limited in at least two ways. First, there is the epistemological warrant for what we believe or claim we understand. Even the best warranted human knowledge has a certainty that is circumscribed by the limits of our knowledge. Beliefs are weaker in only being guaranteed the commitment of individuals or groups, and information has no guaranteed warrant at all.

Second, there is the breadth of representational types we weave together in our understandings of anything and everything. Although information and beliefs add to our repertoire of representations, there is no reason to believe that we have access to all possible representations of what can or might be understood. We are finite and limited in our understanding. We can never know whether all that makes up reality, whatever that might be, is susceptible to representation, making it accessible to human understanding. There is no reason to think that the universe, all that exists, is neatly bundled and labelled so that it is susceptible to human representation and understanding. Indeed, given the limitations of our scientific and mathematical knowledge acknowledged above, there is every reason to believe that like knowledge, human understanding is limited and fallible. Why should all that exists, including all the patterns and connections that might possibly be conjectured but have not been formulated or posited be discernable and comprehensible to human understanding? Although we can distinguish knowledge and understanding, they share similar limitations, they are both bounded. Even our best knowledge and understandings are limited, constrained within the boundaries of what is humanly representable, understandable and knowable.

20.5 Where Is Mathematics? When?

There is an irony in considering mathematics in the context of space-time and the limits of human understanding. For although mathematics provides the language in which theories of space-time are formulated it is widely regarded itself to be outside of space and time. From the time of Plato mathematical concepts, objects and knowledge have been regarded as existing in a timeless world of their own, quite distinct from the material world we live in. The argument is that a mathematical circle is an ideal object and that no circle we draw on earth can be such a perfect and ideal object. Thus mathematical objects are not to be found in everyday or scientific space, but somewhere else. Likewise, mathematical objects are perfect and beyond the ravages and decay that time brings. So where is mathematics to be found? Which space, albeit metaphorical, do the problems, truths, theories and the objects of mathematics inhabit? If they exist in an ideal other space disconnected from us, how can we know them and how can they influence our world? But if somehow they exist in the material human world how can they be ideal and perfect? This is the paradox of mathematical existence.

My solution to this paradox is to say that the objects and truths of mathematics exist in cultural space, they are part of social reality. We meet and come to understand and believe in social reality through our interactions with other people and with institutions. Language and other signs are used to point to and describe the elements of social reality, but beyond this, they create social reality [26]. Through conversational exchanges with other people we learn about the unvarying properties of the objects that exist in social reality. This covers laws, customs, property, money and countless other things. But my concern here is with mathematics. We learn, for example, that a quadrilateral is a plane figure with four straight line sides. We have objects in the environment around us pointed out as approximately quadrilateral in shape, including squares, rectangles, parallelograms, trapeziums, etc. We learn that most things with quadrilateral shapes maintain such shapes, and that there is general agreement in the identification of shapes as quadrilaterals. If we misuse such terms, such as claiming that oblongs have rounded corners, we are soon corrected by those more expert than ourselves. The rough edges of our concepts and their definitions are rounded off, through conversation. We learn about elements of social reality such as these, and all the many others, through conversation, even acknowledging that in formal studies these are largely one-sided conversations. But all formal studies include learner responses via the media of written work submitted and answers produced in testing. In such assessments we show what we know, and where we misunderstand properties of the mathematical objects found in social reality, we are corrected. If we do not take these corrections on board we fail to progress to the next level of study. Thus there are safeguards in the form of social institutions that ensure that we learn the same standardized meanings and properties, in short, the same mathematical objects. Even when we have idiosyncratic personal associations in our understandings, social mechanisms ensure we use the terms describing mathematical objects in the same way. This is a very simplified account of the mechanisms that ensure that the objects of social reality are the same for all of us, and these include the objects of mathematics.

Although mathematical objects appear to be universal and timeless, existing beyond the material human world, this is an illusion. Mathematical objects are social or cultural objects that have to created anew in the mind of any learner. Irrespective of whether or not mathematics acquisition is hard-wired into our brains, it takes an extended exposure to mathematical representations, tasks and conversations for mathematical objects to populate our imaginations. Repeatedly approaching and using representations of mathematical facts, concepts and rules in different ways but with an underlying standardization and constancy convinces us of the invariance and objectivity of mathematical objects. Cognitive psychologist have documented the human capacity for mentally reifying mathematical processes into more abstract mathematical objects, and we observe this across the years of learning. So the gradually increasing abstraction of the mathematical concepts we encounter across the years of statutory schooling results in personal universes populated with abstract processes and entities, and these are of course the mathematical objects under discussion.

Thus the paradox of the position of mathematics in space and time is at least partially resolved. Mathematics is a set of representations and practices embedded in human culture, in social reality. Although mathematical knowledge and objects appear to exist beyond time and space this is an illusion. It takes time for the knowledge and objects of mathematics to develop. We all acknowledge the extensive period of time it took for mathematics to develop historically. The invention and discovery of all of the concepts, problems, results and symbolism of mathematics can be located in specific historical episodes. Sometimes within an already elaborate discipline of mathematics the problems and methods of mathematics have cried out for certain solutions, connections and new abstractions, leading to simultaneous discovery of new ideas and methods, such as the concept of zero, or the calculus [29]. Paralleling but not recapitulating the disciplinary development is the growth of mathematical knowledge in the learner. This takes over a decade of immersion in mathematical activities totalling thousands of hours and thousand of tasks. Once the objects of mathematics have become real for us we forget the long processes, both historical and personal, that were necessary for us to build our understandings. It is illusory for us to ask where mathematics in the geographical sense. Looking for mathematics in some other world is also illusory unless we understand this other world as social reality. Of course the Platonist view that there is an ontologically separate and independent ideal world populated by mathematical objects is irrefutable, just as is an imagined heaven populated by millions of angels. Such answers posit imaginary places where mathematics (and angels) are found. But they leave unanswered the question of how we access them, how they become real to us, at least in this mortal life.

20.6 Personal Time and Space

All sentient beings including humans live in the moment, in the here and now. This is the bedrock, the fulcrum, the pivot point of our lived experience. All else is unreal and secondary, either summoned from memory or imagined for the as yet unreal future. Living in the here and now is the clarion call of some forms of the spiritual life, such as Buddhism. This is illustrated in Aldous Huxley's [13] book *Island*, a story of about awakening consciousness in which a Minah bird flutters from tree to tree calling "here and now", "here and now" to remind the characters to be fully awake in their daily lives. But *here and now* really is the originary point of all thought, all consciousness and all action. Like the origin point in the Cartesian plane, (0, 0), or even in the Einsteinian four dimensional space-time, (0, 0, 0, t), this is the reference point from which all flows or with respect to which all is measured. Weyl and Rotman liken the Cartesian origin point to the vestigial sign of the self or ego in mathematical space. But what is the actual link between this moment, the here and now, and the worlds of mathematics, of science and indeed of all human knowledge and understanding? The link is the human imagination. Consciousness allows us in the here and now to weave rich and ornate tapestries of the imagination. Through weaving the here and now back and forth as a single strand through such a rich many layered cloth we weave vast empires of the imagination. These may be intricate castles floating in the air, but they are shared castles in the air, for we join with others, through conversation, to weave our individual strands, the passage of our private present moments, our 'here and now's into the great shared tapestry of human knowledge. These make up the 'great conversation' of human culture and ideas [12]. Our imaginations spin our webs of thought moment by moment into the rich interlocking web of ideas, more complex than any spider's web, and this makes up our shared knowledge and understanding.

Here we have another paradox. Through the tiny atom of experience, the moment of consciousness, emerges the vastness of all human understanding and constructed artefacts. Like the tip of a needle, coursing inch-by-inch, up and down, in and out, within the tiniest of times and spaces, moment by moment, emerges the vast tapestry of human knowledge and all the material constructions that surround, shape and enable us to live, eat, travel and work in comfort. Imagination and planning string these moments of endeavour together, like beads on a never ending necklace, making the vast intricate cathedrals of human achievement. Many thinkers have used the experience of spatial constraint, with almost unlimited time, to create some of the great cultural artefacts. Cervantes wrote Don Quixote while imprisoned, as did Dostoevsky with some of his greatest novels. Descartes claims to have thought best within the confines of a heated oven, and Gramsci developed his deep political philosophy in his Prison Notebooks, before his execution. Russell wrote An introduction to mathematical philosophy in prison. Some great thinkers lost their visual perception of space, going blind, but went on to create great works, like Milton's Paradise Lost and Beethoven's late symphonies and string quartets, although confined by the denial of visually perceived space.

Here the analogy between space and time breaks down. Although mystics have written of a moment of enlightenment, in which all things are seen as joined together as in Indra's net, and a golden glow of love suffuses the entire universe, such an experience is not something that can be shared. The intensity felt in the moment can inspire subsequent life choices and indeed eloquent communications. But it can only be sketched dimly through analogy, metaphor, parable, pictures, and so on, that can at best depict some adjunct features of the experience. They can never capture its transcendent core, the feeling as the spiritual self overflows the bounds of the ego. The Aha! moment, Satori, Nirvana or whatever else the instant of enlightenment is called, takes place in time, but is beyond time; it takes place in space, but is beyond space. You do not have to believe in god to accept that such experiences real events in human consciousness, and that the material basis of mind does not preclude human spiritual experience.

Among our shared understanding, underpinning its elaborate tapestry of woven ideas and knowledge, even framing its possibilities, according to Kant [16], are our ideas of space and time. Our consciousness of the here and now contains within it the fundamental categories of space and time, we could not experience or perceive anything unless it was framed by and presented through the categories of space and time. In addition, our consciousness and experience also contains within it distinction between space and time. Consider, the bedrock of our experience: it is in and of the moment. But the moment that we experience is not an instant, it is not the imaginary zero magnitude point, so useful a fiction in mathematics and science. This moment has a definite duration of a few seconds, and in our consciousness of that moment we experience both succession and multiplicity. Succession, as I discuss above, gives us the awareness of time and number. It is progression across time. Simultaneity, as in our experience of multiple objects in the realm of our senses, gives us the awareness of space. For a multiplicity that is simultaneously present must not only be an array of distinct objects, but logically these must also be spread out in one or more dimensions of space in order for us to perceive them simultaneously. Thus a fundamental distinction emerges from our awareness of the moment, a dichotomy between our awareness of succession and of simultaneity. These concern the serial relationships of succession (one after the other) versus the parallel functioning of simultaneity (the coexistence of different parts or processes). These give rise, in their basic forms, to the distinction between geometry and arithmetic. This is a deep distinction, resting, as it does on our different and distinct experience of space and time.

Analogous distinctions are in use in many areas of knowledge, such as linguistics. There is diachrony (historical disposition, with events spread in succession over time) versus synchrony (simultaneous events). There is the syntagmatic, the relationship between them in successive sequence of signs, versus paradigmatic, in which signs are contrasted with other substitutable signs [4]. As literary tropes there is also metonymy, which works by the contiguity between two terms or concepts, including taking an attribute to stand for the whole, versus metaphor, which is based upon the analogous similarity of two concepts. It has been argued that different portions of the human brain are devoted to metonymy and metaphor [14].[8] So the difference between space and time, if it is not too much of a stretch to point to these parallel dichotomies as evidence, is a fundamental and deep-seated duality that runs through almost the whole of human ideas, including mathematics.

Kant's great insight about consciousness and indeed all knowing and knowledge, is that it is conditioned and constrained by inbuilt cognitive and epistemological limits. He argued that the categories of time and space are part of the conditions of cognition, and are not something that we experience.

> Space is not something objective and real, nor a substance, nor an accident, nor a relation; instead, it is subjective and ideal, and originates from the mind's nature in accord with a stable law as a scheme, as it were, for coordinating everything sensed externally. [15, p. 403]

Every experience takes place within space and time, but is not of space and time. We are unaware of the airy space that surrounds us just as fish are unaware of the sea. Only when a change of pressure occurs, when a current within the medium brushes against us, stimulating our sense of touch, do we become aware of the 'airiness' (or 'wateriness') of our medium. Space is the invisible medium throughout which tangible objects are present and arrayed.

20.7 Conclusion: The Limits of Human Understanding and Knowledge

In this chapter I explore the paradoxes that emerge from the mathematics of space and time with a focus on the limits of human understanding. For example, I consider how mathematics provides the language of theories of space-time, and yet is itself understood to be outside space and time. Geometry is the science of space, and yet is itself beyond space. Mathematics is believed by many to transcend the limits of human understanding by expressing truths that apply beyond and across all space and time. Mathematical knowledge is understood to be forced by logical necessity on all conscious beings. But mathematics is itself human knowledge, and the history of its creation is there for all to see. So how can it exceed the knowledge capacities of its authors?

I suggest a resolution to this and the other paradoxes. But this solution depends on the Promethean task of bringing mathematics from its remote Platonic heaven back down to earth. My proposed solution is a social constructivist one. Namely, that mathematics is humanly created, and it rests on a social reality. Thus mathematics is not otherworldly but is embodied in the personal knowledge of persons, in cultural practices and in texts containing symbols. These are part of what help

---

[8] The eminent French psychoanalyst Jacques Lacan also argues that the unconscious mind has the same structure as language, and that Freud's mechanisms of condensation and displacement in the unconscious are equivalent to the poetic functions of metaphor and metonymy, respectively.

support and create its social reality, although mathematics is not reducible to any combination of these. Social reality is real and it supports and constrains our understandings and actions. But it does not exist apart, beyond space and time.

There is another paradox in considering mathematics in the context of space-time and the limits of human understanding. Drawing on a social constructivist perspective I argue that mathematical knowledge cannot exceed the knowledge capacities of its authors. The certainty and objectivity of mathematical knowledge is bounded by the limits of human knowledge and understanding. The solidity of the objects of mathematics is a social reality which is made real by their constant properties which are consistent throughout the multiple human presentations and representations of them. This offers one solution to the mystery of how mathematical knowledge can be pure and exist in a superhuman domain, yet be so 'unreasonably effective' in its worldly applications [28]. For mathematical concepts and knowledge originate in our attempts to understand the world, and our attempts to weave patterns to describe our experiences. We abstract and generalize the concepts, symbolic processes and patterns that we so derive, and should not be amazed that they fit our human experiences when we reapply them to nature.

Reflections on the mathematics of space-time reveals the great breadth of our knowledge, but examined closely also bring to light a number of paradoxes. A final irony is that mathematics, the most certain and infallible of all of the disciplines of knowledge, is, its certainty notwithstanding, beset with paradoxes and uncertainties that very well show up very strikingly the limits of human understanding.

References

1. Bell, E. T. (1952). *Mathematics queen and servant of science*. London: G. Bell and Sons.
2. Bloor D. (1984). A sociological theory of objectivity. In Brown SC. (ed.) *Objectivity and cultural divergence, Royal Institute of Philosophy lecture series* (Vol 17, pp. 229–245). Cambridge: Cambridge University Press.
3. Cole, J. (2013). Towards an institutional account of the objectivity, necessity, and atemporality of mathematics. *Philosophia Mathematica (III), 21*, 9–36.
4. Crystal, D. (1987). *The Cambridge encyclopedia of language*. Cambridge, UK: Cambridge University Press.
5. Ernest, P. (1991). *The philosophy of mathematics education*. London: Routledge.
6. Ernest, P. (1998). *Social constructivism as a philosophy of mathematics*. Albany, New York: State University of New York Press.
7. Fuller, S. (1988). *Social epistemology*. Bloomington, Indiana: Indiana University Press.
8. Gentzen, G. (1936). Die widerspruchfreiheit der reinen zahlentheorie. *Mathematische Annalen, 112*, 493–565.
9. Gödel, K. (1931/1967). On the formal incompleteness of arithmetic (J. van Heijenoort ed) (1967) From Frege to Gödel: a source book in mathematical logic, pp. 592–617. Cambridge, Massachusetts: Harvard University Press.
10. Harding, S. (1986). *The science question in feminism*. Milton Keynes: Open University Press.
11. Hersh, R. (1997). *What is mathematics, really?*. London: Jonathon Cape.
12. Hutchins, R. M. (Ed.) (1959) The great conversation (Great Books of the Western World, vol 1). London: Encyclopaedia Britannica.

13. Huxley, A. (1962). Island. New York: Harper & Brothers.
14. Jakobson, R. (1956). Two aspects of language and two types of aphasic disturbances. In: R. Jakobson, M. Halle (Eds.) Fundamentals of language (pp. 55–82). The Hague: Mouton & Company.
15. Kant, I. (1770/1992). On the form and principles of the sensible and the intelligible world (D. Walford, R. Meerbote) Inaugural Dissertation in Theoretical Philosophy (pp. 1755−1770). Cambridge: Cambridge University Press.
16. Kant, I. (1781/1961). Critique of pure reason (Smith NK, Trans.). London: Macmillan.
17. Kitcher, P., & Aspray, W. (1988). An opinionated introduction. In W. Aspray & P. Kitcher (Eds.), *History and philosophy of modern mathematics* (pp. 3–57). Minneapolis: University of Minnesota Press.
18. Kline, M. (1980). *Mathematics: The loss of certainty*. Oxford: Oxford University Press.
19. Knuth, D. E. (1985). Algorithmic thinking and mathematical thinking. *American Mathematical Monthly, 92*, 170–181.
20. Lakatos, I. (1962). Infinite regress and the foundations of mathematics. *Aristotelian Society Proceedings, 36*, 155–184.
21. Lakatos, I. (1976). *Proofs and refutations: the logic of mathematical discovery*. Cambridge: Cambridge University Press.
22. Lakatos, I. (1978). A renaissance of empiricism in the recent philosophy of mathematics? In I. Lakatos (Ed.), *Philosophical papers* (Vol. 2, pp. 24–42). Cambridge: Cambridge University Press.
23. MacKenzie, D. (1993). Negotiating arithmetic, constructing proof: The sociology of mathematics and information technology. *Social Studies of Science, 23*, 37–65.
24. Paris, J., & Harrington, L. (1977). A mathematical incompleteness in Peano arithmetic. In J. Barwise (Ed.) Handbook of mathematical logic (pp. 1133–1142). Amsterdam: North Holland.
25. Popper, K. R. (1979). Objective knowledge (Revised Edition). Oxford: Oxford University Press.
26. Searle, J. R. (1995). *The construction of social reality*. London: Allen Lane, The Penguin Press.
27. Tymoczko, T. (Ed.). (1986). *New directions in the philosophy of mathematics*. Boston: Birkhäuser.
28. Wigner, E. P. (1960). The unreasonable effectiveness of mathematics in the physical sciences. In *The spirit and uses of the mathematical sciences* (pp. 123−140). New York: McGraw-Hill. Reprinted in T. L. Saaty, and F. J. Weyl, Eds. (1969).
29. Wilder, R. L. (1981). *Mathematics as a cultural system*. Oxford: Pergamon Press.

Chapter 21
“Now” Has an Infinitesimal Positive Duration

Reuben Hersh

Abstract This article offers the nonstandard or “hyperreal” line as a model for Time, thereby to resolve a persistent controversy of the meaning of “Now.” As a “monad” in the Leibnitzian time axis, “Now” is a time interval shorter than any standard positive interval, yet longer than any infinitesimal.

I start out with Aristotle’s *Physics.*[1] In order to separate the past and the future, a single point “now” suffices. (This automatically invokes an image, the “straight line”, and brings in the mathematics of Euclid, as modernized of course.)

On the other hand, in order to be faithful to the unbroken connectedness of experience, we require a “now” that overlaps the past and future, a positive duration for what is happening “right now.”

Automatically and uncritically we think of time as a straight line, with a positive direction, infinitely long in both directions. It’s split between the future, a ray stretching to the right, and the past, a ray stretching to the left. This pair of disjoint half-lines is separated or joined by a single point, “the present”, “the Now.”

This image, this time line, is not *Time*. Time itself is not a line, it is the general phenomenon of change. The time line is just a *model* of Time, a mathematical model. It helps us to conceptualize or visualize experience. Like any model, it is not *identical* to the object being modeled, which is the Time we live through, either in ordinary life or in experimental science.

“Now” is a primitive of human language, as when I yell at you, “Do it now!” It is not a problem in common conversation.

But when Logic sticks in its head, a problem does arise. The book edited by Durie [5] presents arguments about the Now by distinguished authors, including

---

[1] Aristotle—“Aristotle’s Physics A Guided Study” Joe Sachs, Rutgers U Press [12, pp 120–124 Aristotle chaps. 10–12].

R. Hersh (✉)
Department of Mathematics and Statistics, University of New Mexico, Humanities 415, 87131 Albuquerque, NM, Mexico
e-mail: rhersh@gmail.com

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_21

Henri Bergson [2] and Gaston Bachelard [1]. Bergson tells us the Present must have a duration. For anything at all to happen, it takes at least a little bit of time. In zero time (without duration) we'd be with Zeno [6], stuck in a "place," unable to move. But things are happening, we are living and breathing, now! For Bergson, "now" holds together the past and the future, a duration where the future becomes the past.

To Bachelard, "Now" is merely a point separating the past and the future.

Bachelard claims to settle the argument by pointing to Albert Einstein. The "event" in Relativity Theory has a time coordinate, a number. End of argument.

And in this regard, not just Einstein, but all of experimental science can be cited. Isaac Newton wrote that he was calculating velocity *at the very instant* when the falling stone hits the ground. Since the work of Newton, observable, measurable motions have been modeled as "functions" of time. We record observations using two lines. One is a time axis. The other is an axis of distance, temperature, pressure, or some other "state variable". The state variable is called a function of time, which means merely that a point on the state variable axis corresponds to some point on the time axis. The use of points or "instants" is taken for granted.

This methodology is so universal, so well established for centuries, that it is taken as an objective reality, as a real truth, not just a successful model. But it doesn't fit well with our subjective time sense, like how long we've been waiting in the doctor's office. Without bringing in any mathematics, we know which things have already happened ("are in the past"), which ones are going on now ("in the present"), and which ones haven't happened yet (they're "in the future"). That kind of "experiential time" is not quite the same as Newton's "everywhere always equable flow". Yet they are not unrelated. They have to be compatible in practice. We manage to shift back and forth between the "independent variable" of physics and the past, present and future of daily life.

There is the objective, scientific duration or time, the independent coordinate of mechanics. And over against it to be reconciled with it, is an internal, subjective sense of time, a consciousness of time passing either slowly or quickly, the flow of time that we feel even with our eyes closed in a silent room, that seems entirely different from the time by which we describe the rising and setting of the sun and the moon.

We seem to be facing a conflict between our lived Now, a brief duration where Past and Future seamlessly merge, and our scientific practice, which evaluates an observable at any given instant [3, 4, 9, 11, 13, 14].

But this opposition is delusional. It is based on the delusion of immaterial consciousness. Our minds are embedded in our bodies, they are activities or functions of brains and nervous systems which are not free floating, they are in our flesh, our muscles and guts. My heart is beating. My breath goes in and out. My stomach empties and needs to be refilled. My buttocks weary from long sitting. All this information available to my brain is somehow combined or condensed into a sense of time passing. It is time to get up, or to lie down, or to find a snack.

The coming and going of daylight, the coming and going of the seasons of the year, the aging of ourselves and others, impose their own measure of lived experience, before or without any intervention of clocks. Muscle fatigue, hunger, heartbeats and breaths in and out, all are different clocks sending signals to our brain where they are compared, coordinated, and synthesized to give us a subjective sense of time passing. How long have I been sitting here waiting? How long has it been since I have heard from so and so? As I stand I feel the passage of time in my muscle fatigue. Or as I sit, I feel the passage of time in my buttocks and back. Maybe I even get a backache. All the while of course my breath is entering and leaving my lungs and my heartbeat is continuing, perhaps at a steady pace, perhaps with speedups and slowdowns. Fatigue not only of the muscles but of the nervous system. The detached observer is a myth. We ourselves, as real observers, are embodied, with stomachs and brains.

There is keeping time, marking time, making time, losing time, but that doesn’t mean that there is an actual thing that you are marking, or keeping, or losing, or making, as the case may be.

The “Now” is an old problem in experimental psychology. How long must a sound or a flash of light persist, in order to be perceived? William James [7] called it “the specious present” and reported experiments in Germany. The empirical reality of the “specious present” is the number of milliseconds that a sight or a sound must endure in order to be perceived.

The empirical Now shrinks as we invent more powerful observing instruments.

On my desk is a calendar with illustrations from high-speed photography. A bullet is shown in mid-air, just as it has passed through an apple. It was in that position for “an instant.” But the camera shutter had to stay open long enough for the photographic chemical process to take place. No photography can be instantaneous. If Now takes some time, then how much? Milliseconds? Nanoseconds? Microseconds?

This empirical and experiential fact seems to ignore a certain logic of past and future. That little interval on the line, the Durational Now, will have boundary points where it meets the Future and where it meets the Past. The perceptive time interval must extend between two instants, a beginning and an end, and we are still left with the question, what we should mean by “now”. When does the Now begin? How long does it last? When does it end? In trying to answer these questions, we seem forced to narrrow the Now back down to a mere point.

To someone not indoctrinated in modern mathematics, not ashamed to use the word “infinitesimal”, it would be tempting to propose a duration of “Now” that is neither zero nor positive. Something in between. “Infinitesimal.”

The standard “real line” taught in school is subject to the Axiom of Archimedes:

> Any interval, no matter how short, will become longer than any other interval, no matter how long, if added to itself sufficiently many times

This axiom amounts to saying, “There is no infinitesimal,” if by “infinitesimal” we mean something so short that no matter how many times we add it to itself, the result will always be less than one unit (inch, mile, whatever you want to choose as your unit.)

Is this Axiom *true*? It is true of the standard real line, because it is part of the definition of the standard real line. Circular reasoning, irrefutable. Is there any other real line that is not the standard one? Yes! It is called the Nonstandard line! (Also called the "hyperreal" line.) It has been recognized and accepted into established mathematics since 1966, when the logician Abraham Robinson published his famous book *Nonstandard Analysis* [11].

Already at the foundation of the calculus, in the 17th century, Leibnitz and his predecessors Cavalieri and others used infinitesimals to calculate areas and volumes. To do so required a certain finesse, because they were unable to explain exactly the meaning of the word "infinitesimal". Roughly speaking, "An infinitesimal is greater than zero, but smaller than every positive number." This definition tells us that an infinitesimal is smaller than itself! Blatantly self-contradictory. Leibnitz explained that although infinitesimals do not actually exist, still we can think about them as if they did exist. In the 19th century, Cauchy, Dedekind and Weierstrass showed how to do calculus without saying "infinitesimal." We use a couple of little extra variables, usually called epsilon and delta. Against the rules of mathematics, geometers, physicists and engineers continue to think and talk with infinitesimals.

Robinson the logician was experienced in applied mathematics. He used logic to construct a continuum, a "line," where every standard number is surrounded by a little cloud of infinitely close nonstandard numbers. How did he get away with it? He had a new tool—modern formal logic, thanks to Frege, Whitehead and Russell. Mathematical logic describes mathematics formally with precisely stated symbols, grammar, and deductions. The "standard real line", that is, the set of standard numbers, subject to the Axiom of Archimedes, becomes expressed in a definite formula. Then it makes sense to talk about nonstandard numbers. An infinitesimal is a number greater than zero but smaller than every *standard* positive number. No more self-contradiction!

I propose the hyperreal or Leibnitzian line as a suitable mathematical model of time.

Nonstandard analysis is no longer a startling novelty. Half a century has passed. Nonstandard analysis is in elementary calculus books by Howard Jerome Keisler and others, that have been used successfully in courses at several universities. It is a powerful research tool. Although most people go through grad school without meeting it, it isn't controversial any more. It is a well established part of mathematics, and a respected research methodology.

In Leibniz's calculus, velocity was a ratio, an infinitesimal distance divided by an infinitesimal time interval. Robinson uses this nonstandard ratio, and obtains as a standard velocity the unique standard number infinitely close to this nonstandard ratio. In this calculation, time is being modeled by a nonstandard real line. In the present paper, I point out that this model of time is useful, not only for calculating velocity, but also for endowing the Present, the Now, with a positive duration, thereby relieving the tension between the experiential now, which has duration, and the scientific now, which must have a definite location.

The nonstandard real line has a zero, which is a standard real number, and around it a set of infinitesimals, both positive and negative. Robinson took a word from Leibnitz and named the cloud of nonstandard numbers infinitely close to any standard real number a "monad."

In a hyperreal model of time, we could choose to say, without contradiction, that past and future overlap or intersect in the present. The present would be a monad, it would be all the nonstandard points infinitely close to some standard point. "Now" would be defined as the intersection of past and future. It would have a finite positive duration. That duration would not equal any number, either standard or nonstandard, for it would be smaller than any standard number, yet greater than any infinitesimal.

The subjective "present" does not have any definite beginning or end. This feature is shared by the nonstandard monad, which does not have a first or last element.

One of the difficulties of describing time phenomenologically is the impossibility of assigning any number to the duration of the instant. The monad does not possess a numerical magnitude. As a model of any instant, including the present instant, it has positive duration, greater than any infinitesimal, yet still less than any standard. This paradoxical feature of the hyperreal model fits nicely with phenomenological introspection.

Intuitively, it makes more sense to think of past and future overlapping than to think of them as disjoint, separated by an instant (the present) through which they can never touch each other. How can the future flow into the past, as tomorrow turns into yesterday, without tomorrow ever being allowed to touch yesterday? The mystery about the future flowing into the past is an artifact of the standard mathematical model. It falls away in the hyperreal model. The future and the past overlap infinitesimally. The infinitesimal overlap of future and past is the region of actuality, of things happening and changing.

Of the three pieces of time—past, present, and future—it is the now that is really real. "Be here now!" That reality requires duration.

Miller [8] presented an exposition of the *phenomenological now*, according to Edmund Husserl. Consider how we hear a melody. The melody is a sequence of tones. We hear one tone at a time, yet we listen to the melody as a whole, not as a succession of unrelated tones. We remember the tones we have just heard, and we anticipate the next tone which we expect to come soon. Before the "primal impression", then, we have a "protention", and in passing it leaves behind a "retention". The connections between these experiential aspects of time permit us to have a connected experience of the whole melody. In fact, the same analysis applies even to our experience of each separate tone of the melody, for even a single tone has a duration, and a "primal experience", a "protention" and a "retention".

This analysis of our experience depends on our already accepting the objective fact–that one tone does really precede or follow another. This amounts to accepting the objective reality of Time, accepting that things do happen out in the World, that one happening does precede or follow another.

The difficulty is with our picture of time as a Euclidean or Newtonian line. These lines carry along with them the Archimedean axiom. "Any interval, no matter how short, can serve to count any epoch of time, no matter how long." There is another kind of line that people have thought with, where infinitesimal intervals are allowed–"infinitely short," so to speak. They were part of calculus when calculus was first invented. Velocity at an instant was just the ratio of the infinitesimal distance traveled to the time elapsed in an infinitesimal time interval.

This made sense, it worked, it was understood, but it was not Aristotelian. It did not fit into the "yes or no" framework of Logic. With effort, it was abandoned for an Archimedean framework, that relied on inequalities of two small variables called epsilon and delta. In modern times, a way has been found to marry the Infinitesimal to Aristotle. The key was mathematical logic. Once logic became precise and powerful enough to have its own theorems, one could talk about language and logic themselves as part of mathematics. Then we could refer to the Archimedean number system of Euclid and Newton, and call it Standard. We can talk about another kind of number, a nonstandard one, that is smaller than every standard one, and still positive. An infinitesimal. That is Robinson's nonstandard analysis, which is now established as part of pure mathematics, and tool of applied math. In contrast to the standard Archimedean or Newtonian line, we call it Leibnitzian, because non-standard analysis is a rigorous re-establishment of Leibnitz's infinitesimal calculus.

The nonstandard real line is a better model for Time than the standard one. We can still think of the present instant as the zero on our time axis. We can associate to it a duration which is less than any standard positive number no matter how small, yet still greater than any infinitesimal. This would be what Robinson called the monad of zero, the set of positive and negative infinitesimals.

It satisfies the phenomenological requirement for a Now that is associated with an instant in time yet has a duration greater than zero. The seeming contradiction is not in the experience of Time, it is only in the attempt to model time by the standard line of Euclid and Newton. Robinson's hyperreal or nonstandard line is a better model, a better fit or description of time as experienced.

For the purpose of representing physical processes, we can choose either version of the line. The infinitesimal is closer to naive intuition, and was never really expelled from applied mathematics.

Aristotle on Now is descriptive, and presents two different understandings of the word Now. Taking the notions of past and future as basic or elementary, the Now is constructed in two different ways. It is the instant that SEPARATES the past from the future. But also, it is the instant that CONNECTS the past to the future.

The first meaning we could call the Newtonian Now. In Newton's mechanics, it is the initial or final time of a process, say when a stone is released to fall or when it hits the ground. The customary scientific Now.

We could call the second Now the Bergsonian or Leibnizian Now. It is the experiential Now, the Now that sees time as a continuous flow, where the future is connected seamlessly to the past, and is constantly being transformed into the past.

To clarify the second Now, where the future is turning into the past, we could say it sticks INTO the future and sticks ONTO the past. It comprises not just a single moment, of duration zero, but also the moments immediately before and after it, the preceding and succeeding moments. To make this mathematical, we could say "those moments infinitely close to it, both before and after".

This is an experiential Now with positive duration. Abraham Robinson dubbed it the "monad". It has duration greater than any infinitesimal but smaller than any standard positive number. It has no definite beginning or ending, no first or last. This is what we intuitively feel, that the experiential Now should be: open without boundary.

Thus the hyperreals or nonstandard reals of Abraham Robinson provide a mathematical model to accommodate the experiential Now.

References

1. Bachelard, G. (1932). The Instant in R. Durie (pp. 64–95).
2. Bergson, H. (1911). The cinematographic view of becoming. In: Holt, Rinehart, Winson (eds) Creative evolution (pp 304−311).
3. Capek, M. (1973). Time. In: Wiener, P. P (Ed.) Dictionary of the history of ideas (vol. 4, pp 389–398). Charles Scribner's Sons.
4. Dorato, M. (2015). Presentism and the experience of time. *Topoi, 34*, 265–275.
5. Durie, R. (2000). *Time and the instant*. Manchester: Clinamen Press.
6. Grunbaum, A. (1969). In: Salmon, W. (Ed.) Modern science and Zeno's paradoxes of motion (pp. 200–250).
7. James, W. (1890). The principles of psychology. Henry Holt & Co.
8. Miller, I. (1982). Husserl's *Account* of our Temporal Awareness. In: H. L. Dreyfus (Ed.) Husserl, intentionality and cognitive science. The MIT Press.
9. Pockett, S. (2003). How long is 'now'? Phenomenology and the specious present. *Phenomenology and the Cognitive Sciences, 2*, 55–68.
10. Robinson, A. (1966). Nonstandard analysis. North-Holland Publishing Company.
11. Robinson, M. (2015). Humanism. The spirit of our time. The Nation 301(19):27–32.
12. Sachs, J. (Ed.) (1998). Aristotle's Physics. Rutgers University Press.
13. Toulmin, S., & Goodfield, J. (1965). The discovery of Time, Harper & Row.
14. "Perception of time," Encyclopedia of Philosophy.

Chapter 22
What's Wrong with the Platonic Ideal of Space and Time?

Lorenzo Sadun

To our senses, space is smooth, 3-dimensional, and flat. We move in a continuum where all points are equal (space is "homogeneous") and all directions are equal ("isotropic", or "round"). If we head off in any direction, we keep on going, with no curving back on ourselves ("flat"). In short, we seem to live in a universe governed by Euclidean geometry.

Actually, almost everything I wrote in the previous paragraph was a lie. Our senses continually detect the difference between different directions. Things fall down, not up. The sun rises in the East, not in the West. We *know* that not all points are equal, and that Hawaii is a lot more pleasant than Antarctica. Curvature is all around us, from the hill my house sits on to the fact that we can fly around the world.

All the same, most of us still believe in 3-dimensional space, plus an added dimension (time) that describes how things change. We rely on a mathematical model of reality in which space is a Platonic ideal: smooth, homogeneous, isotropic, and flat. Everything that breaks that underlying symmetry is attributed to objects: a planet whose gravity causes things to fall in a preferred direction, and whose rotation makes another object appear to move through the sky, the hills and valleys of my home town, and lovely tropical islands. We think that space is simple, that objects are complicated, and that the job of scientists is to understand the messy behavior of objects against the perfect backdrop of space.

This idealization of perfect space and imperfect contents makes for a lovely theory, but is it correct? It was accepted almost without question for over 20 centuries, from the ancient Greeks through the Middle Ages and the Enlightenment. In the late 19th century, however, it started to break down. While the theory works very well to describe physics on many length scales, it gives results that are nonsensical, or

---

L. Sadun (✉)
Department of Mathematics, University of Texas, Austin, TX 78712, USA
e-mail: sadun@math.utexas.edu

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_22

at least that contradict experiment, when dealing with very big things, very small things, and very fast things.

In this essay, I'll touch briefly on the new theories that were developed to deal with these discrepancies—special relativity, general relativity, quantum mechanics and string theory. I'll then turn to the suggestion, popularized by Stephen Wolfram [1], that space and time aren't smooth at all, but come in essentially discrete chunks. Using recent results from the theory of aperiodic tilings, I'll argue that this last suggestion is not realistic, and will defend the conventional wisdom that Euclidean space and time, as modified slightly by 20th century physics, is still the best way to describe reality.

22.1 Relativity and the Fall of Euclidean Space

If space and time are absolutes, then how fast are we moving? After all, the earth rotates on its axis and revolves around the sun, the sun revolves around the center of the galaxy, and the galaxy tumbles through the universe. We *must* be moving, but in what direction, and how fast? In 1887, Michelson and Morley [2] tried to find out. They figured that light moving in the same direction as us would appear to be moving slower (since it has to catch up with us), that light moving in the opposite direction would appear to be moving faster, and that light moving perpendicular to our motion would have an intermediate speed. With a clever interferometry experiment, they measured these differences and got exactly zero, suggesting that we were not moving at all! How could that be?

Several complicated mechanisms were proposed for why the speed of light *appeared* to be the same in all directions. It took almost 20 years, until Einstein's 1905 Special Theory of Relativity [3], for mankind to realize that the laws of mechanics and electromagnetism, and hence the speed of light, really were the same relative to the earth, to the sun, and to the distant galaxies. Different observers moving relative to one another have different notions of space and different notions of time, but the same laws of physics. Space and time are not absolutes, but are only defined relative to an observer. Moreover, they are closely linked, and switching from one reference frame to another is like a rotation in a 4-dimensional space-time[1] As such, it is impossible to speak of the nature of space without also considering the nature of time, and vice-versa. To know one is to know both.

Einstein took things a step further in his 1915 General Theory of Relativity (GR) [4]. He proposed that space-time is not flat. Rather, the presence of mass, momentum and energy causes space to bend, and we perceive this bending as gravity. We can no longer place perfect space and imperfect matter in separate categories. Matter bends space and the geometry of space affects matter. If the distribution of matter isn't

---

[1] When you rotate in the x-y plane, the new value of x depends on both the old values of x and the old value of y. Likewise, when you do a "Lorentz transformation" from one reference frame to another, the new position depends on both the old position and the old time, as does the new time.

uniform and isotropic, then neither is the geometry of space-time. *Matter is lumpy, so space-time is bumpy.*

Since that time, GR has been tested in numerous experiments, and has performed extremely well, most recently in the 2016 observation of gravitational radiation. GR may not be the ultimate theory and may require tweaking in the future (In particular, Sakharov [5] has argued that it is just the first term in an infinite series of corrections to Newton's Laws), but it is hard to avoid the conclusion that, on extremely large length scales, the Platonic ideal of space-time just doesn't work.

22.2 Quantum Mechanics and the Very Small

A different challenge to classical physics came when studying very small distances. According to classical physics, a glowing hot object should emit a certain amount of long-wavelength infrared light. It should emit more shorter-wavelength visible light, still more ultraviolet light, yet more x-rays, and so on. Not only is the bulk of the radiation supposed to be of such high frequency that a coal from your backyard grill would kill you, but the total amount of energy emitted per unit time is supposed to be *infinite*.

To explain why glowing coals aren't lethal, Planck [6] proposed that light energy can only be emitted or absorbed in discrete chunks, called quanta. This theory of light, called quantum mechanics, was soon extended to all forms of matter and energy and then generalized to fields that describe the creation and annihilation of particles. This body of work took care of the "ultraviolet catastrophe" that puzzled Planck, but created other mysteries.

For one, Heisenberg [7] observed that quantities that were once thought to be precise, like the position and momentum of a particle, are actually a bit fuzzy. There is uncertainty to position, there is uncertainly to momentum, and the product of the two uncertainties is at least Planck's constant divided by 4π. Likewise, there is uncertainly in energy and uncertainty in the time when things happen. By general relativity, the curvature of space is a function of mass and energy and momentum, but these quantities can't be nailed down. So not only are particles fuzzy, but space-time itself is fuzzy.

Worse still, the infinities that appeared in the ultraviolet catastrophe aren't completely tamed. By the uncertainty principle in energy, particles can blink in and out of existence. In many problems in quantum field theory, the effect of all these "virtual particles" could be infinite, which doesn't make sense. To avoid these infinites, the laws of physics, and of geometry, have to become qualitatively different at the scale of the so-called "Planck length". This is an incredibly small length of around 10^{-35} m, or about a septillionth the radius of an atomic nucleus. (You could fit more Planck-length sized particles into a single proton than you could fit protons within a million-mile diameter ball.)

According to string theory, space-time isn't 4-dimensional. It's actually 10-dimensional (or 11-dimensional in some versions), with all but 4 of the dimensions wrapped up in a higher dimensional analogue of a surface, of size comparable to the Planck length. Just as we can treat a thin 3-dimensional filament, such as a human hair, as being effectively 1-dimensional, our thin 10 or 11-dimensional universe is effectively 4-dimensional.

A very different solution has been advocated by Wolfram [1]. He suggests that at very small length scales the universe is really 0-dimensional! His theory is that space and time are actually discrete, with the possible points ordered in a neat array. At each new time step, what is happening at each point in space depends only on what was happening at that point, and at all adjacent points, an instant earlier. That is, the universe is like a gigantic array of computers, each one updating based on what its neighbors are doing.

22.3 Life on the Grid?

Such an array is called a “cellular automaton”. The past decades have seen an explosion of work on cellular automatons, including notable advances by Wolfram himself. The most famous example of a cellular automaton is John Conway’s Game of Life [8]. This game operates on a 2-dimensional grid of square “cells”, and time advances in discrete steps called “ticks”. At any given time, each cell is either alive or dead. At each tick of the clock, each live cell either survives or dies, depending on the number of live cells in the 8 positions around it, and each dead cell either stays dead or comes to life (“birth”) by a slightly different rule. This simple game exhibits amazingly complicated behavior, with intricate patterns propagating across the screen.

Could a 3-dimensional version of this sort of game be a model for the complex behavior of the real world? Wolfram says yes, but I say no. In the Game of Life, signals propagate at a maximum speed, just like the speed of light, but this speed depends on direction. Whether a cell at (0, 0) is alive or dead at time 0 affects all the neighboring cells at time 1, all the cells around those at time 2, and so on. After n time steps, the cells that are potentially influenced by the initial situation form a square with vertices at (n, n), $(-n, n)$, $(-n, -n)$ and $(n, -n)$. Signals propagate fastest in the diagonal directions and slowest sideways or up-and-down. This contradicts the experimental fact that the speed of light is the same in all directions.

You might argue that this contradiction resulted from the details of the Game of Life, and that different rules might give a different speed of light. It’s true that more complicated rules can make things a *bit* more isotropic, but they can’t make things completely round. As long as each cell has a finite number of neighbors, there will always be a finite number of directions in which information runs fastest, and intermediate directions in which information runs slower. Put another way, if the underlying geometry of space-time is a grid, then there will always be physical phenomena that reveal the underlying axes of the grid, in the same way that the facets of a crystal reveal the underlying arrangement of the atoms inside.

(An important caveat: Computers use grids to model continuous and isotropic systems all the time. However, these numerical models only work well when looking at patterns that move much slower than the maximum transmission speed of information, a.k.a. the speed of light. Cellular automata can accurately model a world governed by Newton's laws, and can be very useful in understanding a cold weather front that is moving at 15 miles per hour, but they can't handle extreme relativistic motion.)

22.4 Life in a Raindrop?

The universe isn't a grid, but can it still be discrete? Just because a crystal can't be round doesn't mean that we can't make something round (or at least round to the naked eye) out of atoms. The raindrops falling outside my window say that you can! If you take a bunch of building blocks and assemble them randomly, as with the grains of sand in a sand pile or the water molecules in a raindrop, the resulting structure is unlikely to have any preferred directions.

However, random structures have their own problems. Imagine a small explosion in the middle of a sand pile. The sound from that explosion wouldn't go straight to our ears, but instead would ricochet off of the various grains of sound in random directions. The sound *would* go in all directions at essentially the same speed, but different paths would take different amounts of time to reach us. What started out as a sharp BANG! would be heard as a not-so-sharp roar. The wave properties of sound (constructive and destructive interference) could reduce this effect but cannot eliminate it. Waves of different frequencies work their way through the maze at slightly different rates; in raindrops, this distortion causes rainbows. Waves propagating through random media *always* get distorted and smeared.

However, signals from distant galaxies do *not* get blurred as they travel to us through empty space. The neutrino bursts from a supernova, or the gravitational waves from the merging of two black holes, travel for billions of years across the universe and then hit us in an instant. We don't see any of the fuzziness that would be expected from random space-time.

In addition to the experimental evidence against random discrete space-time, such a model would raise as many additional metaphysical questions as it would answer. What determines the random structure at each point in space-time? The random arrangements of sand in a sand pile reflect the details of how the grains of sand were dropped and mixed, but there is no *process* by which space and time are created. Space and time just *are*. Einstein famously objected to the role of pure chance in quantum mechanics; this would be far worse.

Fig. 22.1 The author's bathroom floor. Note the light level-1 supertile sitting in the center of a mostly dark level-2 supertile, which is itself in the center of a level-3 supertile that extends beyond the frame

22.5 Life in an Aperiodic Tiling?

Finally, we consider a possibility intermediate between random space-time and a regular periodic grid. It is possible to have order without periodicity. For instance, imagine a sequence of $+$ signs and $-$ signs. We start with a $+$ sign and follow it with its opposite to get $+-$. We then follow this with the opposite of the pair, namely $-+$, to get $+--+$. We then follow this with the opposite of $+--+$, namely $-++-$, to get $+--+-++-$. Continuing the process forever, we get a infinite sequence, called the Thue-Morse sequence, with the magical property that no pattern within it (e.g., $+--+$) ever repeats itself 3 times in a row. The Thue-Morse sequence is an example of *aperiodic order*, in which the arrangements follow precise rules but are not just the same pattern repeated over and over and over.

An interesting 2-dimensional aperiodic tiling is the *pinwheel tiling* [9] invented by John Conway and Charles Radin. The basic tiles are right triangles with sides of length 1 and 2 and hypotenuse of length $\sqrt{5}$. You can arrange five such tiles to make a bigger triangle of the same shape, which we call a *supertile of level 1*. We can then arrange five supertiles of level 1 to make a supertile of level 2, 5 of those to make a supertile of level 3, and so on. This design is featured architecturally in Federation Square in Melbourne, Australia, and in the author's home (Fig. 22.1).

The center tile of a supertile of level 1 is the same shape as the supertile, but is rotated by the angle $\theta_0 = \tan^{-1}(1/2)$. If the supertile of level 1 is in the center of a supertile of level 2, then the center tile is rotated by $2\theta_0$ relative to the level 2 supertile. Continuing the process, we get rotations by arbitrary multiples of θ_0.

However, θ_0 is an irrational number of degrees, so no multiple of θ_0 will ever take you back exactly to the direction you started in. The pinwheel tiling has tiles pointing in infinitely many different directions, and all directions are equally likely. (In technical language, the distribution of directions is *uniform* on the circle.) While the tiles themselves are pointy triangles, the statistical properties of the pinwheel tiling are rotationally invariant, with no directions preferred over any others. Could the pinwheel tiling, or something like it, be a discrete model for a seemingly isotropic universe?

The problem is that the rotational invariance only manifests itself in the limit of infinite size, and develops *incredibly* slowly. An nth level supertile has 5^n tiles that appear in only $8n$ different directions, with a still smaller number of directions accounting for the vast majority of the tiles. If the tiles were the size of the Planck length, then a Milky Way Galaxy-sized supertile might have $10^{110} \sim 5^{160}$ tiles in it, but the bulk of those tiles would only be pointing in about 100 different directions. Even at astronomical length scales, space would not look isotropic.

Things are qualitatively the same for *all* 2-dimensional hierarchical tilings, and only slightly better in 3 dimensions. To get around the 2-dimensional limitations, Conway and Radin devised a 3-dimensional generalization of the pinwheel tiling, called the *quaquaversal* tiling (Latin for "every which way") [10]. The number of relevant directions does grow faster than for the pinwheel, but a galaxy-sized supertile would still only feature a few thousand relevant directions [11].

22.6 Conclusions

The Platonic ideal of perfectly uniform and symmetric 3-dimensional space, coupled with perfectly uniform 1-dimensional time, did not stand up to 20th century physics. Special relativity shows that we can't study space and time separately, but must instead think about 4-dimensional space-time. General relativity shows that space is not flat, but bends and curves in response to the matter that is in it. Quantum mechanics says that this matter is fundamentally uncertain, making the structure of space-time uncertain. Furthermore, something fundamentally different has to happen at the ultra-microscopic Planck length.

String theory says that, at the Planck scale, space-time is actually 10 or 11-dimensional, with all but 4 dimensions curled up into a tight ball. Many of us are very skeptical of string theory and open to alternatives, since there is absolutely no experimental evidence in string theory's favor. (To be fair, there is almost no experimental evidence against it, either. We simply don't know how to probe things that small.) However, the suggestion that the universe is a gigantic automaton, with space-time being essentially discrete, doesn't hold up, either.

If the universe were built on a lattice, then the directions of that lattice could be detected from physical phenomena occurring near the speed of light, and in particular by the propagation of light itself. If the universe were built with random local geometry, then light would not have a precise speed, and different parts of a signal would travel at slightly different speeds and directions, much as a prism splits light into differently colored beams. If the universe were modeled on an aperiodic tiling with rotational symmetry in a statistical sense, there would still be preferred directions at the scale of actual experiments.

All of the simple explanations have failed us. Platonic space-time works very well for day-to-day life, but the details of the actual universe are more complicated and mysterious that our human intelligences can currently fathom. Not because humans are stupid, but because we have the privilege of living in a universe of awe-inspiring subtlety and splendor.

Enjoy the ride.

References

1. Wolfram, S. (2002). *A new kind of science*. ISBN-13: 978-1579550080.
2. Michelson, A., & Morley, E. (1887). On the relative motion of the earth and the luminiferous ether. *American Journal of Science*, *34*, 333–345.
3. Einstein, A. *Zur Elektrodynamik bewegter Körper*, Annalen Den Physik, Bern (2005). An English translation (*On the Electrodynamics of Moving Bodies*) is found in Einstein's 1923 book "The Principle of Relativity".
4. Einstein, A. (1915). *Die Feldgleichungen der Gravitation* (p. 844847). Sitzungsberichte der Preussischen Akademie der Wissenschaften zu. Berlin.
5. Sakharov, A. D. (1968). *Vacuum quantum fluctuations in curved space and the theory of gravitation*. *Soviet physics, Doklady*, *12*, 1040 [Dokl. Akad. Nauk Ser. Fiz. 177 (1968) 70]. Reprinted in *General Relativity and Gravitation*, *32*, 365–367.
6. Planck, M. (1900). Zur theorie des gesetzes der energieverteilung im normalskektrum (On the theory of the energy distribution law of the normal spectrum). *Verhandlungen der Deutschen Physikalischen Gesselschaft*, *2*, 237.
7. Heisenberg, W. (1927). *Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik* (About the intutive content of quantum mechanics and kinematics). *Zeitschrift für Physik*, *43*, 172–198.
8. *For a simple description of Conway's game of life*. https://en.wikipedia.org/wiki/Game_of_Life.
9. Radin, C. (1994). The pinwheel tilings of the plane. *Annals of Mathematics*, *139*, 661–702.
10. Conway, J., & Radin, C. (1998). Quaquaversal tilings and rotations. *Inventiones Mathematicae*, *132*, 179–188.
11. Draco, B., Sadun, L., & Van Wieren, D. (2000). Growth rates in the quaquaversal tiling. *Discrete and Computational Geometry*, *23*, 419–435. The first author is a purple stuffed dragon, but the other two authors are human.

Chapter 23
The Fundamental Problem of Dynamics

Julian Barbour

Abstract In a world in which all objects are in relative motion, there arises the problem of equilocality: the identification of points in space that have the same position at different times. Newton recognized this as the fundamental problem of dynamics and to solve it introduced absolute space. Inspired by Mach, Einstein created general relativity in the hope of eliminating this controversial concept, but his indirect approach left the issue unresolved. I will explain how the general method of best matching always leads to dynamical theories with an unambiguous notion of equilocality. Applied to the dynamics of Riemannian 3-geometry, it leads to a radical rederivation of general relativity in which relativity of local scale replaces replaces relativity of simultaneity as a foundational principle. Whereas in the standard spacetime picture there is no unique notion of simultaneity or history, if this alternative derivation leads to the physically correct picture both are fixed in the minutest detail. New approaches to several outstanding problems, including singularities and the origin of time's arrows, are suggested.

23.1 Introduction

In his unpublished *De Gravitatione* [1], Newton addressed what might be called *the fundamental problem of dynamics*: if all motion is relative, how can one identify a point in space that has the same location at different times? This is the problem of *equilocality*. Because he did not present the problem or repeat his arguments in the *Principia*, the issue has attracted little attention. In this paper, I will take direct resolution of the problem as the basis of an alternative derivation of general relativity (for a complementary account, see [2]). The main justification for this are new research avenues that are opened up. It is also interesting to see how Einstein's the-

J. Barbour (✉)
College Farm, South Newington, Banbury, Oxon OX15 4JG, UK
e-mail: julian.barbour@physics.ox.ac.uk

J. Barbour
University of Oxford, Oxford, UK

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_23

ory can be derived and radically reformulated following essentially one single idea: the definition of position at different instants of time.

My starting point is Leibniz's notion of space. Although it does not solve the problem, it is a first step. I then describe the creation of dynamical theories by *best matching*, which always leads to a notion of equilocality. Equations that govern the evolution of Riemannian 3-geometries are obtained more or less directly by application of best matching under the condition that only angles, but not lengths, can be compared at spatially separated points. This leads to *shape dynamics* [2]. Remarkably, one recovers not only Einstein's evolution equations in a distinguished foliation but also, in a single package, a *prescription* for how the initial-value problem of general relativity is to be solved and the resulting Cauchy data are to be evolved. To the extent the evolution can be continued, this leads to construction of an Einsteinian spacetime in a foliation by spacelike hypersurfaces of constant mean extrinsic curvature and simultaneously a fibration of the spacetime by timelike curves that pass through equilocal points defined by best matching. It is in this sense that history is fixed in minutest detail.

23.2 The Relational Definition of Position

In his famous correspondence [3] with Clarke, Leibniz rejected Newton's absolute space and time, mainly on the basis the *principle of the identity of indiscernibles*: if two supposedly distinct things or states are in fact indistinguishable, then they are in fact one and the same. This led Leibniz to argue that

> if space was an absolute being, there would something happen for which it would be impossible there should be a sufficient reason… Space is something absolutely uniform; and, without the things placed in it, one point of space does not absolutely differ in any respect whatsoever from another point of space. Now from hence it follows, (supposing space to be something in itself, besides the order of bodies among themselves,) that 'tis impossible there should be a reason, why God, preserving the same situations of bodies among themselves, should have placed them in space after one certain particular manner, and not otherwise; why everything was not placed quite the contrary way, for instance, by changing East into West.

When Clarke objected that "space and time are quantities; which situation and order are not" Leibniz responded

> I will here show, how men come to form to themselves the notion of space. They consider that many things exist at once and they observe in them a certain order of co-existence, according to which the relation of one thing to another is more or less simple. This order, is their *situation* or distance. When it happens that one of those co-existent things changes its relation to a multitude of others, which do not change their relation among themselves; and that another thing, newly come, acquires the same relation to the others, as the former had; we then say, it is come into the place of the former; and this change, we call a motion in that body… And though many, or even all the co-existent things, should change according to certain known rules of direction and swiftness; yet one may always determine the relation of situation, which every co-existent acquires with respect to every other co-existent… And

> supposing, or feigning, that among those co-existents, there is a sufficient number of them, which have undergone no change; then we say, that those which have such a relation to those fixed existents, as others had to them before, have now the *same place* which those others had. And that which comprehends all those places, is called *space*.

Note, first, that Leibniz equates situation with *distance* without saying how that is determined and, second, his definition of space requires "a multitude of others, which do not change their relation among themselves". This means that he had not given a definition of space applicable to the realistic situation in which all bodies of the universe are in motion relative to each other. As I noted, Newton had introduced absolute space precisely to overcome this problem (without explaining the difficulty in the *Principia*). The comment that if all bodies move relative to each other "yet one may always determine the relation of situation, which every co-existent acquires with respect to every other co-existent" is correct *at a given instant* but does not solve the real problem: how can one pair up points whose positions are defined relationally at different times and say they *are at the same place*.

This is the problem of equilocality. Unless it is solved, dynamics (understood as the evolution of relative configurations) has no firm foundation. Consider the principle of least action, which plays a truly essential role in both classical and quantum dynamics. The calculation of the action is impossible if one cannot quantify displacements of particles, which in turn is impossible without a notion of equilocality. The Leibniz–Clarke correspondence gives no guidance on this. In Sect. 23.4, I will show how equilocality can be defined provided certain conditions are met. First it will be helpful to present a notion of space somewhat different from Leibniz's.

23.3 Space as the Order of Coexisting Facts

As we have seen, Leibniz claimed that space is the order of coexisting things and, when pushed, defined order as the *distance* between things. However, Leibniz did not say how distance, which plays a primary role in his notion of space, is to be determined. I have not researched the history of distance determination, which clearly involves measurement and is part of the beautiful discipline of metrology. I will merely note that in the famous lecture given in 1854 in which he introduced his generalization of Euclidean geometry, Riemann said that "measurement consists of placing the quantities that are to be compared on top of each other".

At least since the dawn of agriculture, distance measurement has been been important and remarkably easy thanks to one of what I call 'the gifts of nature', by which I mean the ready availability of measuring rods in the form of straight sticks like bamboo canes or ropes. These have an empirical property of the utmost importance: to a high degree they remain mutually congruent. I can take one short cane as unit, use it to mark notches on as many other long canes as I like, move them around individually over large distances and then bring them back together. The ratios of their lengths, as measured by the notches, will not have changed perceptibly. It was surely thanks to this basic property of rods that Pythagoras' theorem was discovered by the ancient Egyptians at the latest about 4000 years ago.

Now consider this scenario. Suppose $N, N \geq 5$, fixed points in our familiar three-dimensional space. Let the $N(N-1)/2$ distances between them be measured, yielding a corresponding number of positive numbers. These numbers, the measured distances, are *empirical facts*. A priori, there is no reason why they should bear any relation to each other. However, it turns out that, provided $N \geq 5$, they will satisfy certain algebraic relations that will hold to the accuracy with which the measurements have been made and our space is Euclidean. In other words, certain combinations of these distances will, to the corresponding accuracy, be zero. Such a state of affairs is a profound fact and indicates the existence of a controlling law.

The consequences of the law are remarkable. It makes *data compression* possible. Instead of representing the geometrical arrangement of the N points by means of the $N(N-1)/2$ positive numbers, one can express it by means of $3N$ coordinates, conveniently taken to be Cartesian ($\mathbf{r}_a, a = 1, \ldots, N$). The measured distances, the separations $r_{ab} := |\mathbf{r}_a - \mathbf{r}_b|$, are invariant under the Euclidean translations and rotations that can be applied to the $\mathbf{r}_a$. If N is large, the data compression is very significant since the number of distances grows as the square of the number of fixed points (Leibniz's coexisting things), whereas the number of coordinates grows only linearly.

The essential geometry revealed by the possibility of data compression can be taken one step further by the introduction of the root-mean-square length of the system:

$$\ell_{\text{rms}} := \sqrt{\sum_{a<b} r_{ab}^2}, \quad r_{ab} := |\mathbf{r}_a - \mathbf{r}_b|. \tag{23.1}$$

If we now divide all the r_{ab} by ℓ_{rms}, the resulting $\tilde{r}_{ab} = r_{ab}/\ell_{\text{rms}}$ still satisfy algebraic relations analogous to the ones I have already described. However, they are now 'liberated' from the arbitrary the unit of length and therefore scale-invariant. The scale-free separations $\tilde{r}_{ab}$ are invariants of the similarity group (Euclidean translations and rotations augmented by dilatations). In modern terms, Leibniz's case against Newton, *as applied to a single configuration*, is that it is the invariants $\tilde{r}_{ab}$ which define reality.

In the light of this discussion, what then is space? Intuitively, many people (including Newton one suspects) think of it as something like a perfectly translucent block of ice. I would argue that this is a mistake. It reifies the data compression of empirical facts found in observable relations into space. The empirical facts and the relations they satisfy are all we need. They ensure the data compression and our intuitive understanding. Here it is worth quoting Piaget [4], who comments that "space is often conceived as an empty box into which bodies are fitted" but says

> space is not a container. It is the totality of the relationships between the bodies we perceive or imagine, or rather, the totality of the relationships we use to endow these bodies with a structure. Space is in fact the logic of the apparent world or at least one of the two essential aspects (the other being time).

This seems to me very close to the scenario I described, the only difference being the quantitative sharpening made possible by the 'gifts' nature gives us in the form of near perfect measuring rods.

In summary, I replace Leibniz's aphorism "space is the order of coexisting things" with space is "the order of coexisting *facts*". There is a 'totality of relationships'. Our intuition gives us a *conceptual space* that enables us to understand the logic of the world and predict its consequences.

23.4 Best Matching

As we have seen, Leibniz failed to give a satisfactory definition of motion in a universe in which all things are in motion relative to each other. However, he did point out that, in any given instant, the distances between all the bodies in the universe will be well defined. I now want to show how equilocality, and with it motion, can defined relationally if the number of bodies in the universe is *finite*. Ironically, the key to this is to use the very thing that Leibniz employed to argue against the reality of absolute space: the possibility, in imagination, to place one and the same relative configuration of the universe in different positions in conceptual space without changing anything observable. The 'moving to different positions and orientations' is achieved mathematically by means of the generators of Euclidean translations and rotations. Scaling (dilation) brings in fascinating issues which I will discuss later.

Best matching does not use the Euclidean generators to move relative configurations in space but *relative to each other*. For simplicity, let us suppose the conceptual space is two dimensional, so we can picture it as a flat table. Let us also consider the simplest possible non-trivial dynamical situation of three distinguishable point particles of masses $m_a, a = 1, 2, 3$, interacting through Newtonian gravity.

We start with a single configuration: a triangle with the particles at its vertices. We can lay the triangle on the table wherever we please and then, like Leibniz, use the generators to move it anywhere else. We choose one position. Now we take another, slightly different triangle. We can lay it on the table in any position we choose. Each position will correspond to certain displacements of the particles. Because of the freedom in the different placings, it seems we cannot say there have been any definite motions. This is the problem of relative motion. But there is one placing of the second triangle relative to the first that is uniquely singled out.

To see that, suppose Cartesian coordinates on the table and let the positions of the particles in the first triangle be $\mathbf{r}_a$ and those of the second in an arbitrary placing be $\bar{\mathbf{r}}_a$. Now consider the quantity

$$\mathrm{d}s_{\text{trial}} = \sqrt{\sum_a m_a |\mathbf{r}_a - \bar{\mathbf{r}}_a|^2}. \tag{23.2}$$

This is a positive definite quantity and, for some position of the second triangle relative to the first, must have a minimum. Let this *best matched* (bm) position be $\mathbf{r}_a + \mathrm{d}\mathbf{r}_a^{\,\mathrm{bm}}$. Then

$$\mathrm{d}s^{\mathrm{bm}} = \sqrt{\sum_a m_a \mathrm{d}\mathbf{r}_a^{\,\mathrm{bm}} \cdot \mathrm{d}\mathbf{r}_a^{\,\mathrm{bm}}} \tag{23.3}$$

defines a metric on the space of relative configurations (relative configuration space: RCS) [5, 6]. As I said, best matching keeps Leibnizian displacement of relative configurations but not to place them differently in space but relative to each other. The freedom that created the problem becomes the solution to it. Best matching applies to any finite number of particles and has important properties:

- $\mathrm{d}s^{\mathrm{bm}}$ (23.3) is independent of the position of the first triangle in the conceptual space. The best-matched pair can be moved around in that space in exactly the way Leibniz imagined moving a single configuration around in absolute space without changing anything observable. The $\mathrm{d}s^{\mathrm{bm}}$ are invariants of the Euclidean group.
- $\mathrm{d}s^{\mathrm{bm}}$ (23.3) is unchanged under swapping of the first for the second triangle. The resulting dynamics is time-reversal symmetric.
- Best matching establishes a unique pairing of any one point on the first triangle with a best-matched point on the second triangle. A notion of equilocality is well defined.
- Best matching brings the centres of mass of the configurations to coincidence and ‘squeezes’ the relative rotation out of the pair. The instantaneous state of the best-matched system has vanishing momentum **P** and angular momentum **L**.

We can now define best-matched N-body dynamics on the timeless relative configuration space. Suppose two such configurations A and B and any continuous curve joining them in the RCS and for it calculate

$$A_{\mathrm{trial}} = \int_A^B \sqrt{(E - V)\sum_a m_a(\mathbf{r}_a + \mathrm{d}\mathbf{r}_a^{\,\mathrm{bm}}) \cdot (\mathbf{r}_a + \mathrm{d}\mathbf{r}_a^{\,\mathrm{bm}})}, \tag{23.4}$$

where E is a constant and V, a potential, is a function on the RCS. For all such curves, one seeks (as in the standard procedure of the calculus of variations) the one that extremalizes (23.4).

There is now a very interesting way to ‘stack’ the successive configurations in the conceptual space. Place A anywhere. Then move all the configurations, one after another, into their best-matched position relative to their predecessors (for one of the two possible directions chosen for the advance of time). This is called *horizontal stacking* in [5] and leads to dynamical best-matched evolution in the conceptual space with moreover a uniquely preferred time labelling obtained by *vertical stacking* [5, 7].

When this is all done, it is found, first, that the particles evolve in the stacked conceptual space, which is infinitely many copies of the one needed for a single configuration, exactly as would a system in absolute space and time. Newton’s framework is not presupposed but *derived*. Second, the system will have energy E and vanishing angular momentum: $\mathbf{L} = 0$. This latter condition does not follow from Newton’s equations and is a *prediction* of the theory. The total momentum $\mathbf{P}$ will also be zero in the stacked frame, but one can always find an inertial frame in which that happens in Newtonian theory. In [5] it was asserted that in such an approach the constant E must also vanish, but the argument for that was flawed.

An argument for $E = 0$ is scale invariance. In Newtonian dynamics both E and $\mathbf{L}$ are conserved. If therefore they vanish at some initial time, they will vanish at all subsequent times. Now to give magnitude to E and $\mathbf{L}$ one needs an external scale, which Leibniz would surely reject. However, vanishing of E and $\mathbf{L}$ remains true whatever the scale (choice of unit). This argument for scale invariance is supported by the principle of sufficient reason: if the energy is to have some value, what reason can one give for it to have one value rather than another? A reason for zero is that it alone is independent of the choice of unit.

Although such an argument is not decisive—it would also require a vanishing cosmological constant—scale invariance comes into consideration in another way. As noted in [8, 9], the N-body problem with $E = \mathbf{L} = 0$ has a very interesting property. In all of its solutions, except for a set of measure zero, there is a point J at which the system's size, as measured by its centre-of-mass moment of inertia, passes through a unique minimum and rises to infinity in both time directions. In [9], this point is called the *Janus point* J by analogy with the Roman god because the two halves of the evolution curve are qualitatively the same either side of J and define arrows of time that point in opposite directions away from J. A further striking property of the point J is that at it one can specify fully scale invariant 'mid-point' data that determine the evolution in either direction away from J [9]. Thus, all the solution-determining information that is encoded in the mid-point data (and conserved by the dynamical evolution) is represented in a form invariant under the action of the similarity group. We recall that this group expresses the essence of Euclidean geometry and leads to the construction of the conceptual space from the 'totality of relationships' that Piaget identified as the true basis of our notion of space.

It may also be mentioned that throughout the 20th century many physicists, including Schrödinger, repeatedly rediscovered a relational mechanics of N mass points based on replacement of the kinetic term $\sum_a m_a \dot{\mathbf{r}}_a \cdot \dot{\mathbf{r}}_a$ in the Newtonian action by

$$W = \sum_{a<b} \frac{\dot{r}_{ab}^2}{r_{ab}}, \quad r_{ab} := |\mathbf{r}_a - \mathbf{r}_b|. \tag{23.5}$$

Such an action, augmented by the Newton potential, leads to a very interesting relational theory, see [10]. However, it suffers from a fatal defect: it predicts anisotropy of effective inertial masses at a level ruled out to many orders of magnitude by the most accurate null experiments, of Hughes–Drever type [11], so far performed in physics. This led Bertotti and myself to abandon our original Leibnizian/Machian proposal based on (23.5) and replace it by the theory of [5] based on best matching, in which there is no mass anisotropy. It is well known that Einstein sought to employ the equivalence principle to implement Mach's call for the replacement of absolute motion by relative motion. It is interesting that isotropy of inertial mass, and with it certain aspects of Lorentz invariance, is now confirmed by generalizations of the Hughes–Drever experiment to many orders of magnitude better than the equivalence principle.

23.5 Equilocality in Dynamical Geometry

It is striking that in creating metric geometry in 1854 Riemann did not take into account his words I cited in Sect. 23.3: "measurement consists of placing the quantities that are to be compared on top of each other". Indeed, the central section of his paper is headed *Metrical relationships that a n-dimensional manifold can have under the assumption that any interval can be measured by any other.* The final words here mean that intervals have a definite length whatever their position in the considered manifold. In other words, intervals at spatially separated points can be said to have the same length even though there is obviously no way in which they can be laid on top of each other to confirm that fact.

The analogy between Riemann's assumption and the implicit assumption of a universal notion of simultaneity at spatially separated points is obvious. The difficulty with simultaneity was first clearly noted by Poincaré in 1898 [12] and resolved in 1905 by him and Einstein. So far as I know, the first person to note the significance of Riemann's assumption was Weyl in 1918 [13]. In 1916 Levi-Civita (soon followed independently by Weyl) had discovered parallel transport. Weyl noted that parallel transport of a vector in a Riemannian space brings it back to its original position with a changed angle but the same length. Weyl called this *rigidity of length* and the last vestige of Euclidean 'distance geometry' (*Ferngeometrie*). To eliminate it, he introduced the notion of parallel transport of length by means of a new 1-form field, for which he coined the term *gauge*. Although his idea was later to play a key role in the discovery of the various gauge theories that underlie the standard model of particle physics, Weyl's initial belief in the identity of his 1-form field and the analogous gauge field in electromagnetism ran into the well known difficulties that Einstein noted.

In fact, Weyl's desire to eliminate 'distance geometry' can be realized, without introduction of any auxiliary field, at the level of three-dimensional Riemannian, i.e., with $+++$ signature, geometry as opposed to the $-+++$ Lorentzian four-geometry with which Weyl worked, no doubt because, as he emphasized in the strongest terms in his book *Space–Time–Matter*, he believed there could be no way back from the four-dimensional world of Einstein and Minkowski.

However, there is a case for taking a step back if one can then take two forward or, as the French say, *reculer pour mieux sauter*. The 'jumping off point' to shape dynamics [2] is that though lengths at spatially separated points cannot be directly compared (any more than clock readings can) *angles are absolute*. Their determination is purely local. Thus, a radian is the angle subtended at the centre of a circle by an arc equal in length to the radius, which can be taken infinitesimally small. Such an angle emerges from an 'order of coexisting facts' and truly belongs to a point.

Bearing this in mind, consider now a Riemannian metric g_{ab} in a three-dimensional manifold. At any point g_{ab} is represented by a symmetric 3×3 matrix. Three of its six coordinates encode coordinate information, two encode information about the angle between curves in the manifold that meet at the considered point, and one is a local scale factor. Following Weyl's argument and by analogy with the objection to simultaneity at spatially separated points, this is the one datum that needs to be questioned: two such scales at spatially separated points cannot be compared. It may already be noted that the two angle degrees of freedom in g_{ab}, which constitute the *conformal* part of the geometry, match the two degrees of freedom per space point associated with the gravitational field in general relativity.

It is well known that Clifford, who had translated Riemann's 1854 paper, mooted the idea that three-dimensional Riemannian geometry could be dynamical (see [14], p. 1202). If we say that only position-independent aspects of geometry are real, as opposed to gauge, then we should look to construct dynamics of conformal 3-metrics (defined as equivalence classes of a Riemannian 3-metrics with respect to conformal transformations). If we assume a spatially closed universe, the dynamical arena will be the space of all conformal 3-geometries on a closed 3-manifold: *conformal superspace*, which is obtained by quotienting Riem (the space of Riemannian 3-geometries) by three-dimensional spatial and conformal transformations. The resulting group is analogous to the similarity group of Euclidean geometry and may be called the *Riemann group*.

The question then arises of whether one can create a dynamics of conformal 3-geometries by best matching with respect to the Riemann group. The answer is yes [15]. The theory turns out to be vacuum general relativity derived in a manner that bears only a remote connection with Einstein's derivation and has some remarkable additions and restrictions that I will list shortly. The basic idea is already clearly suggested by the manner in which we imagined slightly different triangles 'placed on top of each other' and moved relative to each other into their best-matching position. In dynamical geometry, we suppose two 3-metrics $g_{ab}(x)$ and $\bar{g}_{ab}(x)$,

$$\bar{g}_{ab}(x) = g_{ab}(x) + \frac{\partial g_{ab}(x)}{\partial \tau}$$

that differ slightly and imagine them initially placed 'on top of each other' by saying that points in the two metrics with the same coordinate x are equilocal. As quantity to be extremalized by best matching, it is natural, without at this stage worrying about simultaneity at spatially separated points, to take

$$A_{\text{trial}} = \int \mathrm{d}\tau \int \mathrm{d}^3x \sqrt{R\, G^{abcd} \frac{\partial g_{ab}}{\partial \tau} \frac{\partial g_{cd}}{\partial \tau}}, \tag{23.6}$$

where R is the (three-dimensional) scalar curvature and $G^{abcd} = g^{ac}g^{bd} - \lambda g^{ab}g^{cd}$ (λ is an as yet undetermined parameter and τ is a time label). I won't attempt to give a detailed first-principles derivation of the *ansatz* (23.6) except to say that the square root ensures reparametrization invariance and hence the absence of an external time.

What is critical is the taking of the square root before the integration over space. This leads to one quadratic constraint per space point. Its interplay with the constraints that arise from the diffeomorphism and conformal best matching ensures that the geometry has the two expected dynamical degrees of freedom. Another critical point is that the conformal best matching is marginally restricted to transformations that preserve the spatial volume and merely redistribute the local scale factor det g_{ab}. The restriction makes it possible for the universe to expand and necessitates the inclusion of λ in G^{abcd}.

As regards the main things that emerge from this shape-dynamic approach, I simply give the main results with references to their derivations:

- Best matching creates a succession of conformal 3-geometries that, at least in an open neighbourhood, stacks by equilocality into a four-dimensional spacetime that satisfies the Einstein equations [15].
- A point and tangent vector in conformal superspace determine such a succession of conformal 3-geometries [16].
- Best matching also *prescribes* solution of the initial-value problem of general relativity by the method that York [17] found by trial and error in 1972. The restriction to *volume-preserving* conformal transformations explains York's hitherto unexplained scaling law for the trace of the extrinsic curvature [15].
- Best matching imposes a distinguished foliation of the emergent spacetime by surfaces of constant mean extrinsic curvature (CMC surfaces) and ensures its propagation by also requiring a lapse-fixing equation to be satisfied [15].
- The attempt to couple matter fields to the evolving conformal geometry enforces a universal light cone (and with it the value -1 of the DeWitt supermetric in (23.6) [18]. The gauge principle for 1-form fields is also enforced. The taking of the square root at each space point in (23.6) is crucial for these results.

I think it must be agreed that the solution to the equilocality problem, which Newton so clearly formulated in *De Gravitazione* perhaps already 20 years before he wrote the *Principia* (and which Leibniz manifestly failed to solve), is thought provoking. As Clifford's reaction showed, once Riemann had at least partially 'loosened up' geometry, so that it is only locally Euclidean, the idea of making geometry dynamical was very natural. In fact, Riemann effectively created the ADM phase space of dynamical geometry and with it the two infinite-dimensional Lie groups (diffeomorphic and conformal) that act on it. It is especially striking that a theory designed to ensure that at spatially separated points only angles can be compared, ensuring *relativity of local scale*, simultaneously enforces *relativity of simultaneity*. One gets two for the price of one—and the gauge principle for good measure. Note that Lorentz invariance emerges late in the programme and, in contrast to Einstein's route to general relativity, is not a derivational postulate. The status of the equivalence principle is interesting. Both it and isotropy of inertial mass are strongly suggested on empirical grounds, but the extraordinarily high accuracy of Hughes–Drever type experiments make them an even more powerful guide to theory construction—by best matching—than the equivalence principle.

23.6 Caveats and Conclusions

The attentive reader will have noted the caveats “if this alternative derivation leads to the physically correct picture” (in the abstract), “to the extent the evolution can be continued” (in the introduction), and “at least in an open neighbourhood” (first of the final set of bullet points). The fact is that the results of [15], including the crucial unique solvability of the lapse-fixing equation, ensure evolution in conformal superspace and an emergent CMC-foliated spacetime *only in an open neighbourhood*. It is well known that CMC foliations have ‘singularity-avoiding’ tendencies, but there are solutions of general relativity in which the complete spacetime cannot be covered by a CMC foliation. The best known example is the Schwarzschild solution.

However, this is not yet a failure of shape dynamics, which rules out all solutions of general relativity for which space, as in a single Schwarzschild solution, is not closed. It is obvious that the universe contains many collapsed objects. It also appears to have begun very smooth, without any such objects. If shape dynamics is to supplant the spacetime representation of gravity, a major (clearly daunting) research project for it is to establish the extent to which the evolution in conformal superspace, and with it CMC foliation of an emergent spacetime, can be continued. However, it is encouraging that shape dynamics and the solution of what I have called the fundamental problem of dynamics suggest promising new directions of research, some more immediately tractable:

- The various arrows of time may have a dynamical origin and be nothing to do with special conditions at the big bang [8, 9].
- Since only shape degrees of freedom are regarded as physical, while scale is gauge, this suggests reconsideration of the singularity theorems in general relativity. They are generally held to signal the demise of classical spacetime, but that will not be so if the shape evolution remains well behaved.
- Most approaches to quantum gravity assume that space and time become discrete at the Planck length. If best matching and the underlying assumption of continuity that goes with it are foundational, the belief in discreteness may be unfounded.
- In quantization, symmetry with respect to four-dimensional diffeo-morphisms may be inappropriate. Instead, symmetry with respect to three-dimensional diffeomorphisms and conformal transformations is suggested.
- If the approach based on best matching is correct, many solutions allowed in the spacetime representation are ruled out. For example, spatial closure is required and could lead to testable predictions.

References

1. Newton, I. (1962). *De gravitatione*. In A.R. Hall & M.B. Hall (Eds.), *Unpublished scientific papers of Isaac Newton*. Cambridge: Cambridge University Press; discussed in Barbour, J. B. (1989). *Absolute or relative motion?* (Vol. 1, p. 612 ff). Cambridge: Cambridge University Press.

2. Mercati, F. *A shape dynamics tutorial*. arXiv:1409.0105.
3. Alexander, H. G. (Ed.). (1956). *The Leibniz–Clarke correspondence*. Manchester University Press.
4. Piaget, J. (1971) *The child's conception of time*. New York: Ballantine Books, Inc. (translation of French original on 1927).
5. Barbour, J., & Bertotti, B. (1982). Mach's principle and the structure of dynamical theories. *Proceedings of the Royal Society of London A*, *382*, 295.
6. Barbour, J. (1994). The timelessness of quantum gravity: I. *Classical and Quantum Gravity*, *11*, 2853.
7. Barbour, J. The nature of time. arXiv:0903.3489.
8. Barbour, J., Koslowski, T., & Mercati, F. (2014) Identification of a gravitational arrow of time. *Physical Review Letters*, *113*(18), 181101. arXiv:1409.0917.
9. Barbour, J., Koslowski, T., & Mercati, F. Entropy and the typicality of universes. arXiv:1507.06498.
10. Assis, A. K. T. (1995). Weber's law and Mach's principle. In J. Barbour & H. Pfister (Eds.), *Mach's Principle*. *From Newton's bucket to quantum gravity*. Boston: Birkhäuser.
11. Hughes, V. W., et al. (1961). Upper limit for the anisotropy of inertial mass from nuclear magnetic resonance. *Physical Review Letters*, *4*, 342 (1960); Drever, R. W. P. A search for anisotropy of inertial mass using a free precession technique. *Philosophical Magazine*, *6*, 683.
12. Poincarè, H. (1913). La mesure du temps. *Review of Mètaphysics Morale*, *6*, 1 (1898); English translation in Poincaré, H. *The value of science*. New York: Science Press.
13. Weyl, H. (1918). Gravitation und elektrizität. *Sitzungsber. d.Preuss. Akad. d. Wissensch.* (p. 465ff).
14. Misner, C. W., Thorne, K. S., & Wheeler, J. A. (1973). *Gravitation*. San Francisco: WH Freeman and Co.
15. Anderson, E., Barbour, J., Foster, B., Kelleher, B., & Murchadha, N. Ó. (2008). The physical gravitational degrees of freedom. *Classical and Quantum Gravity*, *22*, 1795. arXiv:gr-qc/0407104.
16. Barbour, J., & Murchadha, N. Ó. Conformal superspace: The configuration space of general relativity. arXiv:1009.3559.
17. York, J. W. (1972). The role of conformal 3-geometry in the dynamics of gravitation. *Physical Review Letters*, *28*, 1082.
18. Barbour, J., Foster, B., & Murchadha, N. Ó. (2002). Relativity without relativity. *Classical and Quantum Gravity*, *19*, 3217. arXiv:gr-qc/0012089.

Chapter 24
General Relativity, Time, and Determinism

James Isenberg

Abstract Einstein's theory of general relativity models the physical universe using spacetimes which satisfy Einstein's gravitational field equations. To date, Einstein's theory has been enormously successful in modeling observed gravitational phenomena, both at the astrophysical and the cosmological levels. The collection of spacetime solutions of Einstein's equations which have been effectively used for modeling the physical universe is a very small subset of the full set of solutions. Among this larger set, there are many spacetimes in which strange phenomena related to time are present: There are solutions containing regions in which determinism and the predictability of experimental outcomes breaks down (the Taub-NUT spacetimes), and there others in which the breakdown of determinism occurs everywhere (the Gödel universe). Should the existence of these strange solutions lead us to question the usefulness of Einstein's theory in modeling physical phenomena? Should it instead lead us to seriously search for strange time phenomena in physics? Or should we simply treat these solutions as anomalous (if embarrassing) distractions which we can ignore? In this essay, after introducing some basic ideas of special and general relativity and discussing what it means for a spacetime to be a solution of Einstein's equations, we explore the use of spacetime solutions for modeling astrophysical events and cosmology. We then examine some of the spacetime solutions in which determinism and causal relationships break down, we relate such phenomena to Penrose's "Strong Cosmic Censorship Conjecture", and finally we discuss the questions noted above.

J. Isenberg (✉)
Department of Mathematics and Institute of Theoretical Science,
University of Oregon, Eugene, USA
e-mail: isenberg@uoregon.edu

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_24

24.1 Introduction

Two of the most widely-known effects of relativity involve time. The "Twin Paradox" describes the difference in elapsed time which a pair of twins will record if one of them travels very fast (close to the speed of light) relative to the other. The "Gravitational Time Dilation" also describes a difference in elapsed time recorded by a pair of twins; in this case, between meetings, one of them lives in a much stronger gravitational field than the other.

Both of these effects are fairly straightforward predictions of relativistic mechanics. As such, they are necessarily present in any model of the universe which is consistent with the broad principles of special and general relativity. Not surprisingly, both of these effects have been directly observed in our universe. As well, although these relativistic time measurement effects were originally perceived as counter-intuitive and even disturbing (note the name "Twin *Paradox*"), physicists now are quite reconciled to their presence in our physical universe.

General relativity predicts the possibility of other strange effects related to time. Specifically, there are models of the cosmos consistent with the principles of general relativity in which observers can move cyclically in time. As well, there are models in which causal relationships and determinism breakdown. Should such possibilities be judged as realistic (and fascinating) predictions? Should they instead be judged as an indication that Einstein's theory of general relativity is seriously flawed?

To be able to discuss these two alternatives, and to possibly choose between them, it is important to understand what general relativity is, and what its role is in modeling the physical universe. Although a full understanding of general relativity requires one to know some differential geometry and to know how to work with partial differential equations (PDEs), some of its key ideas can be understood without these mathematical tools. We present some of these key ideas in Sect. 24.2, including the notion of a spacetime solution of Einstein's equations. These ideas are discussed within the context of the conceptual development of special and general relativity. In Sect. 24.3, we explore the use of spacetime solutions of Einstein's equations for modeling gravitational physics, noting the difference between the modeling of astrophysical events and the modeling of the full cosmos. We proceed in Sect. 24.4 to consider spacetime solutions in which causality and determinism break down. We note the sense in which these effects are consistent with general relativity, and we describe the Strong Cosmic Censorship Conjecture and the sense in which this conjecture argues that these effects are essentially irrelevant for modeling our universe. We make concluding remarks in Sect. 24.5.

24.2 Special Relativity and General Relativity

The conceptual framework used by scientists to model the physical world at extreme scales changed profoundly during the first two decades of the Twentieth Century. The fairly intuitive ideas of Newtonian physics (developed during the late 1600s) work very well for human-scale physical experiments such as tossing balls and sending space probes to Pluto. However, to understand the interactions of subatomic particles, to study the dynamics of objects moving anywhere near the speed of light, and to predict what happens in the presence of extremely dense concentrations of mass and energy, it was found that radically new ways of thinking are needed.

We do not explore here the new ideas—quantum mechanics and quantum field theory—that are needed to work with subatomic and elementary particle physics. Rather, we focus on special relativity and general relativity, which are needed to study the behavior of objects which move very fast, and objects which are very massive and concentrated.

Gedanken (thought) experiments heavily influenced Einstein's development of both special and general relativity. In the case of special relativity (SR), it was thinking about electromagnetism that was most influential. Einstein was driven to understand how the outcomes of Gendaken experiments with moving magnets and moving conducting coils could be consistent with Maxwell's theory of the electromagnetic field (published in 1865). Einstein found that he could obtain this consistency only if (a) he treated the speed of light as an observer-independent physical constant, and (b) he dispensed with absolute measures of space and time, replacing them with locally determined (observer-dependent) measures of distances and time intervals. Both of these principles are completely at odds with the foundations of Newtonian theory, which presumes that all observers (regardless of their relative motion) measure distances and time intervals identically; and presumes that if two observers measure the motion of a light ray (or anything else), then their measurements must differ in accord with their relative motion.

Einstein based his formulation of special relativity on these two principles, together with the empirically-based idea that to make measurements, an observer must rely on a personal clock along with a device for emitting and detecting directed light rays. Thus, to measure the length of a rod some distance away, an observer bounces light rays off each end of the rod and uses the clock-measured time interval between the return of the reflected rays to determine this length. Based on this measurement procedure, along with the assumption that the speed of light is identical for all observers, it is straightforward to calculate familiar SR phenomena such as the Lorentz length contraction, which relates the measured length of a rod as seen by two observers who are moving relative to each other.

Special relativity provides scientists with a new way to think about measurements of time and space which fits beautifully with Maxwell's theory of electromagnetism. SR does this by prescribing how different (relatively moving) observers $\mathcal{O}_1$ and $\mathcal{O}_2$ measure different electric and magnetic fields, and then showing that the resulting observations by $\mathcal{O}_1$ and $\mathcal{O}_2$ of these fields in the presence of moving magnets and conducting coils (as per one of Einstein's Gendaken experiments) are each consistent with Maxwell's theory.

Special relativity is *not*, however, consistent with Newton's theory of the gravitational field. This inconsistency is evident, since Newton's theory predicts that variations in the gravitational field which are generated by the motion of massive objects are manifest everywhere immediately, which contradicts the SR principle that no signals can be transmitted faster than light-speed. As well, there are no transformations of the Newtonian gravitational fields such that the experimental measurements made by observers in different frames (with lengths and time intervals transforming according to the rules of SR) are each consistent with Newton's theory.

This inconsistency led Einstein to focus for the next ten years—from 1905 to 1915—on finding a new theory of the gravitational field. Besides requiring this new theory to be consistent with the principles of special relativity—no signals traveling faster than the speed of light, and all measurements based on local considerations—Einstein believed it to be essential that his new theory of gravity incorporate the *Equivalence Principle*. In its simplest form, the Equivalence Principle embodies the experimental fact that if a pair of bodies move in a fixed gravitational field—say, that of the earth—with no other forces present (e.g., no friction or electromagnetic forces), then their motion is identical, regardless of their masses or composition. Newton's theory incorporates the Equivalence Principle in a somewhat ad hoc way: According to Newton, the acceleration a of a given body is determined by the imposed force F divided by the mass m of that body, and if the force F is gravitational, then F is proportional to m. Hence for a body moving in a gravitational field the mass factor m cancels, and consequently the acceleration induced by the gravitational force is independent of m. This works. However, Einstein believed it to be crucial for his new theory of gravity that the Equivalence Principle be built into the theory in a more essential way.

The most striking feature of general relativity—the name Einstein chose to label his new theory of gravity—is its introduction of *curved geometry* into the setup it uses to model the universe and the motion of bodies contained in it. The idea is simple, but revolutionary: Instead of thinking about the universe as a flat, featureless, static background stage in which bodies move in response to imposed forces, one thinks of it as a dynamic, curved space + time geometry in which the motion of bodies is determined by this geometry. More specifically, through each point in the spacetime and for each choice of a (local) velocity, the geometry determines a unique path. This path, which is called a *geodesic*, is the free-fall path that a body with the prescribed initial velocity passing through the prescribed point will follow, *regardless of its mass or composition*. In this way, the use of curved spacetimes allows general relativity to incorporate the Equivalence Principle in an essential way.

It is important to note that the geometry we are discussing here characterizes the *spacetime* as a unified entity, not just the space as something separate from time. One of the important innovations of special relativity is the unification of space and time into spacetime. Doing this allows one to recognize the physical equivalence of different frames of reference related by the motion of one of the frames relative to the other. It also provides a very useful way of visualizing such things as how different observers perceive the simultaneity of spatially-separated events in different ways. The combination of three-dimensional space and time into four-dimensional spacetime is a key feature of general relativity as well as special relativity. Indeed, spacetimes with specified curved geometries (with their corresponding arrays of geodesic paths) are the fundamental objects which are studied in general relativity, and are used to model gravitational physics.

What determines the curvature of a spacetime used to model the universe (or portions of it), according to Einstein's theory of general relativity? It is often stated that while the spacetime curvature determines how matter moves, the matter determines how the spacetime curves. This is only partially true. The key to understanding this is the Einstein gravitational field equation, which takes the form $G_{\mu\nu} = \kappa T_{\mu\nu}$. The object on the left hand side of this equation represents the spacetime curvature at any given point in spacetime (it is known as the Einstein curvature tensor field). The important thing to note about $G_{\mu\nu}$ is that it controls only a portion of the spacetime curvature—roughly half of it at each point in the spacetime. Even if the Einstein tensor is zero everywhere, the spacetime can be very curved. As for what appears on the right hand side of this equation, besides the constant factor κ (which depends on the speed of the light and the universal gravitational constant from Newton's theory) one has the stress energy tensor field $T_{\mu\nu}$. This object represents the localized mass density and momentum density and angular momentum density of matter and non-gravitational fields (including, e.g., the electromagnetic field) at each point in the spacetime. So, while the matter and the fields and the curvature in a general relativistic spacetime must satisfy Einstein's equation $G_{\mu\nu} = \kappa T_{\mu\nu}$, this relation by no means implies that matter determines curvature. If it did, then general relativity would predict that gravitational radiation does not exist; this of course would be inconsistent with the recent LIGO observations of gravitational waves.

24.3 General Relativity and the Modeling of Our Universe

General relativity is used to model gravitational effects at three very different scales: (a) Near-earth phenomena, such as the gravitational effects of the earth on GPS signals; (b) astrophysical events, such as the collision of a pair of black holes; and (c) cosmological features, such as the production of the cosmic microwave background by the Big Bang. There is an important conceptual difference between the modeling used for cosmological studies as opposed to that used for near-earth and astrophysical phenomena. Regarding the latter two cases, there is a very wide variety of different physical systems of interest which are expected to exist somewhere in the universe, and one studies a particular one by finding a spacetime solution of the Einstein equations which corresponds to that system. Note that such a spacetime solution is not expected to describe the entire universe; it is designed to model a local (relatively isolated) physical system in a very small portion of the universe. Each such spacetime solution may well be physically relevant, describing physical phenomena in widely separated portions of our universe.

In contrast, in modeling cosmological phenomena using general relativity, one works with solutions of Einstein's equations which are supposed to represent the entire universe. Since we live in just one universe,[1] in principle only one spacetime

---

[1] In this essay, we ignore the possibility that we live in a "multiverse", with regions that will never be observable.

solution is needed for cosmology, and only one is completely accurate. The catch is, we don't know enough about our universe to narrow down which spacetime solution to use for cosmological modeling. Consequently, in doing cosmology, we are led to consider many solutions, hoping that such a wide-ranging study can be useful for learning about the cosmology of our particular unique universe.

To illustrate the difference between these two types of modeling, it is useful to discuss an example of each kind: (i) modeling the gravitational radiation produced by the collision of two black holes, and (ii) modeling the universe immediately after the Big Bang.

One of the most exciting developments in physics in the 21st Century thus far is the direct detection of gravitational radiation for the very first time, by LIGO (the Laser Interferometry Gravitational Observatory, located both in Washington state and in Louisiana) [1]. Gravitational radiation is effectively ripples in the curvature in the spacetime, generated primarily by accelerating concentrations of matter. Though ubiquitous, such radiation is generally extremely weak, and consequently very difficult to detect. To enable it to be detected, as well as to be interpreted, it is crucial to be able to accurately model the gravitational radiation which is expected to be produced by very strong sources such as a pair of colliding black holes. Based on its remarkable success in modeling other gravitational effects such as the observed changes in the light signals emitted from the Hulse-Taylor binary pulsar [2], general relativity is used to carry out this modeling.

Conceptually, the modeling of the collision of a pair of two black holes is simple: First, one chooses the distinguishing parameters of the collision: the masses and the spins of each of the black holes, their initial separation and initial relative velocity, and the relative directions of the spins and velocities. This choice picks out the one particular black-hole collision of interest. Next, in accord with the choice of these parameters, one designates the initial data for the collision. This consists of the snapshot initial geometry and the initial rate of change of the geometry. Besides matching the choice of the parameters of the particular collision being modeled, the designation of the initial data must also satisfy a set of initial-data-constraint equations; corresponding to four of the ten Einstein gravitational field equations, these are analogous to the Maxwell constraint equations $\nabla \cdot B = 0$ and $\nabla \cdot E = 4\pi\rho_{charge}$ which the electric and magnetic fields must satisfy. After the designation of the initial data is made, one uses the remaining six of the Einstein gravitational field equations to evolve the geometry into a spacetime which satisfies the full system $G_{\mu\nu} = \kappa T_{\mu\nu}$ everywhere. From this spacetime, with a bit of straightforward work, one deduces such things as how long it takes for the black holes to collide and merge, how much gravitational radiation is emitted, and what the particular profile of the emitted radiation is.

As noted above, modeling black-hole collisions in the way just described is crucial to the success of LIGO in detecting and in analyzing gravitational radiation. While it took well over thirty years to work out the details of how to carry out this sort of modeling numerically, the process is now to a large extent routine.[2] It provides

---

[2] Routine, but very time consuming: Numerical runs can take hundreds of hours.

a wonderful example of the role that solutions of Einstein's equations can play in astrophysics.

At the time of this writing, Only two black hole collisions have been detected (and confirmed) by LIGO. One expects, however, that many more such collisions will soon be detected. Consequently it is very likely that spacetime solutions of Einstein's equations corresponding to the full range of the parameter space of black-hole collisions will each be useful in astrophysical modeling.

The use of general relativity to construct models of the full cosmos is very different from its use in modeling astrophysical events like black hole collisions. We live in a unique universe, so in principle just one solution of Einstein's equations is useful for modeling it in detail. However, as noted above, since we know so little about the full cosmos, we can not hope to know which is the specific spacetime solution which most accurately models our universe. Consequently we are led to construct a wide variety of solutions, not knowing which may be useful and which are not.

For example, say we wish to consider the question of how it is that the cosmic microwave background (CMB) radiation [3], which is believed to be a relic of the Big Bang over thirteen billion years ago, is observed to be very nearly the same in all directions, yet not exactly the same in all directions (i.e., nearly isotropic, but not exactly isotropic). One way to explore this question is to consider all spacetime solutions which evolve from a Big Bang and are broadly consistent with other features of the universe such as its age and apparent matter content, and then try to show that some large portion of these solutions produce nearly isotropic CMB radiation. It is of course impossible to construct all such solutions, even through numerical simulations, One can, however, focus on a subset of them—characterized, for example, by some symmetry—and examine the generic behavior of solutions in this subset. This approach played an important role in convincing many that Einstein's field equations with standard matter fields are not enough to model our universe. Rather, it appears that some mechanism for producing inflation in such models is likely needed [4].

It is important to note that there are many spacetime solutions of the Einstein equations which are clearly *not* expected to be of any direct use for modeling our universe and the gravitational phenomena which may occur in it. For example, cosmological solutions which are static, or which go from a big bang to a big crunch in a very short time, are useless for modeling. This feature distinguishes general relativity from other classical field theories such as Maxwell's theory of electromagnetism. In the case of Maxwell's theory, one can plausibly argue that essentially any solution might serve to model electromagnetic phenomena somewhere in the universe. Indeed, if one focuses on solutions of Einstein's equations which are designed to model localized astrophysical events, then the same argument for the potential usefulness of all such solutions might be made. The feature of general relativity that leads to clear disqualification of some solutions is its service for modeling the entire cosmos, and not just localized phenomena.

24.4 Causality, Determinism, and Solutions of Einstein's Equations

Deeply ingrained in our concept of how science—at least, physics—works is the idea of the *deterministic experiment*: One specifies the initial state of the system—say, the initial position and velocity of a ball near the surface of the earth—and then the system is compelled by the "laws of physics" to evolve in a unique, prescribed way. For example, for the ball near the surface of the earth, Newton's theory prescribes the acceleration of the ball, and hence (with the initial position and velocity specified) it determines a unique subsequent path for the ball.[3]

While there is nothing that tells us that physics *has* to work this way, it is a measure of the success of our science that we have been able to find theories which—at least within the realm of physical phenomena for which quantum theory is not needed—tell us exactly what "initial data" it is sufficient for us to know for a given system so that we can use the theory to calculate the future evolution of that system accurately. Notably, this works perfectly for electromagnetic phenomena as modeled by Maxwell's theory in the context of special relativity: Presuming that there is no charged matter around,[4] if we choose a global inertial frame[5] and if we know the electric and magnetic fields everywhere in space at a given moment of time (relative to this chosen frame), then Maxwell's equations determine these fields everywhere to the future as well as to the past.

If, instead of wanting to determine the electromagnetic fields everywhere for all time, we only seek to determine those fields at some particular point in space[6] x and at some particular time T in the future, do we need to know what the fields are now *everywhere* in space? Presuming for the moment that we are considering this problem in the flat spacetime of special relativity (known as the "Minkowski spacetime"), then in fact we only need to know the values of the fields "now" (which we label as time $t = 0$) in a particular region. This region, which we label $\mathcal{P}_{[x,-T]}$, consists of all those points y such that a light ray or a material object might travel from y at time $t = 0$ to x at time $t = T$ (a path in spacetime which might in principle be traversed by either a light ray or by a material object going slower than the speed of light is called a *causal path*). The values of the fields at time $t = 0$ (now) which are outside $\mathcal{P}_{[x,-T]}$ are completely irrelevant to determining the fields at (x, T), because no signal from this outside region can travel fast enough (faster than the speed of light) to get to the point x at time T.

Correspondingly, still restricting ourselves to the physics of special relativity, we see that if we choose a region in space Σ at a time $t = 0$, then there is a collection of spacetime points (z, t) to the future of $(\Sigma, t = 0)$ such that every causal path which

---

[3] Einstein's theory prescribes essentially the same path.

[4] The presence of charges does not change this, so long as a theory modeling the behavior of charges is prescribed along with Maxwell's theory.

[5] Special relativity allows this.

[6] Here and below, we use single latin letters such as "x" to label spatial points, even though in terms of coordinates, one needs three letters to label such points.

hits one of the points (z, t) *must* pass through $(\Sigma, t = 0)$, and also such that no causal path which hits these points (z, t) may pass through any points outside of $(\Sigma, t = 0)$ at time $t = 0$. This collection of points is called the future *domain of dependence*[7] of Σ, and is labeled $\mathcal{D}^+(\Sigma_{t=0})$. It follows from special relativity and Maxwell's theory that the electromagnetic fields in $\mathcal{D}^+(\Sigma_{t=0})$ are completely determined by the initial data of the fields on $(\Sigma, t = 0)$.

The ability to determine the future evolution of physical systems from initial data, and in particular the ability to do this in a localized way as described above via such constructs as $\mathcal{P}_{[x,-T]}$ and $\mathcal{D}^+(\Sigma_t)$, is a key feature of special relativity. The language used to affirm this feature is that Minkowski spacetime is *globally hyperbolic* and does not violate *causality*.

Does this same sort of thing work with physical phenomena for which general relativity is needed? One of the fascinating features of general relativity is that for certain classes of spacetime solutions it does, while for others it does not.

It is easy to see that there are spacetimes which satisfy Einstein's equations of general relativity, yet fail to be globally hyperbolic. To construct an example (which we label "identified-Minkowski spacetime"), we take the standard flat Minkowski spacetime with a standard set of coordinates (x, t), we throw out all of the spacetime with $t > 1$ or with $t < 0$, and then for each choice of the spatial coordinates (x), we identify the spacetime points $(x, 0)$ and $(x, 1)$. Since the curvature is zero everywhere and since there is no matter around anywhere, the equations $G_{\mu\nu} = T_{\mu\nu} = 0$ are certainly satisfied everywhere. Furthermore, despite the somewhat bizarre identification of spacetime points which has been made in constructing this spacetime solution, it does not violate any explicit rules for general relativity. Yet, with a bit of thought we see that if we choose a point, say $(x, 3/4)$, in the spacetime and then seek to identify the region $\mathcal{P}_{[x,-1/2]}$ as a subset of the spacetime with $t = 1/4$, we are forced to include *all* spatial points at that value of t. This is true because for *any* point $(y, 1/4)$, there is a causal path in this bizarre spacetime which connects $(y, 1/4)$ and $(x, 3/4)$ (it may have to pass through the $t = 0 \leftrightarrow t = 1$ identification several times). Similar considerations show that for any choice of a spatial region Σ at any time t_0, the domain of dependence $\mathcal{D}^+(\Sigma_{t_0})$ in this spacetime is empty. It follows from these strange features that deterministic experiments do not make sense in this spacetime, since there are no sets of initial conditions for the electromagnetic field (or for any other field) in some region Σ at some time t_0 which determine the behavior of that field anywhere into the future. We also see that this spacetime contains causal paths which close on themselves; hence the notions of "future" and "past" lose their meaning in this spacetime.

To avoid determinism and causality problems of the sort just described, one might require that spacetime solutions of Einstein's equations have the topology of $\mathbb{R}^4$. One of the exciting features of general relativity from its very beginnings [5], however, has been its opening up of the possibility of working with spacetimes with topologies more general than $\mathbb{R}^4$. Indeed, until fairly recently, many cosmologists believed that the most useful spacetimes for modeling our universe were likely to

---

[7] There is a corresponding past domain of dependence $\mathcal{D}^-(\Sigma_{t=0})$, defined analogously.

be spatially closed, with a spacetime topology of the form $S^3 \times \mathbb{R}$ (where S^3 is the three-dimensional sphere).

To allow spacetimes with interesting topology, while disallowing fairly contrived spacetimes such as identified-Minkowski, one might decide to include in the stipulations of the theory of general relativity the requirement that the spacetime topology be $\Xi \times \mathbb{R}$, where Ξ could be any three-dimensional manifold.[8] One might then ask if this restriction to the theory prevents not just identified-Minkowski spacetime, but also throws out every spacetime solution with causality and global hyperbolicity problems.

In fact it does not. To illustrate this, we discuss two very different archetypal examples here: the Taub-NUT spacetime and the Gödel spacetime. The Taub-NUT spacetime [6] which has the topology $S^3 \times \mathbb{R}$, is a solution of the Einstein equations with no matter present. It contains a spacetime region (the "Taub region") in which the spacetime is fully deterministic: Within the Taub region, the spacetime has no closed (or almost closed) causal paths and therefore does not violate causality. As well, in the Taub region, domains of dependence can be localized as in special relativity, and since the Taub region lies inside the union of the future and the past domains of dependence of any S^3 hypersurface labeled with a fixed choice of Taub-NUT time, the gravitational field of the Taub region is determined by gravitational initial data on such a hypersurface. The same would hold for other fields (such as electromagnetic fields) on a Taub-NUT spacetime background. The Taub region is thus labeled as "globally hyperbolic".

The Taub region is bounded by a particular S^3 hypersurface (called a "Cauchy horizon"); passing beyond it into the NUT region, one finds that causality, global hyperbolically, and determinism all break down. There are closed causal paths, domains of dependence become empty, and the evolution of fields is not determined by specified sets of initial data. Indeed there are multiple possible NUT regions which can be smoothly attached (as solutions) to the Taub region, thus constituting multiple possible "futures" of the Taub region.

By contrast with the Taub-NUT spacetime, the Gödel spacetime [7] has no region in which it is causal or globally hyperbolic or deterministic in any sense. A solution of the Einstein equations with "dust"-type matter (with non-vanishing vorticity) and with a cosmological constant, the Gödel spacetime is topologically simple: $\mathbb{R}^4$. However, through every one of its points, there are many closed causal paths. Past and future make little sense, and deterministic experiments cannot be carried out in a Gödel spacetime.

Are the Gödel and Taub-NUT spacetimes anomalous examples, which should best be hidden in the closet and ignored? While it is not at all clear if the Gödel spacetime in any sense exemplifies a class of solutions with similar properties, in fact there is a wide class of solutions with properties very similar to the Taub-NUT solution [8]. Notably, these "generalized Taub-NUT solutions" are considerably less specialized than the Taub-NUT spacetime itself, since they are much less symmet-

---

[8] An n-dimensional manifold is a space which locally looks like $\mathbb{R}^n$. The three-dimensional sphere is an example.

ric.[9] There is an infinite dimensional family of them, each one containing a globally hyperbolic region in which determinism holds, and each one extendible (in multiple ways) across a Cauchy horizon into a non globally hyperbolic region with closed causal paths.

Back in the 1960s, after he and Stephen Hawking had proven their celebrated spacetime incompleteness theorems, Penrose [9] proposed a pair of conjectures which have become known as Weak Cosmic Censorship (WCC) and Strong Cosmic Censorship (SCC). The spacetime incompleteness theorems, often called "singularity theorems", show that if a spacetime solution satisfies a fairly general set of hypotheses, then it necessarily contains causal paths which are forced to stop within a finite period of (local) time.[10] The reason such paths are forced to stop could be because the spacetime contains a region in which the curvature blows up; the presence of a Cauchy horizon could cause this as well. Penrose's Strong Cosmic Censorship conjecture proposes that in almost all such cases, it is curvature blowup rather the presence of a Cauchy horizon which causes the demise of causal paths.[11]

If Strong Cosmic Censorship is true, then spacetimes (such as the generalized Taub-NUT) which contain Cauchy horizons (and the other attendant difficulties with determinism) are a very small subset of the collection of all solutions of Einstein's equations. Is SCC in fact true?

Although Strong Cosmic Censorship has often over the past fifty years been cited as one of the major questions in the mathematical study of general relativity, it is far from clear whether SCC is true or not. Model versions of the conjecture have been proven in small families of solutions [10]; there are also recent results which suggest that SCC is not likely to hold in its strongest form. The verity of the conjecture remains a wide open question.

24.5 Conclusion

We know two important things about general relativity: (1) Observation and experiment have shown that it is extremely effective for modeling gravitational physics, from the scale of the solar system to the scale of astrophysical phenomena. (2) It includes among its spacetime solutions a number of them in which causality and determinism break down, at least in certain regions.

---

[9] The Taub-NUT spacetime is invariant under the action of a three-dimensional isometric group, while there are known generalized Taub-NUT solutions with only a one dimensional isometric group.

[10] A spacetime containing such paths is called "geodesically incomplete".

[11] Weak Cosmic Censorship is essentially unrelated to Strong Cosmic Censorship. WCC conjectures that in essentially all astrophysical-type spacetime solutions which contain unbounded curvature, distant observers cannot see signals from the region in which this occurs; that is, such regions must be contained inside black holes.

Based on these two facts, one might be led to believe one of the following assertions:
(A) We should expect to find causality violations and the failure of determinism somewhere in our universe.
(B) Since we have not detected any breakdown of causality or determinism in our universe, and since they are so fundamental to our way of doing science, general relativity must be seriously flawed as a physical theory to be used for modeling our universe.
(C) While general relativity includes among its array of solutions some in which causality and determinism fail, these solutions are irrelevant for modeling physics, and can be more or less ignored.

All three of these statements are consistent with what we know about general relativity. Which of them makes the most sense scientifically?

Since we know already that there are many solutions of Einstein's equations which are of no use for modeling physics (see Sect. 24.3), and since we in fact live in just one universe, it is hard to support statement B. Whether or not one chooses to invoke a selection principle (e.g., restrictions on topology, restriction to globally hyperbolic solutions, etc.) to determine which spacetime solutions might be used for modeling physics, it is straightforward to focus on certain solutions and ignore others in carrying out modeling. That this must be done in using general relativity does not, I believe, harm the usefulness of the theory. One might argue that this makes it more difficult to believe that Einstein's theory is *right*. However, we already know that Einstein's theory cannot be used to model systems in which quantum ideas are needed, so this argument is a red herring. As well, it is very likely that *no theory* that we know now (or will ever know?) will prove to be *right* in the sense that it explains and models phenomena at all scales.

Regarding statement A, it of course makes sense on a grand scientific scale to search for situations (as "predicted" by solutions of Einstein's equations) in which phenomena corresponding to the breakdown of causality and determinism might be detected. The predictive solutions could be used as a guide to finding such phenomena. On the other hand, one might argue that while it would be fascinating to find such phenomena, they are likely to be very difficult to find; hence such searches should be of very low priority.

Statement C is likely the most practical scientific position to take. A wide range of spacetime solutions of Einstein's equations are extremely effective for modeling and predicting gravitational physics. The existence of other, fairly strange, solutions is interesting, but perhaps irrelevant scientifically.

Perhaps; we shall see.

Acknowledgements This work was partially supported by NSF grant PHY-1306441 at the University of Oregon. I thank the Mathematical Sciences Research Institute for hospitality during the course of the writing of portions of this review.

References

1. Abbott, B., et al. (2016). Observation of gravitational waves from a binary black hole merger. *Physical Review Letters*, *116*, 061102.
2. Taylor, J., & Weisberg, J. (1982), A new test of general relativity—gravitational radiation and the binary pulsar PSR 1913+16. *The Astrophysical Journal, 253*, 908–920.
3. Penzias, A., & Wilson, R. (1965). A measurement of excess antenna temperature at 4080 Mc/s. *The Astrophysical Journal, 142*, 419–421.
4. Guth, A. (1981). Inflationary universe: A possible solution to the horizon and flatness problems. *Physical Review D*, *23*, 347.
5. Einstein, A. (1917). Kosmologische Betrachtungen zur allgemeinen Relativitaetstheorie. Sitz. K. Preuss. Ak. Wiss. Berlin. part 1: 142–152.
6. Misner, C. (1967). Taub-nut space as a counter example to almost anything. In J. Ehlers (Ed.), *Relativity theory and astrophysics 1: relativity and cosmology*. Lectures in Applied Mathematics (Vol. 8, p. 160). Providence: AMS Publications
7. Gödel, K. (1949). An example of a new type of cosmological solution of Einstein's field equations of gravitation. *Reviews of Modern Physics*, *2*, 447–450.
8. Moncrief, V. (1984). The space of (generalized) Taub-nut spacetimes. *Journal of Geometry and Physics*, *1*, 107–130.
9. Penrose, R. (1968). Structure of spacetime. In Wheeler, J. (Ed.), *Batteles Rencontre*. New York: Benjamin.
10. Isenberg, J. (2015). On strong cosmic censorship. In L. Bieri & S-T. Yau (Eds.), *Surveys in differential geometry, Vol. 20 (100 Years of General Relativity)*

Chapter 25
Topos Theoretic Approach to Space and Time

Goro C. Kato

Abstract At the level of the Planck scale (around 10^{-33} cm) and beyond, i.e., a sub-Planckian domain (less than 10^{-43} s after so called big bang), the usual concept of space and time becomes uncertain where the gravitational field might be simply a quantum fluctuation of a vacuum. Elementary particles or strings, the fundamental entities for our universe, are difficult to be considered at such a microcosm level. An emergence occurs when a physical phenomenon is the consequence of organization from the given local information data. Namely, we cannot tell any difference between electrons in a human brain and in an apple. This is just as we cannot tell any difference between a note in a piece by Mozart and a note in a piece by Bach. We use the concept of a sheaf as the device from local to global transition. In order to formulate *space* and *time* for those microcosm domains in terms of sheaves providing a background free notion in the sense of *quantum gravity, the notions of the associated (pre)sheaves* of time, space, and matter are introduced in the following sense. For a particle $\overline{m}$, we assign an *associated presheaf* m with $\overline{m}$. A presheaf is by definition a *contravariant functor* from a site (i.e., a category with a *Grothendieck topology*) to a product category. This is the notion of the *temporal topos* theory abbreviated as *t-topos* theory developed in [1–5]. For space and time, we associate a combined sheaf $\omega = (\kappa, \tau)$ where space sheaf κ and time sheaf τ are considered to be *t-entangled* in the sense that both sheaves behave as one sheaf. With the notions of sheaves and categories, we will give the formulations for the *uncertainty principle*, *particle-wave duality*, and *t-entanglement* together with the relativistic concept of a *t-light cone* (or an *ur-light cone*) valid in macrocosm and microcosm. As a consequence of the topos theoretic formulations, the possible scenario of pre and primitive stages of a universe, i.e., *ur-big bangs* in terms of *t*-topos theory will be provided. The main concepts to formulate these notions are coming from categorical notions of a micro-decomposition of a presheaf and a micro-covering of a t-site object.

G.C. Kato (✉)
Mathematics Department, California Polytechnic State University,
San Luis Obispo, CA 93407, USA
e-mail: gkato@calpoly.edu

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_25

25.1 Introduction

Our goal is to study and formulate space, time, and matter in terms of the notions of categories and sheaves. Hopefully, our attempt to place all the known phenomena of the macrocosm and microcosm universe within a topos theoretic structural framework will deepen our understanding of space and time. We will focus particularly on the situation when gravitational fields are extremely strong and also on the situation where distance scales are extremely short. How to capture the nature of space and time is an old problem first for philosophers and later for scientists equipped with sufficient mathematics.

During the twentieth and twenty-first centuries, human beings have made huge progress in cosmology. Even though we cannot answer such questions as why the creation of a universe happened and what caused the creation at a particular moment, we can try to answer how, i.e., in what way such a creation happened. However, for modern science, it is not enough to just argue about the nature of space and time. It is necessary for one to give explicit formulations without ambiguity so that people with critical minds can point out the flaws in the formulations if they exist. This level of explicitness is a requirement for twentieth and twenty-first century science. During the last five hundred years, mathematics has been the device to satisfy the requirement of explicitness and exactness for formulating theories. Especially from the early twentieth century, macrocosm and microcosm have been studied and formulated in terms of differential geometry for a general theory of relativity, and in terms of functional analysis, in particular Hilbert space and group theory for quantum physics. However accurate those main pillars, relativity theory and quantum physics of modern sciences might be in macrocosm and microcosm, respectively, we still live in this undivided world consisting of macro and micro worlds. A new question appears whether there exists a valid formulation both in macrocosm and microcosm. Namely, we seek *one theory* that gives formulations about space and time of the universe and microscopic world. Meanwhile in modern mathematics of post *WW II*, new theories developed in pure mathematics have become available in the studies of topology and complex analysis in several complex variables. These new notions are category theory and sheaf theory. For physical applications, we list the following: the works of Isham and Butterfield on Kochen-Specker Theorem and consistent histories [6, 7], the works of Mallios, Raptis, and Zafiris on Lorentzian quantum gravity as applications of abstract differential geometry [8, 9], the works of Guts and Grinkevich on the topos theoretic general relativity [10], and the work of the current author on temporal topos. See Bibliography in [8] for further papers on topos and categorical approaches in quantum gravity research.

The concepts in categories and sheaves will be used as unifying methods to formulate the physical space and time of macrocosm and microcosm in this paper. On the other hand, another aspect of our universe is the concept of consciousness. One may say it is a matter of self-awareness. The question is how a collection of elementary particles, or strings, become organized enough to form a human being or a brain. The *organization and emergence* can be captured as a process from local entities to global entities. The notion of a sheaf provides the global (conscious) entities. But we will focus on the physical aspects of a sheaf theory in this paper. In what follows we give our sheaf theoretic approach called the *temporal topos theory*, abbreviated as the theory of *t-topos*. The first postulate in the t-topos theory is that for every particle, whether it is in the particle state or wave state, there exists a corresponding presheaf associated with the particle. A presheaf is an assignment like a function or a mapping. Generally, a presheaf is an assignment from objects in a category to objects in another category. On the other hand, space and time are associated with sheaves, which are special presheaves satisfying the so called sheaf condition. One of the reasons to use sheaves instead of presheaves for space and time is that local data can be pasted together to obtain a familiar smooth global space and time.

25.2 "Things" Are Presheaves; The Category of Presheaves over a Site

The sets $\mathbb{R}$ and $\mathbb{C}$ of real numbers and complex numbers, respectively, have been used very successfully as foundations for differentiable functions (and more recently for differentiable manifolds) in established (macrocosm) physics since Newton's time. I. V. Volovich suggested the p-adic number approach to quantum physics in [11] instead of $\mathbb{R}$ and $\mathbb{C}$. Several people during the last century including Albert Einstein and Richard Feynman were aware of the necessity of the introduction of algebraic methods to deal with the discrete nature of the quantum microcosm. Namely, space and time seem to be discrete objects in the microcosm near the Planck scale.

Fundamental questions like, "What is the most fundamental entity in the universe?" and "What is space and what is time?" are the main issues. Philosophers and scientists over the last two or even three thousand years have given their thoughts at the levels of available scientific knowledge.

> "We can't solve problems by using the same kind of thinking we used when we created them." (Einstein)

As mathematical advances have been made throughout history, the languages used for describing physical phenomena were calculus and differential equations, and in the twentieth century, functional analysis, group theory, differential geometry, and near the end of twentieth century, algebraic geometry and algebraic analysis. As W.W.II ended, a new theory appeared in the area of pure mathematics called category theory developed by Samuel Eilenberg and Saunders Mac Lane. Another theory which plays a significant role in our theory is the concept of a sheaf. It first appeared in algebraic topology in the work of Jean Leray and also in function theory in several complex variables in the work of Kiyoshi Oka. However, in algebraic geometry, it was Alexander Grothendieck who pushed the notion of sheaves to the ultimate level of simplicity and purity.

Before discussing space and time, we will address matter since space, time, and matter play roles in this universal treatment. We have a big question: "What is reality?" It is our belief that reality cannot be described either by waves or particles as is done in the quantum world, but rather reality should be described in terms of presheaves over a site. Namely, there is a t-topos behind reality. Let us focus on particles, e.g. electrons and photons. For such a question as "What are things?" some people would answer that basic entities should be strings. Pythagoras (572–492 B.C.) said that things consist of numbers. Our answer would be, "Things are presheaves." The notion of a presheaf gives the wave-particle duality of e.g., an electron, which depends upon the presheaf being reified or not on an object of a t-site. The definitions of a presheaf and "being reified" will be given in what follows. The concept of a presheaf is coming from category theory. A category consists of objects and morphisms satisfying certain conditions. As an example of a category, let us consider the collection of sets. In the category (*Sets*) of sets, objects in (*Sets*) are sets and morphisms in (*Sets*) are set-theoretic mappings (functions) from a set to a set. Our fundamental device is coming from the notion of the category $\widetilde{S}$ of sheaves over a certain site. A general site is a category with a Grothendieck topology, a vast generalization of the notion of a topological space, e.g., a Euclidean space. This category $\widetilde{S}$ of sheaves is said to be a topos over a site S. For the definitions of a site and a Grothendieck topology, see [12–14].

However, as is often the case of an application of pure mathematics to physics, the theory of toposes cannot be used without additional conditions. Our t-topos theory is no exception. We first define a notion of a presheaf over a site or a category S. A presheaf m is a *contravariant functor* from S to a category K. We first need to explain what a functor is. A functor m from S to K is an assignment; for an object V in S, there is the assigned object $m(V)$ in K. This assignment must satisfy the following. For a morphism $V \xrightarrow{\alpha} U$, there is induced the morphism by m in K as

$$m(V) \xleftarrow{m(\alpha)} m(U).$$

Notice that the direction of the morphism $V \xrightarrow{\alpha} U$ is reversed in K to $m(V) \xleftarrow{m(\alpha)} m(U)$. Let t be the usual time corresponding to $\tau(V)$, and e.g., 1 s later $t + 1$ corresponds to $\tau(U)$. Then the above $V \xrightarrow{\alpha} U$ is said to be a *linearly* t-*ordered morphism*, or simply a t-*linear morphism*. In t-topos theory, we consider *non*-t-*linear* morphisms of t-site, especially for the situations sufficiently close to a creation of a universe and a black hole. Namely, there is no concept of "before and after" in the usual time sense. The concept of a t-*light cone* will be defined later. When $V \xrightarrow{\alpha} U$ is non-t-linear, $m(V)$ and $m(U)$ are not mutually in t-light cones. That is, $m(U)$ is outside the light cone of $m(V)$ in the usual sense. The change of direction from right to left of the induced morphism $m(\alpha)$ by a functor from the original direction of α, from left to right in the above, is the contravariantness. Note that for the sake of consistency only, one can define that a morphism from V to U does not exist for a non-t-linear situation; a morphism exists only for a t-linear morphism. A general presheaf, i.e., a contravariant functor, is defined for any object of a site V; however, first an additional condition for our theory postulates that a presheaf m is only defined for certain objects of the site. Such a restricted category of presheaves over a temporal site S is said to be a temporal topos. We denote the category as $\widehat{S}$. As mentioned already, our first axiom on the particle associated presheaf m with a particle $\overline{m}$ is the following.

Axiom 1 *For every particle $\overline{m}$, there exists the associated presheaf m in the t-topos $\widehat{S}$.*

Then we also say that m is representing a particle $\overline{m}$. Our approach is that all the physical phenomena, e.g., uncertainty principle, entanglement, etc., should be formulated in terms of the *t*-topos $\widehat{S}$ and the t-site S, namely in terms of categories and sheaves. Our theory is an observer-dependent theory in the following sense. When a particle associated presheaf m is observed (measured) by an observer presheaf P, we capture the observation (measurement) with the following diagram:

$$m(U) \xrightarrow{S^U_V} P(V). \tag{25.1}$$

Then the induced morphism S^U_V is regarded as a flow of information from the ur-particle state $m(U)$ of m to the observer in the ur-particle state $P(V)$. We consider both effects of m and P on e.g., spacetime sheaf ω. For the particle-wave duality, we introduce the following definition. A particle associated presheaf m is said to be in an *ur-particle state* if there exists an object V in the t-site S so that $m(V)$ is defined. We also say that the presheaf m is *reified* at a t-site object V. An object of the t-site is also called a *generalized time period*. For commonly used terminology in this paper, we refer to [5]. On the other hand, when m is not reified, i.e., not defined, by any object of the t-site S, presheaf m is said to be in an *ur-wave state*. Note that ur-particle and ur-wave states correspond to the usual particle and wave states. When a particle is observed (measured), the associated presheaf needs to be in an ur-particle state. Note that for a reified presheaf m over V, there does not exist a unique t-linear morphism from V to another object U at which m is reified. Namely, t-topos theory is not deterministic. In a way, the t-topos view is very non-deterministic since for a reified object V, there may not exist any morphism at all. In such a case, m will never again be in an ur-particle state. When there exists a non-t-linear morphism from V to W, the ur-particle state $m(W)$ is not within the t-light cone with respect to $m(V)$.

Since we define the sheaf $\omega = (\kappa, \tau)$ associated with spacetime as a final object in the t-topos $\widehat{S}$, for any presheaf m, there exists a unique morphism $m \xrightarrow{\sigma_m} \omega$. This unique morphism σ_m is regarded as the effect of a particle on spacetime. For example, when spacetime is measured by P over a generalized time period V, we have the following commutative diagram:

$$\begin{tikzcd} m(W) \arrow[r, "{}^{W}_{V}\sigma_m"] & \omega(V) \arrow[d, "S^{V}_{V'}"] \\ & P(V') \end{tikzcd} \qquad (25.2)$$

The above diagram may be read as follows. By measuring the ur-state of spacetime sheaf ω over V by the observer P over V', P receives information of the ur-state of m over W. This formulation, i.e., W, V, V' and are all different, is under consideration when the relativistic situation is non-trivial. Namely, for the case when non-global consideration is sufficient, all the t-site objects W, V, V' may be the same. Since ω is a final object of the t-topos, there exists a unique morphim $\sigma_P : P \longrightarrow \omega$ as well.

$$\begin{tikzcd}
m(W) \arrow[r, "{}^{W}_{V}\sigma_m"] & \omega(V) \arrow[d, "S^{V}_{V'}"] & P(V'') \arrow[l, "{}^{V''}_{V}\sigma_m"'] \\
 & P(V') &
\end{tikzcd} \tag{25.3}$$

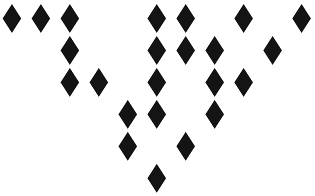
The diagram (25.3) above is the observer-dependent version of (25.2) in the following sense. If P measures the ur-state of spacetime over V, by composing morphisms from m and P over generalized time periods W and V'', respectively, P receives information over about m and P during the generalized time periods W and V''. Note that as for the ur-state of P, V'' precedes the ur-state over V' in the above t-linear sense. Note also that above formulations in (25.2) and (25.3) are in the relativistic form in the sense that all the states considered are within t-light cones. Here we say that two ur-particle states $m(V)$ and $l(U)$ of m and l over generalized time periods V and U are *mutually in* t-*light cones* if

1. there exist finite series of t-linear morphisms

   $V \longrightarrow V_1 \longrightarrow - - - - \longrightarrow V_i \longrightarrow - - -$ from V and

   $U \longrightarrow U_1 \longrightarrow - - - - \longrightarrow U_i \longrightarrow - - -$ from U, respectively, and
2. there exist i_0 and j_0 with a t-linear morphism $V_{i_0} \longrightarrow U_{j_0}$.

Note that the concept of a t-light cone can be defined in the t-site only. See also an equivalent definition of a t-light cone in [5]. As we will explain in Sect. 25.3, the notion of a micromorphism will show that a t-light cone in microcosm looks like a discontinuous cone as shown below.



One can formulate the deterministic classical aspect of the universe in terms of t-topos. Suppose that a particle associated presheaf m is reified at the generalized time period V. Then the deterministic view would be that there exists a uniquely determined t-linear morphism from V to U at which m will be in an ur-particle state

$m(U)$. In t-topos, for such a V, there may not even exist any morphism from V. Even when there is a morphism from V, the morphism may not be t-linear. The strict deterministic view would insist that such a t-linear morphism $V \xrightarrow{\approx} U$ is an isomorphism. Namely, knowing the complete information of the state $m(V)$ is equivalent to knowing the future state $m(U)$ and also its past state, i.e., deterministic. However, our t-topos view is that the behavior of a particle is fundamentally unpredictable due to the non-uniqueness of a morphism from V.

Another important notion in the theory of t-topos is that of t-entanglement. For the t-topos application to a double slit experiment, we refer to [15]. Let p and q be presheaves representing particles. Then p and q are said to be t-entangled when a pair (p, q) behaves as one presheaf. That is, when p is in an ur-particle state over a generalized time period V, then q is reified over the same V. Namely, we have the ur-particle state of the pair $(p, q)(V) = (p(V), q(V))$. For instance, if p is observed by P over V, then the ur-state of q is determined whether q is observed or not. (Note that in the relativistic formulation, we need to replace V with V' for P satisfying the t-light cone condition as we saw earlier.) This temporal topos simultaneity is because an observation, i.e., (measurement) $p(V) \xrightarrow{S_V^V} P(V)$ by P over V forces p to be in an ur-particle state as in (25.2). Then with the same V, presheaf q is reified at V, namely, the state of q over V is determined. Notice also that the notion of t-entanglement is independent of e.g., the distance between the two particles. Some of the consequences of our t-entanglement are the following. Let p and q be t-entangled over V and let $V \longrightarrow V'$ be a non-t-linear morphism. Then p and q are not in mutually t-light cones. In spite of that, an ur-particle state of one of the entangled entities can determine the ur-particle states of all other entangled entities. However, when one of the entangled entities is in an ur-wave state, one cannot tell whether the entities are entangled or not since all are in an ur-wave state.

25.3 The Large and the Small

Next we will focus on the microcosm aspect of t-topos theory so that we can discuss both macro- and micro- aspects of t-topos for space and time during the early universe. For a morphism $V \xrightarrow{f} U$ in the t-site S, we consider a factorization of $V \xrightarrow{f} U$ in the following sense. There exists an object W in the t-site S so that in the diagram below

$$\begin{tikzcd} V \arrow[r, "f_0"] & W \arrow[r, "f_1"] & U \end{tikzcd} \quad \text{or} \quad \begin{tikzcd} V \arrow[rr, "f"] \arrow[dr, "f_0"'] & & U \arrow[dl, "f_1"] \\ & W & \end{tikzcd}$$

we have $f = f_1 \circ f_0$. When such a factorization does not exist, $V \xrightarrow{f} U$ is said to be a *micromorphism*. This concept of a micromorphism is the underlying reason for the *uncertainty principle*. For example, let γ be a presheaf associated with a photon and let $V \xrightarrow{f} U$ be a t-linear micromorphism where photon presheaf γ is in ur-particle states at V and U. Namely we have the following diagram:

$$\gamma(V) \xleftarrow{\gamma_f} \gamma(U).$$

When $V \xrightarrow{f} U$ is a micromorphism, by definition, it is impossible for γ to be in a ur-particle state, i.e., between the time periods from $\tau(V)$ to $\tau(U)$, the photon presheaf γ is always in a ur-wave state. In terms of the position between V and U, positions can only take place at $\kappa(V)$ and $\kappa(U)$. Notice that if $V \xrightarrow{f} U$ is not a micromorphism, i.e., f can be factored into many micromorphisms, and much more accurate information about speed can be obtained.

Let m be any presheaf associated with a particle $\overline{m}$. Let us assume that m is measured twice: first at V and later at U. Namely, we have a t-linear morphism $V \xrightarrow{f} U$ which induces $m(V) \xleftarrow{m(f)} m(U)$ as before. Now we consider a mirofactorization of $V \xrightarrow{f} U$. Let $V \xrightarrow{f_0} V_1 \xrightarrow{f_1} - - - \xrightarrow{f_{n-2}} V_{n-1} \xrightarrow{f_{n-1}} U$ be a microfactorizaiton of $V \xrightarrow{f} U$. Then the *microfactorization number* of $V \xrightarrow{f} U$ is said to be n. Macrocosm and microcosm are distinguished by the microfactorization number of $V \xrightarrow{f} U$. That is, if n is large, $V \xrightarrow{f} U$ defines a macro event. If the *microfactorization number* $n = 1$, then $V \xrightarrow{f} U$ is a micromorphism.

Next we will introduce the dynamical aspect of t-topos theory. Quantum mechanics is flawless in microcosm even though it ignores the gravitational effect, and general relativity is practically flawless in macrocosm. At the beginning of the universe when the universe was a microcosm, gravity was not to be ignored. What we need is the notion of the universal gravitational effect in macrocosm and microcosm. For a given linearly t-ordered morphism $V \xrightarrow{f} U$, let n be the microfactorization number without the gravity effect. Then the *t. g. hypothesis (t-topos theoretic gravitational hypothesis)* says:

> the microfactorization number n decreases under a stronger gravitational effect.

This inequality can be viewed as a *micro-effect* in the sense that for any particle associated presheaf m, m becomes more likely to be in an ur-wave state. The *t. g. hypothesis* implies that entropy deceases. See [3, 5] for the notion of t-entropy. Namely, m behaves more like a microcosm entity under a stronger gravity. We will return to this topic. This g. t. hypothesis allows us to describe possible scenarios for the creation of a universe.

Before we go to the topics of space and time of ultra-early (and possible pre-) universes and back holes, let us focus on microlocalization of a presheaf in terms of the notion of a decomposition of presheaves into subpresheaves. For a macro particle associated presheaf M, consider its decomposition into presheaves associated with particles in a microcosm like elementary particles. In terms of categorical notions, consider a direct product of (sub-) presheaves of M. I.e.,

$$M = \prod_{k \in K} m_k, \tag{25.4}$$

where K is a finite index set. Let

$$M = \prod_{k \in K} m_k \xrightarrow{p_k} m_k$$

be the projection morphism. Notice that when a micro object presheaf m_k is measured over a generalized time period V, by an observer P over U (for non-relativistic consideration one may take $V = U$), as in (25.1), we have a morphism $m_k(V) \xrightarrow{S_U^V} P(U)$. Since M need not be defined over V (where m_k is reified at V), we cannot compose these two morphisms as $S_U^V \circ p_k(V)$ in the following: $M(V) = \prod_{k \in K} m_k(V) \xrightarrow{p_k(V)} m_k(V) \xrightarrow{S_U^V} P(U)$. Such an impossibility of the composition of the morphisms from $M(V)$ to $P(U)$ indicates that local information of M, i.e., m_k over V, gives no information to the observer P about the global object M.

On the other hand, we can consider a refinement for an object V of a t-site. For a covering $\left\{ V \xleftarrow{g_i} V_i \right\}_{i \in I}$ of V, we can further consider a covering $\left\{ V_i \xleftarrow{g_{i,j}} V_{i,j} \right\}_{j \in J}$ of V_i. We define a microcovering when such a covering morphism $g_{i,j,\ldots,k}$ becomes an isomorphism after a finite number of the above processes. See [5] for details.

The main assumption of our theory of temporal topos for our universe is the following:

> *Temporal Topos Postulate* (referred to as t-t-P): the category of presheaves over a t-site exists independently of time.

The appearance of infinity as a singularity at the very beginning of the universe is inadequate and unwelcome news for any theory including the big bang model. Neither density nor temperature is finite, i.e., the value ∞ at the singularity. As one of the motivations to get rid of meaningless infinity from physical theories, topos-categorical approaches have been proposed. This theory of t-topos is no exception. Space and time exist as sheaves even when they are not reified. A big bang hypothesis as a cosmic singularity would be a beginning of reified space, time, and matter. In t-t-P, the word "existence" is used in the non-physical context, i.e., in the mathematical context. Namely, "existence" in t-t-P is used without referring to time and location (i.e., when and where). Nothing is observable until a presheaf is reified, including space and time sheaves $\omega = (\kappa, \tau)$. Namely, before any object of the t-topos (including space and time) has potential to be measured, temporal topos exists in the above sense.

"Before" the beginning of our universe, all the presheaves are (were) non-reified, where "before" is defined as "not a single reified exists." Namely, this 'before' is in the sense of t-topos, which one may call "Temporal topos theoretic 'before' based on t-t-P."

The next state would be the case where some presheaves are reified. Let m be one of the reified presheaves over a t-site object V. Several possibilities may be considered following this state.

Case 1: the presheaf m may never be reified again. This case will not develop into any universe.

Case 2: m is reified at U besides V. Then we can consider two following cases. Subcase 2.1: $V \longrightarrow U$ is a non-t-linear morphism. In this case, $m(V)$ and $m(U)$ cannot be in the same universe. Or at least they cannot be mutually in t-light cones. Since there did not yet exist a measurable universe, this type of fluctuation can provide multiple universes. Subcase 2.2: $V \longrightarrow U$ is a t-linear morphism. This second subcase has potential to lead to a genuine universe provided that series of t-linear morphisms will continue to this day. We do not know what determined the right dimensions and just right forces for our universe we live in to have evolved as it has.

Particularly, as for space and time $\omega = (\kappa, \tau)$ as a final object of temporal topos $\widehat{S}$, we do not know how to postulate:

1. when space and time $\omega = (\kappa, \tau)$ turned from a presheaf to a sheaf, or
2. whether space and time are sheaves all the time, and
3. when space and time became t-entangled

so that we can still have consistent and rich enough axiomatic structures in the t-topos theory. For a presheaf m, there is a unique morphism $m \xrightarrow{\sigma_m} \omega = (\kappa, \tau)$ since the spacetime sheaf $\omega = (\kappa, \tau)$ is a final object of *t*-topos $\widehat{S}$. During those above cases (corresponding to ultra early stages of the universe), even if m and $\omega = (\kappa, \tau)$ are reified at objects X and Y of the t-site S, the induced morphism $m(X) \xrightarrow{(\sigma_m)_Y^X} \omega(Y)$ need not be t-linear. When spacetime sheaf $\omega = (\kappa, \tau)$ is to be affected by the reified ur-particle state of $m(X)$, the morphism $X \longrightarrow Y$ must be t-linear, i.e., $m(X) \xrightarrow{(\sigma_m)_Y^X} \omega(Y)$ needs to be in mutually in t-light cones.

In the formulation of the temporal topos, for such a question as where did space, time, and matter come from, our answer would be: they came from the (un-reified) t-topos $\widehat{S}$.

25.4 Epilogue

What is an adequate definition of time after all? Our definition of time is a sheaf as an object of t-topos. Our t.t.p. says that space and time associated sheaves always exist independently of time: maybe not reified *before* in the sense of t-topos as described in the above at the beginning of our universe. We know that space and time are very different entities. We can move from one place A at time t to another place B at time s, and then we can come back from B at time $t\prime$ to A at time $s\prime$ again. But we cannot return to the previous time, even though during the ultra early universe, as we saw in the above via t-topos, the linearity of time is not guaranteed. When time loses the linearly t-orderedness, a beginning of time becomes ambiguous. A before-after based big bang singularity relying upon the completeness of the set of real numbers has a different meaning. Before creation, time sheaf was (is) not reified. Consequently, there is no reified time $\tau(V)$ to be measured. Just as an image, our beginning of space and time is similar to the scenario proposed by Alexander Vilenkin, i.e., the creation of our universe occurred as a process of quantum tunneling of nothing in [16]. Namely, our universe may have come out of a quantum vacuum. Our theory may be phrased as follows. Our universe occurred as a process of reification of topos of presheaves. By interpreting "Nothing" as un-reified presheaves (hence non-measurable in a t-topos sense), including spacetime sheaf ω, a phrase like a universe is created from "nothing" is a consequence of our method as well. Ultra locally, our universe looks like a fluctuating collection of reified and un-reified presheaves associated with elementary particles and spacetime lasting only sub-Planck length time by keeping the total energy at almost zero.

References

1. Kato, G. C. (2004). Elemental principle of t-Topos. *Europhysics Letters*, *68*(4), 477–482.
2. Kato, G. C. (2005). Elemental t. g. principle of relativistic t-topos. *Europhysics Letters*, *71*(2), 172–178.
3. Kato, G. C. (2010). U-singularities and t-topos theoretic entropy. *International Journal of Theoretical Physics*, *49*, 1952–1960.
4. Kato, G. C. (2011). Temporal topos and U-singularities. In J. R. O'Cornnell & A. L. Hale (Eds.), *The Big Bang: Theory, Assumptions and Problems* (pp. 169–179). Nova Science Publishers Inc.
5. Kato, G. C. (2013). *Elements of temporal topos, abramis academic*. Suffolk, U.K.: Arima Publishing.
6. Butterfield, J., & Isham, C. (1998). A topos perspective on the Kochen-Specker theorem: I. quantum states as generalized valuations. *International Journal of Theoretical Physics, 37*, 2669.
7. Butterfield, J., & Isham, C. (1999). A topos perspective on the Kochen-Specker Theorem:II. Conceptual aspects and classical analogues. *International Journal of Theoretical Physics*, *38*, 827.
8. Mallios, A., & Zafiris, E. (2015). *Differential sheaves and connections: A natural approach to physical geometry*. World Scientific.
9. Raptis, I. (2006). Finitary-Algebraic resolution of the inner Schwarzschild singularity. *International Journal of Theoretical Physics, 45*(1).
10. Guts, A. K., & Grinkevich, E. B. (1996). *Toposes in general relativity*. arXiv:gr-qc/9610073.
11. Vladimirov, V. S., Volovich, I. V., & Zelenov, E. I. (1994). *p-Adic analysis and mathematical physics*. World Scientific (See also Volovich, I. V. (1987). Number theory as the ultimate physical theory. Preprint CERN-TH. 87, pp. 4781–4786).
12. Kashiwara, M., & Schapira, P. (2006). *Sheaves and categories*. Springer.
13. Gelfand, S. I.,. & Manin, Y. I. (1996). *Methods of homological algebra*. Springer.

14. Kato, C. G. (2006). *The heart of cohomology*. Springer.
15. Kato, C. G., & Tanaka, T. (2006). Double slit interference and temporal topos. *Foundations of Physics*, *36*(11), 1681–1700.
16. Vilenkin, A. (1982). Creation of universes from nothing. *Physical Letters*, *117B*(1), 2.

Part IV
Biology/Cognitive Science

Chapter 26
Syntactic Space

Rajesh Kasturirangan

26.1 Introduction

Space, as they say, is everywhere. Since we are visual creatures and visual experience is spatial, space (with a visual tinge) dominates our world. While visual perception is enormously useful—it's the only sensory modality that transcends our earthly existence; no other sense can grasp stars and galaxies—it also prevents us from understanding the world of other creatures. In a now infamous article,[1] Thomas Nagel asked what it's like to be a bat, suggesting that nothing we know about the anatomy or physiology of a bat prepares us to understand the subjective experience of bats or, for that matter, whether bats have any subjective experience at all.

This paper starts with a question in response to Nagel: is there a form of subjective space that transcends the particularities of vision or audition? More generally, *what is subjective space*?

Of course, vision and audition aren't the only mutually incomprehensible mental capacities. Most of us have heard a language we don't understand. Some of us have been exposed to languages that we don't even *begin* to understand—say, a native English speaker who travels to Namibia and overhears! Kung speakers in conversation. Today it's increasingly clear that the languages of the world have much in common.[2] What if there is a spatial instinct that was much more primeval than a language instinct? What would such an instinct look like? One possibility, once again in analogy with the situation in language, is that there's spatial syntax that serves as the ground upon which the different sensory modalities encode space.

---

[1] Nagel, Thomas. "What is it like to be a bat?" *The philosophical review* 83.4 (1974): 435–450.
[2] Steven, Pinker. "The language instinct." *How the Mind Creates Language. NY* (1994).

R. Kasturirangan (✉)
Mind and Society Initiative, Azim Premji University, Bangalore, India
e-mail: rkasturi@gmail.com

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_26

The topic of spatial syntax is too vast to address in a single paper or even a book. My goal here is much more limited: to introduce the idea of syntactic space, to argue that it's a viable hypothesis for modeling space that transcends sensory modalities while being rich enough to capture important aspects of the *geometry of experience*. This paper has three parts.

In part one, I introduce the idea of syntactic space with an emphasis on heuristics that might be the target of future formalization. Part two is an outline of three well known accounts of perception—Gestaltist, Gibsonian and Enactive—that have theorized space from an observer's perspective. In my opinion, these three theories are our richest sources of insights about the geometry of experience within a scientific framework. Finally, in part three, I come back to the original question, "what is subjective space" and speculate on what a synthesis might look like.

26.2 Syntactic Space

I borrow the notion of syntax from the study of language,[3] where syntax denotes grammatical features that have little semantic content. Syntactic constructions are compatible with a range of semantic content. For example, the construction "(noun) X did (verb) V to (noun) Y" can be used to express "John kissed Mary" as well as "Mary greeted Meera" and any number of other sentences.[4]

What is spatial syntax? Let me introduce the concept with a couple of examples. Normal visual experience is rich in content. Consider an image like Fig. 26.1.

We easily recognize a city street with people, buses, flowering trees, street signs and storefronts of international brands. It's also clear that some people are crossing the street and others are walking along the sidewalks. Human beings recognize such scenes effortlessly and make their way through them when compelled to do so. Now compare that scene with this one (Fig. 26.2).

Figure 26.2 has as many objects as the previous one, but with much less meaning. While you can't permute the objects in Fig. 26.1 (just imagine replacing the people on the sidewalk with buses and filling the street with people), there's no reason why we wouldn't accept a field of rocks with the individual stones switched around.

At the same time, we would know how to act when faced with a scene as in Fig. 26.2. We would know how to pick up a small rock that catches our fancy, avoid a sharp stone while walking over the terrain or throw a flat one into the neighboring lake to watch it skip. The skeletal information we perceive from Fig. 26.2 is enough for many purposes. Most scenes have that information, even if

---

[3] Chomsky, Noam. *Syntactic structures*. Walter de Gruyter, 2002.

[4] The analogy to space ends with our emphasis on spatial constructions that have wide application, for the formal study of syntax focuses on features such as recursion that will not concern us here.

Fig. 26.1 A busy street

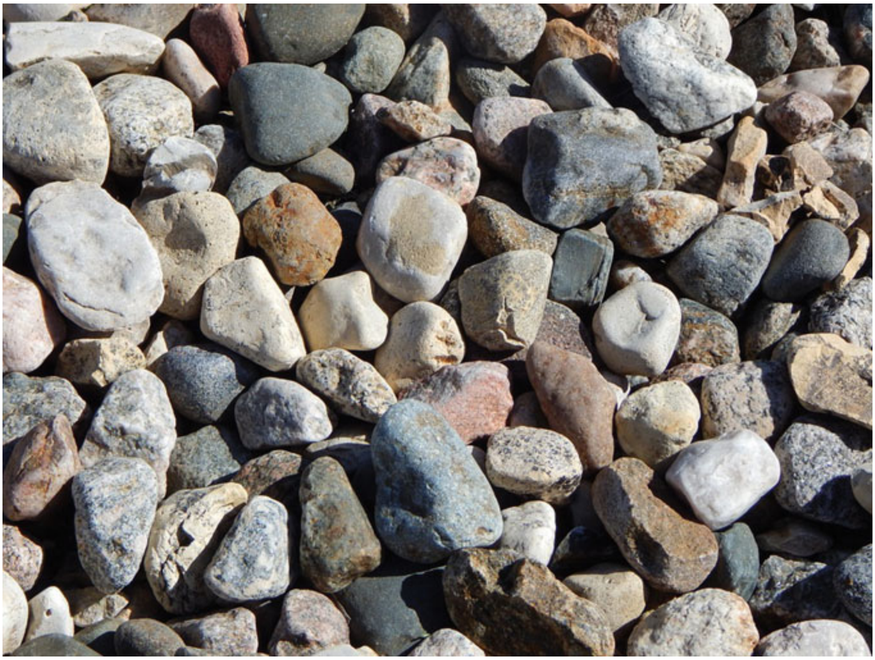


Fig. 26.2 Texture without meaning

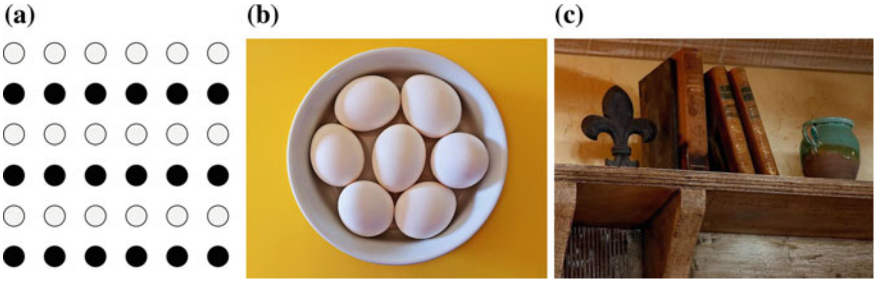


Fig. 26.3 Universal spatial properties

they don't have rich content. By *spatial syntax*, I mean universal spatial properties that are available in every scene.

These include features such as visual groups (Fig. 26.3a), topological relations such as containment (Fig. 26.3b) and features that track physicality such as support (Fig. 26.3c). Some of these properties are more like to be universal than others; for example, a deep sea fish or bacteria in a viscous liquid are unlikely to care much about support, but containment and grouping ought to matter to them. To summarize, spatial syntax consists of clusters of features that occur repeatedly in organismic environments. In that, they are like a pattern language for life in a spatial world.[5]

Is there a complete catalog of these patterns? Can they be deduced from underlying principles? I think it's too early to offer a definite answer. These properties don't have an obvious origin—they mix metric properties (grouping) with topological ones (containment) and when you add physical-functional properties such as support, you're mixing physics, geometry and topology in a manner that doesn't have much precedent. The closest we have to a theory of these features is the notion of an image schema[6] but that's a descriptive account rather than an explanatory one.

If that isn't enough reason to be cautious, it becomes clear that we are at an early stage of our investigation when we start probing how spatial syntactic features come together to form complex layouts. For example, unlike the case of grammar, there's no obvious analog of the sentence. What's a believable layout? What feels unacceptable? Our answers to these questions are likely to be statistical.[7]

Another direction of theoretical progress concerns the origins of spatial organization. Where do these syntactic features come from? The theories addressing this question have been among the most influential in the study of the mind, even

---

[5] Alexander, Christopher, Sara Ishikawa, and Murray Silverstein. "A pattern language."

[6] Hampe, Beate, and Joseph E Grady. *From perception to meaning: Image schemas in cognitive linguistics*. Walter de Gruyter, 2005.

[7] Geisler, Wilson S. "Visual perception and the statistical properties of natural scenes." *Annu. Rev. Psychol.* 59 (2008): 167–192.

influencing modern technology via design.[8] I am going to focus on three accounts of perception: gestalt, gibsonian and enactive.

26.3 Three Theories of Perception: Gestalt, Direct and Enactive

There[9] are many problems of perceptual organization, but I am going to emphasize a core question: *why do humans (and other creatures) live in a complete and coherent world*?

All of us have heard of the blind spot in the retina. The input to our eyes is incomplete, yet we don't walk around with a hole in our visual field. More generally, the stimulus that enters our sensory receptors is limited, noisy and unstable while our experience is holistic, coherent and stable. What explains that discrepancy?

Why isn't our experience a blooming buzzing confusion that William James imputed to babies?[10] Mobile creatures have very limited exposure to the environment; we can take in only so much in a single gaze and every time we move our head or saccade our eyes, the input changes, sometimes dramatically. Yet, our perception of the world is stable. Why is that so?

The gestalt psychologists were the first to provide a convincing alternative to the blooming buzzing confusion. In their view, the mind organizes experience according to its own principles. Consider the image in Fig. 26.4.

Doesn't it look like a triangle with white borders obscuring a triangle with black borders? In principle it could be rather different, i.e., a triangle with three pac-man figures surrounding it, but we don't see it that way. In the Gestaltist account, that because our minds automatically *completes* the input to form two figures. That completion—called amodal completion in the literature[11]—is a property of the mind, not of the stimulus. So if we are to capture the gestalt theory of complete coherent experience in one line, it is:

Experience is complete because our mind imposes structure upon the incoming stimulus. Completeness is a property of the mind.

The perceptual scientist J.J. Gibson was greatly influenced by the gestaltists, but he ended up with the exact opposite conclusion, that *completeness is a property of the world*. How did he come to that conclusion?

---

[8] Norman, Donald A. *The design of everyday things: Revised and expanded edition*. Basic books, 2013.

[9] My treatment of these three theories is sketchy at best; the goal is to introduce you to three alternate accounts of organized experience rather than a full articulation of these theories.

[10] James, William. "The Principles of Psychology. 2 vols. New York: Henry Holt and Co. Reprint." (1950).

[11] Kanisza, G, and Walter Gerbino. "Amodal completion: Seeing or thinking." *Organization and representation in perception. Erlbaum, Hillsdale, NJ* (1982).

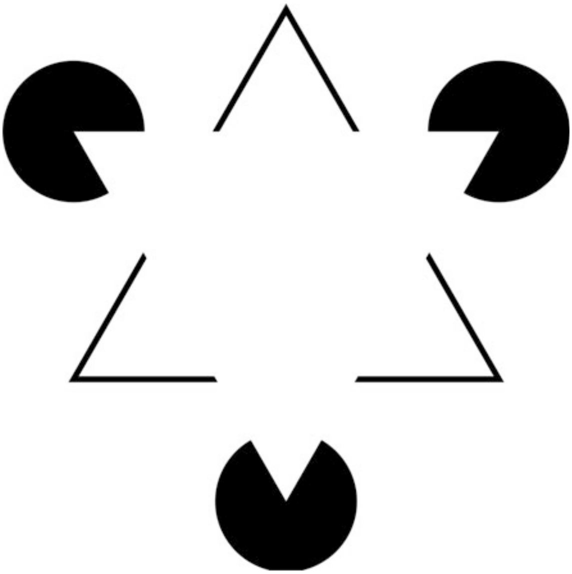


Fig. 26.4 The kanisza triangle

Gibson starts by noticing that the black bordered triangle is hidden behind the white bordered triangle only if you're rooted to one spot which no animal ever expects to do so in an ecologically realistic situation. Moving your head and body around will quickly convince you that it's no pac-man biting a triangle, for as you move, you will notice the hidden parts of the circle. Even better, moving will help you disambiguate whether the hidden object is a circle or some other continuous object whose contour matches that of a circle. Gibson's view[12] is radically externalist, i.e., his theory of perception is about the external world, not our mind/brain's representation of it. He's also the first scientist to point out that the basic unit of perception is an animal that's moving its eyes, head and body rather than a static pinhole camera. In summary, here's Gibson's explanation for the completeness of visual perception:

Perceptual experience is complete because the world is complete. The circle is perceived as a circle because it's really a circle and it reveals itself to be one.

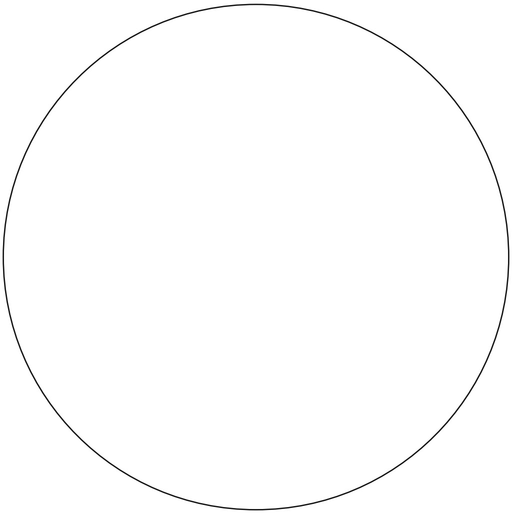
In other words, Gibson says the world is both coherent and transparent. The enactive view of perception[13] fleshes out Gibson's emphasis on the mobile animal; while the enactivists agree with his premise that an animal isn't a pinhole camera, they place less emphasis on the informational structure of the external environment and more emphasis on the sensorimotor coupling of the organism with its environment.

---

[12] Gibson, James J. *The ecological approach to visual perception: classic edition*. Psychology Press, 2014.

[13] Rosch, Eleanor, Francisco Varela, and Evan Thompson. "The embodied mind." *Cognitive Science and Human Experience* (1991).

Fig. 26.5 A simple circle



What does that mean for our understanding of perceptual organization? Let's take the circle (Fig. 26.5) that reveals itself according to Gibson. Enactivists point out that in most views, the circle projects to an ellipse (unless you're looking straight at it), and they also argue that *looks* like a circle precisely because it's *viewed* as an ellipse.[14] While Gibson would only focus on the circle, the enactivists claim that we can explain perception only through the sensorimotor understanding of how ellipses are generated by an underlying circle as the animal moves around. The world doesn't reveal itself in one go; it reveals itself only to the go-getting creature.

To the enactivist, perception is complete because the world is accessible to a mobile creature.

The Gibsonian and enactivist views are close to each other—in both cases, the shape of an object in the world is an invariant revealed by the motions of an animal. Gibson emphasizes the acquisition of that environment in one shot, what he calls direct perception.[15] The enactivists emphasize the agents motions through the environment and the sensorimotor coupling between the two; the global invariant, i..e, shape, may never be fully revealed but as long as it's there *in the world*, relevant aspects can be accessed as necessary. I don't have any idea of the shape of a mountain but I do have an understanding of its local contours as I hike up its trails. That's good enough for me.

So we have covered three accounts of completeness: it's imposed by the mind of the animal (Gestalt), the world is already complete (Gibson) and the world is potentially complete and every aspect is available to the probing creature (enactive). Note how all three theories are theories of how the world presents itself to an animal. The

---

[14] Noë, Alva. *Action in perception*. MIT press, 2004.

[15] Gibson, James J. *The ecological approach to visual perception: classic edition*. Psychology Press, 2014.

animal world isn't the same as the material world of physics, for it comes with a point of view—the animal's perspective. Where the three disagree is in their explanation of the source of the organization. The gestaltist believes that the organization is imposed by the mind of the animal; it's radically anti-physical. Gibson believes that the world is intrinsically organized, but even for him, that organization doesn't consist of the laws of physics; instead the organism picks up features relevant to its ecological needs. The fact that chairs are *sit-table* isn't a fact of physics but a fact of the human environment. Finally, the enactivists locates the world's organization in the interaction between the organism and its environment. I know that chair is sit-table because I know that if I position my body just right I will land on the chair's surface.

They are all compatible with our idea of spatial syntax, though they would each put a different spin on it; the gestaltists would say that these features are imposed by the mind, the Gibsonians would say they are directly perceptible and the enactivists would attribute them to sensorimotor invariants. So would any other theory that imputes some intrinsic structure to the organismic world. The main contrast is with the blooming buzzing confusion, i.e., a disorganized chaos. James and other empiricists thought that humans are born with a blank slate and acquire knowledge through their contact with the world. That hypothesis about the fragmented nature of children's experience has been convincingly rebutted by child psychologists such as Elizabeth Spelke.[16] We don't need to enter into that debate except to say that *whether we start with a blank slate or a full plate, our experience is never confusing; it's always whole and complete.*

The baby lives in a baby's world. The adult lives in an adult's world and the path between the two is always from one complete world to another. By the way, the completeness of our world isn't due to the completeness of our perception; researchers are increasingly aware that we represent the world in fragments,[17] lacking awareness of major changes if they don't interfere with our task at hand. Why wouldn't it be so when the world is there to supply whatever is missing? Perception isn't complete but the *world* we inhabit is complete.

Indeed it's the completeness of our world that makes it hard for us to imagine ourselves into the bat's world. The visual array of colour, shape and form is so obviously *the world* for us that we have a hard time understanding what it's like to be an organism whose world is not made out of visual shape and form. The 'what it's like for us to be human' is an impediment toward understanding what it's like for a bat to be a bat.

26.4 Conclusion: The Universality of Spatial Syntax

Every animal lives in a world of its own. The world of ticks may be limited, signaled entirely by butyric acid. The human world is richer, with signs coming from several senses. However, both the tick and the tyrant are content in their

---

[16] Spelke, Elizabeth. "Initial knowledge: Six suggestions." *Cognition* 50.1 (1994): 431–445.

[17] Simons, Daniel J, and Ronald A Rensink. "Change blindness: Past, present, and future." *Trends in cognitive sciences* 9.1 (2005): 16–20.

respective worlds, believing that it has everything one might desire. These worlds aren't separate; the tick can suck my blood and I can squash it in return, but there might not be much else in common. Indeed, we might worry with Nagel that only features our world shares with the tick's word are *objective* features of physics, the features that make it possible for the tick to bite me and for me to crush the tick.

Does spatial syntax offer a way out of Nagel's cave? Are there universal features of the subjective experience that transcend species boundaries? I can only speculate but let me do so with our account of spatial syntax in the background. If these features are indeed universal, depending only on being a creature moving in a spatial environment, then a first step toward understanding the world of the bat would be in testing whether it responds to these features in *some* sensory modality.

Both bats and humans are mobile, tactile creatures, features we share with every animal on earth. Tactility is even more general, since we share it with plants as well. Our exploration of perceptual organization suggests the possibility that the world of every creature is an organized whole, as hypothesized by Von Uexkull a while ago.[18] He didn't have the benefit of a formal account that would underlie his theory, but we can hypothesize that something like spatial syntax must be a building block of any animal world. Here's our heuristic:

- Every mobile, tactile creature shares a world with other mobile, tactile creatures. Put another way, spatial syntax is available to every mobile, tactile creature.

Let me reiterate the claim once again: what it is to be a mobile, tactile creature is—at some elemental level—like what it's to be another mobile, tactile creature.

I might be guilty of overreach here—a bacterium chemotaxing its way to food is as much a mobile tactile creature as I am—but at this preliminary stage it's better to be generous and draw the boundary as widely as possible. Then there are other worries: some robots are also mobile, tactile devices. Is it the same to be me as it is to be a robot at some level?

These are real worries though not ones we can address in this paper or anywhere else at this moment. We share a long evolutionary history with flies and ticks and it might be wise to address the question "what is it like to be a mobile, tactile creature conditional upon a billion plus years of shared history?" than without assuming that shared heritage. Then again, intuition suggests that evolutionary history shouldn't enter the question at all—if I was imaged in a molecular 3D scanner and that image was used to print out a replica that moved and talked, I suspect the replica would share much of my consciousness.

Too many conflicting intuitions. All we know is that tactility and motility unifies much of life from bacteria to blue whales and that's as good a starting point for an investigation of shared experience as anything else. A skeptic wouldn't believe these arguments, but if you're inclined to be generous when it comes to a shared sense of being alive, then here's a way for us to theorize our common position.

---

[18] Von Uexkull, J. "A Stroll through the Worlds of Animals and Men (C. Schiller, Trans.)." *New York: International Universities.* (*Original work published 1934*) (1957).

Chapter 27
Time Measurement in Living Systems: Human Understanding and Health Implications

Lakshman Abhilash and Vijay Kumar Sharma

> *"Everything comes to us that belongs to us if we create the capacity to receive it".*
>
> —Rabindranath Tagore
>
> (Limits of our understanding are stunted, but only by the capacity of human faculty)

Abstract It is now very well accepted that over the history of life on Earth, organisms ranging from bacteria to humans have evolved mechanisms that can quite efficiently measure passage of time that help in restricting behavior, physiology and metabolism to certain times of the day, thereby generating overt rhythms. In this article, we discuss answers to questions that a naïve but interested reader may ask regarding such intricate biological timing systems. Do we really have biological clocks? And if so, where are they located? How do they work? More importantly, why do we have them, and why is it important to study them? In an attempt to answer these questions, we first discuss studies that demonstrate the endogenous and innate nature of biological rhythms. We then discuss how the mechanisms that generate such rhythms serve as our very own biological clocks to keep time efficiently. Further, we elaborate on the anatomical location of biological clocks, and describe the molecular processes underlying them. Penultimately, we discuss the origins and possible reasons for the existence of such clocks, and finally describe the implications of aberrant and misaligned (out-of-sync with the external environment) clocks on human health.

Keywords Circadian rhythms · Biological clocks · Jet-lag · Sleep disorders · Seasonal affective disorder (SAD) · Bright light therapy

---

L. Abhilash · V.K. Sharma (✉)
Chronobiology Laboratory, Evolutionary and Organismal Biology Unit,
Jawaharlal Nehru Centre for Advanced Scientific Research,
Jakkur, Bangalore 560064, Karnataka, India
e-mail: vsharma@jncasr.ac.in

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_27

27.1 Dawn of ‘Time’ in Biology

The concept of ‘time’ has been of profound interest and a topic of debate for mankind with deep roots in religion and philosophy from ancient times to the present day, with some philosophers even dismissing time as being a mental construct and ‘unreal’ [1, 2]. While a debate on the ‘nature of time’ is worthy of a chapter in itself, here, however, we will discuss ‘time *in* nature’. Have we ever wondered why is it that we sleep at certain times of the day and eat at certain times? Why we work more efficiently at some times of the day and less efficiently at other times? Is such a time-dependent phenomenon restricted only to us humans or is it common to other organisms? This article will discuss efforts of mankind in understanding time-keeping and temporal order in living systems and its clinical implications.

Aristotle, as early as ~4th century BC, observed that several other animals also sleep at night just like us humans [3]. Although such descriptions of temporal processes in biology were among the first of its kind, eventually a major concentration of the discipline steadily moved into the domain of spatial biology wherein researchers were busy studying anatomical location of organs, location and function of cellular components (with the advent of microscopy), geographical distribution of organisms, and theories of organismal development [4].

The study of ‘time’ in modern biology, a topic that will remain the focus of this article, however, gained impetus only by ~18th century AD with a French astronomer, Jean-Jacques d’Ortous de Mairan, discovering that leaves of the touch-me-not plant (*Mimosa pudica*) open and close at certain fixed times of the day [5] through a mechanism that appeared to be internal to the organism rather than a mere response to external cycles, thus paving the way for the study of yet another extraordinary feature of living systems—a biological clock!

27.2 Rhythms of Life

Following de Mairan’s discovery, systematic studies revealed rhythmic processes in a range of behavioral, physiological and metabolic processes with many different periodicities, and across several organisms (Table 27.1). For instance, rhythms with very short periodicities (on the order of seconds to minutes) are observed in respiration, electrocardiograms (ECG), electroencephalograms (EEG) and neuronal firing. On the other hand, there are rhythms in sleep-wake, foraging, and several nephrological variables in mammals that have intermediate periodicities (on the order of hours); while estrous cycles in non-primate mammals and menstrual cycles in humans, migration and reproduction in birds, flowering of bamboos, and quite intriguingly suicides in humans have relatively longer periodicities (on the order of months to years) [3, 6, 7]. Among such rhythms, the most widely distributed and commonly studied are the ones that have near 24-h periodicities (‘*circadian*’, meaning ‘*approximately-a-day*’) [5, 8].

Table 27.1 The spectrum of rhythmicity observed in life-forms

Rhythm Type	**Period range**	**Examples**
ULTRADIAN *ultra*, 'beyond' *dies*, 'day' Rhythms with frequencies much higher than that of a day	$\tau < 20$-h	1. Respiration rate 2. Electrocardiogram 3. Electroencephalogram 4. Neuronal firing 5. Cell-cycle 6. Reproductive hormones 7. Muscle contractions 8. Human impatience 9. Foraging in crabs
CIRCADIAN *circa*, 'about' *dies*, 'day'	20-h $< \tau <$ 28-h	1. Sleep-wake 2. Feeding 3. Urinary variables 4. Eclosion[a] 5. Reproductive activity 6. Blood pressure 7. Body temperature 8. Locomotion 9. Egg-laying in insects
INFRADIAN *infra*, 'below' *dies*, 'day' Rhythms with frequencies much lower than that of a day	$\tau > 28$-h	1. Menstrual cycle 2. Flowering of bamboo 3. Hibernation 4. Body temperature 5. Suicides 6. Bird migration 7. Breeding 8. Weight gain 9. Melatonin secretion

'τ' refers to the inherent periodicity of the rhythm that is expressed under constant (non-cyclic) conditions. [a]Eclosion is the act of adults emerging from pupal cases in certain types of insects

Circadian rhythms are thought to have evolved in almost all life-forms on Earth as a consequence of the 24-h cycling of its physical environment (such as light and temperature) caused by Earth's rotation about its own axis. Skeptical researchers, for a very long time, dismissed the notion that such rhythms were indeed endogenously generated by arguing that they are passive responses to unidentified geophysical cycles arising due to Earth's rotation [3]. However, experiments in a spacelab using a fungus (*Neurospora*) revealed that the circadian rhythm in conidiation (generation of a fungal reproductive unit) persists in the absence of any earthly influences thereby suggesting that such rhythms are not passive responses to physical cycles on Earth [3]. These rhythms, moreover, were also observed to persist under constant (non-cyclic) conditions of the laboratory with an inherent periodicity that is close to but significantly deviating from 24-h indicating that they are endogenously generated. Additionally, breeding experiments using ‘pure’ bred lines with deviant periodicities revealed that such rhythms are indeed innate (genetically inherited) [3].

The presence of such endogenous and innate rhythms that persist under non-cyclic conditions justifiably warrants the notion that there are mechanisms within organisms that are capable of keeping track of time on a daily scale (circadian timing system) and drive timed expression of behavior, physiology and metabolism.

27.3 Circadian Timing Systems as Biological Clocks: A Time-Keeper Within?

Temporal order can be maintained only if time-keeping is efficient, and this raised two major questions that concerned researchers regarding the nature of time-keeping in organisms, (i) ‘do organisms have the ability to measure passage of time?’ and (ii) ‘can circadian timing systems serve as biological clocks?’

Gustav Kramer, a well-known German zoologist and ornithologist, while studying bird navigation stumbled upon some remarkable evidence which suggested that organisms may have the ability to measure time. After training birds to feed at an artificial feeder, located at a particular angle relative to the sun, he demonstrated that birds navigate to the same feeder even after several hours despite the sun being in a different position relative to the feeder, suggesting that birds have the ability to correct their navigation system for the motion of the sun by measuring passage of time between the training and test events. This provided evidence supporting the notion that birds can measure passage of time [5]. Additionally, experiments by Karl von Frisch, an Austrian ethologist and Nobel laureate demonstrated that organisms can keep time and that circadian timing systems may play a key role in doing so. von Frisch and his student Ingeborg Beling marked individual bees and trained them to feed on sugar solution at an artificial feeder at the same time every day (once every 24-h), and on the test day did not provide the sugar solution, and noted that most individuals arrived at the feeder within the training time, thereby suggesting time-keeping ability in honeybees. Such training was not successful when the bees were trained to feed once every 19 or 48-h, implying that a 24-h timing system was in use for time-measurement [5].

However, rate of biochemical reactions (that may govern circadian rhythms) vary with changes in temperature akin to early mechanical watches, which had metal balance springs that would expand or contract depending on the temperature, thereby providing incorrect estimates of time. Colin Pittendrigh (one of the pioneers of the study of circadian rhythms) asked if natural selection has solved this problem of accurate time-keeping under different ambient temperatures in organisms, and since this aspect would be more relevant to poikilotherms (individuals whose body temperature changes with changes in environmental temperatures), and not so much to mammals such as ourselves, he used the fruit fly (*Drosophila*) model to address this question. Pittendrigh demonstrated that the period of circadian rhythms in *Drosophila* (and therefore its ability to measure time) under constant conditions was indeed maintained stably at high and low ambient temperatures, a phenomenon referred to as temperature compensation [5].

Several such experiments revealed, for the first time, that the biochemical systems that generate persistent, endogenous and innate circadian rhythms could measure passage of time robustly across temperature fluctuations, thereby suggesting that the functional significance of such systems is to appropriately schedule behavioral, physiological and metabolic activities in-sync with cycling abiotic and biotic factors of the environment. This synchronization of endogenous rhythms to the environmental cycles by daily resetting of the underlying 'biological/circadian clocks' in response to time-cues is known as entrainment (described elaborately in [5]), and is considered to be one of the clock's most crucial functions.

27.4 Where Are Biological Clocks Located and How Do They Govern Rhythms?

After identifying the major properties of biological clocks, studies were conducted in order to locate them and understand how such clocks function at molecular and physiological levels. However, due to space constraints, we will highlight only the key evolutionarily conserved features of the anatomy and molecular-physiological cogs and gears of biological clocks hereon and offer our sincere apologies to others whose spectacular work we are unable to discuss.

Early studies in cockroaches revealed that their biological clocks comprise a group of neurons located in the optic lobes, ventrally between the medulla and lobula, and in similar regions in crickets, grasshoppers and desert beetles [9]. Furthermore, in *Drosophila*, transplantation experiments along with brain mutants, mosaics and transgenic flies led to the identification of a network of ~150 lateral and dorsal neurons in the fly brain that serve as biological clocks [9]. In these insect models, one of the remarkable features is that similar neurochemicals such as PDH and PDF (pigment dispersing hormone and factor, respectively) mediate communication among constituent neurons of the biological clock circuit to bring about coherent overt rhythms [9].

Additionally, lesioning various parts of the mammalian brain proved to be useful in identifying the location of biological clocks in mammals, wherein a paired neuronal structure in the antero-ventral hypothalamus, just above the optic-chiasm on either side of the third ventricle (Suprachiasmatic Nuclei/SCN) of the brain was discovered to be the mammalian biological clock [10]. Each unilateral SCN of the mouse contains ~10,000 neurons, and is divided into a ventral 'core' and a dorsal 'shell.' Each of these neurons in the SCN is known to function autonomously to produce circadian oscillations in spontaneous firing as indicated by multi-electrode recordings from cultured tissues. Coupling of these individual SCN neurons via neuropeptide signaling and/or gap-junctions, gives rise to robust circadian rhythms [10]. At a more gross level, the core receives direct light-input from the eye (perceived by photoreceptors such as melanopsin) [10], via the retino-hypothalamic tract (RHT) and produces several neurotransmitters including VIP (vasoactive intestinal polypeptide), CAL (calretinin), NT (neurotensin) and GRP (gastrin releasing peptide). The shell, on the other hand, receives inputs from the core and expresses the neurotransmitters AVP (arginine vasopressin), angiotensin-II and met-enkephalin [10]. Although it is thought that the SCN core and shell talk to each other primarily using GABA (γ-aminobutyric acid) to produce coherent rhythms, this model is heuristic and the physiology of biological clocks in mammals may be much more complex than what is described here [10].

These studies point towards the fact that biological clocks are nervous tissue-based systems and the fact that, in most cases, biological clocks are in physical proximity to the eye, or have direct connections to it, suggest that light may be the primary source of temporal information for them.

Subsequently, several studies demonstrated the presence of not one, but many biological clocks within an organism, and we will discuss results of some such experiments here. Hamsters are nocturnal, and display rhythmic activity with higher and consolidated activity during night-time. Curiously, it was observed that when these hamsters were moved from light-dark cycles to constant illumination, their activity bouts tend to split into two or more components with each component often exhibiting a different period (suggestive of different clocks) [5]. In another study, human subjects living in isolation in Norway (during the Arctic summer, therefore constant illumination) were subjected to artificial light-dark cycles with 16.5-h of light and 7.5-h of darkness and the rhythms in sodium, potassium and creatinine excretion were regularly monitored. Once the rhythms stabilized, the light-dark cycle was reversed. It was observed that the rate at which each of these rhythms resynchronized to the newly imposed light-dark cycles was different, suggesting independent control of these rhythms [5]. Such experiments revealed that there might be more than just one biological clock within organisms such that each of these activity components in case of the hamsters, and rhythms in case of humans are regulated by different biological clocks. Astounding progress has been made over the years, from ancient organisms all the way up to humans, and it has been demonstrated that there are clocks in multiple peripheral tissues (for instance, adrenal gland, heart, intestine and liver) that are synchronized by the 'central' biological clock in the brain, each of which regulates local rhythmic functions [5], thereby illustrating the extent to which one can push the frontiers of human understanding of 'time' and 'time-keepers' in biological systems.

The most basic thematic understanding of biological clock organization is described in what is popularly called an 'Eskinogram', wherein temporal information from the local environment reaches the core rhythm generators (central biological clocks) via input pathways and the rhythm generator regulates overt rhythms in behavioral, physiological and metabolic processes via downstream output pathways (Fig. 27.1, Top). However, at the intracellular level, only

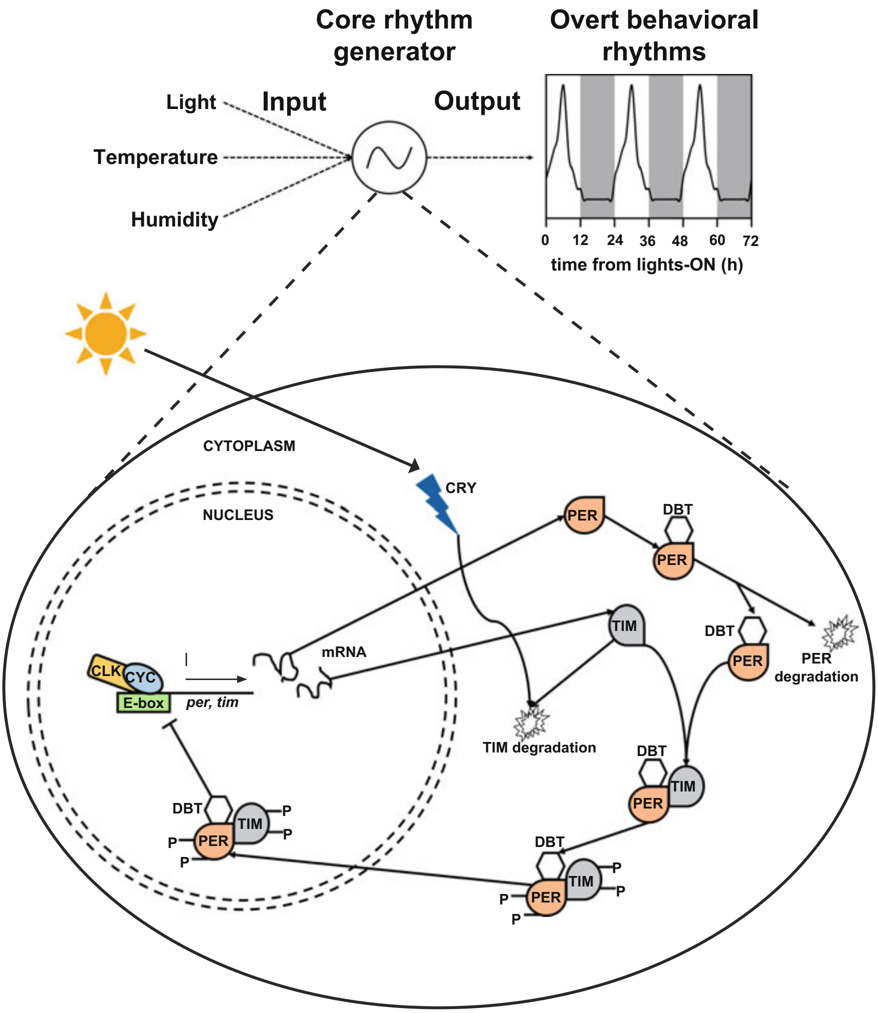


Fig. 27.1 Organization of biological clocks and its molecular mechanisms in *Drosophila melanogaster*. (*Top*) Environmental time information via light, temperature and humidity among others is received by the core rhythm generator through input pathways. This core rhythm generator then uses temporal information and regulates overt behavioral rhythmicity via output pathways. Grey shaded regions of the rhythm in the extreme right indicate dark phases of a laboratory light-dark cycle. Such a schematic of biological clock organization is often referred to as an 'Eskinogram.' (*Bottom*) Transcription-translation based feedback loop that generates rhythmicity in the core rhythm generator of *D. melanogaster* (see text for details)

input and core rhythm generators are very well-studied and will be discussed here. From theoretical predictions on oscillatory systems, it was postulated that positive and negative feedback loops are at the heart of mechanisms that generate rhythms. Thus, the wait was only to discover molecules that complete the feedback loop puzzle in living systems.

Here, we briefly describe the well-studied molecular underpinning of biological clocks in *Drosophila* to give the readers a general impression of the molecular mechanisms that drive circadian rhythms (Fig. 27.1, Bottom). Interestingly, studies ranging from cyanobacteria to mammals have revealed remarkably conserved feedback mechanisms (only with different names for the participating genes) comprising transcription-translation based loops along with post-transcriptional and post-translational modifications, and these mechanisms are discussed at great length in [11]. In *Drosophila*, the transcription factors CLK (CLOCK) and CYC (CYCLE) form a heterodimer complex that binds to the promoter region (E-boxes) of the *per* (*period*) and *tim* (*timeless*) genes, thereby transcribing *per* and *tim* such that the peak mRNA levels are reached around early-to-midnight. Subsequently, the respective PER and TIM proteins peak in the cytoplasm. As the PER and TIM proteins accumulate in the cytoplasm, another protein DBT (DOUBLETIME) kinase associates with and degrades PER. This degradation of PER is hindered by interaction of TIM with PER which forms a stable PER-DBT-TIM complex. The stable PER-DBT-TIM complex is acted upon by yet another set of proteins SGG (SHAGGY) and CK2α (CASEIN KINASE 2α) which promotes the entry of PER and TIM into the nucleus. Once inside the nucleus, PER and TIM displace the CLK-CYC heterodimer thereby repressing their own transcription. This entire loop is completed in ~24-h thus giving rise to ~24-h (circadian) oscillations in mRNA and protein levels of many of these genes. Subsequently, it was found that an additional feedback loop involving *vri* (*vrille*), *pdp* (*PAR domain protein*) and *cwo* (*clockwork orange*) genes in addition to the one just described also plays a key role in the orchestrated generation of rhythms [12]. Moreover, the core loop must be capable of receiving environmental information that allows biological clocks to synchronize to the external light-dark cycles. In *Drosophila*, a dedicated circadian photoreceptor known as CRY (CRYPTOCHROME) degrades TIM in the cytoplasm in the presence of light. Therefore, depending on the time of the day at which flies are exposed to light, formation of the PER-DBT-TIM complex will be accelerated or decelerated thereby facilitating synchronization of the molecular loop with the environmental light-dark cycles. Although, the output pathways are not characterized as well as the input and core rhythm generators, it is believed that the core clock genes *per*, *tim*, *clk*, *cyc* and *vri* along with other components may play a role in mediating overt rhythmicity in behavioral, physiological and metabolic processes (Fig. 27.1, Bottom) [12].

A biological clock can, therefore, be defined as a network of input and output pathways along with the core rhythm generators that interact in a complex feedback, coupled via several neuropeptides/transmitters that ultimately gives rise to rhythms in different biological processes with the emergent property of the network being its period (~24-h) that can efficiently keep track of local time.

27.5 Why Do We Have Biological Clocks?—Origins and Evolution

Although justifying the origins and evolution of biological clocks is a herculean task in itself, here we briefly touch upon tentative answers to the three following questions, and redirect the interested reader to [13] for a detailed discussion of the same.

27.5.1 When Did Biological Clocks Originate?

The period of the Earth's rotation has undergone changes, and it is estimated that when cyanobacteria (the earliest organism whose biological clocks are well-studied) originated ~3.5 billion years ago, the length of one day was ~8-h. Our current understanding of biological clocks in cyanobacteria suggests a transcription-translation feedback loop based mechanism comprising the *KaiABC* gene cluster with a periodicity of ~22-h. However, the *KaiA*, *KaiB* and *KaiC* genes do not appear to have evolved at the same time i.e., during the origin of cyanobacteria (see Fig. 27.2). Comparative and gene homology studies across several species suggest that *KaiC* was the first to appear (perhaps providing by itself

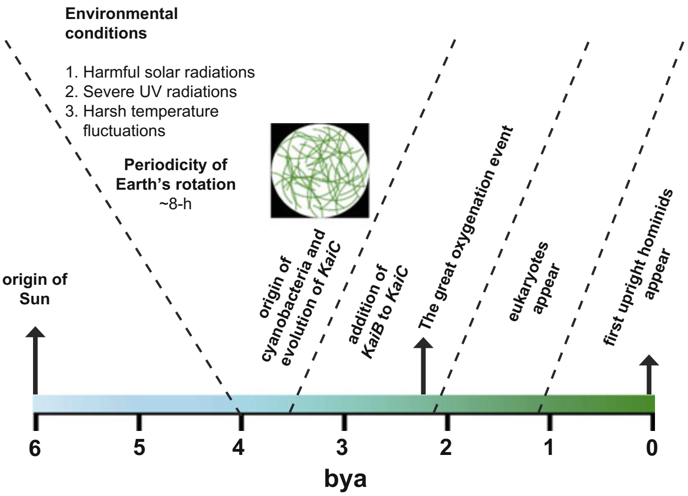


Fig. 27.2 A timeline for the origin of biological clocks with respect to important events in the history of Earth. The *x*-axis label, 'bya' refers to billion years ago. See text for details on *Kai* genes

a negative feedback loop generating rhythms with ~4-h periodicity), followed quickly by *KaiB* (that may have lengthened periodicity of the clock as day-length on Earth became longer), and *KaiA* is thought to be the most recent addition to the loop (that may have given rise to cyanobacteria's present day periodicity) and is found only in *Synechococcus* spp. (one genus of cyanobacteria). Therefore, our best guess as to when biological clocks originated is at ~3.5 billion years ago (Fig. 27.2) [13]; however, this is only a conjecture.

27.5.2 Why Did Biological Clocks First Originate?

Charles Darwin is thought to have provided the first recorded explanation for why organisms may need biological clocks, and what may have driven the evolution of the same in the first place. Around the late 19th century, he systematically recorded movements in plants and made detailed observations of several plants that open and close their leaves at certain times of the day [14]. This led him to believe that some features in the environment may be deleterious to the plant's survival, in this case prolonged duration of light. He argued that in order to avoid such damage, plants 'choose' to open and close their leaves at relevant times of the day. Subsequently, two major theories have been proposed to understand the first origins of biological clocks, both of which are discussed at length elsewhere [13].

Organisms such as cyanobacteria, whose biological clocks are extensively studied, originated much before the great oxygenation event that occurred ~2.5 billion years ago (Fig. 27.2) [13]. It is conceivable that organisms that lived on Earth much before this event would have been exposed to extreme temperatures and solar radiations (see Fig. 27.2), and these may have been potent factors driving the origin of biological clocks; the idea being that clocks would help time behavior and physiology appropriately to avoid harsh conditions to the best possible extent. In this regard, we have the 'escape from light' hypothesis, where it is theorized that factors such as exposure to UV-radiations (an effect of which is increased errors in DNA replication and photochemical RNA and protein reactions) may have led to the origin of biological clocks that may help organisms avoid such harmful effects of environmental factors.

On the other hand, we have the 'endosymbiotic coordination' theory. It is believed that the origin of eukaryotic life-forms (see Fig. 27.2) lay in multiple prokaryotes coming together as endosymbionts that form something of a precursor for the present day cellular organelles. It was thought that along with cellular compartmentalization of such proto-organelles, temporal coordination of interacting sub-cellular processes is essential to avoid intracellular chaos [13]. However, neither of these hypotheses have very convincing data to support them, hence remain open questions to date.

27.5.3 Why Do We Need Biological Clocks in the Present Day?

Natural factors that may have driven the origin of biological clocks may not be so obvious to us today. Then, why do we still have biological clocks? Do we really need them? What functional role may it have to be still retained by natural selection? Two major hypotheses have been raised to understand the same.

27.5.3.1 Intrinsic Advantage Hypothesis

As mentioned earlier, although the central rhythm generator for most animals is in the brain, clocks in peripheral tissues such as adrenal gland, liver, heart and intestines have also been discovered over the years [5, 8]. An intuitive explanation for why biological clocks are required can be drawn from the very fact that multiple clocks co-exist, and temporal harmony within an organism can be maintained only via coordinated functioning of each of these clocks. This idea is popularly referred to as the intrinsic advantage hypothesis.

Early experiments were done in order to understand if indeed harmony among constituent clocks had any noticeable advantage for the organism. Cockroaches were used elegantly to demonstrate that the suboesophageal ganglia (the clock regulating activity-rest rhythm) when out-of-sync with other constituent clocks of the body led to the development of tumors. This was later attributed to desynchrony among various biochemical oscillations [15]. Additionally, in case of intracellular metabolism, components of different biochemical pathways may be incompatible with each other, as is the case with nitrogen fixation and photosynthesis in cyanobacteria, and it is thought that biological clocks may serve to create separate temporal niches to avoid such biochemical clashes [13]. Modern lifestyles which include rapid travel across multiple time-zones, or shift-work schedules have revealed that different biological rhythms (such as those in body temperature and psychomotor performance) take different times to re-adjust to the new time-zone, thereby creating disharmony among constituent physiological processes in humans, leading to much discomfort, a condition known as jet-lag (discussed in the next section) [5], and this therefore suggests that synchrony among constituent biological clocks is essential for well-being.

Unequivocal evidence for such a hypothesis comes from studies performed in the laboratory that monitors trajectories of biological clock evolution in real-time [13]. If the intrinsic advantage hypothesis were not true, one may hypothesize that biological clocks in populations living under constant (non-cyclic) environments would gradually regress. Two studies using *Drosophila* populations that were reared under constant illumination and darkness for over ~700 and ~1000 generations respectively tested this hypothesis. Both these studies reported the persistence of robust circadian rhythms in their respective *Drosophila* populations despite being reared under non-cyclic environments for several hundreds of generations [13], suggesting that their biological clocks had not regressed over time.

Such data from multiple experiments compel us to believe that an evolutionary advantage of possessing circadian clocks in the present day would be that they could synchronize and temporally partition internal and incompatible processes, respectively.

27.5.3.2 Extrinsic Advantage Hypothesis

Variation is the essence of natural systems. Daily variations in physical parameters of the earth due to its rotation include cycles in light intensity and spectral composition, temperature, humidity and barometric pressure among several others. Such concerted cycling of several parameters gives rise to a complex network of 'favorable' and 'unfavorable' times of the day for individuals [15]. It is thought that biological clocks exist to help organisms time their various behaviors to such ecologically favorable times of the day so that it is in-sync with its environment, thereby enhancing survival and reproduction.

Evidence for this hypothesis comes from several lines of investigations. Early studies indicated that in cyanobacteria appropriate timing of events with respect to laboratory light-dark cycles provided competitive advantage over other strains with inappropriate timing [16]. Additionally, some interesting experiments were carried out in the wild to test the extrinsic advantage hypothesis. Antelope ground squirrels are generally diurnal, but it was observed that when their biological clocks are removed they become more active in the night, and mortality due to predation in these animals was 30 % more than the ones with intact clocks [8]. Similar experiments were performed with free-living chipmunks and the results were strikingly similar [8]. Furthermore, several studies have shown that aberrant or misaligned biological clocks lead to severe reduction in survival and Darwinian fitness (measure of an individual's contribution to the gene pool of the next generation) in *Drosophila* [13]. Additionally, several studies have also reported the adaptive evolution of timing of behavior in real-time [13], thereby suggesting that appropriate timing of behavior may be important for survival and reproduction, and biological clocks help organisms choose favorable times to perform functionally relevant tasks.

Studies described in this section have led to our understanding of why biological clocks may have originated, when they may have originated and what functions they serve today. Although, the present day adaptive functions of biological clocks are well-studied and agreed upon, the why and when of origins of biological clocks are still open questions awaiting rigorous experimentation and unequivocal data.

27.6 Biological Clocks (Rhythms) and Their Clinical Implications

Understanding biological clocks and thereby rhythms in such great detail with a fine grasp on their molecular underpinnings brings upon us the responsibility of disseminating the clinical relevance of such research. It is now known that nearly half of the mammalian genome is expressed rhythmically in one or more tissues [10], suggesting that most processes in living systems are likely to be associated with biological clocks. Therefore, the clinical relevance of studying such systems is enormous. As discussed earlier, the primary function of biological clocks is to synchronize internal processes with external environments to promote well-being. It has now been shown that misalignments and/or disorders of biological clocks have serious consequences on health, one of the most common being the effect on sleep.

Two common pathologies that stem from dysfunctional biological clocks are advanced and delayed sleep phase syndromes (ASPS and DSPS, respectively) [8]. Patients with ASPS, a disorder known to stem from a mutation of a single dominant autosomal gene on the 2nd chromosome, sleep (~8 p.m.) and wake (~3 a.m.) very early in the day, while patients with DSPS sleep (~4 a.m.) and wake (afternoon) very late during the day. Such disorders can be treated both behaviorally and by using medication. Sleep-time readjustment therapies by controlled sleep-wake schedules that gradually progress towards the desired sleep-time have been found to be successful [5]. Moreover, the use of pharmacological agents such as melatonin that is known to reset biological clocks can also be administered at specific times of the day to achieve appropriate sleep-times [8].

Additionally, misalignment of biological clocks (often caused due to rapid travel from one time zone to another, thereby disturbing the synchrony between internal and external cycles) induces discomfort associated with disruption of sleep, gastro-intestinal functions and decreased vigilance and attention-span, collectively known as jet-lag. A more common form of jet-lag is now referred to as social jet-lag [17]. This is another common example of misalignment of biological clocks, not with the environmental cycles but with social time [17]. Social time can be defined as a collectively agreed upon time of performing certain kind of activity, for instance, school timing and work-place timing. Human populations are characterized by a continuous distribution of chronotypes with the extremes being lark and owl chronotypes, or in other words, extremely ‘morning’ and extremely ‘evening’ people. Chronotypes are naturally occurring variation in timing of events among organisms that are mediated by an interaction of underlying biological clocks and the environmental cycles [17]. It is, therefore intuitive that different chronotypes are likely to have problems adjusting to different social times [17]. Several forms of misalignment of biological clocks exist, and these could commonly occur due to repeated shift-work as is the case with long-distance drivers, factory workers and flight attendants, and extended work-hours as is the case with nurses, hospital interns, resident physicians, surgeons and experimental researchers. Such circadian misalignments are known to have dire consequences such as road and factory accidents, improper care for patients, and reduced work-efficiency which are discussed in great detail elsewhere [8, 17]. Thus, circadian discord is a serious health concern and must be treated with appropriate sleep-time readjustment therapies via imposed sleep-wake schedules combined with, in some cases, appropriate administration of drugs such as melatonin.

Furthermore, certain associations have been made between biological clocks and psychiatric disorders albeit the link between the two is still unclear. For instance, seasonal affective disorder (SAD), a depression like state that often occurs due to reduced exposure to light and low ambient temperatures (characteristic of winters and hence referred to as winter depression), is known to be associated with delayed biological clocks and reduced amplitude of circadian oscillations [8]. Moreover, biological clock malfunction is associated with several neurodegenerative disorders such as Alzheimer's, Parkinson's and Huntington's disease causing sleep disruption in patients [18]. Effects of SAD and psychiatric disorders such as Alzheimer's can be ameliorated to some extent by imposed sleep-wake schedules and exposure to bright light (bright light therapy), with good success rates [8].

The study of biological time is at an exciting stage now. Recent work has focused on understanding the associations between biological clocks and immunity, metabolism and obesity, cancer chemotherapy, cardiovascular events, and DNA repair mechanisms. It is now known that several processes of the immune system such as phagocytosis, proliferative responses to antigenic substances, cytolytic activity, synthesis and release of cytokines and chemokines, are all rhythmic, and under the influence of biological clocks [19]. Moreover, it has been reported that biological clocks regulate metabolism and energy homoeostasis in the liver, and have associations with lipogenic and adipogenic pathways, also linking obesity and other metabolic disorders such as diabetes with inappropriate times of eating [20]. Additionally, many drugs used by humans are known to show daily variation in efficacy, which has led clinicians to optimize the timing of dosage such that maximum effectiveness can be complemented with minimum toxicity, in treatments such as cancer chemotherapy [8]. Furthermore, it is well established that adverse cardiovascular events take place primarily during early morning, and recent work has linked higher cardiovascular pathology to misaligned biological clocks [21]. Aziz Sancar, a Nobel laureate, who is recognized for his work on DNA repair mechanisms also showed that a molecular component of biological clocks (*cryptochrome*) is closely associated with DNA repair (during replication), thereby striking an interesting link between biological clocks and DNA repair, a case that was used to understand the first origins of biological clocks (see Sect. 27.5.2) [22]. Such studies thereby highlight the fundamental and integral role of biological clocks in sustaining health, and that the repercussions of disorders and misalignment of biological clocks may be catastrophic.

Thus, investigations into the nature of biological rhythms have revealed how rhythm generators serve as biological clocks that are innate. Further, these clocks are primarily located in neural tissue in metazoans and their molecular physiology is remarkably conserved across the living kingdom, thereby suggesting that biological clocks are adaptations to the cyclic nature of life on Earth. Finally, the plethora of functions that biological clocks perform necessitates further research on not only the medical implications of clock dysfunction but also the nature of treatment of other non-clock related disorders. We hope that the readers of this discourse imbibe awareness of such a system into their outlook on biology such that we may better discern the contours of the interaction of living systems with time. Further, we hope to ignite interest in the field to help resolve the several challenges that still lay ahead of us in understanding these fascinating timing systems we possess.

Acknowledgements We thank K.L. Nikhil, Vishwanath Varma, Abhinay Lakshman and Dr. Sheeba Vasu for carefully reading the article and suggesting some very useful changes. We would also like to acknowledge financial support from Science and Engineering Research Board (SERB), New Delhi.

References

1. Kant, I. (1781). *Critique of Pure Reason*. New York: Macmillan.
2. McTaggart, J. E. (1908). The unreality of time. *Mind* 457–474.
3. Daan, S. (2010). A history of chronobiological concepts. *The circadian clock*. New York: Springer.
4. Serafini, A. (1993). *The epic history of biology*. Berlin: Springer.
5. Moore-Ede, M. C., Sulzman, F. M., & Fuller, C. A. (1994). *The clocks that time us*. MA: Harvard University Press.
6. Halberg, F. (1969). Chronobiology. *Annual Review of Physiology, 31*, 675–726.
7. Aschoff, J. (1981). A survey on biological rhythms. *Biological rhythms*. Berlin: Springer.
8. Dunlap, J. C., Loros, J. J., & DeCoursey, P. J. (2004). *Chronobiology: Biological timekeeping*. MA: Sinauer Associates.
9. Helfrich-Förster, C., Stengl, M., & Homberg, U. (1998). Organization of the circadian system in insects. *Chronobiology International, 15*, 567–594.
10. Welsh, D. K., Takahashi, J. S., & Kay, S. A. (2010). Suprachiasmatic nucleus: cell autonomy and network properties. *Annual Review of Physiology, 72*, 551–577.
11. Sehgal, A. (2004). *Molecular biology of circadian rhythms*. New York: Wiley.
12. Hardin, P. E. (2011). Molecular genetic analysis of circadian timekeeping in Drosophila. *Advances in Genetics, 74*, 141–173.
13. Nikhil, K. L., & Sharma, V. K. (2017). On the origin and implications of circadian time-keeping: an evolutionary perspective. In V. Kumar (Ed.), *Biological timekeeping: Clocks, rhythms and behaviour*. Berlin: Springer.
14. Darwin, C., & Darwin, F. (1881). *The power of movement in plants*. New York: D. Appleton.
15. Vaze, K. M., & Sharma, V. K. (2013). On the adaptive significance of circadian clocks for their owners. *Chronobiology International, 30*, 413–433.
16. Ouyang, Y., Andersson, C. R., Kondo, T., Golden, S. S., & Johnson, C. H. (1998). Resonating circadian clocks enhance fitness in cyanobacteria. *Proceedings of the National Academy of Sciences, USA, 95*, 8660–8664.
17. Roenneberg, T. (2012). *Internal time: Chronotypes, social jet-lag, and why you're so tired*. MA: Harvard University Press.
18. Mattis, J., & Sehgal, A. (2016). Circadian rhythms, sleep, and disorders of aging. *Trends in Endocrinology and Metabolism, 27*, 192–203.
19. Labrecque, N., & Cermakian, N. (2015). Circadian clocks in the immune system. *Journal of Biological Rhythms, 30*, 277–290.

20. Froy, O. (2010). Metabolism and circadian rhythms-implications for obesity. *Endocrine Reviews, 31*, 1–24.
21. Morris, C. J., Yang, J. N., & Scheer, F. A. (2012). The impact of the circadian timing system on cardiovascular and metabolic function. *Progress in brain research*. Oxford: Elsevier.
22. Thompson, C. L., & Sancar, A. (2002). Photolyase/cryptochrome blue-light photoreceptors use photon energy to repair DNA and reset the circadian clock. *Oncogene, 21*, 9043–9056.

Chapter 28
The Cellular Space—The Space of Life

Pier Luigi Luisi

It is well known that life is cellular and only cellular: all tissues and organs of all animals and plants are organized assemblies of cells—so that we can consider the cell as the elemental constituent of life on this planet. Cells are (more or less) spherical structures, closed by a membrane, as illustrated by the simplified Fig. 28.1. It is in fact a very particular form of space, and in this short article we would like to discuss why it is so and highlight the particular relation between this form of space-and life.

The first thing that comes to mind is that the cellular membrane, which closes the spherical structure, permits to clearly discriminate the inside from the outside. This discrimination between inside and outside, between my world and the environment outside, is the first operational definition of life' individuality.

But this is not all. The closed membrane also means a protection, a warranty that the organized inside is not going to be diluted or poisoned by all agents swimming outside; and in this way, all what happens inside is going to remain a property of the internal world. This is important also at the level of evolution: if some positive mutation or any other chemical change takes place inside, this is going to be of benefit only for the internal world.

However, the membrane is, and must be, *semi-permeable*: this means, that some particular compounds will be allowed to enter, so as to activate the internal mechanisms of metabolism, while other compounds, no more useful, are going to

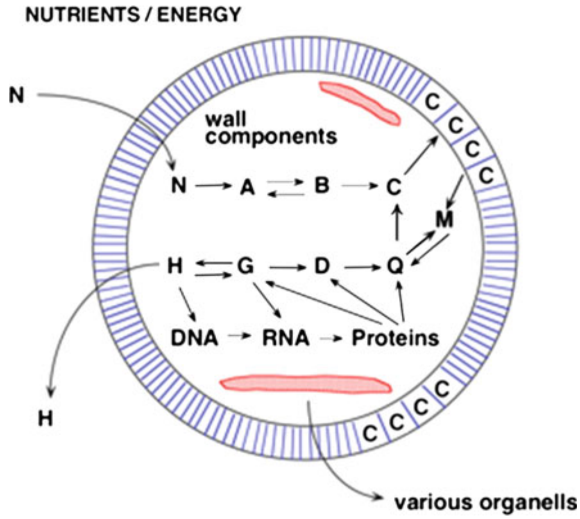
---

P.L. Luisi (✉)
Emeritus of the Swiss Federal Institute of Technology Zurich (ETHZ), Zürich, Switzerland
e-mail: luisi@mat.ethz.ch

P.L. Luisi
Dipartimento di Biologia, Università degli Studi di Roma, Tre Viale Guglielmo Marconi, 446, 00146 Rome, Italy

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_28

Fig. 28.1 A simplified representation of a cell (cross section), with the semipermeable membrane which permits the entrance of certain selected nutrients (N), the expulsion of certain catabolites (H); and then notice the web of reactions inside the cell



be expelled through the membrane. The selection of what may enter, and what not, is determined by very fine biochemical structural details of the membrane—and of course different organisms will have different cellular selection criteria.

28.1 Being Open, Being Closed

This entrance of particular substances indicates a general principle: that the cell—and this is true for any larger living organism—is a system which is *thermodynamically open*. The living remains such, living, only if it acts as an open system, permitting nutrients and energy to enter. In turn, this indicates that the living space and the environmental space are necessarily linked to each other, and form, in a kind of complementarity, the entire space of life.

The notion of thermodynamic openness is in apparent contrast with another general principle of cellular life, commonly referred to as "*operational closure*"—as particularly emphasized by the Santiago school of Maturana and Varela.

This means the following: that all what is needed to the liver cell to be a liver cell, is contained in the cell itself. Or, in other words: the ant does not need any information from the outside world to be an ant. It is the internal, specific organization of the cellular structure' inside, that fully determines the properties and behaviour of the organism. There are of course in life all kind of stimuli from the outside-but, and this is the important point—these act as triggers, but not as modifiers of the internal structure of the living. Thus, if I kick a dog, or I kick a snake, the dog-and the snake-are going to react on the basis of their dog-ness or snake-ness, from their internal circuit, *from within*.

This is in agreement with what we were saying before about the protection of the internal world, as operated by the membrane. It is indeed an important property of the circumscribed space: to be such, that the behaviour, namely the "way of doing", is due to the organization of the internal world.

This consideration is the basis of the *theory of cognition*, of Maturana and Varela, see for that the references indicated below [1, 2, 4].

28.2 The Paradox of the Cellular Space

How large is a cell? The unit of measure is the micron (μm), one thousand of a millimetre, and the smallest unicellular organisms—bacteria—may have a typical diameter of a few microns (Escherichia Coli in our intestine is about 2 μm) with the red blood cell we are at about 9 μm, with the amoeba we are at 90 μm, a human egg cell is around 100 μm, giant bacteria can reach 600 μm, without forgetting that nerve cells-the neurons-can reach one meter of length.

Inside this relative small space, there is an incredible chemical activity. In each of our liver cell, every second there are thousands of transformations: sugars being oxidized, proteins being hydrolysed, hormones being synthesized, proteins being made thanks of the information of nucleic acids, nucleic acids being synthesized thanks to the catalytic power of enzymes (special proteins)…

Figure 28.1, in its simplicity, then, does not yield the real complexity of a cell. And as soon as one considers the continuous series of transformations taking place inside each cell of each living organism, one realizes an apparent paradox. The paradox is given by the fact, that despite all these transformations, a liver cell remains a liver cell, an amoeba remains an amoeba, and so on… (at least for a long period of observation, say during the so-called homeostasis).

How is this contradiction possible? How can we have self-maintenance, and continuous chemical transformations at the same time in the same place?

Here lies the very essence and definition of cellular life, and life in general. In fact, the main function of the cell is self-maintenance: the cell maintains its integrity by re-generating from within all the compounds which are being destroyed. The cell, as an open system, does so thanks to the nutrients and energy that comes from the external milieu, so that there is a continuous conversion of these nutrients into the cell components-so as to maintain a constant concentration of all reagents inside. The cell, and each living organism in general, can be seen as a factory that re-makes itself from within. Autopoiesis (self-production) is the term that Maturana and Varela coined to illustrate this very important principle.

And the collateral principle is the following: that all living cells work in the same way, autopoietically. *Autopoiesis is the invariant property of life*. Conversely, the structure of the various cells in the different organisms is *the variable property*: the membrane constituents can be different, there can be in that particular cell more or less lipids, more or less saccharides, and the concentration of some reagents may vary accordingly—but there is no cell that does not comply to the general principle of autopoiesis-the invariant mechanism. The invariant autopoietic mechanism, and the variable structure, form the two complementary aspects of life. Note that what it has been said for a single cell, is also valid for an entire organism—for the whole elephant or for any human being. I am re-making continuously my haemoglobin, or my hair, or my tissues, from within. And note also that this “from within”, which means from within the boundary of the membrane, is equivalent to the notion of operational closure, in the sense that all takes place thanks to the internal organization. Autopoiesis is thus a systems property, in the sense that what is important is the organization of the entire cell, namely the interactions among all cell constituents. This systemic view is in sharp contrast with the reductionist, DNA/RNA centred view, according to which life is essentially based and due to on one type of molecule [3, 4].

28.3 Space and Localization

The cell as pictured in Fig. 28.1 does not give a good idea of the great complexity of the chemical transformations inside the cellular space.

For that, better to consider Fig. 28.2, which gives part-only part-of the metabolism of a bacterial cell: thousands and thousands of reactions, all linked with each

Fig. 28.2 The metabolic scheme (part of it) of a bacterium

other, each reaction being catalysed—i.e. made possible at room temperature—by the action of a specific protein—an enzyme; whereby each enzyme originates from a specific gene, which in turn is synthetized by a specific family of enzymes. Really a picture of great complexity.

But the question that now we would like to ask is the following: where is life localized?

Well, the answer is obvious: life is not localized-there is not a single spot, a single reaction, a single metabolic cycle, which may represent that specific bacterial life. Life is the entire net of reactions. Life is not, and cannot be localized, as it is a global, distributed property.

And the same can be said for every living organism. Is there a point, a place, where the life of an elephant, or, for that matter, the life of a person, is localized? Of course not.

And this concept, of a systemic space complexity with a distributed quality without a centre of direction, is an important concept in the modern theory of complexity. Where is New York localized? Where is the centre of localization of a bee hive, of a termites' nest, of a migrating bird formation?

Again, we have a situation as in the cellular space, which reflects life, but that space encompasses the entire network of relations of all components with each other. In a way, this is a direct consequence of the system' view of life, whose main tenet, is the following: that is the entire web of relations that makes up the main general properties.

28.4 Finale

Let us go back to the principle of operational closure, and to the view from within of each cellular space, and, from that respect, from the view from within of each living organism. This view from within has important consequences from the epistemic point of view, and one has to do with the relation with the environment, which we have already touched upon, both talking about the thermodynamic openness, and about cognition.

More generally, the consideration that the environment is seen "from within", implies that each organism sees the world in its own way. Clearly the world seen by the fish, thanks to its specific sensorial tools, is different from the world of the bat, and different from the world of the earth warm. There are all different kinds of cognition, which means, that the world is not seen as an objective reality, equal for everyone. There are then as many worlds, as many different organisms.

Now, if we translate this at the level of humankind, we have a similar situation. My view of a rose, as a western man well versed in romanticism and traditional European poetry, is different from the view of an Eskimo who may have never seen a rose. Thus, it is not that the sensorial tools take a picture of the outside world (what we call "Representationalism") which is equal for all. Rather, there is a co-emergence between the inner world and the environment, to form what

Maturana and Varela call "enacting", namely the construction of a world. The world outside is due to the particular internal organization of the observer—but we are not dealing however with a simple form of constructivism, rather with a co-emergence between two factual realities.

Clearly, this view from within determines a crisis, or at least a critical discussion, of the notion of objectivity, and brings in the question of the subjectivity in science —a theme which we can only point to here. And related to that, still at the level of human kind, there is the entire question of the consciousness-how and why do we know that we exist?—which again I like to skip in this particular short assay.

Rather, I would like, on ending, put the attention to another aspect of the interaction between the living and the environment, which has become so important in the last decades: the questions of pollution, global warming, deforestation, loss of biodiversity, the ozone hole, …all phenomena created by humankind, and which threaten our own existence-or at least that of our future generations. One can say, at this point, that they are all problems created by our hands and minds, namely created mostly by us—and then we should be capable to remediate to all these insults to Nature by going back to a healthier relation to our mind and hands.

References

1. Maturana, H., & Varela, F. (1998). *The tree of knowledge*. Shambala (revised from the original 1987 version).
2. Varela, F. (1979). *Principles of biological autonomy*. North Holland: Elsevier.
3. Luisi, P. L. (2016). *The emergence of life* (2nd ed.). Cambridge: Cambridge University Press.
4. Capra, F., & Luisi, P. L. (2014). *The systems view of life*. Cambridge: Cambridge University Press.

Chapter 29
The Consciousness of Space, the Space of Consciousness

Mauro Bergonzi and Pier Luigi Luisi

In this article we intend to discuss some basic aspects of consciousness, focusing at the end in the relation between space and consciousness. The notion of consciousness in the last couple of decades has been the subject of lively academic and mass media debates, and is certainly not our aim to review all this. However, preliminarily, is necessary to somehow characterize the term. To this aim, it is still useful to refer to the work of the philosopher Chalmers [6].

Accordingly, one should make a discrimination between the 'hard' and 'easy' problem of consciousness. The easy problem is that of sensorial consciousness (also called 'access consciousness', see for example Thompson [14]), which refers to the five perception senses (seeing, hearing, touching, etc.) plus mind/thinking. It is considered 'easy' only in the sense that at least we can see a relation with our brain–neurons, whose connections are implied in all these aspects of cognition. The hard problem instead is related to the *subjective experience*: my feeling of blue, or my wonder to beauty, which cannot be shared with others—my blue is only my blue. This cannot be linked to the sensorial perception or to the brain activity: actually we don't even know why it is there. So, according to Chalmers' view, it is as if there were two different states of consciousness.

Later on we will argue that we do not necessarily agree with this dichotomy, but for the time being let us go on, focusing on this subjective aspect of consciousness—and following broadly in that the outline of Bergonzi's book [1].

M. Bergonzi (✉)
University "L'Orientale", Naples, Italy
e-mail: bergma@libero.it

P.L. Luisi
University of Roma3, Rome, Italy
e-mail: luisi@mat.ethz.ch

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_29

29.1 The Intuitive Evidence of 'Being-Awareness'

In order to introduce the problem with a phenomenological approach, we would like to ask a question to our readers:

> "Right now, are you being aware?". Or, more simply: "Are you aware?"
>
> Your answer is of course positive: "Yes, I am aware".

Here is, for us, a very first and important point: your answer "Yes, I am aware" is a thought, but this thought is a consequence of a checking up *prior* to your thought. Did you need to think about it, or was the evidence of being aware already there before any thought?

This elementary, preliminary awareness—namely, the pure and simple fact of being aware, independently from what one is aware of—is for us the basic essence of consciousness. In fact, we could have asked "Are you conscious?"

On this topic, many Western philosophers assume that consciousness or awareness can only be 'intentional', i.e. consciousness *of* something, while the idea of a contentless awareness is just a delusion. However, Forman [7], in his studies on mysticism, quotes many reports of meditative experiences that furnish some evidence of a 'pure consciousness event' (PCE) without any content at all, which suggests a wider definition of 'consciousness' ([7], 112):

> Thus, a synonym of consciousness should not be something like "awareness of anything (intentional) at all", but rather "awareness *per se*" which can (and usually does) become aware of things.

Does that mean that you are conscious of your being too? Namely, that there is a being, distinct from consciousness, which consciousness is aware of?

In order to find an answer, let us revert to our phenomenological exploration by asking our readers another simple question:

> "Right now, are you sure that you exist?"
>
> After a short pause, the answer will be "Yes", beyond any doubt.

Both the question and the answer were thoughts, but in the short gap between them (that occurred in order to verify directly whether you are present or not), the undeniable evidence of your being became immediately plain.

Did you need to ponder on the answer, or was the evidence of 'being' *previously* there, before any thought? If you were not already present, how could you even think at all?

The very fact that our being is intuitively *evident*, proves not only that we exist, but also that we are *aware*, otherwise our being would be unknown.

So, whenever we ask ourselves "Do I exist? Am I aware?", we can immediately ascertain that the answer is "Yes". During that instantaneous intuitive inspection, our being-awareness arises not as an object to be known, but as a self-evident fact that, like light, does not need to be illuminated, because it shines of its own. In this respect, there is a close similarity between our phenomenological analysis of being-awareness and some current Western offsprings of the ancient Indian philosophy of *advaita* or non-duality (see for example [12, 13, 15]).

Is being-awareness something that I have, or rather something that I *am*?

Could I ever exist *apart* from being-awareness?

Could I ever experience two 'myselves' at the same time?

Not at all, there is always only one 'I' at a time, which is the same as being-awareness.

Therefore, if I am only one and I am being-awareness, then being and awareness cannot be two separate things, but just two aspects of the same identity, or rather two different ways to describe it.

Consequently, consciousness cannot be aware of 'being' as an object, because, in order to do so, there should be two different things—consciousness (namely the knowing subject) and being (namely the known object)—while they are just two aspects of the same thing.

Someone could rebut: "But I *am* clearly conscious of my existence!". This is not entirely correct: one can only be aware of the *thought* "I exist", which is not the actual being, but rather its translation into a mental object made of words, that is a mere *content* of consciousness. Being-awareness is not knowable as an object: it shines as a self-evident actuality.

Note that also the expression "I know to know", as a reflexive level of consciousness, is due to our thinking, and as such is only an object of consciousness. The sense of being (which the mind translates into the sentence "I am") is actually the precondition for everything to appear: if first of all I am not here, then no perception, no sensation, no action, and no thought can be experienced.

So consciousness as being-awareness—i.e. the very root of our identity—is epistemically and existentially *prior* to the appearance of anything else. Of course, here the primacy of consciousness is restricted only to the epistemological and experiencial domain and does not convey any ontological stance.

This sense of being is an undeniable actuality: even in order to negate it, one must firstly be there. We cannot avoid even for one single moment the fact of existing and being aware, since we do not 'have' existence-awareness: we *are* it. The sense of being or consciousness as such is a too simple and immediate evidence for the thought to grasp: it is a non-conceptual awareness.

According to our view, this immediate *aware presence* is precisely what the term 'consciousness as such' points to: namely, the very ground and precondition of any specific content that arises in experience.

This basic awareness is at the root itself of our 'I-experience', but nevertheless it cannot properly be regarded as 'personal', due to the fact that we use the term 'I' as a pointer to very different aspects of our experience.

For example, when we say "I am tall, fat, sick, old" and so on, the term 'I' clearly refers to the physical body. When we say "I am sad, happy, anxious, intelligent" and so on, the term 'I' refers to the mind. At the same time, the term 'I' is always pointing to the 'first person' subject of experience, i.e. consciousness.

These three different meanings of the term 'I' cannot be confused with each other or compacted into one single entity, because all bodily sensations and all mental thoughts are mere contents of consciousness, namely objects the consciousness is aware of as a subject.

Therefore, since we usually employ the term ‘I’ to mean our ‘personality’, which is made up of recurring mental patterns that are mere contents of awareness, it would not be correct to equate *tout court* consciousness to the ‘I’. Consequently, the basic consciousness, or presence, does not exactly correspond to the ‘I’, to the ‘me’. We may in fact agree that it appears as a subjective experience, but this does not necessarily mean that I can call it ‘my’ consciousness, since it cannot be confused with the objects which it is aware of. This very point—whether and to what extent one can talk in terms of ‘my consciousness’—is another intriguing question, to which we will come back later on in this article.

29.2 Consciousness: A ‘Blind Spot’ for Science

Over the last years, several schools of neurobiology and cognitive science have turned into the study of consciousness, at the aim of getting a deeper, scientific understanding of this concept.

However, as already pointed out by some researchers [2, 3, 4, 5, 14], in our opinion this kind of approach may be quite problematic.

In fact, most of the explanations or interpretations of consciousness assume that it is something ‘out there’ that can be studied objectively—for example an emergent property of the brain, or a quantum state, or a particularly complex brain resonating structure (for a critical review, see [14]).

To some extent, it is possible to inquire experimentally some correlations between the brain activities and the *contents* of consciousness, i.e. the experiences which we are aware *of*´(like perceptions, bodily sensations or thoughts): this is what Chalmers refers to as ‘the easy problem’.

However, *consciousness as such*—namely the very fact of being aware, a sort of ‘awakeness’, or ‘first person aliveness’, or ‘aware presence’ which allows every experience to appear and as such should not be confused with its own specific ‘contents’—is, so to speak, the ‘bottom line’ which cannot be consistently explained as the end result of any physical or mental cause, since no ‘explanation’ nor ‘cause’ could appear without consciousness *already* being there as a precondition. Using a different language [5] consciousness is not a phenomenon, is the phenomenology itself.

Therefore, since all phenomena can be observed, studied or explained *if and only if* consciousness is already there, it is impossible to regard consciousness as the end product of any other phenomenon without stumbling upon an epistemological paradox. This is what Chalmers refers to as ‘the hard problem’.

As already mentioned, David Chalmers’ view of consciousness is essentially ‘dualistic’. Based on the notion of consciousness as discussed above, we tend to disagree with it. The sensorial aspects have to do with physical perception (e.g. the vibrations in our ears when we hear a sound), but ultimately consciousness—the 'luminosity' of awareness, to use the terminology of ancient Indian philosophy—is always the final point. In this sense, there is no difference with the subjective state of consciousness, and it is confusing to invoke two different kinds of consciousness.

Let us go back to the general question of the scientific study of consciousness, as presented mostly by neurobiologists and cognitive scientists. At this regard, we can highlight that, despite the variety of theories and interpretations, all these scientific views generally rely on one basic principle: that consciousness is a 'secondary' property that 'emerges' from an organic support of biological matter, namely the brain (once that it reaches a critical level of complexity).

We maintain that this hypothesis is epistemically weak in many respects.

Firstly, as mentioned already, due to some lack of philosophical accuracy, neuroscientists often confuse the *objects* or *contents* of consciousness—thoughts, sensations, perceptions, images or memories which can be observed and studied experimentally in connection with the functions of the brain—with the *very fact of being aware* ('consciousness as such'). The latter, as said before, cannot be observed as an object by science, since it is the source itself of any observation.

Secondly, let us consider the following: the theory of emergence can only ascertain that from complex interactions between multiple components, some new 'properties' can appear that were not previously owned by the original components and that can be detected by the observation of new qualities or behaviours. However, consciousness as such is neither a 'quality' nor a 'behaviour', but rather the 'first person' awareness that *observes* the emergence of any possible quality or behaviour, and as such it eludes the entire range of the observable objects: it eludes the very phenomena that are consistent with an emergentist interpretation.

Thirdly, the cause/effect theory that regards consciousness as derivative from the brain presumes a hierarchical causal order ranging from biological matter (the primary factor) to consciousness (the secondary factor). But how can we assert the priority of the brain over consciousness, if it is only through the medium of consciousness that it becomes possible to perceive, know and study the brain?

Furthermore, in order to preserve its own objectivity and coherence, the standard scientific method must resort to the conventional artifice of studying what is observed *as if* it were really independent and separate from the observer: so the observer must always be kept out of the picture.

However, the price of this unavoidable abstraction is a self-evident epistemological limitation: its resulting world-view will always be *incomplete*, exclusively confined as it is within the range of the observable objects, from which the observer (i.e. consciousness) remains in any case excluded. In order to preserve its own validation, any procedure of scientific investigation can only occur inside the boundaries of consciousness, which is always epistemically *prior* to any possible objectification carried out by the scientific research.

In other words, any observation or theory about the so called 'objective' reality can never include consciousness, that is the very background of all observing and theorizing activity: consciousness can never be 'observed' or 'known', because it is always prior to whatever object is observable and knowable.

The notion of the primacy of consciousness is not yet very popular in our scientific world. The main reason is probably due to the fact that in such a case consciousness is not accessible epistemically to an objective scientific knowledge, and therefore it is like a ‘blind spot’ for science: just as the eye cannot see itself, so knowledge cannot grasp the very source of knowledge. This point has been already emphasized in the literature (see for example Bitbol [4], Bitbol and Luisi [5], Thompson [14], Bergonzi [1]). However, already many centuries ago, the Indian philosopher Śaṅkara asserted with keen philosophical arguments that, just as fire cannot burn itself or a sword cut itself, so consciousness cannot know itself as an object, since it is always on the side of the subject.

It appears then, that science may be unable to study consciousness ‘objectively’ (i.e. by describing it in ‘third person’), because it is always given to us in ‘first person’, and this is precisely its only distinctive feature: inevitably, whenever science tries to study consciousness, it is obliged to force it into the range of observable objects, so as to falsify its very essence of observing subject through an epistemological inconsistency.

So the scientific study of consciousness as such seems to be not even a ‘hard problem’, but actually an *unsolvable* problem.

How can we avoid this epistemological loop? Chalmers, in addressing this question, puts forwards the idea that consciousness may be a fundamental property, just like space or time. Space and time cannot be defined, and we cannot say that they derive from some other property—they are there and form the fundament of all our science and our thinking. So should it be for consciousness?

Time does not seem mature enough for science to accept this. However, strangely enough, some authoritative thinkers of quantum physics, coming from a very particular way of arguments and scientific knowledge, have been more willing to take seriously this hypothesis—and already several years ago.

For example, Plank [9] said:

> I regard consciousness as fundamental. I regard matter as derivative from consciousness. We cannot get behind consciousness. Everything that we talk about, everything that we regard as existing, postulates consciousness.

Schrödinger [10] was in total agreement with him:

> Consciousness cannot be accounted for in physical terms. For consciousness is absolutely fundamental. It cannot be accounted for in terms of anything else.

To some extent, even the simple act of *thinking* about consciousness (which is *not* an object), by involving words that are always ‘object-oriented’, unavoidably engenders a misrepresentation of it as an object.

However, to say that consciousness cannot be known as an object (since it is the knowing subject) does not mean that subject and object are two separate things, because both of them entail each other and always appear linked together in experience.

There is a sound, and you may say: “I hear a sound”. Namely, there is me, and there is the sound—a duality. However, let us ask you: can you say when or where is it, that the sound ends, and the hearing begins? Can you make a distinction between these two things? Or is it rather that we should say: "There is a hearing"?.

Actually, one and the same experience can be defined either as 'hearing' if described in terms of the subject who hears something, or as 'sound' if described in terms of the object heard. However, in the actual experience of hearing, one cannot establish a precise 'edge' where sound ends 'out there' and hearing begins 'in here': in fact, there is just one, immediate experience and only later on, in order to describe it, the thinking mind says "*I* heard a *sound*", creating a deceptive subject/object duality that is only due to some rules of grammar ([12], 67–69). In fact, the distinction between 'I' and 'external world' is rather a creation of our mental habits.

So the 'boundary line' that splits experience into two apparent 'halves' (subject and object) is not 'out there', but only in our minds: the terms 'consciousness' and 'world' are just two different *descriptions* of one and the same indivisible experience (respectively in terms of the 'first' or of the 'third' person), while the alleged separation between 'subject' and 'object' is nothing but an illusory mental construct, just as 'ascent' and 'descent' are only two different words for the same slope, depending which way one is going. Schrödinger ([11], 127) expresses this revolutionary idea as follows:

> The world is given to me only once, not one existing and one perceived. Subject and object are only one. The barrier between them cannot be said to have broken down as a result of recent experience in the physical sciences, for this barrier does not exist.

As a matter of fact, in our phenomenological exploration of consciousness (that partakes to every experience as a knowing subject, but is not a knowable object), we can safely rely on experience itself, but as soon as we begin to reflect on it, we must never forget that, due to its own intrinsic limitations, thought is to be regarded only as a preparatory tool for a sort of *intuitive* understanding of consciousness.

29.3 Consciousness and the Metaphor of Space

In this respect, employing space as a metaphor can perhaps be of some help to get an intuitive understanding of consciousness.

On considering the similarities between space and consciousness, an interesting point is this: though we cannot see space (since it is invisible), nonetheless we can somehow clearly 'feel' it as a sort of 'unlimited openness' all around us. Actually the same can be said for silence—again something that nobody can 'hear', but that we all perceive distinctly. How does this perception/non-perception come about? Is it all this fruit of our thinking mind, or are at work some unknown biological innate mechanisms?

Those are probably questions we cannot answer, but in this respect the similarity with consciousness is quite apparent: we cannot know consciousness as an object, but we can 'sense' it as a sort of 'awake spaciousness' or 'alive brilliance'. We could say that we 'know' consciousness by *being* it, so this is not a subject/object kind of knowledge, but rather what Forman ([7], 118) calls a 'knowledge by identity':

> In knowledge-by-identity the subject knows something by virtue of being it. […] Being something in this case carries within itself a non-inferential sense of what it is to be it. It is a reflexive or self-referential form of knowing.

Another aspect of the space metaphor which can be useful, concerns the concept of consciousness' contents. Someone can try to explain space by pointing at the sky full of stars, or indicating a far away chain of mountains: there is the space!

But in fact those are only contents of space, as actually the space is what contains all of that. Likewise, things I can point out to are only contents of consciousness, which is the all-inclusive dimension of experience.

There may be another interesting reason why the metaphor of space can be properly drawn when talking about consciousness: although we do not have a direct experience of space as such, this notion is 'embodied' in us. In this regard, Lackoff and Johnson argue [8] that we perceive our own body as having clearly a front/back, right/left, up/down, or in/out side, and this internal knowledge conditions us to see the external world with the same categories, including the container/contained one. Likewise, if the notion of consciousness is 'embodied', this can give rise to the metaphor of inner/out, which, in turn, brings about the dualistic perception of self/other, subject/object and consciousness/external world.

At this regard, one can pose the question: is consciousness localized?

Before examining this question—mentioned already before—we must consider the following interesting point: most of us, on thinking about consciousness, are spontaneously brought to assume that its 'centre' abides somewhere behind the eyes, inside the space of the head. However, this cannot be regarded as a proof of its localization, as there is a simple explanation of this belief. In fact the head contains, besides the brain (that is unperceivable), all the receptors of the five senses, plus the thinking, and is therefore the seat of all perceptions. So it is the common confusion between perception and consciousness which leads one to assume that consciousness abides in the head, and in particular behind the eyes.

In order to clarify this point, again the metaphor of space can be of some help.

Our thinking mind needs to order events through pairs of opposite concepts, like here/there or before/after: space is the common ground that enables the difference between here/there to appear, just as time is the common ground that enables the difference between before/after to appear.

More specifically, space allows localization: all objects are located here and there through the common ground of space. However, space itself is not localized: it is neither here nor there, because every 'here' and 'there' can appear only *in* it. So space transcends the dualism here/there: a localized object can only be either here or there, while space encompasses both of them.

Similarly, consciousness is the precondition for everything to appear in experience, but it is not limited by any specific object, since every experience arises *in* it, as a *content* of it. So consciousness is the common ground that transcends any difference or separation between one object of experience and another: like space, it is not localized in a specific place or time, because every place and time appear in it.

Our exploration of space as a metaphor for consciousness can proceed further by investigating a simple and ordinary everyday situation: imagine that, while preparing some tea, besides the noise of the boiling water you happen to hear also the roar of an airplane passing by. At first, just for one instant, there are only two bare auditory frequencies, but right afterwards the thinking mind identifies their different sources as ‘boiling water’ and ‘airplane’. Then you would probably think: “The sound of the boiling water is closer to me than the sound of the airplane”.

Is that statement really true?

It is, if the word ‘me’ means ‘my body’ and the word ‘sound’ means ‘the source of the sound’: quite obviously the boiling water is closer to my body than the far airplane, since all the distances between objects that are located in space can be precisely measured.

But what about the distance between each object and space? Can an object be closer to space than another one? Indeed, we should say that the distance between any object and space amounts to zero.

Likewise, getting back to the above-mentioned sentence (“The sound of the boiling water is closer to me than the sound of the airplane”), if the word ‘me’ means ‘my consciousness’ and the word ‘sound’ means a bare auditory perception, how can we assert that one hearing perception is closer to consciousness than another one, since both of them appear *in* and *as* consciousness? Just as each localized object is at a zero distance from space, so the distance between any perception and consciousness amounts to zero as well ([12], 68).

In the Indian philosophical tradition of advaita-vedãnta, the relation between the body-mind organism and consciousness is often elucidated by the metaphoric image of a vase (symbolizing the mind-body organism), from which the thinking mind easily gets the odd notion that the space ‘inside’ it (the individual consciousness) is separate from the space outside, and that only when the vase cracks open, the space inside will join up with the space outside.

However, space cannot actually split into separate parts: there are not *two* spaces (one inside the vase and another outside it), but rather one and the same indivisible space incorporates the vase itself. The very notion of a space ‘inside’ the vase is unsuitable, since it is not so much the space to be in the vase, as the vase to be totally *immersed* in space.

Strangely enough, if we rely exclusively on our *direct experience* rather than on thought, there is no evidence whatsoever that consciousness should abide inside the body-mind organism: on the contrary, all the thoughts that we call ‘mind’, all the physical sensations that we call ‘body’ and all the perceptions that we call ‘external world’ appear *inside* the ‘awake space’ that we call ‘consciousness’, and no aspect whatsoever of reality can arise outside it [1, 12].

From a phenomenological perspective, space appears changeless, and this is the reason why it can allow any movement or change of the localized objects that arise and pass away in it. Here the similarity between space and consciousness is quite close. Time is the measure of change: what we usually call ‘time’ is just a conventional tool for our thought to ‘translate’ change into numbers. Actually in our direct experience we do not perceive seconds, hours or years, but only movement and change. However, to notice change, consciousness must be steady, as any movement requires a constant background in order to be detected.

For example, when an itch appears for a few seconds, to note the beginning of the itch, consciousness must be there both before and after its arising, while to note the end of the itch, consciousness must be there both before and after its passing away. So in our experience consciousness as such is the changeless background that allows any perception of movement measured by time. This phenomenological observation could suggest an intriguing philosophical hypothesis: that the very perception of time implies a timeless consciousness as an immutable background, since it is not consciousness to be in time, but rather it is time to appear in consciousness.

This sheds some light on the question, whether consciousness exists only in instantaneous flashes, or whether instead is having a continuity. This problem is in fact linked to the relation between consciousness and time. If one would accept the view that consciousness lives in time, then the above-mentioned question about continuous *vs.* discrete mode of being would have a *raison d’être*. But if we instead consider that it is time that lives in consciousness—time as a ‘content’ of consciousness—then the above question is no longer arguable.

In our experience, indeed, consciousness as such is always here and now, but it is not localized in space and time, because the ‘here’ of consciousness incorporates any possible ‘there’, and the ‘now’ of consciousness includes all the thoughts about past and future: here the word-play between ‘no-where’ and ‘now-here’ is quite appropriate.

The simile between space and consciousness reaches its limit when considering that space itself, though not localized, is still an object which consciousness can be aware of: therefore consciousness encompasses even space, putting an end to the analogy between them.

When still using the metaphor of space, we could say that consciousness is the ‘aware space’ that ‘contains’ every experience. But, if even space appears only *in* and *as* consciousness, then actually consciousness, far from being a ‘container’ (as the container/content relationship still implies the metaphor of space), is rather the *intrinsic essence* of experience, namely the ‘first person aliveness’ that permeates it: as such, consciousness is a sort of ‘singularity’ without any extension in space or time.

Unlike space, consciousness cannot be perceived as an object, and here lies its paradoxical mystery: on the one hand its evidence is undeniable for the very fact that we are aware of objects, on the other hand it is utterly unknowable, just as the eye cannot see itself, but its existence is undeniable for the very fact that we can see objects.

So, at the end of our short inquiry into what we call ‘consciousness’, we are ultimately confronted with the unknown ‘bottom line’ of any human knowledge.

References

1. Bergonzi, M. (2011). *Il sorriso segreto dell'essere*. Milano: Mondadori.
2. Bitbol, M. (1996). *Schrödinger's philosophy of quantum mechanics*. Boston Studies in the Philosophy of Science: Kluwer.
3. Bitbol, M. (2008). Is consciousness primary? *NeuroQuantology, 6*(1), 53–72.
4. Bitbol, M. (2016). À propos du point aveugle de la science. In G. Hess & D. Bourg (Eds.), *Science, conscience et environnement*. Paris: Presses Universitaires de France.
5. Bitbol, M., & Luisi, P.-L. (2011). Science and the self-referentiality of consciousness. *Journal of Cosmology, 14*, 4728–4743.
6. Chalmers, D. (1995). Facing up the problem of consciousness. *Journal of Consciousness Studies, 2*(3), 200–219.
7. Forman, R. K. C. (1999). *Mysticism mind, consciousness*. Albany: State University of New York Press.
8. Lackoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
9. Plank, M. (1931). *The Observer*, 25 January.
10. Schrödinger, E. (1931). *The Observer*, 11 January.
11. Schrödinger, E. (1964). *My view of the world*. London: Cambridge University Press.
12. Spira, R. (2008). *The transparency of things*. Salesbury: Non-Duality Press.
13. Spira, R. (2011). *Presence*. Salisbury: Non-Duality Press.
14. Thompson, E. (2015). *Waking, dreaming, being*. New York: Columbia University Press.
15. Wheeler, J. (2007). *You were never born*. Salisbury: Non-Duality Press.

Chapter 30
Time and Suffering: False Metaphors, (De-)Synchronous Times, and Internal Dynamics

Norman Sieroka

Abstract At times, people feel something which they would describe as "time pressure". But is time really a substance which could put pressure on us? I shall argue that it is not and that indeed a philosophical revisionism is sought for in this context; namely the avoidance of the permanent use of such false physical and economic metaphors like "time pressure", "time costs", "loss of time". Having said that, the real phenomenon underlying this metaphorical talk is a discrepancy or dissonance between different time scales: the individual time scale of a person as structured by her intentions, aims, and goals; and an objective time scale or intersubjective time scale as structured by (natural or social) events of the world around her. I will argue that also the so-named "fear of death"—that is, the suffering resulting from the worry about one's finitude—can be classified as a special type of such a dissonance; namely as an irreversible dissonance or desynchronization based on the overall stoppage of one's individual time. This stoppage, in turn, will be characterised as the termination point of the ever new and tensed division or divergence between what is experienced as happening now and as being just past. Towards the end of the paper I will come back to the question of a philosophical revisionism—that is, to the cases where discrepancies or dissonances between individual time and intersubjective time do allow for a "re-synchronization" and, thus, for an avoidance of suffering.

30.1 Introduction: Sufferings and Time Scales

Time can make one suffer. On a general level, one might realise one's own finitude and one might suffer from the fact that there are good inductive reasons to suppose that the world around will carry on after one's death. On a more everyday level, one may suffer from what is often described as "stress" and "time pressure": myriads of

---

N. Sieroka (✉)
Professur für Philosophie, Eidgenössische Technische Hochschule Zürich (ETH),
CLW C3 (Clausiusstrasse 49), 8092 Zürich, Switzerland
e-mail: sieroka@phil.gess.ethz.ch

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_30

appointments, deadlines, commitments keeping one busy to an unhealthy extent. In what follows I will analyse the origins of these two kinds of individual suffering. (I use the term "*individual* suffering" because it is related to personal life-worldly limitations and has to be distinguished from a notion of formal suffering, as related to the human capacity of receptiveness quite generally, and from a notion of social suffering, as resulting from shared traumatic experiences of a whole community; cf. [19, 22]).

Apart from analysing the origins of these time-induced sufferings, I will attempt to answer the question whether or to what extent they can be avoided or overcome—that is, I will look at the possibilities and limits of a philosophical revisionism regarding our awareness and understanding of time.

In order to explicate these phenomena in an adequate fashion, it is helpful to first distinguish between different time scales. Generally speaking, time is a fundamental dimension of the world we live in. Given that there are different ways in which we encounter this world, there are also different ways in which one may ascribe a temporal order to (the objects of) these encounters (cf. [20, 21]). One obvious way to give a temporal order to events or things is by means of clocks—that is, by ordering them in relation to periodical physical processes and, thus, by what one might call physical or "objective" time. Not determined by clocks, but given by the way we experience the enduring presence and absence of things, is a different time scale, namely "individual time". This second time scale might also be described as "time as perceived" (by an individual subject) and is to be distinguished from a third scale which might be called "intersubjective" time and which governs our daily social life—that is, our interaction with others.

All three time scales—physical, individual, and intersubjective—are experienced as real and usually we are not aware of any dissonances or discrepancies between them. In fact, the true core of expressions such as "time pressure" is related to the (unusual or pathological) case of experiencing such a discrepancy or dissonance. However, before explicating this "true core", let me start by explaining why expressions such as "time pressure" and "time costs" are indeed misleading and might themselves intensify a potential suffering.

30.2 Overcoming False Metaphors Regarding Time

As already mentioned, time is a fundamental dimension of the world, but time is not a substance. We experience ourselves as living *in* time (both physically as well as mentally), but time, especially physical time, is nothing one could collect, gather or accumulate. Even though this point might seem trivial, it is important to emphasise that time is not a material or economic resource. It is important because a lot of contemporary talk about time works as if time was such a resource and because a lot of suffering is related to taking these false metaphors seriously.

Given that time is not a material substance, time cannot put pressure on something or someone. Similarly, there can be no “lack of time” in any literal sense. Every day lasts 24 h and no day has a leak through which some of these 24 h might seep away. By the same token, time is not like money. To speak of “time savings” and “time costs” is to employ bad metaphors, for there is no way in which one might, for instance, salt away a couple of minutes on Tuesday to then withdraw them on Saturday.

This is not to deny that there are real phenomena related to the use of expressions like “time pressure” and “time costs”. But these real phenomena are related to the durations and relative arrangements of actions and events (and I will come back to this in more detail below). At this point I merely want to emphasise that this talk about *time itself* being a material and economic substance or resource is misleading in a systematic fashion. It suggests that there are issues with time itself whereas the issues are really about us and about our organisation of events in time. Today as well as yesterday and tomorrow we have “all the time in the world”, as it were—that is, 24 h a day. No more, no less. But as soon as one starts to think that time works like money, one might start to shove more and more appointments into one’s datebook because this seems similar to being rich and important. However, whereas a wallet might indeed stretch as I become richer and put more and more coins and banknotes into it, a day does not do anything similar … and instead of becoming rich I merely start being stressed out.

This example might suffice to illustrate how such economic metaphors can have a negative influence on human life by decreasing one’s overall well-being. Accordingly, one would do better by avoiding or revising the use of such metaphors. This might be a first and indeed important goal of philosophical inquiries about time: to reveal bad metaphors and to encourage and help overcoming them. Of course, it is by no means a new idea that a central aim of philosophy is to do revisionary work on metaphors—this has been argued for in various other contexts by, for instance, Nietzsche [16] and, more recently, by Lakoff and Johnson [12].

This kind of revisionism, however, can only be a first step. Even if one successfully overcomes all material and economic metaphors regarding time, one might still feel some kind of temporal disquietness or urgency. So let me explicate this urgency more precisely and relate it more specifically to the different times scales mentioned above.

30.3 Desynchronized Time Scales

The (temporal) relations on the individual time scale are fundamentally tensed. We experience things to happen *now*, we remember certain events to have happened *in the past*, and we anticipate certain things to occur *in the future*. That is, the temporal order of our immediate encounter with the world is ordered in tensed terms of "being present", "being (more or less) past", or "being (more or less) future". In contrast, objective or physical time is usually described in tenseless terms. That is, the fundamental ordering relation here is that of "being earlier" or "being later".

There is a long and ongoing debate about which kind of temporal ordering is more fundamental on a metaphysical level: the tensed or the tenseless one? This debate was triggered largely by McTaggart [13] and recent advocates of tenseless and tensed views include Mellor [14] and Bourne [4], respectively. In any case, what is important in the present context is the fact that even those philosophers who defend a tenseless view on the metaphysical level admit that our experiences come in a tensed order and that there is something like the inescapable presence of experience (see, e.g., [14]).

When it comes to everyday life, actually both kinds of temporal orderings, tensed and tenseless, are important. Tenseless orderings are particularly important for making appointments with friends, authorities, medical practitioners, etc. Calling my dentist's office, I will arrange for an appointment on April 25th at 8 a.m., say. And I will talk about this appointment with the receptionist in exactly those tenseless terms: "April 25th at 8 a.m.". Our conversation will not include tensed classifications about the appointment—that is, there will be no phrasings such as "in thirty-seven days, twenty-one hours and sixteen minutes from now" … and one obvious reason for avoiding such tensed talk is that we would have to permanently update this talk every passing minute.

In contrast, tensed orderings are particularly important in situations where an immediate action is in demand. To take a simple example: if I intend to cross the street it is extremely important for me to realise that the black SUV is approaching me from the left *right now* … whereas, at that very moment, the tenseless classification of the looming accident (i.e. whether it is 3:13 p.m. or 3:16 p.m., say) appears irrelevant. Similarly, also the (tensed) difference between being past and being future is relevant in everyday life and makes a huge difference regarding suffering and overall well-being. Usually we prefer good things to happen in the future rather than the past, whereas the opposite holds true for adverse events. Parfit [17] calls this the "bias towards the future" and an, again simple, example might be my painful visit at the dentist which, if I had the choice, I would prefer to have already happened in the past instead of being the first thing to happen tomorrow morning.

To sum up and to relate things to the terminology from above: tensed orderings are fundamental to what I have called the individual time scale, whereas tenseless ordering relations provide the basis for what I have called the intersubjective time scale and the physical (objective) time scale. Note that this is not to claim that intersubjective time and physical time are the same (cf. [21]). However, at this point nothing much hinges on this difference. In fact, the "true core" of those aforementioned bad metaphors such as "time pressure" can be uncovered independently from that distinction.

What really happens in cases where people claim to be under (temporal) pressure is a desynchronization between one's individual time and the time of one's environment; whereby "environment" might refer either to the social reality, and hence to intersubjective time, or to nature, and hence to physical time. Due to an upcoming storm-front a fisherman might have to hurry up in order to reach the coast "in good time". However, it is not that physical time would put pressure on him. The concern is about the rhythm of his individual actions in relation to physical time. And even if there is no upcoming storm: the fisherman still needs to be synchronized with certain natural rhythms such as given by the cycle of day and night or maybe the tides.

Structurally, the same holds true for the relation between individual time and intersubjective time. There is no pressure originating from intersubjective time itself. Again, the problems of an alleged "time pressure" arise from desynchronization (see [2]). More precisely, the desynchronization is a "running after" of the individual time: I feel "under pressure" because the others around me are all ahead of me. As a simple example: having lunch with my colleagues and realizing that I am the only one who hasn't finished his pasta yet makes me feel uncomfortable and I start to wolf down my pasta because I know that everyone would like to return to work as soon as possible. This feeling might indeed be described as a kind of *social* pressure. However, just to reemphasise this point, this feeling results from a discrepancy or mismatch between my expectations and actions as compared to my colleagues' expectations and actions—it is by no means time itself (neither individual nor intersubjective time) which is pushing me.

Of course, the pasta example is "harmless" in the sense that one would hardly speak of a case of (severe) suffering. However, things come in degree here and the general mechanism is always the same. A repeated or even continuous running behind of my individual time, as compared to intersubjective time, makes me feel that I am always too late, that all the important things happen ahead of me, and all the important decisions are made before I made up my mind. It is for this reason that scholars from philosophy and psychopathology identify this (pathological) case of desynchronization between individual time and intersubjective time with depression (see [7, 24]).

In turn, the pathological occurrence of the opposed desynchronization—that is, the case in which one's individual time is permanently ahead of the intersubjective time—has been identified with mania (cf. again [7, 24]). So mania is the pathological case of what might start out as a kind of boredom. To use the above example again: this time I am faster at lunch than my colleagues are and now I am bored because I have to wait for them to finish their pasta. Next, my boredom might change to impatience and then I might become fussy, calling the waiter, carelessly flicking through the newspaper from the next table, doing a quick call on my cell phone, starting to agitate my colleagues, counting the cash in my wallet, calling the waiter again, getting my jacket, doing another quick phone call, etc.

It is not my point whether this characterisation in terms of doing a lot of unrelated short-term actions is the correct way to describe the borderline between agitation and mania (for more details cf. [1, 8, 22]). My aim is simply to illustrate the fact that both types of desynchronization between individual time and intersubjective time can lead to suffering.

30.4 Re-synchronization and Escapism

If the explication from Sect. 30.3 is correct, then it is obvious what a revisionism has to aim for in order to escape from boredom, impatience, and agitation on the one hand and stress, grief, and depressiveness on the other: namely to overcome the desynchronization between individual time and intersubjective time. Here two broad strategies might come to mind and might be adopted.

First, one might aim at a direct re-synchronization of the two times involved. In the contemporary literature this has been described as a return to the “virtue of patience” [7]. Here “patience” does not mean that one simply “waits”. Instead it means that one feels in accordance—or for that matter: “in resonance”—with one’s social environment. The aim is to re-achieve a kind of common rhythm in which to experience the world and the interaction with the people around.

Second, one might aim to overcome the desynchronization between individual time and intersubjective time by trying to escape from the latter time scale altogether. Surely, as soon as I care only about my individual time, there is no way in which I could feel a desynchronization. Usually, however, such an escapism will provide only a short-term solution to the problem. In daily life one depends very much on other people and, in practice, one might “uncouple” from intersubjective time only occasionally—maybe by doing meditations in which one forgets about the world around one.

From a higher point of view, such a meditative uncoupling might also be viewed as a special case of the first strategy. Successful meditations are often described in terms of an experience of an “eternal now” (cf. also [11]). The experienced present, as it were, widens beyond all limits. That is, within one’s individual time there aren’t any tensed distinctions left between events being more or less future or past. All is experienced to be ordered in tenseless terms and hence no room is left for any dissonance between subjective and intersubjective (or objective) time.

Various revisionist claims from the history of philosophy hint in a similar direction. Some variant of the idea of an “eternal now” underlies, for instance, Spinoza’s claim that (at least “under the dictates of reason”) we should ideally be always “affected equally, whether the idea be of a thing future, past, or present” ([23]: 4p62). Or think of medieval mysticism and claims such as the following one by Jacob Böhme ([3]: 20, my transl.): “He to whom time is like eternity, and eternity is like time, is freed from all strife.” There is no space left for any kind of dissonance or discrepancy because the present as experienced has become all-embracing.

30.5 Finitude and Divergence

A further kind of time-induced or time-related suffering was mentioned at the very beginning of this essay, namely the worry about one’s own finitude. This worry might also be described as the “fear of death”, and its revisionist overcoming, or at least some kind of adaptation, might be considered, once more, as a central aim of philosophy. Ever since the days of Plato ([18]: 67e) it has been claimed that philosophy is meant to help us in coping with our finitude and with the inescapability of death. Quoting Montaigne ([15]: 179): “to philosophize is to learn [how] to die”.

In the light of Sects. 30.3 and 30.4, this worry about one’s finitude might be described as a variant of the desynchronization between intersubjective time and individual time; namely as the fear of an irreversible desynchronization based on the overall stoppage of one’s individual time. The coming into existence of ever new and tensedly ordered individual experience supposedly stops, whereas the intersubjective world of all those other people will continue without any serious interruption. (Though this is not the topic of the present paper, it is worth mentioning that there also exists a complementary worry—namely about the stopping of intersubjective time due to dramatic changes in the environment, such as through a nuclear war or severe climate change.)

Thus, in order to explore possible revisionisms, it is important to better understand the internal dynamics of individual time—that is, to better understand the exact processes involved in the formation of experience and of its inescapable presentness. The most fundamental characteristic or ingredient here is the ever new falling-apart or the ever new (tensed) division between what is now and what is just past. Kant [10] famously described this in terms of what he called the pure form of inner intuition, Husserl [9] described and analysed it in terms of his concept of a time halo, and Brouwer [5] termed this ever new divergence of past and present a “two-oneness” and made it the cornerstone of his intuitionistic foundation of mathematics.

Given the limited scope and length of this essay, I cannot go into much detail here (see [21, 22] for that). But maybe I can motivate at least some important connotations with the help of the following quote from Feuerbach [6]: “Where there is no time, there is no individual; where there is no individual, there is no experience, and vice versa”. The claim is that, without time, there is nothing which is not divisible; without time, there is nothing in-dividual. Or, to put it the other way round and a little more precisely: where there is something that continuously falls apart—something that divides again and again—there is a tensed temporal order. In fact, the German original makes this close relation obvious even on a verbal level, since the German term for "experience" is *Empfinden* which derives from *entfinden*, where the latter means to find oneself to be different, to be distinct, from something. Thus, searching for a single noun describing the experience of a continuous process of falling-apart, once more words such as "di-vergence", "di-remption", "di-chotomy" or "di-visiveness" might come to mind.

This leads back to the claim from the beginning of Sect. 30.3: individual time or time as experienced is fundamentally tensed—and this tensedness is a consequence of the ongoing divergences *within* this time scale. These scale-internal divergences are not to be confused with the dissonances or discrepancies between different time scales which induce suffering. Instead, these scale-internal divergences provide our encounter with the world with a directedness and it is from here that the talk about an "orientation" or "sense" of an individual life arises.

Returning to the worry about one's finitude, the last question to address then is that of a possible revisionism. How, if at all, might the fear of death be attenuated? Some comfort might be gained occasionally by the aforementioned part-time escapism of meditation. Further relief could be found in the, again previously mentioned, Spinozist attempt to view all things tenselessly "under the dictates of reason"—that is, to be affected equally, regardless of whether these things are past, present or future.

Before criticising this Spinozist attempt, let me briefly describe a related but milder kind of revisionism—and then turn to a joint criticism afterwards. Instead of suggesting that one should be equally affected by *all* things, Parfit [17] discusses a kind of pointwise symmetrisation between the past and the future. That is, one would aim to overcome the aforementioned bias towards the future and aim to be equally affected by things which are, say, two weeks past and two weeks future. (Thus, other than the Spinozist attempt, this Parfitian revisionism would still allow for different attitudes towards what is, say, 5 h in the future in contrast to what is 5 years in the future.) Parfit suggests that this would attenuate the fear of death for the following reason: since one does not fear the event of one's birth or the time before, there is then no reason to fear death or being dead.

However, one might object that the latter step in Parfit's argumentation does not necessarily follow. Even if one admits that such a symmetrisation of past and future would be possible, it is not at all clear whether the lack of fearing one's birth will actually override one's fear of death. Why might not the exact opposite happen? Why might not the fear of death override one's easiness with one's origins and thereby lead to an analogous "fear of birth"?

After all, both revisionisms, the Spinozist as well as the Parfitian, make very high demands on human rationality. Even though these revisions may be realizable "under the dictates of reason", it seems questionable whether they will ever occur "under the dictates of immediate experience (or intuition)", as one might call it. Fortunately, this is not bad news, since a full neglect or overcoming of the differences between past, present, and future is not desirable anyway. For this would dampen, if not abandon, experience as such, given that experience itself is fundamentally tensed.

30.6 Conclusion

There are different types of time-induced suffering. First, there is a deceptive one which has nothing to do with time itself but with the use of false metaphors. Then there are different genuine (conceptually related) types of time-induced suffering: stress, impatience, and fear of death. The first two, stress and impatience (and their pathological variants depression and mania), are related to the dissonance or desynchronization between individual time and either intersubjective time or physical time. Such desynchronizations are possible because the different time scales have their individual periodic processes or their particular chains of events which mark their individual "eigenzeit" (their proper or appropriate time scale). Thus, what a philosophical revisionism has to seek, in order to avoid stress and impatience, are strategies for re-synchronizing the different processes and time scales involved.

The fear of death might be described as a more specific worry about the stoppage of one's individual time, about there being a termination point in one's continuous and tensed experience of a falling-apart of things (called "divergence" above). To attenuate this, several strategies are on offer. First, based on beliefs about, for instance, metempsychosis or maybe some other religious convictions, one might simply negate the existence of such a stoppage. Second, and less radically, one might aim at the aforementioned kind of short-term transcendence of one's finitude by means of the experience of an "eternal now". Third, there are the aforementioned attempts of turning humans into perfect rational agents which do not distinguish between past, present, and future. And fourth, one might attenuate the upcoming stoppage of one's individual time by fostering and facilitating an ongoing kind of "derivative existence" in intersubjective or physical time—for instance by writing books and articles in order to make one's ideas and thoughts survive or by having children in order to also propagate one's genes.

This is not claimed to be a complete list. The important point is just that, given that human experience *qua experience* (*ent-finden*) is fundamentally tensed, there seem to be two general ways in which to attenuate the fear of death: either one has to deny that there is a stoppage of one's individual time (first strategy) or one has to follow a rather indirect or short-term strategy. In the latter case (second to fourth strategy) no full dissociation from the suffering from finitude can be gained … at least not without giving up one's existence as a conscious individual human being.

Acknowledgements I would like to thank Jonathan Lorand for his very helpful comments on an earlier draft of this paper.

References

1. Binswanger, L. (1960). *Melancholie und Manie – Phänomenologische Studien*. Pfullingen: Neske.
2. Blumenberg, H. (1986). *Lebenszeit und Weltzeit*. Frankfurt (Main): Suhrkamp.
3. Böhme, J. (1961). *Sämtliche Schriften* (Vol. 10). Stuttgart: Frommann.
4. Bourne, C. (2007). *A future for presentism*. Oxford: Oxford University Press.
5. Brouwer, L. E. J. (1975). *Collected works* (Vol. 1). Amsterdam/Oxford: North-Holland.
6. Feuerbach, L. (1980). *Thoughts on death and immortality* (J. A. Massey, Trans.). Berkeley: University of California Press.
7. Fuchs, T. (2001). Die Zeitlichkeit des Leidens. *Phänomenologische Forschungen, 2001*, 59–77.
8. Fuchs, T. (2013). Temporality and psychopathology. *Phenomenology and the cognitive sciences, 12*, 75–104.
9. Husserl, E. (2008). *On the phenomenology of the consciousness of internal time (1893–1917)* (J. B. Brough, Trans.). Berlin: Springer.
10. Kant, I. (1996). *Critique of pure reason* (W. S. Pluhar, Trans.). Hackett: Indianapolis.
11. Koch, A. (2010). Glück und Zeit. *Neue Zeitschrift für Systematische Theologie und Religionsphilosophie, 52*, 219–231.
12. Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.
13. McTaggart, J. M. E. (1908). The unreality of time. *Mind, 17*, 457–474.
14. Mellor, D. H. (1998). *Real time II*. Cambridge: Cambridge University Press.
15. de Montaigne, M. (1910). *Essays* (Vol 1.) (C. Cotton, Trans.). New York: Hill.
16. Nietzsche, F. (1982). On truth and lies in a nonmoral sense. In *The Portable Nietzsche*. (W. Kaufmann, Ed. and Trans.) (pp. 42–46). New York: Penguin.
17. Parfit, D. (1984). *Reasons and persons*. Oxford: Clarendon Press.
18. Platon. (1977). *Phaedo* (G. M. A. Grube, Trans.). Indianapolis: Hackett.
19. Renault, E. (2010). A critical theory of social suffering. *Critical Horizons, 11*, 221–241.
20. Sieroka, N. (2009). Ist ein Zeithof schon genug? – Neurophänomenologische Überlegungen zum Zeitbewusstsein und zur Rolle des Auditiven. *Philosophia Naturalis, 46*, 213–249.
21. Sieroka, N. (2015). *Leibniz, Husserl, and the brain*. Basingstoke: Palgrave Macmillan.
22. Sieroka, N. (2016). Phänomenale Zeit, Leid und Leidvermeidung – Eine Bestandsaufnahme. *Zeitschrift für philosophische Forschung, 70*, 47–73.
23. Spinoza, B. (1992). *Ethics* (S. Shirley, Trans.). Indianapolis: Hackett.
24. Theunissen, M. (1991). *Negative Theologie der Zeit*. Frankfurt (Main): Suhrkamp.

Chapter 31
Evolutionary Time and the Creation of the Space of Life

Randall E. Auxier

31.1 Introduction: Two Bad Habits

I want to make it clear that I have no quarrel with most scientists, but even many good scientists in general have developed some bad habits in dealing with time during the last century. These habits are built into the practice of science, but do not have much bearing on actual results or argumentation. Still, there is room for discussion about how to think about the way even good science handles time and temporal unfolding. But scientists in general have an even worse habit in dealing with the relationship between genuinely scientific claims and patently unscientific philosophical claims, well beyond the limits of the problem of time. It is the former habit that is destroying our prospects for a greater scientific progress and I will address it in this essay. The latter is destroying our public discourse regarding science, especially in the area of evolutionary theory, and a word or two is due on that score first. As John C. Greene describes the situation:

This paper has benefitted from the criticism of a number of people, and in no way does my acknowledgement of their help imply their agreement with what I assert here. Joungbin Lim of Troy University, Jim Shelton of the University of Central Arkansas, John Bickel of Mississippi State University, and my friend and regular collaborator Gary L. Herstein all criticized an early version of the paper. I have also benefitted from the comments of the members of the Philosophy and Biology Departments of Luther College in Decorah, Iowa, and must thank Greg Jesson and Laura Mueller for arranging the opportunity to pass these ideas through their pleasant gauntlet. A version of the paper was criticized in its late stages by Bogdan Ogrodnik and Łukasz Lamża, to the great benefit and improvement of the result.

R.E. Auxier (✉)
Southern Illinois University, 1263 Lincoln Drive, Carbondale, IL 62901-6899, USA
e-mail: personalist61@gmail.com

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_31

> The predominately deductive approach of the Greek philosophers, says [Ernst] Mayr, "helped to raise questions which no one had asked before, it led to an ever more precise formulation of these questions, and it thereby set the stage for a purely scientific approach which ultimately replaced philosophizing." At what point or period in history, one may ask, did a purely scientific approach in either physics or biology replace philosophizing? When Darwin stated that "as natural selection works only by and for the good of each being, all corporeal and mental endowments will tend to progress toward perfection," was he writing as a scientist or as a philosopher or as an evolutionary theist?[1]

Wanting a pure science that replaces philosophizing and *having it* are two different things. No one is well served when we pretend, to the public or among ourselves, that we possess a kind of knowledge and understanding we simply do not have. Scientists themselves cannot afford to let charlatans like Richard Dawkins (who now apparently believes himself a competent theologian as well as a philosopher) run amuck in the public sphere, misrepresenting the important work that science does. It is time that scientists disavowed car magnets showing fish with feet and "Darwin Loves You" bumper stickers.

31.2 Time Won't Give Me Time

Formal modeling in evolutionary science, such as Ernst Mayr developed in the early 1940s, and Francis Crick brought with him from his study of physics to the DNA project, has removed our thinking from immediate observation and direct (unmediated) application. To reason upon a formal model is to risk making errors that may not be either self-correcting or even discoverable within the confines of the model and the thinking done upon it as a hypothesis about the world. The reason we take these risks in reasoning (through the creation, adaptation, and rearrangement of formal models) is that the payoffs have been great in the past and there is reason to expect more such "progress" in the future. (This assumes that payoff, in terms of greater control, *is* progress—an assumption I would question.[2]) But treating time as a divisible sequence, T_1, T_2, … T_n, on a graph (or the like) is, after all, not time at all. It is a spatialization more closely akin to a wall calendar than even to a grandfather clock. It is a risky formalization of a flux that may not yield its core to such a habit of division.

---

[1] Greene, *Debating Darwin*, p. 154. The Mayr quote is from p. 86 of *Toward a New Philosophy of Biology*.

[2] The metaphor of "progress" is the one that receives the most attention in the debate between Greene and Mayr. See *Debating Darwin*, chs. 2 and 8–11.

The problem of time in modern science traces back to the days of the first improved chronometers. The ancient world had believed in the immutability and cyclical character of the structures of nature. But the engagement of scientists with the flux, as the Renaissance began, was renewed, and with consequences that have been both enabling and crippling ever since that time. Greene says:

> What, then, brought about the gradual erosion of belief in the absolute stability of the basic structures of nature, species included? Scientific observations—Tycho Brahe's new star, the discovery of extinct volcanoes in the Auvergne region of France, the growing awareness of fossil remains of organisms apparently no longer extant—played their part, but they were not sufficient in and of themselves to overthrow the dominant belief in stability. The scientists themselves found ways of explaining away or sweeping under the rug these inconvenient evidences of universal mutability. Something more was needed to accomplish the overthrow of established belief. What was needed was an alternative conception of nature, which, if pushed to its speculative limits, could make sense of these anomalous facts.[3]

The reigning dogmas today in both evolutionary theory (neo-Darwinism) and physics (the so-called Standard Model of Gravitational Cosmology, i.e., Einstein's Spinozistic universe) are cases of scientists still clinging to the stability view. Most philosophers are just as bad about propagating this unempirical, unscientific view. Progress in thought and in science requires that responsible scientists and philosophers overthrow that dogma. We are in possession of some ideas that help in framing an alternative, more empirical, and more promising conception of nature. These alternatives are very serious (even perennial) philosophical ideas and they will not go away. One day they *will* possess the whole vineyard of human thought and we shall make better wine.

In the meantime, the inertia and dullness of the scientific journalists as they lead, by the nose, an impressionable but impatient public imagination, is complemented rather nicely by the evangelism of reductionist throwbacks like Dawkins and E.O. Wilson—and Hawking, and all the rest who are still dreaming crypto-theological dreams of a nature governed by "laws," or mathematical order, whatever *that* may mean. But what these laws or regularities or patterns (or whatever other cluster of a priori concepts they land upon) can possibly *be*, these dogmatists won't allow a philosopher even to comment upon, these days. To be allowed through the doorway, philosophers must sign a tacit non-aggression pact, or preferably behave like the sycophants of science who just weren't quite smart enough to become scientists

---

[3] Green, *Debating Darwin*, pp. 162–163.

themselves (this is the way Hawking speaks of the philosophers, in any case). Overcoming this anti-intellectual dogma and rejoining spatialized thought (i.e., the results of formal modeling) to experienced and *un*experienced time is even trickier than I have indicated, since it isn't clear what can be meant by a "new" concept of nature.[4] I will offer an approach, but first I have to describe what time is.[5]

---

[4] I will not here go into a critique of the received approaches to formal models of time, but I will say that the classic book by Michael Friedman captures well the habits of inquiry of these "normal scientists" and well summarizes the formal models of time favored by contemporary physics, as well as their development after Einstein. These have not changed in any significant way since the book appeared, but the whole paradigm of inquiry is currently under serious attack and is, in my view, collapsing under the weight of incoherencies that it always had. See Michael Friedman, *Foundations of Space-Time Theories: Relativistic Physics and the Philosophy of Science* (Princeton, NJ: Princeton University Press, 1983). See esp. pp. 309–320 for a summary of the simultaneity problem raised by Minkowski space and its relation to the conventionality thesis of Reichenbach. I believe this part of the received formalizations is effectively destroyed by the recent book by Canales, *The Physicist and the Philosopher* (Princeton: Princeton University Press, 2015). One aspect Canales documents thoroughly was the conscious effort on the part of Einstein and Minkowski to suppress questions they could not answer, a practice carried forward by Hans Reichenbach to philosophy in the USA when he immigrated. It is his theories, and the formal models they embrace, that Friedman concentrates on in his book. But Friedman recognizes the limitations and describes them well. In addition, see Friedman pp. 317–320 for the real gremlin in the space-time model of general relativity. Friedman says:

> The central problem for both Reichenbach and [Adolf] Grünbaum [who have defended Einstein's approach by appealing to conventionalism as the way to understand measurement], then, is that the standard simultaneity relation is much more intimately connected with the rest of the geometrical structure of Minkowski space-time than they have realized. In particular, they have not seen that the standard simultaneity relation is an integral part of the conformal structure of Minkowski space-time. (p. 320)

Friedman is too kind here. I think these philosophers knew very well that the measurement problem of cosmology went right to the heart of the assumptions of general relativity and its spatialization of time. They simply buried this problem in modeling it. People were impressed with the math and were willing to say "close enough," completely ignoring the fundamental assumptions. Canales shows this in historical detail in her book, and Gary L. Herstein shows the structural and philosophical problems with the standard model of space-time cosmology in his book *Whitehead and the Measurement Problem of Cosmology* (Frankfort am Mein: Ontos Verlag, 2006). Unfortunately, the otherwise impressive looking work of Reichenbach, and Einstein's other minions, bequeathed to the world of physics and scientific journalism, and eventually to the life sciences through former physicists such as Francis Crick, habits of modeling and of thinking about models that have been nothing short of disastrous for honest inquiry. The disciplines dependent upon a deep understanding of time have been in the grips of an ideology about the relation of time to space for almost a century.

[5] It may be that my sketch here will have significant overlap with the current project of John Dupré at the University of Exeter. From the description available of his current grant, it would appear that his concerns and sources are quite similar to mine as regards the constructive parts of this paper. See his description here: https://socialsciences.exeter.ac.uk/sociology/research/projects/details/index.php?id=236, accessed April 26, 2016. Thanks to Gary Herstein for directing my attention to this work.

31.3 The Flux

I don't believe anything in the description I will now provide is seriously debatable, at least among those willing to be guided by common sense. The fact that my description runs counter to the prevailing ideology associated with space-time in physics is not of serious concern, to me at least. The current ideology is both anti-philosophical and sub-philosophical, and yet depends up patently philosophical (as opposed to scientific) assertions. That it has persisted for so long is a testament to the narrowness of mind and sometimes intellectual dishonesty of those who are willing to do almost anything rather than acknowledge the legitimate role of experience *as had by humans* in modeling the physical world.[6]

I do not claim my description is perfect or beyond improving, but the basic points seem to me unavoidable. The basics are built from the philosophies of Bergson and Whitehead, although I believe my description is developed and elaborated in ways they did not consider.[7] Nor does accepting their ideas wholesale figure into what I will say. That is not a requirement, but common sense is. Having provided this description, I will set out a procedure for formalizing it, in the broadest sense, that aims to include all the valuable results of past inquiry and to enable scientists and (I hope) the scientific journalists, philosophers of science, and the public to set aside the destructive and needless debates about evolution and time that contribute nothing to our understanding and much to our unhappiness.

Time is a variable flux. It is not a homogeneous medium "in" which events occur. It is also not a fourth "dimension." That is a model. Yes, time is *treated as* a homogeneous medium by those who have already conceptualized, and hence, spatialized time for the purpose of *thinking* about it and *reflecting* upon what they have thought. This treatment is often valuable, but it is only one model and not useful for every purpose, and indeed, for none beyond a few fairly narrow pursuits within science. Such models are removed from our actual experience of the passage of time, which is varied and complex. Yet, the ubiquitous and most damaging assumption made about time by most scientists (and many philosophers) is that *thinking* about time (in one of these spatialized senses) and *thinking* about the flux are commensurable, either in terms of a temporal unit of some sort, from the vague "event" to the precise "nanosecond," or in terms of a homogeneous medium which will support the seamless application of closely observed spatialized relations to the flux itself.

---

[6] One could spend a good deal of time arguing whether human experience is or is not relevant to a scientific conception of time. The claim is a red herring. There is no scientific conception of time that is not dependent upon a group of philosophical ideas and assumptions, and these are, manifestly, thoroughly human. The human experience of time is not all there is to time, surely, but what we experience as temporal change cannot be wisely or even reasonably excluded from the consideration of scientific time. See, for example, Charles M. Sherover, in his headers and footers to *The Human Experience of Time* (Evanston, IL: Northwestern University Press, 1999).

[7] A scholarly problem in combining Whitehead's and Bergson's ideas is addressed in my article "Influence as Confluence: Bergson and Whitehead," *Process Studies*, in the special focus section on "Bergson and Whitehead," 28:3–4 (Fall/Winter 1999), 267; 301–338; 339–345.

The former approach, using units, however vague or precise, is a crypto-metaphysical atomism, and the latter, "homogeneous medium" view is a crypto-metaphysical monism. Both are unempirical and have religious hangovers from the days in which we believed nature was a stable domain of laws. Still, both atomism and monism "work" as base assumptions for purposive formal modeling, within limits. The limits are established by the assumptions made in *conceptualizing* time. Concepts imply limitation and determination of the phenomena subsumed under them. These limits and simplifications, imposed for the sake of thinking, exclude much that might be considered important about the concrete reality of time, whatever it is. *Concepts* of time are thus radical reductions of the flux.

Thinking about time, conceptualizing that activity of thinking, and then reflecting on the results of such conceptualization is a complex and fallible process. Not all of the structure of thinking is captured and preserved in the *result* of thinking, the "thought." Not all of the thought is preserved in *reflection* upon the thought. The cognitive operations we can carry out in reflection are different, and indeed exhibit different structures, than those we carry out in the process of thinking.[8] By the time we make the decision to introduce a "unit" of measure to carry us from reflection back through thinking and into the flux, we have engaged in radical, fallible reduction at least four times (experiencing to thinking, thinking to thought, thought to reflecting upon thought, and results of reflection to modeling), and then we ignore the accumulated hazards and assert the accuracy of our models. The model, even if poorly made, certainly has a constructive and positive relation to the flux, to be sure, and a range of applicability, in all likelihood, but it is not anything like time as experienced. It is a created space—indeed, it is a thrice-spatialized and reduced product of reflection. Such is the nature of, for example, the "minute," or "hour," and certainly the "nanosecond." No one has ever *experienced* such a unit in the flux and no one ever will. Whitehead rightly argues that when we transpose efficacious process into presented space, we always make an "error." But error is the price of creative advance. Presentational spaces are the ground of such creativity.[9]

---

[8] Kant's distinction between determinate and reflective judgment, in judgment's logical function, is just one example of the different structures, but at the very least, in some cases, thinking begins in the particular and searches for generals and universals that will fit the experience. In other cases, we begin with generals or universals and seek particulars that are subsumed under these concepts. The structures are quite different. See Kant, *Critique of Judgment*, trans. Werner Pluhar (Indianapolis: Hackett, 1983), First Introduction.

[9] See Whitehead, *Process and Reality*, corrected ed. by D. Sherburne and D.R. Griffin (New York: Free Press, 1978), pp. 83–129, 168–183, 208–217. The presentational space we create from the causally efficacious transition underway, the "flux," is a symbol of that flux. It is analogous to the flux, but is more stable, obviously, and captures enough of the flux to be useful, at the price of error. But its creation is "creativity" *par excellence*. Following Whitehead, Susanne Langer says:

> There is a profound difference between using symbols and merely using signs. The use of signs is the first manifestation of mind. It arises early in biological history as the famous "conditioned reflex," by which a concomitant of a stimulus takes over the stimulus function. The concomitant becomes sign of the condition to which the reaction is really appropriate. This is the real beginning of mentality, for here is the birthplace of *error*, and

There are temporal units that are less arbitrary than minutes and nanoseconds. The "day," for example, has two fundamental, empirical meanings. One is the product of reflection—24 h and a bit more, as adjusted (as often as needed) for the small variations in the speed of the earth's rotation. The other meaning is, roughly, "and there was evening and there was morning, one day," to put it in the familiar language of Scripture. Obviously this varies by season and latitude. Interestingly, the religious meaning is also closer to the flux. Every human being has experienced this latter as a transition and as a schema (a thought). The more precise offerings, by contrast, are reduced and corrected and subjugated to an abstract unit. Human beings began experiencing reflective "minutes" only after the clock was devised. Before that, it would better be named a "moment." Both descriptions of a "day" *work*, in general, and the abstract version will work better for purposes that require greater accuracy, such as global day trading and science. The less reflectively mediated version of a "day" will work for knowing when to begin the Sabbath or when to come in from the golf course. Both versions are artifacts of human thinking. Some organic beings have "days" while others really don't—sea creatures living a mile beneath the surface, or mosses that grow only in dark caves, and so forth, don't really have "days."

There is no reason to become an absolutist about the "day" as a unit of time, even though it is so generally applicable to life on earth. Nor is there any basis for becoming intellectually wedded to *any other* "unit" of time. No matter how well grounded the unit may be in the regularities we can discern in the flux, there is always more to the flux than our discernments, in sense experience, in intuition, in perception, in thinking, or in reflection. Our bodies are not suited to the discernment of *all* of the temporal patterns in the flux.[10] We have, by indirect means, become

---

(Footnote 9 continued)

> therewith, of truth. If truth and error are to be attributed only to belief, then we must recognize in the earliest misuse of signs, in the inappropriate conditioned reflex, not error, but some prototype of error. We might call it a *mistake*. Every piano player, every typist knows that the hand can make mistakes where consciousness entertains no error…. The use of signs is certainly a *mental* function.

Langer, *Philosophy in a New Key*, 3rd edition (Cambridge, MA: Harvard University Press, 1957 [1942]), p. 29. (Langer dedicated this book to Whitehead.) Langer's ideas come close to a highly empirical morphological theory. Although I will not be able to discuss it here, it is worth noting that botanists, as distinct from zoologists, have long discussed and often favored morphological ideas that follow a different historical course, from Goethe through Schleiden and into the present; see the work of, for example, the Polish dendrologist Paweł Kojs (a list of papers is here: https://scholar.google.pl/citations?user=eAU98osAAAAJ&hl=pl).

These are, in my view, non-Darwinian evolutionary ideas, and the public and even the scientific community has been so focused.

[10] I go into some detail about how our bodied deal with the flux in "*In Vino Veritas*," in *Southwest Philosophy Review*, 30:1 (January 2014), 39–66; and "Image and Act: Bergson's Ontology and Aesthetics," in *Sztuka i Filozofia/Art and Philosophy*, 45 (2014), 64–81.

aware of patterns of change in the flux that have both greater regularity (pulsars, the decay of radioactive isotopes, etc.) than our bodies can confirm, on their own, and also we have become aware of aspects of the cosmos that currently have no *observable* pattern of regularity, even with our most developed instruments of mediation (empty space, black holes, and other unthinkables). Our bodies are unsuited to thinking about nothingness, but we must remember that nothingness cannot mean more than "nothing *to us*."[11]

Our instruments of mediation do make available to our bodies aspects of the cosmic order that we couldn't otherwise perceive, but there is no warrant in imagining that what our instruments, from super-conducting super-colliders to magnifying glasses, deliver to us *the whole flux*. Whether the flux is a plenum or whether there is a universal medium of some sort is a question about whether there is "something" that is "nothing *to us*." It s not a scientific question and will never become one, no matter how much of the flux we manage to draw into our thinking by use of indirect instruments. But the answer is "yes." Things do exist that are nothing to us. When we consider the flux as an *idea* about the cosmos, we have to acknowledge that the idea of the flux is inexhaustible. It can be conceptualized and reduced in infinitely many ways, but the idea itself, as a response to the very real results of our careful observations, thinking, and reflecting, should be general enough to keep us all, and scientists especially, constantly in mind of the fact that we don't know all there is to know, and even if we did, and had reached the completion of human knowing, more would exist than we know, because knowing is a reduction of the real to those parts we can fit into a knowledge scheme, developed according to a method or methods, no matter how broadly conceived and applied.

Now, granting that my *idea* of the flux is *not* the flux itself, I still believe that a constructive, careful, scientifically valuable description of time can be offered to those who have settled for either the monistic or the atomistic picture of the applicability of some unit or common medium across reflecting, thinking, perceiving, intuiting, sensing, and acting. Both the monistic and the atomistic approaches habitually ignore the precognitive experience of time and both exhibit no curiosity about the flux, or even the *idea* of the flux, and whether there may be structure in the flux that is intelligible but, as yet, unthought. Instead, the entrenched defenders of one approach or the other dig in their heels and insist upon the sufficiency, even the *sole* sufficiency in the case of the most ideological individuals, of their units and models. The arrogance is unbecoming of science, and many scientists do not participate, but the loudest and most noxious among them hold forth without surcease, especially to the scientific journalists who will convey their pontifications to the masses, and these prolix prattlers are not muffled by the more humble majority of scientists.

---

[11] See Bergson, *Creative Evolution*, auth. Trans. Arthur Mitchell (Lanham, MD: University Press of America, 1983 [1911]), Part IV.

31.4 Overlapping Temporal Modalities

The flux is, at the very least, highly variable in its temporal modalities.[12] This is not debatable. Relative to our bodily capacity, we could say, without great distortion, that these differing, variable realities within the flux *exist* at different "speeds" or "velocities." When we assert this, we use our bodily capacities (including thinking), and their limitations (including whatever goes unthought), as a unit-marker for what we take to be a more varied flux than we can actively use or even know about. Our bodies are, after all, *in the world*, and as actual as anything else we might use as a measure. Some particles, such as neutrinos, which travel either near the speed of light, or faster than light, depending on whom you listen to, have no observable effect on our bodies, since they leave no energistic traces. It seems that this sort of particle exists without being experienced in any meaningful sense.

Taking ionizing radiation as an example of *experienced* particles, we find other particles, such as alpha and beta particles, passing through without our being able to *perceive* them, but with some effect upon our bodies (hence, "experienced"). Yes, we do gradually *sense* ultra-violet light, for example, by its tendency to cook us, but still without perceiving it. But the point is that the flux has variable modalities that interact differently with the human body, making the flux variably relevant. This process is well described by Ogrodnik applying Feynman to the physical situation:

> We can say metaphorically that quantum interference is a special kind of entanglement of many 'fibers', which constitute the subtle 'tissue' of reality. If an action in a region is stationary then there is a constructive interference and the particle runs through the region along an almost classical trajectory in conformity with the principle of least action. If the action is non-stationary in a given region, then we obtain a very complicated pattern of presence for the particle in the region.[13]

---

[12] The best formal description I know of regarding this is Feynman, et al., 1965, summarized by the physicist Bogdan Ogrodnik in this way:

> In classical mechanics, a particle moves along the trajectory on which a variation of action reaches the minimum. The particle does not 'see' or 'feel' the neighbouring, possible trajectories but at once 'chooses' the optimum. The situation is completely different from a quantum point of view. Is it true—asks Feynman—that a particle does not simply follow the right path but it 'looks around' to 'see' other possible trajectories? And if it finds such obstacles which unable it 'to look', does this result in diffraction? The true miracle is that it is exactly as has been described. It may be seen, however, that the principle of least action has not as yet been formulated completely, for this is not the way that a particle chooses the path of least action. Rather it 'feels' all the nearby paths and chooses the one the action is the least. The right path is the one for which there are plenty of neighbouring paths with the same phase.

See Ogrodnik, "The Metaphysical Dimension of Optimizing Principles," in *Concrescence: Australasian Journal of Process Thought*, 5 (2004), p. 3. He is summarizing Richard Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals* (New York: McGraw-Hill, 1965).

[13] Ogrodnik, "The Metaphysical Dimension of Optimizing Principles," p. 3.

Here Ogrodnik uses the language of "region," which betokens a mereotopological approach that I think is favorable to theorizing the sorts of coordinate spaces I am describing. The point is that an interpretation of the flux in terms of a modified quantum description is adaptable for my purposes.[14]

Further, as perception goes, within the visible spectrum, our perception of light almost seems to *precede* our sensations of light—to have eyes is to reach for light, physically, to anticipate it, whether it is available or not. In short, we temporally extend our bodily processes into what the flux may or may not supply, and hence, we extend our durational feelings and intuitions into the future before it becomes actual.

These examples are offered as ways we might understand the flux as already separated into tremendously variable processes. "Speed" relative to our bodies' capacities for *use* of the flux is not the only way to analyze the flux.[15] It is the easiest, but all sorts of relations exist in the flux that may not be captured by thinking according to variable speed, or by the criterion of use. For example, it doesn't seem that gravitational influence is exactly the same thing as speed, which is usually thought of in terms of electro-magnetic propagation.[16] This variability of the flux is confirmed in a thousand ways and confirmable in a million more. It is not debatable.

---

[14] Other examples of such process-friendly theorizing that would work for a truly general theory of evolution include Łukasz Lamża, "Six Phases of Cosmic Chemistry," in *Hyle: The International Journal for the Philosophy of Chemistry* (2014), http://philsci-archive.pitt.edu/11272/1/Lamza_Six_Phases_of_Cosmic_Evolution_DRAFT.pdf, accessed June 9, 2016.

[15] The "use" criterion is Bergson's favorite landing zone for his flights into intuitive temporal experience. To my mind, he relies on it a bit too heavily. Langer notices, for example, that symbol formation seems to await useless experience. See *Philosophy in a New Key*, pp. 116–117.

[16] Whitehead says:

> If we go below the quanta of time which are the successive vibratory periods of the primate, we find a succession of vibratory electromagnetic fields, each stationary in the space-time of its own duration. Each of these fields exhibits a single complete period of the electromagnetic vibration which constitutes the primate. This vibration is not to be thought of as becoming of reality; it is what the primate is in one of the discontinuous realisations. Also the successive durations in which the primate is realised are contiguous; it follows that the life history of the primate can be exhibited as being the continuous development of occurrences in electromagnetic field.

See *Science and the Modern World* (New York: Macmillan, 1925), pp. 137–138. Here we bring together, of course, what I am calling the space of life and that of mentality as primarily electromagnetic overlapping of temporal modalities. It is true that the electric and magnetic character can be divided and there is displacement, so there is no reason to assert an ultimate simplicity in this mode of the flux, only that the overlap is productive of spaces that include the space of life. See Ogrodnik, "Towards a Metaphysics of Light: Whitehead's Metaphysics of Vibrations".

31.5 The Creation of Space

What I will say now is debatable, but not easily. Space is the dynamic overlap of multiple modes of the flux. There is complementarity, tension, and conflict in the overlap, the effect of which is, overall, to slow down various aspects of the flux and bring them into a transient stability. We call that transient stability "space," in all of its forms. In their mutual interactions, multiple aspects of the flux come into such conflict and complementarity along multiple axes. As far as we can grasp it, space exists *only* as a generality—that is, as a sort of average among competing and overlapping modes of the flux. Some spaces are fleeting and unstable, while others are astonishingly stable, given their complexity. We do not know why some are stable while others, far simpler, are not, but we can probably make some progress in understanding that issue—if we could leave behind dogmas that grow from our failure to understand what we are *doing* when we use ourselves as a physical and physiological measuring rod.

At one level we simply have no alternative to positing ourselves, and our actual experience, as the measuring rod. At some point, no matter what instrument we may consider, apart from our bodies, someone had to calibrate it by hand and eye for our perception. That is as true for a useful concept as for an instrument.[17] Such indirect means indicate the intelligibility of what is *not* part of our experience. We are not justified in asserting that there is, in the flux, anything that is unintelligible in principle, to *some* experiencer, simply because we cannot ourselves imagine that experiencer. But we must not go so far as to imagine that we know *all* of it to be intelligible *to us*, even by means of mediation.

Thus, there are spaces throughout the flux of which we are unaware, but there is no reason to pronounce them undiscoverable in principle. Good physical investigation seeks not just particles and patterns of propagation, but highly defined spaces of intelligibility (overlap in conflict and complementarity of the flux) in which highly specified phenomena *might* appear and leave a trace on our instruments or our perceptions (including, by implication, our senses). What we learn is not, for example, that "gravity waves exist," but rather that a space can be created in which a contrast in the flux we *name* "gravity waves" might be indirectly detected, under extremely limited, very expensive, and highly idiosyncratic circumstances. What that complex space is really like, and whether we have understood its limits and conditions is not provided by the experiments by which the trace of its conjectured

---

[17] For more on the relation between such tools as language and concepts and ordinary physical tools, see Larry Hickman, *Philosophical Tools for Technological Culture* (Bloomington, IN: Indiana University Press, 2001).

actuality is documented. Whether “gravity wave” is an apt name for what has been discovered is also not provided by the indirect observation.[18]

We tend to find that the hypothetical existences we *expect*, just as our eyes “expect” light, and those we *find* require tremendous refinement over our history as we learn more about the limits and conditions of the spaces in which they are observable or which they create—the electron of today bears little resemblance (or none) to its first theoretical origins. Phenomena we believed to be different are gradually seen to be the same, while some phenomena we thought to be the same come to be distinguished. The “names” we have given them often deceive us and

---

[18] In response to Stuart Hameroff’s announcement that LIGO had measured the gravitational waves, I wrote the following to him, and to the full list of academic discussants on the public discussion of science among intellectuals on the list called Sadhu Sanga:

> “With due respect –and I really mean respect, not merely “due” but in all sincerity: They [LIGO] are attempting to re-discover or re-insert genuine time (eliminated by the theory of GR) as wave forms of variable space (also static). They have not even begun to accomplish the task of temporalizing the standard model of gravitational cosmology. These “measurements” (and they are barely that at all, they admit) are not temporal transformations, they are whatever *of* a transformation remaining that can be conformed to wave-*shapes*. This is geometry first, algebra subordinated, just like all of GR. The LIGO people, in interpreting their measurements, fail to consider that (1) there is more to time than the spaces it creates; and (2) all spaces are created by variable time-spans overlapping and inter-nested. They are only looking for the tiny, tiny aspect of time they can conform to the static 4D model. They are not interested in time. This isn’t a “direct detection” of anything, nor does it confirm GR (any more than it confirms any other model of gravitational cosmology that allows the spatialization of gravitational transformations as wave-forms).”
>
> “Even if the LIGO experiments showed what they claim, there would still be massive, massive problems with General Relativity. There is a fundamental problem with the form of the argument itself. Showing x exists in accordance with Theory T's prediction does not imply the truth of Theory T. It shows that T is not eliminated, and it also shows that *and any other Theory T1, T2, T3,* etc., that predicts x exists is *also* not eliminated. Further, even if a gravitational wave “exists,” we do not learn from that fact its full character, implications, effects, or causes. AND, the claims being made completely ignore the measurement problem of cosmology, and the problem of the cosmological constant, and the reconciliation of GR with quantum phenomena. These are not even addressed by this line of experimentation, among other things. They simply ignore the problems and make self-referential, model-centric claims in circles.”

There was some outcry among the scientists at what I said, but nothing that was argued publicly subsequent to my statement by a number of prominent physicists and philosophers gives me any cause to think this statement needs revision. A number of scientists agreed with what I said, in part or whole, including Deepak Chopra of UC San Diego and Stanley Klein of UC Berkeley. The fact is that when we allow the theoretical model to dictate the terms and evaluations of the experiment, we are no longer doing reliable science. We can neither confirm the model itself nor discover its flaws. It is close to what Greene said in the passage I extracted early in this essay. The formal models developed by Mayr and Crick and others for testing selection are not exactly “natural” or “science.” The selection they measure, if it occurs (and it does), is “natural” by courtesy of analogy, tenuous analogy, between any fully formalized model and natural processes.

those names are slow to adapt to our shifting needs. One might use a phrase like "natural selection," for example, in such a way that no one could earnestly tell what it means, even when they think they do. A remark like this applies quite directly to the term "gravity." Scientists themselves are apt to become confused and the scientific journalists and public lag still further behind.

31.6 Evolutionary Time

We are now in a position to address to our main topic: evolutionary time. The fundamental ambiguity in the word "evolution" must be noted—do we mean something as broad as "ordered change," or something as specific as the "progress" of biological beings from whatever they were to whatever they are? The latter of these is a much narrower *space* than the former. In the flux, numerous spaces could be created by the conflict and complementarity of the variable flux that would fairly be called "ordered change." Biological evolution, however, is delicate and, as far as we can tell at present, quite rare in the cosmos. But we might be wrong about its rarity. It may only seem delicate and rare due to our limited perspective. Perhaps it is less unusual than we think, but we are enabled, at all events, to say some things about it here.

It is not difficult to see that if we pushed the orbit of the earth 200 miles closer to the Sun than it currently is, we would probably still be within a dynamic, created space that would permit the development of biological life. But whether it would be the life we see today is very doubtful. The first differences would show up in the flora, and our understanding of the actual delicacy on micro and macro environmental systems that provides just the flora we have would be greatly altered by a 200 mile variance in the constitution of the space. This variance far exceeds the daunting predictions that climate scientists have provided relative to our current climate change. Flora are temporal arrestings of the flux that capture a dynamic tension, a conflict and contrast, as well as a complementarity of a tremendous mix of overlapping modalities in the flux. The flora hold these modalities for a time and then, as it might be poetically described, they must renew by going to seed—nearly die in order to live again.[19] Any significant alteration in the dynamic space might make that renewal less well accommodated by the created space, but obviously there *are* alterations in the space to which the flora respond. It is a dynamic overlap that permits a disequilibrium (and hence concentration) of energies that wouldn't otherwise be concentrated.

Similar thought experiments can be carried out by moving the earth's orbit back by 500 miles, or 1000. As the flora vary so will the fauna. Human beings probably wouldn't be part of those temporary, dynamic overlappings of the flux if we altered

---

[19] I cannot help recommending here Voltairine de Cleyre's moving description of the dead morning glory vine that she saw bloom, and her poetic but not altogether unscientific understanding of how such things happen, how they are possible. See "The Dominant Idea" (1910), http://www.voltairine.org/dominantidea.php, paragraphs 2–4 (accessed May 18, 2016).

any of a billion characteristics of the delicate overlap. And indeed, even tiny variations would endow us with sensation, intuition, perception, cognition, and reflection quite in contrast to that which we currently possess—if those categories could still be maintained, which is doubtful.

The role of evolutionary theory, far from being the imperative to instruct the public on the necessities of biological development (as Crick and Watson and Wilson foolishly hold[20]), is to examine the flux and to understand how such phenomena as DNA are one sort of overlap of temporal processes, creating a dynamic space within a range of conflicts and complementarities that exemplify—without determining or exhausting—what is *possible* for the complex space we call "biological evolution." It always could have proceeded in ways other than it did. There is no predetermined class of variations that are somehow "allowed," or even specified in the presence of our current observations. We know some of what is actual, but very little about what is possible. Becoming adamant and inflexible in the presence of challenges to a pet model helps no one. It strengthens the spirit of dogmatism among all the humans, scientific and religious as well, and blocks the road of inquiry. The humans are not of great consequence from the standpoint of the flux, it appears, but to ourselves we are important. We want to live, to flourish, to thrive, to be at peace without being bored. We want not just a physical or a biological space, but a *cultural* space in which the endowments of the flux are available for our highest aspirations. We want to be good, to enjoy beauty, to tell the truth—well, many of us want that, at least. These spaces of goodness and beauty are rarified spaces indeed, so delicate as to come into existence only for the briefest moments among our hopes and dreams, and never existing at all in our macro-experience, except perhaps in discovery and such moments as childbirth, or death, or revelry in nature, or, rarer still, even in worship.

How, then, you might fairly ask, is the theorist, the scientist, the inquirer, the delicate and temporary wanderer amid the flux, to address himself or herself to the task of theorizing inquiry? If our older approaches to framing an evolutionary theory have led us down blind allies and into ideologies that can be avoided, how is "the avoiding" to be done? I will do my best to explain, but it is important to remember that if the assertions about variable flux and space creation I have described are accepted, there is no reason to think that "evolution," even as applied to the narrowest sphere of life, i.e., the individual organism, let alone broader phenomena, such as populations and species, is *just one process*, to be explained by a single theory. We must bear in mind the possibility that evolution is a complex of processes the explanation of which will never be reducible to a single mechanism or concept or unit of measure.

Not every attempt at sound theorizing has missed the mark as completely as the neo-Darwinian strategy. For quite a long time there have been evolutionary theories

---

[20] See Megan Mustain, *Overcoming Cynicism: William James and the Metaphysics of Engagement* (New York: Bloomsbury, 2011), 62; and see Francis Crick, *Of Molecules and Men* (Amherst, NY: Prometheus Books, 2004), 98; cited in Mustain, *Overcoming Cynicism*, 63.

that were more comprehensive than the problematic "received view." I would mention the "emergence" model of Robert G.B. Reid as a fine example of broader, non-ideological theorizing.[21] There is the work by numerous theorists reviving parts of earlier Lamarckian theories, saltation hypotheses, quantum and thermodynamic processes and other thinking that might easily be described as "process thinking." Beyond evolutionary theories themselves, certainly those who have studied systems theory, bioenergetics, cybernetics, circadian rhythms and other bio-temporal studies, right down to studies of traditional medicine, are thinking along temporalist lines. I would think that much of what has been thought, measured, correlated, and tested from these ways of approaching the question of evolution could be fitted within a broader, genuinely temporal theoretical framework.

Clearly my admonition that the overlap of temporal modalities be considered in theorizing evolution has been done by numerous theorists, including some who see themselves as loyal to neo-Darwinism, such as Mayr.[22] The fact that these theorists do not think of created space in the way I have described above may seem unimportant, so long as they have developed a nuanced sense of the delicacy of the life-sphere and a sense of all the relevant processes that must be maintained within a limited overlap of the actual time-processes they are modeling. That demand is surely met to some degree by the kinds of theories I listed above, as well as by the standard model. But there is a methodological difference that accompanies my description of space-creation, and which would be of significant value to our future theorizing, were it considered.

It has to be admitted that, space being already general in its mode of existing, we *must* spatialize the phenomena further in order to think about space and, hence, frame any theory. It is important to bear in mind as we create the spatial framework for a theory that the theory does not do certain things, and should never be thought of as doing so. No theory is identical with the existences it intends to describe. No theory is identical with the phenomena (experience of those existences, whether

---

[21] See Robert G.B. Reid, *Evolutionary Theory: The Unfinished Synthesis* (New York: Springer, 1985); and *Biological Emergences: Evolution by Natural Experiment* (Cambridge, MA: MIT Press, 2007). I also recommend as a summary of such literature Pete A.Y. Gunter's article, "Darwinism: Six Scientific Alternatives," *Pluralist*, 1:1 (Spring 2006), 13–30. Gunter discusses quantum evolution, thermodynamic evolution, non-linear evolution, neo-Lamarckian theories, the revival of the Baldwinian thesis, and the genome capture hypothesis.

[22] Highly informative in this regard are the debates of John C. Greene with Mayr and Theodozius Dobzhansky. See *Debating Darwin: Adventures of a Scholar* (Claremont, CA: Regina Books, 1999). This is Greene's memoire, dedicated to Ernst Mayr, "in friendship," but Mayr did Greene a bad turn in the late 1980s. After they had been debating genially for more than two decades (and their correspondence bears this out), Mayr published an attack on Greene called "The Death of Darwin?" which stimulated a good bit of acrimony. In that essay, Mayr was not wholly up front about the debate as it had existed in both private and public for that extended number of years, and he labeled Greene with views he knew to be half truths. Greene's reputation was damaged and he was obliged (with permission) to publish their correspondence in full, to defend himself, in a major journal. But then Mayr placed his attack essay, unchanged, in his widely studied collection *Toward a New Philosophy of Biology*. Rather than becoming petulant, Greene worked through the problems with Mayr. Still, in his memoire, Greene gently sets the record straight.

considered at the level of sensing, perceiving, or feeling) it intends to describe. No theory is identical with the actual *thinking* about the experiences or existences it describes. Nor is the theory identical with the process of *reflecting* upon that thinking about those experiences or existences. Rather, the theory presents the *results* of reflection upon the whole of the existences, experiences, thinking, and reflecting. Its applicability is always impaired by these points of remoteness from actual experience, but not to the point of uselessness.

There are significant transformations when existences are experienced—i.e., as I said, there is always more in existence than is experienced. What is *excluded* in the transformation from existence to experience is nevertheless important to consider while framing a good theory. It cannot be left aside or ignored. This is part of the problem with using Mill's logic of induction in assessing evidence. It is not designed to help us consider what we are leaving aside. Its value (which is considerable) is limited to the ways it helps us combine and sort and interpret observations under the assumption that we actually *have* the important positive pieces of evidence before us, and on the assumption of the "uniformity of nature" within the limits of our current inquiry. This sort of uniformity is an undischarged assumption for all our results. Nature is not in fact uniform in the way we assume. If it were, we wouldn't be here. The deviations from such law-like regularity are what produce the space of life.

Nature then is not uniform in the ways that would be required to generalize beyond fairly narrow inquiries. It may be uniform enough to apply Mill's logic, or the contemporary version of it we find in, for example, regression equations. Still, the hypothesis that nature evolves is the precise contrary of the uniformity assumption, and the former is the warranted assumption. To put a finer point on it, I am asserting that "nature evolves" in the sense that we know of no concrete existences that are unaffected by time, in the sense of the flux, and that among those we experience, it appears that there is sufficient organization in the processes of change to use the word "evolution" for all beings we describe as living. We may be justified in using the term beyond the living things, but we are certainly justified in applying it exceptionlessly to everything we call living. To apply the term "evolution" beyond living things requires a serious discussion that can be set aside at present. But the term "nature" is inclusive of some temporal modalities that, overlapping in certain transient and comparatively unstable ways, will evolve. Since nature includes evolution, it is not misleading to say "nature evolves," even if not everything in nature does.

This result may also not seem important at first, but it is. It means that whatever else may be characteristic of nature, nature *as a whole* is compatible with the reality of evolution. We cannot responsibly frame accurate conceptions and theories of physical laws that make evolution a mystery, which, unfortunately, is what all the reductionists do, both those from physics who defend general relativity, and those from life science who defend neo-Darwinism. Evolution is not a mystery; it is a fact. Reductionist views of the sciences that assert the supremacy of physics but which fail to account for the occurrence, and probably the emergence (since it isn't easy to see how life could exist beyond the sorts of transient limitations we currently associate with the incidence of some very complex carbon combinations) are simply bad theories. They are refuted by the facts. Whatever nature is, it does in fact happen in the way we experience it. (This is true even if our experience is an illusion—the illusion is nevertheless part of the process of nature.)

Nature also includes the processes of sensation, perception, thinking, and reflecting. As far as we know, these processes are even more delicate than life, in terms of the temporal modalities that must overlap to bring them about. These processes, as far as we have been able to discover, occur only in living beings; we do not assert it is impossible for them to exist beyond living beings, but I do not see how we can speak of that possibility in a radically empirical, scientifically responsible way, at this point. Even asserting that it is *possible* that such processes as thinking and reflecting could exist without the occurrence of our kind of embodiment is unclear to me. My suspicion is that the word "possible" is being misused when this sort of idea is asserted. But it can be considered *without* being *asserted*, and I see no harm in that. Thus, we conceive of all the temporal modalities associated with sensing, perceiving, thinking, and reflecting as also evolving, at least to the extent that the organisms *exhibiting* those processes evolve.

It follows that an adequate theory of evolution has to be broad enough in its conception of *how* temporal processes overlap and create spaces to include the evolution of ideas, and along with these, the residual, coagulated and enduring products of ideas, such as cultures. This is nothing new, of course—social and psychological evolution, and even cultural evolution, are topics that reach into the very beginnings of evolutionary theory. As we know, these domains bring with them complexities that most life-scientists, and indeed, most neo-Darwinians (but not Crick and Wilson!) would rather avoid. The problems cannot be ignored, however. To have an adequate evolutionary theory is to have a theory about how everything that evolves does what it does.

My assertion that overlapping temporal modalities are creative of spaces, and that at least *some* of these evolve. Such an assertion must include the spaces created by sensing, perceiving, thinking, and reflecting, which are clearly embedded in increasingly rarified modalities of time. When we are thinking, the relevant energy (itself a very peculiar combination of temporal modalities) moves at a certain speed, and the contrast of that speed with the speed of the energies *surrounding* thinking energies creates a space that is the dynamic space of thinking, while the other surrounding dynamic spaces are not thinking, but also not preventing thinking. The same may be said for reflecting. Although we do not understand these spaces very well, we know far more now than we ever have about the conditions that need to be in place for these spaces to emerge. But we still don't know very much about what "thinking" is. We know even less about reflecting, since it seems to require thinking, but not to be reducible to it. We do have fair practice at habituating thinking and ordering reflection, but the relation of these susceptibilities to something like evolution is barely understood at all, in the present. And our understanding of such things will not improve unless we ask the right questions. No Darwinian is asking the right questions, or, if it could be said that some are, I would suggest that the theoretical framework they have preferred will distort the attempt at interpreting whatever answers may come to those questions.

Time creates spaces. As far as we know, if one could remove all time, one would destroy all spaces. Variable temporal modalities create different sorts of spaces. To ask after the origin of variation in temporal modalities is a very important question. Was there less temporal variation in the past and more in the present? It seems like that might be true. Are there just a handful of fundamental temporal modalities that somehow conflict and become diffused and hence varied as the universe unfolds? This is very hard to know. But it is relevant to physics. Could such an unfolding be called "evolution"? Perhaps. That question will have to be deferred for a future inquiry.

31.7 Method

But what about method? How are those who want to theorize about evolution to proceed such that a full appreciation of both the possibilities and limitations, scientifically and culturally, might be kept before those who are inquiring? I think the answer exists in Whitehead's method of genetic and coordinate analysis. With Gary Herstein, I have written extensively on this topic elsewhere,[23] but for this context I can provide a limited description of the method for evolutionary theory. It is decidedly non-Darwinian, and the general field to which it belongs is a species of mereotopology. The sense of the term "genetic" here (and henceforth) includes but is far broader than the typical usages. It certainly is not the adjectival form of the Dawkinsian "gene."

The initial move in a method that can accommodate the creation of spaces from the variable flux is to propose a coordinate whole to which all evolution belongs and to inquire after the "parts." There is no reason at all to assume that the parts will exist in any clean hierarchy of substances and relations or that they will even be commensurable. Given enough time, irreconcilable phenomena exist at distant ends of both temporal duration and the spaces it creates. Look to the spaces. The spatial structure will be a symbol of the temporal structures, no matter how remote from one another they may be. The temporal modalities belonging to the whole called "evolution" will need to include agencies of transition that *could* belong equally to changing one's habitat and changing one's mind. The operational factors will then surely be very general and perhaps, at first, poorly specified. Their relations will be discoverable only in part, at the outset. Just as Whitehead proposed eight categories of existence (entities), some of which were more final than others, nine categoreal obligations and 27 categories of explanation, our ways of creating a genetic

---

[23] See Auxier and Herstein, *The Quantum of Explanation: Whitehead's Radical Empiricism* (London and New York: Routledge, forthcoming), esp. chs. 7–9.

arrangement of the coordinate whole are both organic and nonlinear—but they must be able to accommodate linear simplifications.

In the case of evolution, theorizing the *whole* of evolving existences may prove to be beyond our ken, but however we propose the whole, it must be open to the inclusion of what we have not yet imagined—indeed, it needs to be open to the reality of what does not yet exist at all and could not be imagined by generalizing responsibly from what *does* now exist. I feel confident that our current peculiarities as existences would not be deducible as genuine possibilities even by the smartest beings we can imagine, if we were to retrogress to a time when the earth was young. There is not much in the primal soup the portends The Beatles. But they were a possibility, were they not? And there is not much in the soup of the present that portends whatever is removed from *us* by 4.5 billion years. Yet, here we are, trying to develop a framework that stretches from soup to Beatles and from Beatles to … well, whatever. Natural laws, if there are any, themselves evolve—or if they do not, then we are currently ignorant of them, since we only experience things that do evolve.

Thus, our evolving whole must be inclusive of possibilities not detectable in or derivable from our present. This whole must be inclusive of all the known modes of time which, although they are far too numerous to specify, are not too numerous to categorize. I believe I have in fact done so, broadly, in distinguishing the spaces of reflection, thinking, perception, sensing, and acting. These spaces can and do exist together. The aspects of physical time that support the overlap of complementary and conflicting physical time that appears, at this point, necessary for the subtler overlap and conflict of modalities peculiar to life, supply the needed categories. They are, interestingly, electromagnetic, but not overtly gravitational. The effects of gravitation upon this overlap remain vague in the present, but it certainly seems to be a requirement for life. For the present, we treat it as a dynamic constant—earth gravity, for now, but allowances must be made in the coordinate whole for tremendous variance in gravitational influence on life. How much variance? It is impossible to know for certain. "As wide as possible for our thinking" would be the wise position.

The electromagnetic range of temporal overlap needs, if we would treat it as a primary basis for the conflict and complementarity of the space of life, to include the life of the senses, and then the synesthetic processes that bring us perception, and then the temporal modalities that allow perception to function as a *feeling of the world*. When we have included these, and the widest range of variation thinkable, we need to take a close look at the temporal relations and modes that can be sorted from among those of perception into those of thinking—very difficult work indeed. And hereafter, we must be able to name and situate the temporal character of reflection in relation to thinking. When we have categorized, named, and sorted these levels to the best of our ability, we should re-examine their results. The "ideas" we *think* with are in contrast to those with which we *reflect*. They may deserve more specific terminology, reserving, for example, "images" for the units of active thinking, and "thoughts" for the units of reflecting, both being types of ideas on different levels of generality and determinate formation. The processes that make us "live," including ideas, seem to work against the entropy of physical systems. For example, I have an idea to exercise and become healthier and by transforming the electrochemical processes within my body concentrate a different quality of disequilibrated localized energies in my body that not only slow its entropic diffusion, but in the case of the idea, taken functionally as the commencement of that exercise regimen, positively counter-act entropy. The entire temporal system of life is shot through with contrasts of this kind—anti-entropic, nonlinear. Obviously the work of Ilya Prigogine is to be treated as a great discovery in this kind of theorizing.

We begin to make progress in evolutionary theory in proportion to our willingness to examine, earnestly and honestly, the difference between what we *do* know and what we would *like* to know in order to *have* an advanced theory. We can show serious progress in some areas—for example, in addition to Prigogine's discoveries about nonlinear processes and disequilibrium, the workings of time in the relationship between thinking and reflecting has been very thoroughly studied, by phenomenologists and psychologists (and gurus and wandering monks) for a very long time. We actually have enough in the way of results, herewith, to speculate responsibly on the order and operation of time *in reflecting* as it exists in contrast (that, is, complementarity and conflict) with actual thinking. The relation of time in the life process to sensation is poorly understood, and for all our efforts, most of our work in relating sensation to perception has been ineffectual. There are gleanings in, for instance, the psychology of sensation of J.J. Gibson and the philosophy of tacit knowing of Michael Polanyi, I think. All of their hypotheses and results can be accommodated by a thoroughly temporal hypothesis about evolution. I have found in the unpublished work of Ted Calhoun (the Bell Labs experimentalist) a number of formulae, expressed in Lie algebras, that may enable a fairly rigorous specification of the most interesting temporal relations between physical process and life. I hope that some of these results will be published soon.

I could say more, but the truth is that I am not so far down the road as to be able to offer more than a vague sense of the coordinate whole to which evolutionary theory should belong. Neither is anyone else. But what I think I do grasp is how the genetic specification *of* that whole should proceed, in a general but adequately determinate way. Evolution, if it is to be theorized, presupposes a *quantum of explanation* (this is *not* a unit of measure) that is not itself an actual unit of division in the conflicting and complementary overlap that is the space of life. That space is not divisible in actuality, only in analysis. If we would understand the "time of life," or of living, we must be reconciled to its actual *in*divisibility. We must not try to quantify what cannot serve first as a quantum. That may sound cryptic, but it means that the *interpretation* of a unit in the mereotopological whole precedes the projection of the whole itself. In short, we project the whole before we can comprehend the whole as the sum of its (interpreted) parts. Whether a given division *in analysis* can become a division in the actual continuum of temporal modalities is a question to be tested, but not to be insisted upon. And what is divisible now may not be divisible forever, and what is indivisible now may yield itself to more than mere analysis in some other cosmic situation. History is an example of the former, and language is an example of the latter. Hence, the question of *genesis* (not the gene, not Genesis in the Bible, but in the Whiteheadian sense) *is* the question of evolution, and it is particular, while the question of the coordinate whole is universal in import.

Part V
Logic/Computer Science

Chapter 32

A Computational Mathematics View of Space, Time and Complexity

David H. Bailey and Jonathan M. Borwein

Abstract Modern computational mathematics requires a philosophical perspective largely at odds with that of traditional mathematics, since current computational mathematics (as distinct from computer science) is by its very nature is discrete, not continuous, and tied to the real world in ways that the more theoretical branches of mathematics (and computer science) often are not. Indeed, computational mathematics provides a means to escape the trap feared by John von Neumann when he wrote,

> [T]here is a grave danger that the subject [of mathematics] will develop along the line of least resistance, that the stream so far from its source [in empirical reality] will separate into a multitude of insignificant branches, and that the discipline will become a disorganized mass of details and complexities.

But even a computational approach to mathematics has limits, not the least of which are the uncertainties of errors in hardware, software and algorithms that inevitably are part-and-parcel with computation, although there are ways to limit these uncertainties. In our chapter, bulwarked by concrete examples, we will try to situate past, present and future mathematical views of space, time, infinity and certainty within a computational context in which, for example, error due to quantum effects begins to compete with traditional sources of logical and numerical inaccuracy. We shall also argue that traditional taxonomies of complexity and completeness are not only outmoded but actually destructive of progress.

[†] Borwein sadly passed away on 2nd August 2016

D.H. Bailey (✉)
University of California, Davis, CA 95616, USA
e-mail: david@davidhbailey.com

D.H. Bailey
Lawrence Berkeley National Laboratory (retired), Berkeley, CA 94720, USA

J.M. Borwein
CARMA, University of Newcastle, Callaghan, NSW 2308, Australia

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_32

32.1 Historical Perspective of Computational Mathematics

In traditional pegagogy, mathematics is taught as a sequence of axioms, definitions, theorems and proofs, entirely similar in style to Euclid's 2300-year-old *Elements*. While we in no way wish to disparage the axiomatic approach (and any serious mathematician surely must master it), there are downsides to an overly pedantic focus on purely formal methods.

To begin with, a purely axiomatic approach is historically dishonest, because the real work of mathematical discovery throughout history has almost always involved numerical and algebraic experimentation, from which insight is gained and hypotheses formulated. Only later were these hypotheses turned into precisely worded theorems and proofs that we read today in journals and textbooks.

Carl Fredrich Gauss, arguably the 19th century's greatest mathematician, explained that his way of arriving at mathematical truths was through "systematic experimentation" [1]. When just 14 or 15 years old, after computing long tables of prime numbers, he conjectured that the number of primes less than n is, for large n, approximately $n/\log n$, which is now known as the prime number theorem [2, pg. 13]. On another occasion, while examining tables of integrals provided by James Stirling, he noticed that the reciprocal of one integral agreed numerically with a limit of the arithmetic-geometric mean iteration. This purely computational observation led Gauss to discover and develop elliptic and modular function theory [2, pg. 13].

Indeed many great mathematicians from Archimedes and Galileo—who apparently said *All truths are easy to understand once they are discovered; the point is to discover them.*—to Gauss, Poincaré, and Lennard Carleson, have emphasized how much it helps to "know" the answer.

More importantly, the axiomatic approach fails to train mathematicians, pure or applied, in the real work of 21st century mathematical discovery, which, more than in any previous era, involves substantial amounts of computational experimentation. Although some of the older generation still say that "real mathematicians don't compute," the majority of research mathematicians today are fluent and accustomed to using a variety of computational tools, such as the commercial products *Maple* and *Mathematica*, as well as any number of custom-written tools for both symbolic and numeric computing.

This synergy between humans and computers has produced a new set of mathematical results that would not have been possible (or at least not likely) to have been discovered in an earlier era. As a single example, out of many that could be listed, in 1996 the following formula for π (known as the "BBP" formula [3]) was discovered by a computer program:

$$\pi = \sum_{k=0}^{\infty} \frac{1}{16^k} \left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6} \right). \tag{32.1}$$

This formula has the remarkable property that it permits one to calculate, by means of a surprisingly simple algorithm, binary or base-16 digits of π beginning at an arbitrary starting position, without needing to compute any of the digits before the specified position. The numerical computation that led to the discovery of (32.1) certainly does not constitute a rigorous proof of this identity, but once discovered numerically, the proof turned out to be a relatively simple exercise of calculus [2, pg. 118].

We are hardly the first to observe that discovery is often the more important part of proof. Two millennia ago, Archimedes wrote, in the Introduction to his long-lost and recently reconstituted *Method* manuscript,

> For it is easier to supply the proof when we have previously acquired, by the method, some knowledge of the questions than it is to find it without any previous knowledge.

Similarly, as 2006 Abel Prize winner Lennart Carleson described in his 1966 International Congress of Mathematics speech on his positive resolution of Luzin's 1913 conjecture, only after many years of seeking a counterexample, he finally decided none could exist. He expressed the importance of this confidence as follows [4]:

> The most important aspect in solving a mathematical problem is the conviction of what is the true result. Then it took two or three years using the techniques that had been developed during the past 20 years or so [to prove the result].

32.2 Reliability of Computations

Computing in mathematics (or in any other discipline) raises troubling questions about the reliability of computation. After all, there are many possible sources of error:

- The underlying formulas or algorithms used might be incorrectly deduced or incorrectly transcribed from underlying theory.
- Computer programs implementing these formulas and algorithms, which often employ highly sophisticated techniques to accelerate the computation (e.g., FFTs, parallel programming constructs, etc.), are certainly prone to error.
- Mathematical and scientific computing relies heavily on floating-point calculations, and the results of these calculations may be numerically unreliable, especially when performed on a parallel computer system with uncertain order of operations.
- All computers rely on a vast infrastructure of system software and compilers that inevitably contains errors.
- Hardware errors can and do occur, particularly in large-scale, long-running calculations. Even quantum-level errors are now expected to be a factor in future computation.

So why should anyone believe the results of calculations? The answer is that most important calculations are checked with one or more independent calculations done using a different algorithm, software package, precision level, hardware system or all of the above.

For example, floating-point calculations, as mentioned above, often give slightly different results when run on a different computer system or one with a different number of processors. Indeed, ensuring reproducibility of floating-point results is a growing challenge in computing in general and mathematical and scientific computing in particular [5]. But there are relatively simple ways to ensure reproducibility, such as by changing the default rounding mode (to see what difference this makes in the results), or by employing software-based higher precision. Although one cannot guarantee absolute integrity in this way, if two calculations using high-precision arithmetic agree to many digits beyond the level required to discover the phenomena in the first place, this is a pretty strong confirmation that the computation is numerically sound (although see below in Sect. 32.3).

As a related example, 21st century mathematicians, continuing a centuries-old tradition, have computed large numbers of digits of constants such as π, in an effort to explore the long-standing unanswered question of whether and why these digit expansions are statistically random in a certain sense [6]. As of the present date, the state-of-the-art in this regard is the computation of over 10 *trillion* base-16 digits and 12.1 *trillion* decimal digits of π, by Alexander Yee and Shigeru Kondo [7]. Their main computation employed the following formula, due to David and Gregory Chudnovsky [2, pg. 108]:

$$\frac{1}{\pi} = 12 \sum_{k=0}^{\infty} \frac{(-1)^k\,(6k)!(13591409 + 545140134k)}{(3k)!\,(k!)^3\,640320^{3k+3/2}}. \tag{32.2}$$

Using this formula, they computed 10,048,832,487,050 base-16 digits of π. Then, in a separate computation, they directly computed 64 base-16 digits of π, beginning at position 10,048,832,487,013, using the following variant of the "BBP" formula [2, pg. 124]:

$$\pi = \frac{1}{64} \sum_{k=0}^{\infty} \frac{(-1)^k}{1024^k} \left(\frac{256}{10k+1} + \frac{1}{10k+9} - \frac{64}{10k+3} - \frac{4}{10k+5} - \frac{4}{10k+7} - \frac{32}{4k+1} - \frac{1}{4k+3} \right).$$

In both calculations, the slightest error at any stage almost certainly renders the final results to be utterly and wholly incorrect. Here are the two corresponding sets of results:

```
d9ae13df df0c64d9 49bacf10 f55ae963 254699a8 bb24624b d47aea96
8016b052
d9ae13df cf0c64d9 49bacf10 f55ae963 254699a8 bb24624b d47aea96
8016b052
```

Needless to say, the final results dramatically agree, thus confirming (in a convincing but heuristic sense) that both sets of results are almost certainly correct.

This raises the following question: What is more securely established, the assertion that the base-16 digits of π in positions $10,048,832,487,043$ through $10,048,832,487,050$ are `8016b052`, or the final theorem(s) of some very difficult work of mathematics that required hundreds or thousands of pages, that relied on many results quoted from other sources, and that (as is frequently the case) only a relative handful of mathematicians besides the author can or have carefully read in detail? When is computation more reliable than proof?

As another example, when one uses a credit card to purchase an item online, it is quite likely that the underlying software generates two pseudorandom prime numbers. To establish that a generated integer is prime, one could use the provable polynomial-time algorithm recently discovered by three Indian mathematicians [8, pg. 302], but in practice the Monier-Rabin probabilistic primality test is used instead [8, pg. 301].

The Monier-Rabin algorithm is *not* a "provable" test, in the sense of an absolute guarantee of primality. But if the integer has 500 or more bits (as is typical in e-commerce) and passes this test once, then we may argue that the 'probability' that it is not prime is less than 1.6×10^{-24}. If it passes the test five times, then the probability is much less than 10^{-100}. Note that such tiny probabilities are inconceivably smaller than the chance that an undetected hardware error glitch or quantum-mechanical fluke occurs during the calculation, not to mention the possibility of a computer program bug or a human mathematical error in deriving, transcribing or applying the test. Given these realities, what is the point of distinguishing between a "provable" primality test (performed either by a human or by a computer) and a probabilistic primality test (performed by a computer)?

Such considerations have led a growing number of mathematicians to now regard computation as on a par with formal reasoning, provided calculations are double-checked by sufficiently rigorous tests. They have also drawn into question the value, from a practical point of view, of pursuing algorithms whose only advantage over other algorithms is that they are "provably correct."

32.3 Finite Computations

Computation, by its very definition, is a decidedly *finite* and *discrete* process. There is no place in real-world scientific or mathematical computing for actual infinities, nor can computation deal directly with the uncountably infinite and continuous real line (or 2-D or 3-D space). It is true that symbolic mathematical software can represent infinity, and can, for example, symbolically evaluate the integral of a function over an infinite interval, but beneath the symbolic processing, the software and the computer it runs on are, of course, discrete, finite entities.

While in some cases the integral of a function over an infinite interval can be computed symbolically, evaluation over some finite interval may well not be possible symbolically and will require numerical computation to obtain a concrete answer. As Stanislaw Ulam once said, "The infinite we shall do right away. The finite may take a little longer" [9]. On the other hand, the space of functions that can be numerically integrated to any desired accuracy is far vaster than the space of functions that can be symbolically integrated in closed form. Similarly, the space of ordinary and partial differential equations (which are a mainstay of applied mathematics) that can be numerically solved to any desired accuracy is far vaster than the space of such equations that can be analytically solved by symbolic processing.

It is true that one cannot absolutely rely on numerical computations. For example, consider the innocent-looking integral [10]

$$\int_0^\infty \cos(2x) \prod_{n=1}^\infty \cos\left(\frac{x}{n}\right)\, dx = \tag{32.3}$$
$$0.392699081698724154807830422909937860524645434187231595926812285162\ldots$$

One might first be tempted to think that this is equal to $\pi/8$, namely

$$\frac{\pi}{8} = 0.392699081698724154807830422909937860524646174921888227621868074038\ldots,$$

but note that while two values agree for 42 digits, they differ beginning in the 43rd digit; they are *not* equal. As it turns out, the integral (32.3) is merely the first term of a very rapidly convergent series; if two terms are taken, the sum is $\pi/8$ to over 500 digits; if three are taken, the sum is $\pi/8$ to over 8,000 digits.

A related example if the following. Consider the following "identity" [11, 12]:

$$\sum_{n=-\infty}^{\infty} \operatorname{sinc}(n)\operatorname{sinc}(n/3)\operatorname{sinc}(n/5)\operatorname{sinc}(n/7)\cdots\operatorname{sinc}(n/p)$$
$$\overset{?}{=} \int_{-\infty}^{\infty} \operatorname{sinc}(x)\operatorname{sinc}(x/3)\operatorname{sinc}(x/5)\operatorname{sinc}(x/7)\cdots\operatorname{sinc}(x/p)\, dx, \tag{32.4}$$

where sinc x means $(\sin x)/x$. Provably, the following is true: This identity is precisely valid for prime p among the first at least 10^{176} primes; but stops holding after some larger prime. Thereafter the sum is less than the integral, but they differ by much less than 10^{-100}.

In short, numerical coincidence is no guarantee of mathematical certainty. On the other hand, when is a mathematical identity "close enough for government work"? When can one for all practical purposes replace the infinities in the above formulas with some large finite values? Many mathematicians are quite content, say, with 100-digit agreement, to conclude that the apparent result is worth seeking a proof for, or at least understanding why the near-identity holds.

32.4 A Finite Universe

As mentioned above, much mathematics from calculus forward (with the exception of finite and/or discrete algebraic structures) assumes a continuous, complete, infinite space, so, for example, we can subdivide the real line *ad infinitum*, the limits of convergent sequences are actual points in the real line, and if a continuous real function is positive for one argument and negative for another, then there is at least one real point between these two arguments for which the function is exactly zero. Since the nineteenth century a careful development of countably infinite limiting processes has been central to mathematical analysis.

But, one can ask, is nature, at the most basic, fundamental level, really like this?

From several lines of research, the answer appears to be "no." For example, the Bekenstein bound, derived from quantum theory by Israeli theoretical physicist Jacob Bekenstein in 1981, places an upper limit on the number of quantum states that can be contained within a volume [13]. For a sphere with mass m kg and radius R meters, the limit is

$$I \leq \frac{2\pi c R m}{\bar{h} \log 2} \approx 2.577 \times 10^{43} m R, \tag{32.5}$$

where c is the speed of light and $\bar{h}$ is Planck's constant. Thus, for example, at most 2.6×10^{42} bits of information are sufficient to perfectly recreate any human brain down to the quantum level [14].

But one can also apply the Bekenstein bound to the entire observable universe. If this is done, one obtains approximately 10^{123} bits, so that the maximum number of quantum states is some $10^{10^{123}}$ [15, pg. 108]. These reckonings are tightly connected with the "holographic principle," in which the information within a sphere is inevitably limited to the information that can be held on its boundary [16]. Whatever these values are, they are finite—the real universe is not the same as the infinite, continuous space of classical mathematics.

We may then ask how many digits of π can be stored in the known universe? Surely it cannot be more than the total number of quantum states, which while enormous is still finite.

Along this line, physicist Max Tegmark, in his recent book *Our Mathematical Universe: My Quest for the Ultimate Nature of Reality* [17], argues that at its most fundamental level, our universe is not merely described by a mathematical structure; it *is* a mathematical structure. He points out that only a mathematical structure, defined by axioms and relations, can qualify for a description of reality that is completely free from human (or even extraterrestrial) baggage. He further proposes two additional hypothesis [17, pg. 267]:

- *Computable Universe Hypothesis*: Our external physical reality is a mathematical structure defined by computable functions.
- *Finite Universe Hypothesis*: Our external physical reality is a finite mathematical structure.

These two hypotheses are explored in Chap. 12 of his book. But whichever of these one is more inclined to accept, both suggest that fundamental physical reality is discrete, not continuous.

32.5 Theoretical Versus Practical

As we noted above, computational mathematics, if performed with suitably rigorous double-checks, can produce results that are arguably as reliable, in a practical sense, as formal proofs in many cases, or at least can be thought of as another avenue to approach secure mathematical knowledge.

But there is one other advantage of modern computational mathematics: It provides a means to escape the trap feared by John von Neumann when he wrote in 1947,

> As a mathematical discipline travels far from its empirical source, or still more, if it is a second and third generation only indirectly inspired by "reality" it is beset with very grave dangers. It becomes more and more pure aestheticizing, more and more purely *l'art pour l'art*. This need need not be bad if the field is surrounded by correlated subjects which have still closer empirical connections, or if the discipline is under the influence of men with exceptionally well-developed taste. But there is a grave danger that the subject will develop along the line of least resistance, that the stream so far from its source will separate into a multitude of insignificant branches, and that the discipline will become a disorganized mass of details and complexities. In other words, at a great distance form its empirical source, or after much abstract inbreeding, a mathematical subject is in danger of degeneration. [18, pg. 291].

Modern computational mathematics provides a medium of communication between theoretical and applied mathematics, and a route to empirical reality. What we can compute, we can be fairly certain has some tangible existence and some tangible possibility of leading to greater understanding of both the corpus of modern mathematics and also the universe around us.

For example, one can argue that the traditional hierarchy of rational, algebraic, elementary and "advanced" functions was driven by a pre-computer age. Exponentials, logarithms, sines and cosines are all considered "elementary," yet the elliptic integral functions

$$\begin{aligned} \mathrm{K}(x) &:= \int_0^1 \frac{\mathrm{d}t}{\sqrt{(1-t^2)(1-x^2t^2)}} \\ \mathrm{E}(x) &:= \int_0^1 \frac{\sqrt{1-x^2t^2}}{\sqrt{1-t^2}}\,\mathrm{d}t \end{aligned}$$

are, from a computational point of view, simpler and easier to compute. This is because these functions can be quickly evaluated by quadratically convergent arithmetic-geometric mean calculations, and so for high precision calculations they are faster to compute than the "elementary" functions. Indeed, K admits of a quadratic transformation

$$\mathrm{K}(k) = (1+k_1)\,\mathrm{K}\left(k_1\right), \quad k_1 := \frac{1-\sqrt{1-k^2}}{1+\sqrt{1-k^2}}, \tag{32.6}$$

as was known already to Landen, Legendre and Gauss. To compute $\mathrm{K}(\pi/6) = 1.699075885\ldots$ to five places requires using (32.6) only twice and then estimating the resultant integral by $\pi/2$. A third step gives the ten-digit precision shown.

In fact, the elementary functions can be computed from the elliptic functions. For example, the logarithm to high precision can be computed as an approximation of K [2, 8, 19]. Explicitly, [19, Algorithm 7.1] gives for $1/2 < x < 1$ and $n > 3$ that

$$\left|\log(x) - \mathrm{K}'\left(\frac{1}{10^n}\right) + \mathrm{K}'\left(\frac{x}{10^n}\right)\right| \le \frac{n}{10^{2(n-1)}}, \tag{32.7}$$

where $\mathrm{K}'(\mathrm{x}) = \mathrm{K}(\sqrt{1-\mathrm{x}^2})$.

The same can be said of algebraic functions. Even in the case of cubic algebraic numbers, in practice it is usually more efficient to numerically solve the underlying polynomial, using Newton iterations, than to use the closed form for a solution due to Cardano; for higher-degree polynomials it is almost always faster to use numerical methods (and hardly any analytic solutions exist).

Thus, perhaps the traditional taxonomy of functions and their computational complexity needs to be re-thought in the computer age.

Along this line, it is interesting to compare theoretical mathematics with the field of theoretical computer science. On one hand, some remarkable and significant results have been produced in the field. We now know that there are some computations that are simply not possible—for example, we know that it is not possible to infallibly determine, by means of a finite computer program running on a real computer, whether or not another computer program has an infinite loop (this is a result originally due to Turing). It cannot, because such a program cannot perform this diagnosis on itself.

Similarly, the $P = NP$? problem is as intriguing as it is far-reaching. This problem is described as follows by the Clay Mathematics Institute, which has offered a US$1,000,000 prize for its resolution:

> Suppose that you are organizing housing accommodations for a group of four hundred university students. Space is limited and only one hundred of the students will receive places in the dormitory. To complicate matters, the Dean has provided you with a list of pairs of incompatible students, and requested that no pair from this list appear in your final choice. This is an example of what computer scientists call an NP-problem, since it is easy to check if a given choice of one hundred students proposed by a coworker is satisfactory (i.e., no pair taken from your coworker's list also appears on the list from the Dean's office), however the task of generating such a list from scratch seems to be so hard as to be completely impractical. Indeed, the total number of ways of choosing one hundred students from the four hundred applicants is greater than the number of atoms in the known universe!

Some problems now known to be in the category *NP*-complete (and thus presumed to be not solvable in reasonable computer time) include [20]:

- The *traveling salesman problem*. Given a list of cities and distances between each pair, find the shortest route that visits each city exactly once and returns to the start.

- The *knapsack problem*. Given a set of items, each with a weight (or size) and a value, find how many of each item to include in a collection so that the total weight is less than a specified maximum weight, and the total value is as great as possible.
- The *subgraph isomorphism problem*. Given two graphs, determine whether one contains a subgraph that is isomorphic to the other.

However, these lines of research have only had muted impact on the real world of mathematical and scientific computing. The main difficulty is that *NP*-complete considerations are driven by a worst-case analysis of a measure that does not reflect practical usage. After all, providing $O(n)$ bounds on the cost, polynomial or otherwise, does not say much if we do not know the size of the high-order coefficients of the polynomial. At a minimum, one must distinguish between questions of the possibility or impossibility of solving the most general form of these problems and the practical difficulty of obtaining satisfactory practical solutions to real-world versions of these problems.

For example, while the traveling salesman problem is known to be *NP*-complete, problems of this general type are solved routinely, every day, in both scientific research and private industry. Airlines, for example, use sophisticated computer programs running on highly parallel computer systems to schedule their aircraft, logistics and staff to the ever-changing requirements of their passenger-destination load. This is done using linear or semidefinite programming techniques, which are special cases of mathematical optimization. They are known to have effective algorithms in *P*, while the classical simplex method, which is still heavily used, does not lie in *P*.

Relatedly Goemans and Williamson showed for the *NP*-hard 'min-cut/max-flow' problem that a randomized version of a semidefinite relaxation after dualization—which being a semidefinite program can be solved approximately in polynomial time—has expected performance exceeding 0.87856 % of the exact solution [21, Sect. 3.3].

Even the navigation software now incorporated into virtually all smartphones performs computations of this type to find a shortest-time route. For that matter, the C, C++, Fortran and Java compilers that are used countless times every day by software programmers routinely employ similar algorithms to schedule instructions for optimal performance of the generated code.

It is important to note that in real-world applications, certainly including airline scheduling, navigation and compilers, it is *not* necessary to find the absolutely optimal solution—a solution that is reasonably close to an optimal solution is entirely satisfactory. In this regards, the theoretical problems addressed, say, in studies of *NP*-complete problems, are of a different class and not really applicable.

Similar considerations apply to other arenas of computer science. In parallel computing, for example, thousands of papers have been written presenting algorithms for the parallel random access (PRAM) model of computing. This model presumes that each of n processors can either perform one arithmetic operation or read from or write to any location in a shared memory. However, scientists performing scientific computations on parallel computer systems have not found this body of research very useful, because the model does not match very well the characteristics of real parallel computers. Instead, virtually all computers at the present and foreseeable future time employ a cache memory system, wherein data access to anything but local registers requires many more clock periods than arithmetic operations; and accesses to data in a remote memory node requires many times more clock periods than accessing data in the local node. For these reasons, today most performance analysis and tuning of algorithms and applications are today targeted to more realistic models, such as the $\log P$ model developed by researchers at the University of California, Berkeley [22] and the roofline model developed by Samuel Williams [23].

A related issue in parallel computing is that for many problems of practical, everyday interest, in mathematical, scientific or even business applications, considerations of massive parallelism are moot, because the problem to be solved simply does not possess sufficient concurrency to justify parallel processing (a consequence of *Amdahl's Law* [24, pg. 348]), particularly when the ever-present software and hardware overhead of coordinating large numbers of parallel processors is considered. It is true that some applications do possess very high levels of concurrency; many such applications are running on large-scale parallel supercomputers operated by universities and government laboratories. But for the vast majority of applications, only modest levels of parallelism can be efficiently exploited; and for still others, only single-threaded execution makes practical sense.

One final item that should be mentioned is the ineluctable fact that mathematicians, scientists and others who are using computers today in their research often do their work via heavy-duty software layers, such as the mathematical software environments *Maple* and *Mathematica*. In such environments it is often not possible to know what algorithms are really being performed "under the hood." This fact considerably complicates any attempts by the user to employ advanced algorithms.

32.6 Conclusion

We have observed that computational mathematics is a more honest representation of mathematics, in that it does not hide or obscure the experimental processes by which a mathematical hypothesis is discovered. Only later are such discoveries turned into the polished, axiomatic expositions that we read in textbooks.

We have also observed that computational mathematics leads to interesting and important considerations of reliability—how to define and arrive at secure mathematical knowledge, given that computation is inherently subject to errors of many types, ranging from programming errors to submicroscopic glitches rooted in quantum mechanics. Yet we have also seen that suitably chosen validity tests can mitigate these errors, turning questionable calculations into extremely reliable mathematical and scientific assertions. Indeed, at some point, computations are arguably as reliable, if not more so, than formal reasoning.

Additionally, there is considerable evidence that a computational approach to mathematics permits the field to escape the trap, originally highlighted by John von Neumann, of becoming so isolated from the richer empirical world of modern science that it becomes both sterile and irrelevant. And even in computer science, connections to real-world mathematical and scientific computations can help focus research in the field in useful directions and help it avoid similar traps of isolation and irrelevance.

While we have discussed current computer architecture and not advanced quantum computing, we are relatively sure that even such advances will not entirely change the situation.

Finally, an expanded reliance on computation in mathematics does need to guard against a diminution in the reliability of the mathematical corpus. Proofs still need to be developed when possible and heuristics need to be clearly labeled.

References

1. Epstein, D., & Levy, S. (1995). Experimentation and proof in mathematics. *Notices of the American Mathematical Society*.
2. Borwein, J. M., & Bailey, D. H. (2008). *Mathematics by experiment: Plausible reasoning in the 21st century* (2nd ed.). Natick, MA: A. K. Peters Ltd.
3. Bailey, D. H., Borwein, P. B., & Plouffe, S. (1997). On the rapid computation of various polylogarithmic constants. *Mathematics of Computation*, *66*, 903–913.
4. Carleson, L. A. E. University of St. Andrews, Scotland. http://www-groups.dcs.st-and.ac.uk/history/Biographies/Carleson.html.
5. Bailey, D. H., Borwein, J. M., & Stodden, V. (2016). Facilitating reproducibility in scientific computing: Principles and practice. In H. Atmanspacher & S. Maasen (Eds.), *Reproducibility: Principles, problems, practices*. New York: John Wiley and Sons. To appear.
6. Bailey, D. H., Borwein, J. M., Calude, C. S., Dinneen, M. J., Dumitrescu, M., & Yee, A. (2012). An empirical approach to the normality of pi. *Experimental Mathematics*, *21*, 375–384.
7. Yee, A. J., & Kondo, S. 12.1 trillion digits of pi, and we're out of disk space. http://www.numberworld.org/misc_runs/pi-12t/.
8. Borwein, J. M., Bailey, D. H., & Girgensohn, R. (2004). *Experimentation in mathematics: Computational paths to discovery*. Natick, MA: A. K. Peters Ltd.
9. MacHale, D. (1993). *Comic sections: Book of mathematical jokes*. New York: Humour: Wit and Wisdom, Boole Press Ltd.
10. Bailey, D. H., Borwein, J. M., Kapoor, V., & Weisstein, E. (2006). Ten problems in experimental mathematics. *American Mathematical Monthly*, *113*, 481–509.
11. Bailey, D. H., & Borwein, J. M. (2011). Exploratory experimentation and computation. *Notices of the American Mathematical Society*, *58*, 1410–1419.
12. Baillie, R., Borwein, D., & Borwein, J. (2008). Surprising sinc sums and integrals. *American Mathematical Monthly*, *115*, 888–901.
13. Bekenstein, J. D. (1981). Universal upper bound on the entropy-to-energy ratio for bounded systems. *Physical Review D*, *23*, 287–298.
14. Bekenstein bound. *Wikipedia*. Retrieved January 08, 2016 from https://en.wikipedia.org/wiki/Bekenstein_bound.
15. Vilenkin, A. (2006). *Many words in one: The search for other universe*. New York: Hill and Wang.
16. Holographic principle (2016). *Wikipedia*, Retrieved January 08, 2016 from https://en.wikipedia.org/wiki/Holographic_principle.
17. Tegmark, M. (2014). *Our mathematical universe: My quest for the ultimate nature of reality*. New York: Knopf.
18. Kline, M. (1980). *Mathematics: The loss of certainty*. London: Oxford University Press.

19. Borwein, J. M., & Borwein. (1997). *Pi and the AGM: A study in analytic number theory and computational complexity*. New York: John Wiley., paperback 1998.
20. P versus NP problem. *Wikipedia*. Retrieved January 10, 2016 from https://en.wikipedia.org/wiki/P_versus_NP_problem.
21. Borwein, J. M., & Vanderwerff, J. D. (2010). *Functions*. Convex. Constructions. Characterizations and Counterexamples. London: Cambridge University Press.
22. Culler, D., Karp, R., Patterson, D., Sahay, A., Schauser, K. E., Santos, E., et al. (1993). LogP: Towards a realistic model of parallel computation. In *proceedigns of the fourth ACM SIGPLAN symposium on principles and practice of parallel programming*. CA: San Diego.
23. Williams, S. W. (2011). The roofline model. In D. H. Bailey, R. F. Lucas, & S. W. Williams (Eds.), *Performance tuning of scientific applications*. CRC Press.
24. Bailey, D. H., Wang, L.-W., Shan, H., Zhao, Z., Meza, J., Strohmaier, E., & Lee, B. (2011). Tuning an electronic structure code. In D. H. Bailey., R. F. Lucas., R. F., Williams, S. W. (Eds.) *Performance tuning of scientific applications*. CRC Press.

Chapter 33
'Photographing the Footsteps of Time': Space and Time in Charles Babbage's Calculating Engines

Doron Swade

Computing machines are artefacts of human creativity. Historically such machines were the most complex artefacts of their time. It was not just a question of structural complexity—the intricate geometry of components and their spatial organisation—but time-dependent behavioural complexity. One arena in which time and space entwine is in design—whether time can be traded for space or vice versa. Another is that of representation—meeting the unprecedented demands of description and understanding that complex time-dependent systems present. System and representation are not always separable: the languages we use to represent things often become the terms in which we understand them. We explore here some of the issues of complexity, time-dependence and representation in the design of the earliest automatic computing machines.

33.1 The First Computing Engines

The earliest designs for automatic computing machines were those for the vast mechanical engines of Charles Babbage (1791–1871), English mathematician and polymath, who devoted the best part of his professional life to designing and attempting to build computing engines [1, 2]. He conceived and produced advanced designs for the first fully automatic calculating engine, Difference Engine No. 1, and went on to design the first programmable general-purpose computer, the Analytical Engine, which incorporates just about every logical feature of a modern electronic digital computer [3]. His tale is one of genius and failure. The designs are towering intellectual achievements that represent a quantum leap in logical

---

D. Swade (✉)
Computer Science Department, Royal Holloway, University of London,
Egham Hill, Egham, Surrey TW20 0EX, United Kingdom
e-mail: doron.swade@blueyonder.co.uk

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_33

conception and physical scale in relation to what had gone before. Yet despite personal wealth, social position, and near-obsessive commitment, he failed to complete any of his engines in their entirety. The first complete calculating engine built to Babbage's original designs was not completed till the modern era nearly 150 years after the designs left the drawing board [1, 4].

Babbage's engines were massively ambitious. Their physical scale was unprecedented. Difference Engine No. 1, designed during the 1820s, called for some 25,000 separate parts. A mid-range Analytical Engine called for an estimated 50,000 parts and the complexity of its mechanisms, control, and behaviour continue to defy complete analysis. The fabrication of parts stretched manufacturing capability to its limits: intricate shapes, achievable precision, standardisation and reproducibility were issues that frustrated advancement.

The features embodied in the Analytical Engine designs are simply dazzling. At an operational level the Engine was capable of conditional branching, iterative looping, and microprogramming, though neither Babbage nor his contemporaries used these anachronistic terms. At systems level it had a separate Store (Memory) and Mill (central processor), a serial fetch-execute cycle, punched card input for data and instructions, an internal repertoire of automatically executed operations, and parallel processing using multiple processors. At user-level it was programmable using punched cards [3, 5].

So how did Babbage represent and describe the unprecedented structural complexity of the mechanisms and their time-dependent behaviour? Other devices of the day—clocks, steam engines, locomotives, textile machines—were simple in comparison, and were describable using the established representational forms: textual description, and conventional pen-and-ink drafting.

33.2 Mechanical Notation

In the case of Babbage's engines the shape and size of parts, and their organisation into mechanisms, were depicted in mechanical drawings using contemporary conventions familiar to us today: plan views, front and end elevations in mainly third-angle projections [6]. There are some 500 such drawings describing the Analytical Engine alone [7]. The information content is essentially spatial and Babbage called these drawings *Forms*. There is nothing radical or revolutionary about the *Forms*—they conform to familiar drafting conventions of the day.

Babbage quickly found that the intricate mechanisms and long trains of action were impossible to hold in his mind at once. He wrote that keeping track of the multitude of parts and their relationship to one another 'would have baffled the most tenacious memory' [8]. His solution was the 'Mechanical Notation', a language of signs and symbols of his own devising that he used to formalise the description of his machines and its mechanisms. He described, in 1826, the need for such an aid, and its genesis:

> The difficulty of retaining in the mind all the contemporaneous and successive movements of a complicated machine, and the still greater difficulty of properly timing movements which had already been provided for, induced me to seek for some method by which I might at a glance of the eye select any particular part, and find at any given time its state of motion or rest, its relation to the motions of any other part of the machine, and if necessary trace back the sources of its movement through all its successive stages to the original moving power. I soon felt that the forms of ordinary language were far too diffuse to admit of any expectation of removing the difficulty, and being convinced from experience of the vast power which analysis derives from the great condensation of meaning in the language it employs, I was not long in deciding that the most favourable path to pursue was to have recourse to the language of signs [9].

When it comes to the power and importance of notation, Babbage has form. As a mathematics undergraduate at Cambridge he was a vigorous advocate of Leibniz's notation for differential calculus in preference to Newton's which was the established Cambridge orthodoxy and resolutely resistant to change. His provocative campaign in 1811 for the adoption of Leibniz's form was instrumental in the founding of the Analytical Society the purpose of which was to reform English mathematics so as to benefit from new Continental theories favoured by Babbage and several of his contemporaries [2].

The Mechanical Notation starts with the naming of parts. Each part was assigned a letter of the alphabet. Various typefaces were used—italicised letters for moving parts, upright letters for fixed framing pieces, for example, and a variety of alphabets including Etruscan, Roman, and Script. The kind of part, its intended motion and function, and its spatial relationship to other parts were indicated by a suite of up to six indices—superscripts or subscripts deployed around each identifying letter [10, 11]. Figure 33.1 shows the arrangement—a single identifying letter for the part, the position of the six indices, and some examples of individual symbols. Two of the indices are non-numerical: the Sign of Form, and the Sign of Motion. The Sign of Form indicated the kind or species of part—gear wheel, arm, lever, spring, rack, shaft, pinion…. These symbols are partly pictograms for ready recognition of function. The Sign of Motion indicated the nature of motion—reciprocating, linear or circular, or combinations of these as depicted in a particular view in the *Forms*—plan, elevation, or end view.

In 1851 Babbage listed ten symbols in the Alphabet of Motion, and some 80 symbols in the Alphabet of Form and wrote that as many as 200 might be needed [12]. There are eight different symbols for 'arm' in the Alphabet of Form, and this part appears at first sight to be over-determined—difference without distinction. However, the differences are nuanced and meaningful, and allow differentiation between an arm that is driven or a driver, for example, whether it acts as a stop in compression or as a pawl in tension, and so on.

The remaining four indices are numerical. The Linear Position index indicated the order in which parts might be arranged on a shaft. The Circular Position index indicated the angular relationship of a part to neighbouring parts. So in a regular helix, for example, made up of arms on a shaft and arranged with a fixed angular

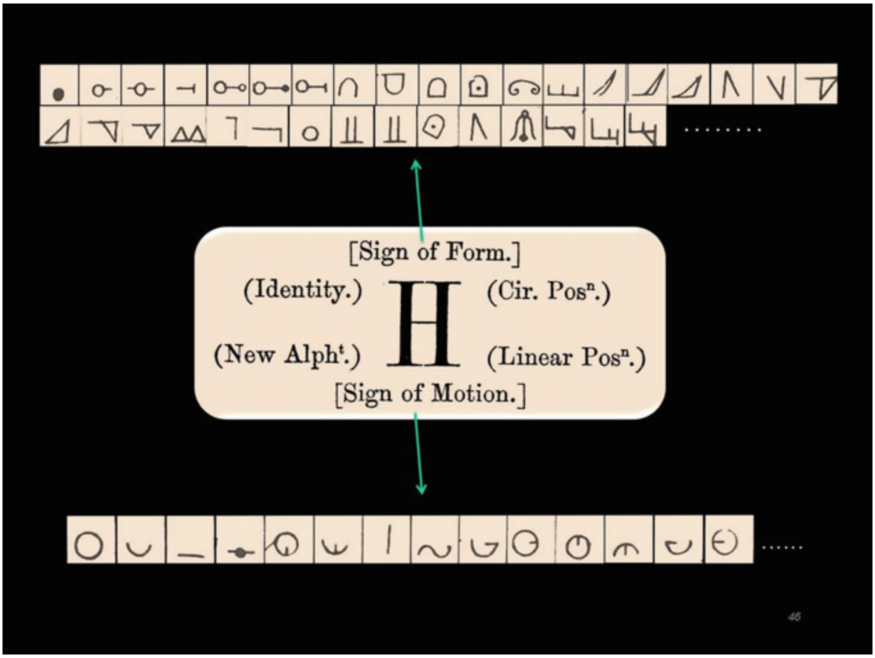


Fig. 33.1 Mechanical Notation: general template for a single part

displacement between them, the Circular Position index and Linear Position index would increment by *1* for each arm as the arms progress along the helix. The index of Identity described whether multiple parts were part of the same ‘piece’ where a ‘piece’ was taken as a single part or combination of parts that together acted on, or were acted upon by, other parts. The New Alphabet index was a contingency in the event that one ran out of identifying letters. It was rarely if ever used.

By combining these indexed letters using syntactical rules Babbage was able to describe the complete causal chain between the prime mover and end result—the species of each part, its physical relationship to other parts, whether driver or driven, its generic function, and the nature of motion. He called these diagrams *Trains* and they represent a symbolic compression of the influence of parts on one another in the transmission of motion. An example is shown in Fig. 33.2.

Here connection and mutual influence of the 4,000 parts required for a printing and stereotyping apparatus was compressed onto a single sheet. Tracing the chain of action is aided by the use of ‘working points’—points of action on each part—represented by lower case letters, italicised for moving parts and upright for framework. Working points could be annotated with the same index of identity as the host part and were used liberally to annotate the *Forms*. The 4,000 parts of the calculating section of Difference Engine No. 2, which calculates and tabulates any 7th order polynomial to 31 decimal places, were reduced to 5 modest manuscript sheets. The *Trains* were clearly an effective tool of representational reduction.

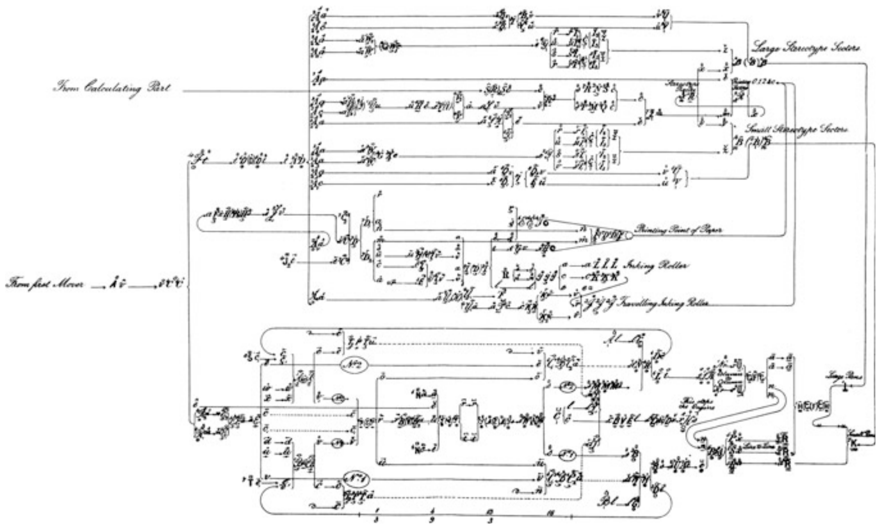


Fig. 33.2 Mechanical Notation: *Train* for 4000-part printing and stereotyping apparatus for the Analytical Engine and Difference Engine No. 2, 1851

The direction of flow in the *Trains* is, in general, from left to right (except where there is feedback). So while time is not explicit, succession in time is implied by the order in which parts appear in the chain. But the *Trains* do not contain an exact description of time-dependent behaviour. There is a third element in the suite of descriptive tools that does this. Babbage called these timing diagrams *Cycles* which show what state any given part is at a particular point in the cycle, what action it is in the process of performing as well as contemporaneous actions of all other parts represented.

An example of a *Cycle* is shown in Fig. 33.3 which depicts the repeated pipelined addition of two multi-digit numbers with conditional carriage of tens. The operational cycle was typically divided into time units along a vertical axis, and the name of parts was arrayed horizontally across the heads of columns.

New notational conventions were used to indicate the action of each part and the timing of actions in relation to other parts. Annotations at the head and tail of an arrow, for example, indicate features of motion: whether the direction of rotation is positive or negative, circular or linear, and whether or not the motion returned the part to its rest position. Other conventions indicate whether motion is continuous or intermittent, conditional or unconditional, in what time window it can occur, and so on. The *Cycles* show the orchestration of the action of individual parts to make a functioning whole.

Babbage’s son, Henry Prevost, who had mastered the Notation, described the difficulty of grasping the actions of parts too numerous and too fast to follow in real time. He wrote that “the Mechanical Notation revealed even its most fleeting movements. It had, as it were, photographed the footsteps of time…” [13].

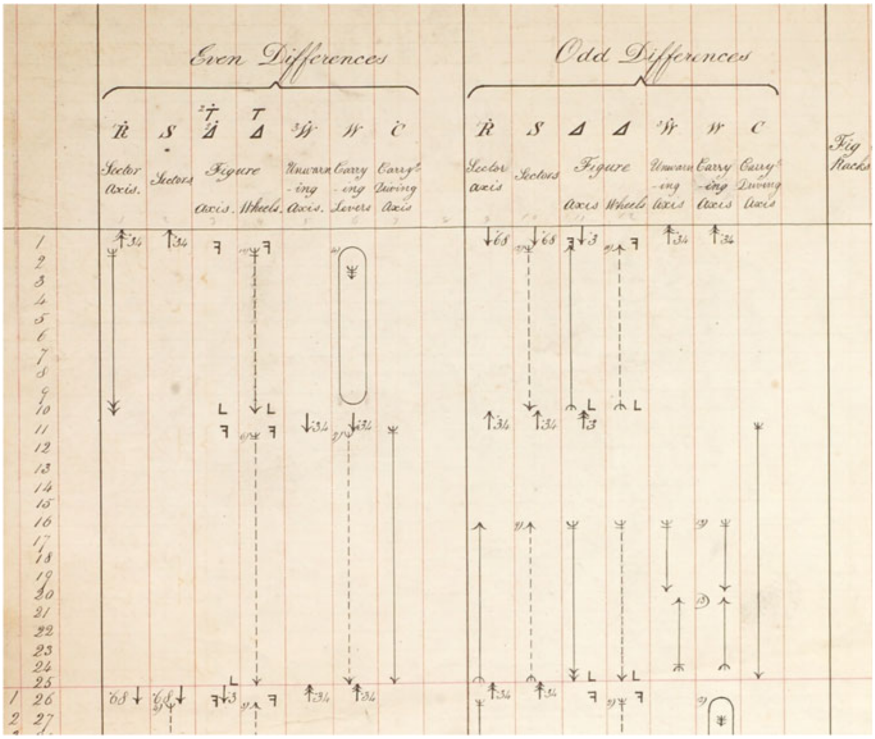



Fig. 33.3 Mechanical Notation: *Cycle* for addition and carriage in Difference Engine No. 2 (part). *Credits* Science Museum/Science and Society Picture Library

Earlier, Dionysius Lardner, an indefatigably colourful populariser of science, captured the idea of the abstractive power of the Notation when he wrote, ‘what algebra is to arithmetic, the notation… is to mechanism’ [14].

The three elements, *Forms, Cycles,* and *Trains* describe, respectively, Space, Time, and Connection and were collectively, though loosely, referred to as the Mechanical Notation. They work as does a three-legged stool, each indispensible to the whole. Babbage made few concessions to his successors, leaving little in the way of explanatory text. What there is in the way of descriptive commentary is fragmented, haphazard, and all too often obscure. Instead, our knowledge of his work is informed largely by the *Forms, Cycles* and *Trains* of his Mechanical Notation and it is on this symbolic encoding of design that we primarily rely for an understanding of his work.

Babbage was inordinately proud of the Notation which he regarded as his greatest contribution to knowledge. He saw it as a universal language with application beyond science and engineering. Two examples he gave of its extended use were the circulation of blood in birds and the deployment of armies in battle. The Notation represented a first level of abstraction and he used it extensively as a design aid—to optimise timing, identify redundancy, derive new motions from existing ones, and manipulate long chains of events in an abstract shorthand all his own. However, the fate of the Notation, baroque in its intricacy and idiosyncratically novel, has been largely one of obscurity. It was used by few others and was not adopted as the indispensable tool of engineering training that Babbage was sure it was destined to be.

The Notation pre-echoes what we would now call a ‘hardware description language’ (HDL). Such languages rose again to prominence in the early 1970s to manage complexity in computer circuit and system design, particularly for integrated circuits where higher-order representation was used to manage vast and complex detail. Babbage had adopted the same route driven by the same imperative —a symbolic language to manage otherwise unmanageable complexity at the component level—the same solution to the challenges of complexity 150 years apart.

33.3 Memory and Time-Dependence

The design of the Analytical Engine was sufficiently advanced by 1837 for Babbage to write ‘programs’ though he did not use that term but referred to his step-wise sequences of operations as ‘Notations of Calculations’. Between 1837 and 1840 Babbage wrote twenty-four such programs for a variety of problems—solutions for simultaneous equations of various kinds, and three examples of series calculated using recurrence relations requiring the iteration of the same set of operations [15].

Number values in Babbage’s machines are represented by the rotation of geared wheels, called figure wheels, engraved with numbers 0–9. These wheels are acted on by the engine’s internal control mechanisms to execute computational rules. It is the ‘physicalisation’ of quantity in a material medium that allows machines to compute at all, which they do by manipulating the physical representations of numbers according to rules. The logical relations in a mathematical statement can be seen as timeless or, arguably, even atemporal. But once physicalised in a machine, actions are necessarily subject to physics and mechanics in ways that logic is not: actions need to be phased in time, measures need to be taken to ensure the integrity of representation and control, and the algorithmic sequence needs to be a correct encoding of the problem.

The ‘physicalisation’ of memory had unexpected implications for the earliest programs. The Store of the Analytical Engine consists of columns of figure wheels that represent multi-digit number values with the least significant digit at the bottom and the most significant digit at the top. Babbage called these stacks of figure-wheel ‘Variables’ to convey that the contents of column changed. It is not the case that the Variable (*qua* column) is assigned to a ‘variable’ in the conventional mathematical sense or that this assignment necessarily holds throughout the computation. Rather, a Variable is closer to what we would now call a ‘register’—a memory location that holds changing data. The Variables in the Store are numbered sequentially, V_1, V_2, V_3, …. Data are directed from the Store to the Mill, operated on, and results returned to the Store. Where in the store the data is placed is determined by a Variable Card—a punched paste-board card that indicates in which Store Variable the data to be operated on is to be found, and in which Variable the results are to be returned [3, 5].

Babbage’s programs consisted of a numbered sequence of instructions from the built-in repertoire of operations (the four arithmetical functions) that could be automatically executed. In using conventional mathematical notation he ran into an immediate problem. The statement $V_4 - V_1 = V_4$, for example, meant ‘subtract the contents of V_1 from contents of V_4 and put the result in V_4’. The operation uses the value in V_4 in a subtraction and overwrites the original value in V_4 with the result. In this way V_4 is reused by replacing its contents. The difficulty was that ‘$V_4 - V_1 = V_4$’ was uncomfortable for mathematicians who would instinctively read it as a mathematical statement that was patently false for all non-zero values of V_1 and trivial in the case of V_1 being zero. To a mathematician the statement seemed an offence to the notion of mathematical identity denoted by the ‘=’ sign. Babbage’s solution was to add a leading index to indicate alteration. The statement would then read, ${}^1V_4 - {}^1V_1 = {}^2V_4$. The leading superscript (index of alteration) indicated that the value of the Variable had changed during the operation. The trailing index (index of location) remained as before, denoting the place in the Store of the Variable in question. Each new reuse of the Variable would increment the index of alteration by *1* and the history of the contents of the Variable could be traced back through the chain of program steps.

The need for solution arose because memory, for the first time in a computing machine, had spatial location, and instructions, expressed in standard mathematical notation, did not reflect time-dependence. The very first program Babbage wrote, dated 4 August 1837, has a sequence of instructions for the solution of two simultaneous equations. It features the double index to show the reuse of Variables though, strangely, he first used Roman numerals for the index of alteration, later changing these to more familiar Arabic numbers [16, 17]. This first program prophetically flagged a more general finding—that coding would require new notational conventions, as indeed turned out to be the case.

33.4 Interplay of Time and Space

The Mechanical Notation, and the notation for the reuse of Variables, were responses to the need to describe complex time-dependent machine behaviour, and their purpose is essentially representational. There is interplay of time and space of a different kind that occurs in design—in the ways in which the physical medium of devices and systems can be manipulated to perform what it is we wish them to do.

In a rare description of the Analytical Engine Babbage wrote:

> It is impossible to construct machine occupying unlimited space; but it is possible to construct finite machinery, and to use it through unlimited time. It is this substitution of the *infinity of time* for the *infinity of space* which I have made use of, to limit the size of the engine and yet to retain its unlimited power [18].

Babbage appears to be claiming that he had succeeded in trading off space and time against each other and that through this exchange the necessary physical limits of a machine did not compromise the scope of its capabilities. He takes some trouble to argue the case. Even though the Variables in the Analytical Engine are limited to 50 digits, he argued, this did not restrict the size of the largest numbers that can be computed: arbitrarily large numbers could be broken up into 50-digit segments and distributed across several Variables and operated on to produce extended-precision results. The price of course is extended computing time and he calculates the increase in execution time for the compound calculation involving large numbers expressed in multiple runs of 50 digits.

He also argued that a physical machine did not need infinite memory for unlimited capacity. With the Store of the Analytical Engine featuring up to 1,000 Variables—a large but finite number—he argued that Variables that were no longer required could have their contents printed or even punched and the Variables then reused—perhaps the first description of a memory dump. He also put the case that the size of a program was not limited by the number of punched cards as the number of Operation Cards, as he called them, could be indefinitely large and 'follow each other in unlimited succession'—a remarkable pre-echoing of the endless tape posited by Turing in his 1936 description of a Turing Machine.

The manner in which Babbage states the space-time trade-off has the ring of general principle though the case he makes in places is contingent on the particular features of the Analytical Engine. There are examples of the space-time trade-off that are less seemingly a priori in several other instances where using a different mechanical principle reduced execution time. In 1837 Babbage wrote of his Analytical Engine, 'the whole history of the invention has been a struggle against time' [19]. He was not referring to permanent overruns in the delivery date for a working machine, but to his unrelenting commitment to reducing the time taken to perform a computational operation, and he went to extraordinary lengths to achieve this.

In the design of Difference Engine 1 and 2 Babbage used a 'successive carriage' technique for the carriage of tens in arithmetical addition. In the difference engines the number value in a complete column of figure wheels is added to an adjacent column with the addition taking place for all digits at the same time i.e. simultaneous parallel multi-digit addition. However, the carriage of tens is not executed during this first phase of the addition. Rather, if a given digit wheel exceeds ten during the addition, a latch is set (a one-bit memory) to warn that for that digit there is a carriage of ten yet to be performed and that the next higher digit needs to be incremented by *1*. In the next stage of the cycle each of the latches is automatically polled in turn. If the latch is warned, then the next higher digit wheel is incremented. If the latch is not warned, no action is taken. The polling of the latches ripples up the figure wheel stack executing carries as needed in strict serial order conditional on whether or not the latch is set. Since in this serial process the polling of the latches is slightly staggered, the time taken for the carriage of tens is proportional to the number of digits in the column i.e. the higher the digit precision, the longer it takes to perform a full addition. The polling arrangement takes account of secondary carries—carries that result from carries—without time penalty [6].

The time demands for successive carriage were tolerable for the difference engines which required only repeated addition to implement the method of finite differences. But in the Analytical Engine, where direct multiplication and division were irreducibly more time-consuming, the successive carriage technique became an unaffordable extravagance. Babbage set out to design a new mechanism on a different principle that would make ‘an unlimited number of carriages in one unit of time’ [20]. The result was his ‘anticipating carriage’ mechanism, akin to a modern ‘carry look-ahead’ technique, which was capable of executing all requisite carries in a single operation. He recalled that the ‘exhilaration of the spirits’ he felt at his success could not be rivalled by the excellent champagne of his host with whom he had later dined on the day he cracked the problem. He regarded the anticipating carriage mechanism as ‘the most important part of the Analytical Engine’ though it is arguable whether this was indeed so, given the prolific number of its other remarkable features [3, 5, 20].

The anticipating carriage mechanism is an example of how a new mechanical principle and a more ingenious spatial configuration of parts produced economies of time. There are many other examples in the engine designs and also of trade-offs in the opposite direction. Babbage’s ‘half zero’ technique, for example, reduced by up to half the time taken to return a figure wheel to zero whatever its initial position. The price was more parts and a more complex control and timing mechanism but the benefit was found in reduced execution time. The printing mechanism in the Scheutz difference engine is an inverse example. Scheutz used an eccentric drum to drive print heads into stiff cardboard to impress numerical results in deep relief to create a mould from which to make printing plates. The forces are excessive as required to make deep cuts and Scheutz’s method takes much longer than does Babbage’s which though quick, is incapable of impressing anything harder that than soft plaster [4].

These examples suggest that the earliest computing engines stretched our ability to represent and effectively manage complexity and time-dependent behaviour. They also remind us that creative ingenuity has a role in trade-offs between space and time—part of the tantalising appeal of engineering design, and one of its great rewards.

Acknowledgements This work is supported in part by the Leverhulme Trust.

References

1. Swade, D. (2000). *The Cogwheel Brain: Charles Babbage and the quest to build the first computer*. London: Little, Brown. In *US as* (2001), *The difference engine: Charles Babbage and the quest to build the first computer*. New York: Viking.
2. Hyman, Anthony. (1982). *Charles Babbage: Pioneer of the computer*. Oxford: Oxford University Press.
3. Bromley, Allan G. (1987). The evolution of Babbage's calculating engines. *Annals of the History of Computing, 9*(2), 113–136.
4. Swade, Doron. (2005). The construction of Charles Babbage's difference engine No. 2. *IEEE Annals of the History of Computing, 27*(3), 70–88.
5. Bromley, Allan G. (1982). Charles Babbage's analytical engine, 1838. *Annals of the History of Computing, 4*(3), 196–217.
6. Swade, D. (1996). *Charles Babbage's difference engine No. 2: Technical description: Science museum papers in the history of technology, No. 5.*
7. Bromley, Allan G. (1991). *The Babbage papers in the science museum: A cross-referenced list*. London: Science Museum.
8. Babbage, C. (1864). *Passages from the Life of a Philosopher* (p. 113). London: Longman, Green, Longman, Roberts & Green. In M. Campbell-Kelly (Ed.) (1989), *The works of Charles Babbage* (Vol. 11). London: William Pickering.
9. Babbage, C. (1989). On a method of expressing by signs the action of machinery. *Philosophical Transactions of the Royal Society, 2*(1826), 217–225. In M. Campbell-Kelly (Ed.), *The works of Charles Babbage* (Vol. 3, pp. 209–223). London: William Pickering.
10. Babbage, H. P. (1856, May). Scheutz's difference engine, and Babbage's mechanical notation. *Minutes of Proceedings of Civil Engineers.* In H. P. Babbage (Ed.) (1889), *Babbage's calculating engines: A collection of papers by Henry Prevost Babbage* (pp. 248–257). London: Spon.
11. Babbage, C. (1851). Laws of mechanical notation. Privately printed. In M. Campbell-Kelly (Ed.) (1989), *The works of Charles Babbage* (Vol. 3, pp. 224–232). London: William Pickering.
12. Babbage, C. (1851). *The exposition of 1851 or views on the industry, the science, and the government of England* (2 ed.). London: John Murray. In M. Campbell-Kelly (Ed.) (1989), *The works of Charles Babbage* (Vol. 5, p. 84) London: William Pickering.
13. Note 10. (1889). p. 257.
14. Lardner, D. (1989). Babbage's calculating engine. *Edinburgh Review, 59*(1834), 263–327. In Campbell-Kelly, M. (Ed.), *The works of Charles Babbage* (Vol. 2, p. 174). London: William Pickering.
15. The Babbage Papers, Science Museum Library, BAB L-series, Notations of Calculations.
16. The Babbage Papers, Science Museum Library, BAB_L_1_0002.
17. Babbage, C. (1864). *Passages from the Life of a Philosopher* (p. 127). London: Longman, Green, Longman, Roberts & Green. In M. Campbell-Kelly (Ed.) (1989), *The works of Charles Babbage* (Vol. 11). London: William Pickering.
18. Babbage, C. (1864). *Passages* from the Life of a Philosopher (p. 124). London: Longman, Green, Longman, Roberts & Green. In M. Campbell-Kelly (Ed.) (1989), *The works of Charles Babbage* (Vol. 11). London: William Pickering.
19. Babbage, C. (1982). On the mathematical powers of the calculating engine. *Papers of Charles Babbage, Museum of the History of Science*, Oxford. MMS Buxton. MSS 7, 26 December 1837. In B. Randell (Ed.), *The origins of digital computers: Selected papers*. Berlin: Springer.
20. Babbage, C. (1864). *Passages* from the Life of a Philosopher (p. 114). London: Longman, Green, Longman, Roberts & Green. In M. Campbell-Kelly (Ed.) (1989), *The works of Charles Babbage* (Vol. 11). London: William Pickering.

Chapter 34
The Black Hole in Mathematics

Alexander Keewatin Dewdney

The main activity of research mathematicians is to search for new and interesting theorems. Until 1930 mathematicians had always assumed that if a mathematical statement was true, there must be a proof for it. But in that year, the logician Kurt Gödel showed that there exist theorems that, while true, could never be proved. They exist in a kind of mathematical black hole from which nothing—no proof—can ever emerge.

Certain mathematical impossibilities are not of this type. For example, squaring the circle, finding a ruler-and-compass construction of a square that has the same area as a given circle, is now known to be impossible.

If a mathematician ever succeeded in squaring the circle, the following statement would soon appear in a leading journal:

Theorem: Pi is a *constructable* number.

The term *constructable* means that the number in question is the length of a line segment that can be constructed with ruler and compass. Unfortunately, no one has been able to square the circle. It was the German mathematician, Ferdinand Lindemann who, in 1882 proved a theorem that forever ruled out the possibility that pi is *constructable*:

Theorem: Pi is a transcendental number.

Another theorem tells us that no transcendental number is *constructable*. Transcendental numbers are not algebraic and therefore not *constructable*, either. Mathematicians tend to accept such limitations philosophically. After all, a theorem is a theorem, and once it has been proved, that's it. But imagine a limitation on our ability to prove theorems. What if not all theorems are provable?

A.K. Dewdney (✉)
Department of Computer Science, University of Western Ontario, London
ON N6A 5B7, Canada
e-mail: akd@uwo.ca

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_34

When mathematicians encounter a statement they think might be true, they call it a conjecture, then attempt to prove it. Or they may try to disprove it by finding a counterexample. If the conjecture says that every object with property A also has property B, they may hunt for an example of an object with property A that does not have property B.

Conjectures are both famous and infamous in mathematics. Our inability to prove them may seem insurmountable until some bright young wizard finds the proof that had eluded everyone else—or finds a counter-example. Conjectures inspire much research and so play a valuable role in the development of mathematics. But should one of the conjectures currently before mathematics happen to be true yet have no proof, then a truckload of young wizards would not suffice to prove it. Yet it would be true.

Perhaps an unprovable theorem would involve concepts far beyond our ability to understand. (Who knows?) Or perhaps an unprovable theorem would involve obscure, grotesque, or utterly unappealing content. (Who cares?) Or perhaps an unprovable theorem might be both easy to understand and involve an interesting proposition. Perhaps the conjecture made by the German mathematician Christian Goldbach in 1742 is both true and unprovable. That would be something!

Goldbach's conjecture declares that every even number greater than 2 is the sum of two odd prime numbers. Going on 300 years, mathematicians have yet to prove the conjecture (which would automatically make it Goldbach's theorem) or to find a counterexample. Yet the conjecture seems to be true. Give me an even number such as 142 and it will not take me long to find two primes that have 142 as their sum. Let's see. How about 59 and 83?

What we used to call "Fermat's Last Theorem" was, until recently, actually unproven. But in 1998, a young Cambridge mathematician named Andrew Wiles proved what we may now call a "theorem:" An equation of the form,

$$x^n + y^n = z^n$$

has no solution for any value of $n > 2$.

Will Goldbach's conjecture go the way of Fermat's? Or does it lurk somewhere in a mathematical black hole? Do true but unprovable theorems exist? If such a thing could be proved, it would be a theorem, to be sure, and a metatheorem, at that. Such a theorem would have seemed incredible to the Greeks, as well as to the Indian and Arab mathematicians who followed them, no less to the Europeans up to the end of the nineteenth century. Yet this is exactly what the twentieth-century mathematician Kurt Gödel proved in 1930.

There is a back door to Gödel's theorem: either there exist unprovable theorems or the standard arithmetic is inconsistent. Gödel's theorem is about the "standard arithmetic," a term I will explain later, but which is merely a formalized portion of the mathematics we all learned in elementary school. In short, either there are theorems that our mathematics is simply not capable of dealing with, or our mathematics is itself inconsistent.

From the Greeks to the Europeans, to all the world's mathematicians in the year 1900, nothing would have been more disturbing than the idea of an unprovable theorem—unless it was an inconsistency within mathematics itself. The story of this amazing result begins in the year 1900. The setting is Paris, the Second International Congress of Mathematicians. It was an ideal time for one of the world's leading researchers to set the agenda for a new century. The German mathematician David Hilbert challenged his worldwide audience with twenty-three problems. The first two of these would have a profound influence on Kurt Gödel, whose birth lay six years in the future.

34.1 The Ghosts of Infinity

The first problem was to prove the continuum hypothesis, formulated some sixteen years earlier by another German mathematician, Georg Cantor. As a result of his groundbreaking conceptual invasion of infinity during the years 1871 to 1884, Cantor had formulated a new system of infinite numbers with a strange arithmetic all their own. The first of these was written $\aleph_0$ and called aleph nought. It was the cardinality (infinite, to be sure) of the natural numbers, or counting integers. If the members of any infinite set could be paired off with the numbers 1, 2, 3, … and so on forever, that set has $\aleph_0$ members. The second number, $\aleph_1$ (aleph one), stood for the cardinality of the set of real numbers. If the members of an infinite set could be paired off with the real numbers, that set would have $\aleph_1$ members.

The continuum hypothesis, as formulated by Cantor, stated that every subset of the real numbers either had cardinality $\aleph_1$ or had cardinality $\aleph_0$. There was nothing in between the two numbers. Every attempt to construct a set of real numbers that was not in one-to-one correspondence with either the integers or with the real numbers met with failure.

Thus was born the new field of transfinite arithmetic, with its first two numbers, $\aleph_0$ and $\aleph_1$. However, it was not until 1891 that Cantor was able to prove that the transfinite numbers $\aleph_0$ and $\aleph_1$ were different! To understand Cantor's proof of this result we need one or two mental tools.

The power set of a set is simply the set of all subsets of the set. For example, the power set of the finite set $A = \{1, 2, 3\}$ consists of eight sets, namely $\{1, 2, 3\}$ itself, as well as $\{1, 2\}$, $\{1, 3\}$, $\{2, 3\}$, $\{1\}$, $\{2\}$, $\{3\}$, and the empty set. We can write the power set of A as 2^A.

The new system of transfinite arithmetic was justified, in part, by Cantor's 1891 theorem, a humble but profound result: Specifically, he showed that no pairing or one-to-one correspondence could exist between the real numbers and the integers themselves. His method of proof involved a type of argument that was new to the mathematics of the day but that later would play a key role in Gödel's theorem. His argument used diagonalization, a process that singled out the main diagonal of entries in an infinite table. His proof is simple enough to be presented here.

Suppose that a one-to-one correspondence could be found between Z and its power set, 2^Z. It could then be written as a function f that, for every integer k, would produce a set f(k) of integers. As the variable k ran through the integers 1, 2, 3, and so on, the function f would run through the subsets of Z, all of them, sooner or later. Cantor then examined a very special set that consisted of all the integers z that were *not* members of their corresponding set f(z).

This was a peculiar thing to do. If we made a vast table with the integers down one side and the subsets of Z across the top, every entry of the table would consist of a pairing between an integer and a subset of Z. It would be natural, in such a table, to place the pairs z, f(z) down the main diagonal. That, at least, accounts for the name "diagonal argument."

Some of the integers z will appear inside their corresponding subsets f(z) and some won't. If we take the set of all integers z that are not members of f(z) and place them in a special set W, we can ask a very serious question about W: To what integer does W correspond under this scheme? If we write that integer as w, we can ask if w belongs to W. If w belonged to W, then w would not be a member of f(w) by the definition of W. But wait! This is a contradiction, since W = f(w). On the other hand, if w were not a member of W, then w must lie in f(w) (i.e., W), another contradiction.

What has gone wrong? We got into this mess by assuming that there was a function f with the stated property. The assumption must therefore be wrong. There is no one-to-one association between integers and all subsets of the integers. Hence there is no one-to-one correspondence between the integers and the real numbers. It immediately follows that $\aleph_0$ and $\aleph_1$ cannot have the same cardinality and that the distinction between them is real. The question that would immediately suggest itself did so to Cantor. Was there yet another transfinite number between $\aleph_0$ and $\aleph_1$? The "continuum hypothesis remains unresolved to this day.

In his 1900 address to the world mathematical community, Hilbert proposed that resolving the continuum hypothesis was problem one. He also wished to resolve a another question that had been simmering on back burners (and a few front ones) for a decade or more: Could arithmetic be axiomatized in such a way as to guarantee the exclusion of any and all inconsistencies?

34.2 Consistency

By the turn of the nineteenth century, mathematicians had become aware that deep questions attended the simplest-seeming subject: arithmetic. In particular, the attempt to axiomatize arithmetic had led, in some cases, to the recognition of logical anomalies.

For example, in 1888 the Italian mathematician Giuseppe Peano had introduced a set of five axioms that characterized the natural numbers (positive integers). They were nothing if not simple.

1. 1 is a natural number.
2. If a is a natural number, so is a + 1.
3. If a and b are natural numbers and a = b, then a + 1 = b + 1.
4. If a is a natural number, then a + 1 $\neq$ 1.

These statements will trouble few readers. From these axioms, with the aid of a fifth axiom that Peano would employ for the deductive process itself, all the properties of the natural numbers could be derived.

For example, from these axioms it was possible to prove the associative law that for all natural numbers a, b, and c, a + (b + c) = (a + b) + c

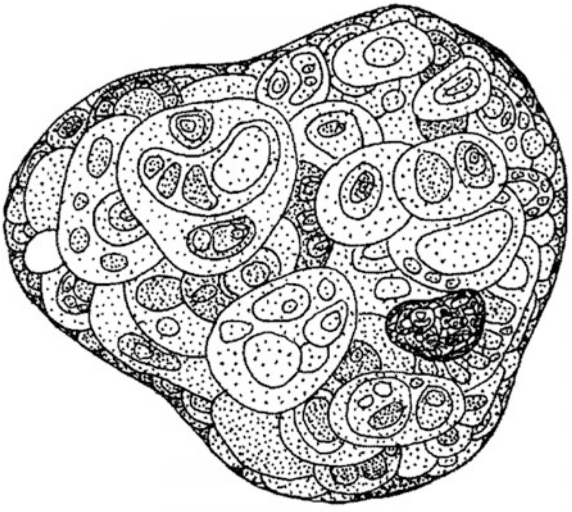
This law justifies something that people do all the time: It makes no difference in what order numbers are added. The fifth axiom would prove troublesome. By allowing a set to be arbitrary, it implicitly included a very nasty set that consisted of all sets whatsoever.

5. If A is a set and 1 lies in A, and if for every natural number a in A, a + 1 also lies in A, then all natural numbers lie in A.

Known as the principle of induction, this axiom was meant simply as a formal statement of a major tool for working with natural numbers. As long as mathematicians using the axiom of induction do not invoke the monstrous set of all sets, they are safe. This bizarre object seemed to spew out anomalies and downright contradictions. The "set of all sets" sounds almost like a spiritual entity.

In 1893, the German logician Gottlieb Frege published the first volume of his Grundgesetze der Arithmetik, a rigorous formulation of arithmetic that appealed to the same concept, the set of all sets. The English logician Bertrand Russell was at first unaware of the anomaly and championed the work of Frege, whose unnecessarily difficult book had attracted few readers. But as Frege's second volume was about to appear, Russell discovered to his horror that a severe logical flaw underlay the entire work.

The set of all sets contains some nasty items. My vain attempt to render the situation graphically appears here.



The set of all sets

It is surely no more difficult to conceive of sets that are not members of themselves than it is to conceive of sets that are. I will use my favorite nonsense word for such sets: “gzernmplatz.” I will go farther and define G to be the set of all gzernmplatz sets. Now I come, as Russell did, to the key question: Is G gzernmplatz? Even an amateur logician could work a proof, perhaps something like this: suppose G is gzernmplatz. Then G must lie within the set of all gzernmplatz sets, namely G. Thus G is a member of itself. Whoops! I guess G can’t be gzernmplatz after all. It follows that G is not in the set G and, therefore, not a member of itself. So G must be gzernmplatz. Whoops again! Has mathematics suddenly become inconsistent under my feet? Does a black hole await me?

The impact on Frege’s second volume was devastating, and the Grundgesetze was nearly forgotten. There was a growing awareness of difficulties in axiom systems, in the tendency for paradoxes to leap out from the shadows of the subject. It also had been extremely difficult to prove even the simplest of fields, such as arithmetic, consistent. No one knew if serious contradictions might someday appear in our reasoning about numbers.

The dream of a purely logical formulation of mathematics itself, became the passion not only of Hilbert but also of Bertrand Russell and his English colleague Alfred North Whitehead. In 1910 Russell and Whitehead published the first volume of *Principia Mathematica*, a strict formulation of mathematics in terms of a purely logical system that involved both propositions and predicates.

In the *Principia* Russell and Whitehead clearly demonstrated that not only could mathematics—or significant portions of it—be reduced to logic, but also that all mathematical truths were ultimately logical truths. Here was the vehicle that mathematicians could ride in search of consistency. Simply demonstrate that the axioms of the *Principia* could never lead to an inconsistency or contradiction.

In 1923, Hilbert rode forth to the lists with a new program of action that he called Beweisstheorie, or the theory of proofs. He published the proposal in a paper and followed up at many talks and conferences to promote the idea. But the current climate of logical uncertainly had rattled him. Speaking at a conference in June of that year, he declared paradoxes as “intolerable.”

In a nutshell, Hilbert proposed the reduction of mathematics to a symbolic script—marks on paper, as it were. A proof would amount to a sequence of formulas, each derivable by purely logical (and symbolic) operations from one or more of its predecessors, each leading inexorably to the final formula, a statement of the theorem being proved. Following in the footsteps of Russell and Whitehead, he cast his ideas in predicate logic. Hilbert’s system began with so-called atomic formulas, the simplest combinations of variables and constants, then proceeded to statements or well-formed formulas that were themselves composed of atomic formulas, logical connectives, and quantifiers of the type discussed above.

An example of a well-formed formula in mathematics would be

$$\forall x\, \exists y \text{ s. t. } (x < y) \,\&\, (x + 1 > y).$$

This rather compact notation may be interpreted as follows:

"For all x there exists a y such that x is less than y and x + 1 is greater than y." The symbols ∀ and ∃ are the same quantifiers we discussed earlier. Their role is to specify for each variable under their jurisdiction, whether the formula is to be true for all values or at least one of them, respectively. The & symbol is an example of a logical connective, and the expressions "$x < y$" and "$x + 1 > y$" are examples of atomic formulas. The whole point of predicate logic was that it amounted to a language in which axiom systems and theorems for various areas of mathematics could be expressed. This language became the focus of interest in the newly emerging field that Hilbert called "metamathematics." It would interest young Gödel profoundly.

34.3 The Troublemaker

Kurt Gödel was born in 1906, the second son of Rudolph and Marianne Gödel in Brno, a city in the Moravian part of Czechoslovakia. Although living in Czechoslovakia, the Gödels always considered themselves German. Rudolph Gödel was a successful businessman in the textile industry. In a privileged but highly structured household, young Kurt flourished academically. He graduated from High School in 1924 and was sent to study at the University of Vienna.

Gödel began in physics but found himself increasingly attracted to mathematics. While taking courses in physics, he would read the classic works of Euclid, Euler, and others. Gödel examined mathematics from a philosophical point of view, reading Kant and Russell as well. Indeed, his interest in mathematics was sparked even more strongly by attending a weekly seminar conducted by the philosopher Moritz Schlick. Sometime that year, Gödel met Hans Hahn, a prominent Viennese mathematician. Hahn had recently turned his attention to the foundations of mathematics. It may have been Hahn who suggested that Gödel attend another special weekly seminar, later to be called Der Wiener Kreis (the Vienna Circle). It was there that Gödel heard the crucial logical issues of the day discussed.

By 1927, Gödel was hopelessly involved in mathematical issues and questions. The Vienna Circle had brought him to the heart of difficult and important mathematical questions. He read and attended lectures. He walked the streets of Vienna alone or with colleagues. He sat for hours in various coffeehouses discussing mathematics. In 1928 Gödel had finished his undergraduate work and by 1929 was already hard at work on his Ph.D. thesis, an attempt to show that predicate logic was complete. In 1929 Hilbert and his colleague Willheim Ackermann had published their *Grundzuge der Theoretischen Logik*. In his thesis, Gödel succeeded in proving that the logical system suggested by Hilbert and Ackermann was complete. Gödel showed that every valid formula (true expression) was derivable within the system.

He began by reducing the problem of proving completeness to showing that each formula within Hilbert's system was either satisfiable or refutable. He established the latter result, in turn, by using induction.

It was an impressive performance from someone so young. But the result surprised no one. Everyone expected Hilbert's system to be complete.

34.4 The Trouble

The "habilitation" denotes a second hurdle that had to be leaped by all aspiring academics in Continental universities. It was not enough to write and defend a thesis. If one expected employment at an institution of higher learning, one had to publish something of note after the thesis. For his habilitation paper Gödel chose to work on Hilbert's second problem, that of showing the consistency of arithmetic. It would undoubtedly be more difficult than the proof that predicate logic was consistent and infinitely more difficult than the proof that propositional logic was consistent. That result had been achieved by Emil Post, a mathematician at New York's City College in 1921. The propositional calculus is the simplest form of logic, essentially a subject first codified by Aristotle in the fourth century B.C. The "propositions" are merely symbols such as a, b, and c, which stand for fixed statements that contain no variables. Two propositions could be conjoined logically by either the "and" operator or the "or" operator.

In their classic *Principia Mathematica,* Russell and Whitehead gave axioms for the propositional calculus. Instead of the "and" connective, however, they used implication, symbolized by the double arrow, $\Rightarrow$. The expression $a \Rightarrow b$, or "a implies b," is understood at the outset to be logically equivalent to $\sim a \vee b$.

1. $(p \vee p) \Rightarrow p$
2. $p \Rightarrow (p \vee p)$
3. $(p \vee q) \Rightarrow (q \vee p)$
4. $(p \Rightarrow q) \Rightarrow [(p \vee r) \Rightarrow (q \vee r)]$

Each axiom in this system seems either harebrained or mildly insane. Axiom 2, for example, says that if a proposition p is true, then either p is true or p is true. There is, in any event, very little to argue with in the axioms.

The consistency of propositional logic may be proved by supposing that it is possible to derive two contradictory statements, T and ~T, from the axioms. Suppose then that we have discovered a disastrous proposition T such that T and ~T are both true. We may use an additional tool of propositional logic, the principle of substitution, replacing a proposition symbol with T.

$$T \Rightarrow (\sim T \Rightarrow q).$$

Since T is true, we may use the rule of detachment to establish that $\sim T \Rightarrow q$ is also true. But ~T is also true, and we may use the rule of detachment to show that q is also true. But q can be any proposition whatever. The assumption of a contradictory pair of propositions had therefore led to the conclusion that all propositions are true, including the negations of the axioms themselves. It follows from this contradiction that no such pair of contradictory statements or propositions can be derived within propositional logic, and the foundations are secured. It is consistent.

This example also serves to illustrate an important distinction between two kinds of reasoning. You will notice that our treatment of propositional logic involved proofs inside the system and proofs outside the system. The derivation of the theorem p ⇒ (~p v q), like all theorems of propositional calculus, was proved within the system by applying the rules of substitution and detachment, but the proof of consistency was proved outside the system. Our reasoning was no less formal, but there was no way to express it as a sequence of propositions, each derived logically from previous propositions in the sequence. In short, there was no way to express the assumption about T. The distinction is fundamental to metamathematics. We will see it made again when we consider Gödel's amazing theorem.

It is still a long way from showing the consistency of propositional logic to that of predicate logic when applied to arithmetic. For one thing, propositional logic is silent on the subject of arithmetic. Numbers cannot be expressed within it. Gödel began by trying to prove that arithmetic, as expressed in the logical framework of the *Principia*, was consistent. In seeking to express the consistency of arithmetic, however, he discovered that he could express this consistency within arithmetic itself. Not only that, but this very expression led directly to a theorem that could not be derived within the logical system of Russell and Whitehead. After setting up the axioms of a predicate logic that embodied what is called the standard arithmetic, Gödel drew up a list of all the symbols used and assigned a special code number to each symbol, as shown in the following table:

Symbol	Code number	Symbol	Code number
0	1	x	9
s	2	l	10
+	3	˥	11
X	4	&	12
=	5	∃	13
(	6	∀	14
)	7	→	15
.	8		

The symbol that resembles the letter L rotated on its head stands for implication in the predicate calculus. Thus A ˥ B means simply that A implies B or that B can be deduced from A. The symbol s is the successor function. When applied to a natural number, it yields the next number in sequence. Because each formula would have to be expressed by a single, unique number, Gödel would need a way of boiling all the numbers that represented a given formula down into one number that would encode not only all the number symbols of the formula simultaneously, but also their order.

Gödel achieved this trick by using consecutive prime numbers such as

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, \ldots$$

each raised to a certain power. The power used would depend on the position of the symbol in the formula. If a symbol x appeared as the seventh symbol in a formula, for example, it would be represented by a 9 (the symbol for x) raised to the 17th power (the 7th prime in the sequence).

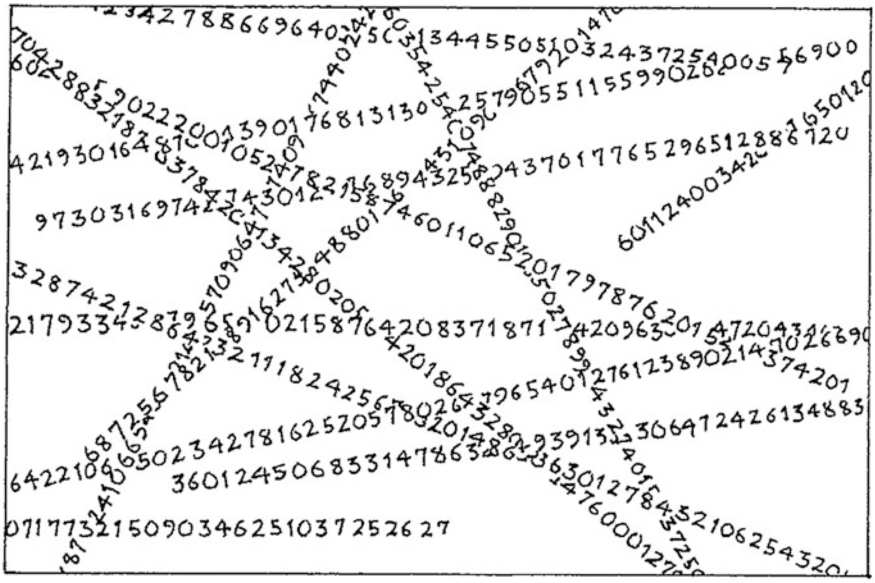
Thus if Gödel wanted to translate the formula (axiom) into a single number, he would first replace every symbol by its numerical code:

$$9, 10, 3, 2, 9, 10, 10, 5, 2, 6, 9, 10, 3, 9, 10, 10, 7.$$

In this case he would obtain the seventeen integers listed above. Next he would raise the first seventeen prime numbers to these powers and multiply them all together, the raised decimal points representing ordinary multiplication.

$$2^9.3^{10}.5^3.7^2.11^9.13^{10}.17^{10}.19^5.23^2.29^6.31^9.37^{10}.41^3.43^9.47^{10}.53^{10}.59^7$$

The resulting number is huge but finite. The encoding procedure could be specified as a finite process that would, given enough time, serve to express any formula whatever by an integer, no matter how large. It was recognized by the early metamathematicians that in explorations of the infinite, the best way to stay out of trouble was to use only finite tools. A proof in Gödel's system could, be encoded by a single integer (called its Gödel number). Conversely, given any positive integer whatever, there was a finite procedure for producing either a corresponding formula or sequence of formulas, complete nonsense, or nothing at all, depending on what symbols were encoded (or not) and in what order. Thanks to an important theorem within arithmetic itself, any positive integer can be written as the unique product of primes, which would represent the given number uniquely. The powers of the primes could then be translated directly into the corresponding symbols—unless they happened to be larger than 15. The fact that an arbitrary integer might not represent a formula was not a problem. But when it did, Gödel knew that no other integer could represent that formula. Gödel had developed a language that linked two seemingly different universes, the universe of formulas and the universe of Gödel numbers, a system of huge integers that each represented one of these expressions. The very process of thought, as embodied in the formulas, had been reduced to a cloud of numbers.



Thoughts as Gödel numbers

Among the formulas that Gödel showed how to construct were ones that would be typical of theorems and proofs. He had shown how formulas could express the process of checking that a proof really was a proof.

The Gödel numbers were all, of course, natural numbers, and as such formed part of the standard arithmetic. This meant that he could construct statements about the Gödel numbers, just as he could for ordinary integers. Within the standard arithmetic he could frame predicates that took the Gödel numbers as their subject matter. Of special interest is a rather complex predicate that I will write in a simple symbolic form:

Proof (x, y, z)

I have used the traditional variable names instead of x_1, x_{11}, and x_{111} to make the path into Gödel's mind a little easier to follow. The interpretation of the predicate called "proof" depends on knowing that X is a sequence of formulas that allegedly prove something, while x is the Gödel number of the proof X. Another formula, Y, has only one variable, and y is its Gödel number. With these elements in mind, the predicate can be described as saying, x is the Gödel number for the proof X of a one-variable formula Y with Gödel number y and with the integer z substituted into it.

In other words, Gödel number aside, X is a proof of Y, with z substituted into it. If the formula Y is true for this value of z, and if the system is complete in that every theorem can be proved in it, the proof X will exist, x being its Gödel number. In this case, of course, the formula Y will be the last formula in the string represented by X. Under these conditions the predicate called "Proof" will be true.

The expression “Proof(x, y, z)” does not belong to the system under study, but to the metalanguage in which truths about the system are expressed. Yet this shorthand, “Proof(x, y, z),” refers to a series of formulas that express what it means for X to be a proof of Y with z substituted into it, an encoding, if you like, of the mental machinery required to check such a proof. For example, the actual expression would have predicates that checked that each formula in the proof sequence was derivable from ones earlier in the same sequence. To spell out the proof-checking procedure, as long as it amounted to a finite process, was tantamount to an actual check. All that machinery within the logical system, along with the proof X of the formula Y, amounted to an extremely long, but finite, string of formulas. Consequently that string would itself be a formula and would have its own Gödel number.

The next and most important step that Gödel took was to realize that the formula Y referred to in the predicate Proof (x, y, z) could have its own Gödel number, y, substituted into it, instead of the more general variable Z- In other words, Gödel’s attention now focused on the predicate

Proof(x, y, y)

In this form the predicate is true if x is the Gödel number of a proof that the formula Y is true when its own Gödel number y is substituted into it. In other words, the predicate automatically symbolizes all formulas Y within the system that happen to have a proof X when y (the Gödel number of Y) is substituted into Y. This seems an odd thing to consider.

At this juncture, Gödel’s breath may have caught as he sensed himself on a collision course with common sense—or about to enter a black hole. What he did next was to form a new predicate that denied the existence of such a proof:

$\sim\exists x$ Proof(x, y, y)

This expression formed an element of the metalanguage that Gödel used to reason about the logical system under examination, the one that contained the standard arithmetic of integers. Yet by merely appending the same logical symbols to the extremely lengthy expression inside the logical system, the one represented by “Proof(x, y, y)” he now had a first-class anomaly on his hands.

The new expression denied the existence of a proof X that the formula Y would be true with its own Gödel number substituted into it. Yet this new predicate, considered as a lengthy expression within the system, would have yet another Gödel number all its own—say, g.

What, asked Gödel, is the status of the predicate $\sim\exists x$ Proof(x, g, g)? This expression asserts that there exists no proof of the predicate symbolized by g, namely the predicate $\sim\exists x$ Proof(x, y, y). If $\sim\exists x$ Proof(x, g, g) were true, then no proof could exist. If $\sim\exists x$ Proof(x, g, g) were false then the expression

$\exists x$ Proof(x, g, g)

would be true and a proof would exist. But a proof of what? It would be a proof of $\sim\exists x$ Proof(x, y, y), because that’s what g stood for. But a proof of this predicate would hold for all possible values of y including g, leading Gödel to the inevitable conclusion that $\sim\exists x$ Proof(x, g, g) was indeed true.

Here was the crunch. If the predicate $\exists x$ Proof(x, g, g) were both true and false, then his logical system—and therefore the standard arithmetic—was inconsistent.

The only way out was to assume that the statement symbolized by g was true. And that statement had no proof. At the time of his discovery of the famous incompleteness theorem, Gödel was twenty-six. Perhaps because he was not then an established mathematician, news of the result did not exactly spread like wildfire. At the same time, the colleagues and contacts to whom Gödel communicated his result most frequently expressed confusion rather than admiration. And yes, resistance grew—to a point.

34.5 The Aftermath

In August 1930, Gödel met his colleague Rudolph Carnap to plan a trip to a conference in Konigsberg (today Kaliningrad) in one month's time. Quietly announcing his result, Gödel was surprised to discover that Carnap did not quite follow the argument. But he was confident that the mathematical world would soon catch up with him. The conference in Konigsberg was attended by no less than John von Neumann, a world-renowned mathematician who had published in a great variety of areas and who took a particular interest in the new metamathematics, having wrestled with the consistency problem himself. After the talk, von Neumann congratulated Gödel and inquired further into the stunning new result. Once he fully understood it, von Neumann suffered a slight fit of pique. Yet his admiration was genuine, and the desire to promote Gödel's new result was sincere.

Hilbert, who also attended the conference, apparently had no idea of what was going on. "For the mathematician there is no ignoramibus, and, in my opinion, not at all for natural science, either - - - The true reason why [no one] has succeeded in finding an unsolvable problem is, in my opinion, that there is no unsolvable problem. In contrast to the foolish ignoramibus, our credo avers: We must know. We shall know."

Hilbert apparently was quite upset about the new theorem and probably somewhat depressed as well. After all, his credo that "we must know" had just been shattered by the knowledge that unless mathematics was inconsistent, there would be some things that we would never know, namely, which conjectures might turn out to be unprovable.

In March of 1931 a formal paper by Gödel appeared in the *Monatshejte fur Mathematik und Physik* titled "Uber Formal Unentscheldbare Satze Principia Mathematica und Verwandter System I" (On Certain Difficulties of Proof in the Principia Mathematica and Related Systems). The new paper slowly made its way into the collective consciousness of the mathematical world. The following September Gödel attended a meeting of the German Mathematical Union in Bad Elster. By now news of the result had made the rounds, and Gödel met his first real opposition in the person of Ernst Zarmelo, the mathematician who had first axiomatized set theory, a subject intimately related to logic. When colleagues proposed getting Zarmelo to lunch with Gödel, Zarmelo at first refused on a suspiciously wide variety of grounds: he didn't like Gödel's looks; he couldn't walk that far; if he attended the lunch, there wouldn't be enough food to go around. His subsequent hour with Gödel seemed outwardly pleasant, but Zarmelo held such different views of logic that he did not fully grasp the import of the iincompleteness theorem.

Other notables also found the result difficult to understand, including the mathematical philosopher Ludwig Wittgenstein and even Bertrand Russell, who expressed gratitude that he no longer worked in mathematical logic. By the mid-1930s there was barely a mathematician alive who did not know of Gödel's theorem and its philosophical implications for mathematics. Most preferred to ignore the other horn of the dilemma in Gödel's theorem, that mathematics was privately plagued by monsters of inconsistency.

Even assuming that mathematics was consistent, it would never be quite the same. The mere possibility of unprovable theorems has added a third potential outcome (or non-outcome) in the pursuit of conjectures. Someone will find a proof, someone will find a counterexample, or no one will find anything, thanks to the black hole.

Chapter 35
Gödel Incompleteness and the Empirical Sciences

N.C.A. da Costa and F.A. Doria

Abstract We show how widespread are metamathematical phenomena in mathematics and in the sciences which rely on mathematics. We will consider specific examples of undecidable sentences in mathematics, physics and economics. Our presentation is informal; rigorous developments can be found in the references.

35.1 Introduction

The usual version of a mathematical formalism requires it to be recursive, that is, computable—at least in principle—by a computer program. In particular, the axioms, logical and specific, have to be a recursive set,[1] and the rules of inference must also be recursive.

In a nutshell: one such formalism is essentially an algorithm or a Turing machine, that is, a program. It involves logic and specific axioms, being, so to say, an inference device; its proofs and syntactic procedures are all recursively decidable, in the sense that we are able to verify if they satisfy the conditions of their (in general) inductive definitions.

Gödel's fundamental idea was to view a (mathematical) formalism as part of arithmetic via the process of arithmetization, and then to take advantage of this move. His first incompleteness theorem says, in outline, that a consistent formalism containing enough elementary arithmetic is always *incomplete*: there are sentences ξ such that

---

[1] Given by a computer program and decidable.

Partially supported by CNPq, Philosophy Section; the authors are members of the Brazilian Academy of Philosophy.

N.C.A. da Costa (✉) · F.A. Doria
Fuzzy Sets Laboratory, Advanced Studies Research Group, Mathematical
Economics Group, HCTE, Production Engineering Program COPPE, UFRJ,
68507, Rio de Janeiro 21945-972, Brazil
e-mail: ncacosta@terra.com.br

F.A. Doria
e-mail: fadoria63@gmail.com

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_35

neither ξ nor $\neg\xi$ (the negation of ξ) are provable in the formalism. His second theorem, that can be proved as a corollary to the first, asserts that a consistent formalism, S, encompassing arithmetic, cannot prove a particular arithmetic sentence that constructively expresses the consistency of S.

In consequence, all classical, strong and consistent mathematical theories are incomplete. As a Gödel says,

> The human mind is incapable of formulating (or mechanizing) all its mathematical intuitions, i.e., if it has succeeded in formulating some of them, this very fact yields new intuitive knowledge, e.g., the consistency of this formalism. This fact may be called the "incompletability" of mathematics. On the other hand, on the basis of what has been proved so far, it remains possible that there may exist (and even be empirically discoverable) a theorem–proving machine which in fact is equivalent to mathematical intuition, but cannot be proved to be so, nor even be proved to yield only correct theorems of finitary number theory ([13], p. 324).

Anyhow, according to Gödel, we are able to discover new mathematical truths and intuitively establish them. Actually we are even able to get to know new reasonable axioms and to continue more or less systematically the search for truth. Gödel wrote that:

> Despite their remoteness from sense experience, we do have something like a perception of the objects of set theory, as is seen from the fact that the axioms force themselves upon us as being true. I dont see any reason why we should have less confidence in this kind of perception, i.e., in mathematical intuition, than in sense perception... The set theoretical paradoxes are hardly any more troublesome for mathematics than deception of the senses are for physics... Evidently, the 'given' underlying mathematics is closely related to the abstract elements contained in our empirical ideas. It by no means follows, however, that the data of this second kind, because they cannot be associated with actions of certain things upon our sense organs, are something purely subjective, as Kant asserted. Rather, they, too, may represent an aspect of objective reality, but, as opposed to the sensations, their presence in us may be due to another kind of relationship between ourselves and reality ([12], p. 272).

So, one could argue that, in the spirit of Gödels stance, it would be acceptable to extend the notion of mathematical formalism, eliminating some of its (constructive) restrictions.

35.1.1 Non-constructive Systems

There is a price one has to pay in order to have complete formal systems.

For example, we may add to the system non-constructive rules of inference that, if intuitively valid, might be acceptable. Let us consider some examples of the use of non-constructive rules of inference. If we add to Peano Arithmetic (*PA*) [16, 21] the ω–rule,

$$\frac{A(0), A(1), A(2), \ldots}{\forall x\, A(x)}$$

where $A(x)$ is a formula and x a variable, then the new formal system PA^* is syntactically complete: given a sentence ξ or its negation $\neg\xi$, one of them will be provable in PA^*.

This extended formalism has, as its theorems, all true sentences of PA, that is, sentences that are true of the standard model of PA and only them.

The two incompleteness theorems proved by Gödel do not apply to PA^*. Although PA^* has an intuitive appeal, its notion of proof is not decidable (proofs may be of infinite length). Since the first incompleteness theorem is not valid for PA^*, its concept of proof cannot be constructive (recursive, computable—written as a computer program). Notwithstanding, PA^* possesses an intuitive interpretation and deserves to be included in the class of (extended) mathematical formalisms.

Let us now present two examples of incompleteness in axiomatic theories, with a clear mathematical meaning. (We apologize for the brief mathematical sections, but we believe that the arguments required are quite simple and easily understood.)

35.2 Gödel Incompleteness: Kleene's Example

It is frequently said that the example given by Gödel of a sentence which is independent of the axioms of arithmetic is contrived, artificial—and without a clearcut mathematical meaning. However it is not so for many examples of undecidable sentences. Gödel published his milestone paper in 1931; in 1936 Stephen Kleene published one of the major papers of that period, "General recursive functions of natural numbers" [15]. In it he offers (among other developments) a simple proof of the Gödel phenomenon (incompleteness of the usual mathematical theories) which moreover has a clearcut, pedestrian, mathematical interpretation.

Kleene's argument goes as follows: suppose that S is a theory based on the classical first–order predicate calculus; we also require that S contains enough arithmetic (it must be able to 'talk' about the natural numbers, and their sums and products), and is such that its theorems can be listed by a computer program (the usual mathematical theories satisfy that last condition).

Start the program that lists all theorems of S. Suppose that f_n, all n, are functions (which can be computed by computer programs) defined on all natural numbers, with values on them. Then among the theorems of S look for those that assert, "f_n is total," that is to say, f_n is defined on all natural numbers. We thus get a list,

$$f_0, f_1, f_2, \ldots$$

of all *S–provably total recursive functions*[2] in the theory.

---

[2] Recursive: they can be computed by some computer program.

Make the following arrangement:

$f_0(0), f_0(1), f_0(2), f_0(3), \ldots$
$f_1(0), f_1(1), f_1(2), f_1(3), \ldots$
$f_2(0), f_2(1), f_2(2), f_2(3), \ldots$
$f_3(0), f_3(1), f_3(2), f_3(3), \ldots$
$\ldots$

Now the trick: define a function F:

$F(0) = f_0(0) + 1.$
$F(1) = f_1(1) + 1.$
$F(2) = f_2(2) + 1$
$\ldots$

Then F is different from f_0 at value 0, from f_1 at 1, from f_2 at 2, and so on.

We can now conclude our reasoning. Recall that the $f_0, f_1, f_2, \ldots$ functions are said to be *provably total* in our theory S, as they appear in the listing of the theory's theorems. However F cannot be provably total, since it differs at least once from the functions we have listed. Yet F is obviously computable, and given programs for the computation of $f_0, f_1, f_2, \ldots$ we can compute F too.

So the sentence "F is total" cannot be proved in our theory.

Also, if we suppose that the theory is *sound*, that is, if it doesn't prove false facts, then the sentence "F isn't total" cannot be proved too. Therefore "F is total" is an undecidable sentence within our theory.

Such is Kleene's 1936 argument.

35.2.1 Theme and Variations

The preceding discussion is almost completely uninformative about F: which are the specific properties, if any, of that function F? Is it—different, peculiar, in some sense?

Does it appear in a conventional mathematical environment?

Yes. We'll learn about it if we take an alternative approach in its construction. Let's split our argument into a series of steps:

- Recall that our theory S, supposed consistent, behaves as a computer program which is able to list all theorems of S—that is, S has a recursively enumerable set of theorems.
- Therefore each proof in S is coded by a natural number, its Gödel number.
- For another function F, and for an integer n, the value of $\mathsf{F}(n)$ is given as follows:

- Consider all $k \leq n$.
- For S–provably total f_e, consider all those whose Gödel number of the proofs of "f_e is total" is smaller or equal to n.
  (That is to say, the sentence "f_e is total" is a theorem of S whose Gödel number $\leq n$.)
- Consider the maximum of the values of those functions, that is the maximum value of $f_e(k)$, all $k \leq n$ and all e with bounded proofs as specified.
- Add 1 to that maximum.
- That value is the value of $\mathsf{F}(n)$.

Such a function is different (by construction) from each one of the S–provably total computable functions. And *it also tops all such functions*. F is an intuitively total computable function that dominates all the S–provably total computable functions.

F is a very fast growing function.[3] Actually it is easy to see that there are infinitely many functions like it—with similar properties when it comes to undecidability; F is a kind of minimal such function with respect to S. However S cannot 'see' that such functions are total and computable. Growth rate for such functions cannot be perceived in S. And this is just the beginning. Out of those functions we can concoct many strange but mathematically meaningful examples.

Yes. Rather weird results that follow out of the existence of functions like F are, for instance:

- There are infinitely many (actually a whole continuum–sized set of) Liouville transcendental numbers which cannot be proved so, or disproved, by the axioms of set theory (Zermelo–Fraenkel set theory with the axiom of choice, or ZFC).
  But a subset of that set of transcendental numbers may be even—intuitively—described by a computer program which however cannot be proved total in ZFC. So, we cannot prove that the Liouville numbers intuitively described by that program are actually described by it!
- There is an infinite set of (intuitively) poly machines—Turing machines whose operation time is bounded by a clock that stops in time bounded by a polynomial on the binary length of the input—which can neither be proved nor disproved, to be a set of exclusively poly machines. That set can even be explicitly described by a computer program.
  When we consider major open questions such as the P versus NP problem, which may require that we deal with a set of poly machines in S like the one just described, we see how catastrophic that particular example of incompleteness can be.

Function F sort of codes lots of information about axiom system S. For instance, S proves the following sentence:

$$[\mathsf{F} \text{ is total}] \rightarrow \text{Consis}(S).$$

---

[3] But it is recursive! Soon we'll meet nonrecursive examples of fast–growing functions.

(Implication is strict; the converse doesn't hold.) Here Consis(S) is the usual Gödel sentence that asserts the consistency of our axiom system S. So, once we add sentence "F is total" to our system S, we learn that the combined system:

$$S + \text{“}\mathsf{F}\text{ is total}''$$

proves the consistency of S, and as such is strictly stronger than S. (The combined system $S +$ "F is total″ is consistent once S is consistent.)

What we see here is the following: *Axiomatic theories like S have a definite bound on the rates of growth that they can recognize,* so to say.

Another point: since "F is total" is independent of the axioms of S, then the system:

$$S + \mathsf{F}\text{ isn't total}$$

is consistent. It is a Σ_1–unsound system, or 1–inconsistent system, with many counterintuitive properties. The interested reader is invited to search for them; he will discover an Alice in Wonderland thematic park in the very heart of classical mathematical logic…

Perhaps someday a gifted and imaginative researcher will guide us through the wilderness of a 1–inconsistent theory—like the so–called Alice–spacetimes, which are now being explored in general relativity.

(A good and colorful discussion appears in https://en.wikipedia.org/wiki/Wikipedia:Reference-desk/Archives/Mathematics/2009-June-1).

35.2.2 The Busy Beaver Function; The Counterexample Function to P = NP

Is there one such fast–growing function with a pedestrian meaning? With relevance to a mainstream question in mathematics? Well, the first example of this kind of mathematical monster in a quite innocent–looking environment is Radò's Busy Beaver function. Why a monster? It is noncomputable, it grows—absurdly, we may say—fast, and it dominates all total recursive functions, in an intuitive context. We refer the interested reader to Radò's 1962 paper [19].

The second example is equally interesting; it is the counterexample function to the $P = NP$ hypothesis. We may expand our presentation a bit more here.

We quote an anecdote [10] to motivate it. A long time ago we decided to take a nonconventional look at the *P versus NP* problem. It begins in a simple, naïve question:

> Mrs. H is a gentle and able lady who has long been the secretary of a large university department. Every semester Mrs. H. is confronted with the following problem: there are courses to be taught, professors to be distributed among different classes of students, large and small classes, and a shortage of classrooms. She fixes a minimum acceptable level of overlap among classes and students and sets down in a tentative way to get the best possible schedule

> given that minimum desired overlap. It's a tiresome task, and in most cases, when there are many new professors or when the dean changes the classroom allocation system, Mrs. H has to redo everything again; again she has to check nearly all conceivable schedulings before she is able to reach a conclusion. In despair she asks a professor whom she knows has a degree in math: "tell me, can't you find in your math a fast way of scheduling our classes with a minimum level of overlap among them?"

Roughly, Mrs. H's problem is a problem where:

- Once we have a solution, it is very easy to test it.
- But the known general algorithms to look for a solution are hard, as Mrs. H knows by experience.

We followed a suggestion given by a top–ranking logician[4]: we should look at the counterexample function to the $P = NP$ hypothesis. Our expert hinted that it might be a fast–growing function.

It is, indeed [4, 6].

$P = NP$ means, there is a time–polynomial Turing machine (a *poly machine*)[5] that settles all instances of a problem in the NP–class—a problem like the one described by Mrs. H, our gentle lady. Then the counterexample function is given as follows:

- List in a computable way all Turing machines. We only require that their Gödel numbering be monotonic on the length of their tables.
- Let f be the counterexample function. If n is the Gödel number of a poly machine, then $f(n)$ equals the first instance of the input to machine M_n which fails to be a solution to the problem we are considering, plus 1.
  If M_n settles all instances, $f(n)$ is undefined, and we have that $P = NP$.
- If M_n isn't a poly machine, then $f(n) = 0$.
- f is total—defined for all n—if and only if $P \neq NP$.

The references give a proof that f tops in its peaks all total recursive functions. f is highly oscillating; but it overtakes all recursive fast–growing functions. Therefore it cannot be a computable function. The main trick in the proof is to show that, given an arbitrary total recursive function, it can always be embedded in f via a primitive recursive map. Therefore, if we suppose that there is a computable bound to f, we can always show (via that embedding) that there are peaks in f that overtake the supposed bound.

Suppose that we want to check whether our theory S proves that f is total. Pick up the function F which cannot be proved total and recursive by the proof machinery in S. F can be embedded in f, and we can restrict f to the domain of F in the counterexample function. If S proves f total, the restricted function is total over its restricted domain.

But as S cannot prove that F is total (it is recursive by construction), we reach a contradiction.[6]

---

[4] Georg Kreisel, in private.

[5] A Turing machine which is total and whose operation time is bounded by a polynomial clock on the length of its binary input.

[6] This is an informal argument; we must be careful with several technical details in our construction.

Now let us go back to Gödel's incompleteness theorem and to examples of incompleteness in arithmetic. We will now relate undecidability and incompleteness.

35.3 Undecidability—and Incompleteness, Its Sibling: Post's Argument

Recall that a question is *undecidable* whenever there is no algorithm to settle it. And a theory S is *incomplete* if there is a sentence ξ that is neither proved nor disproved from the axioms of S. ξ and its negation $\neg\xi$ are said to be *undecidable sentences* within S.

Undecidability (non–existence of algorithms for a given task) and incompleteness (existence of undecidable sentences) are related, as we can see from the next discussion.

But before we present that discussion, let's ponder a short tale ([2], p. 23):

> In the 1950s computers were mainly built by electrical engineers, and once upon a time a team of engineers who were working on one of the brand–new "electronic brains" of the period met some colleagues from the math department at some university's cafeteria. The engineers began to discuss their work with their brand new computer and said that they were trying to develop a kind of test program that would avoid bugs: the test program would know about the program they were running in the machine and out of that knowledge it would test beforehand any input to see whether it resulted in an infinite loop — a never ending succession of operations without any output — or not.
>
> So far they hadn't been successful.
>
> Then one of the math department people chuckled and started to laugh. What are you laughing about? complained the engineers. The math guy's answer: what you are trying to do is impossible; Turing proved it in 1936, two decades ago.
>
> You are trying to write a program that solves the halting problem, and that cannot be done.

OK. But why can't it be done?

Here is an argument that shows why such a program cannot exist:

- List all Turing machines $M_0, M_1, M_2, \ldots$.
- Suppose that there is a program $g(x,y)$ that executes the following tasks:
  - $g(x,y) = 1$ if and only if machine of program y stops over input x and gives some output.
  - $g(x,y) = 0$ if and only if machine of program y enters an infinite loop when it receives input x.

- If g is a program that can be constructed then h which we now describe can also be constructed:
  - $h(x) = 1$ if and only if $g(x,x) = 0$.
  - $h(x)$ diverges if and only if $g(x,x) = 1$.
    (In order to make h enter an infinite loop we can plug a diverging subroutine to it at convenient places.)
- At this point let's go back to our listing of all Turing machines

$$M_0, M_1, M_2, \ldots$$

  If h is a program, it can be implemented as a Turing machine, and there is a k so that $M_k = h$.
- Then:
  - If $h(k) = M_k(k) = 1$ we get that $g(k,k) = 0$, from the definition of h. Now, from the definition of g, $M_k(k)$ must diverge. A contradiction.
  - If $h(k) = M_k(k)$ diverges, then $g(k,k) = 1$, which means that $M_k(k)$ converges. Another contradiction.
- Therefore we can neither write a program like g nor one like h.

Now, the related incompleteness theorem, in E. Post's version.

It's a quite brief argument. Suppose that we can formalize in our theory S (with enough arithmetic) the sentence "Turing machine M_k over input m diverges." (We can do it.)

Suppose that there is a proof of all such sentences in S, any k, m.

To prove is to compute. If there is one such proof procedure, then we can make it into a calculation procedure. Thus we would settle all non–halting instances of the halting problem for k, m.

That's impossible. Therefore there are k_0, m_0 so that while it is true that machine k_0 over m_0 diverges, we cannot prove within S that in fact machine k_0 over input m_0 diverges. Also, as we've supposed that S doesn't prove false assertions, it cannot prove the negation of the sentence:

Turing machine M_{k_0} over input m_0 diverges.

Thus that sentence is undecidable within S, as well as its negation. And it is an undecidable sentence with a clearcut mathematical meaning.

This proof of Gödel's first incompleteness theorem is due to Emil Post [18]. It shows the interdependence between undecidability (the unsolvability of the halting problem, in this case) and incompleteness (the existence of sentences which one can neither prove nor disprove within formal systems like S).

35.4 From Logic to Physics

So far we have established that there are mathematical problems which naturally (if we may say so) arise within logic or computer science (the theory of Turing machines) and which are undecidable, or are related to undecidable sentences.

How about physics? How about classical mechanics, which is a well–established domain? Can we find undecidable questions in the realm of classical mechanics? There are quite simple noncomputable expressions for functions within mathematics; those expressions creep up even within languages close to arithmetic, and lead to very simple questions such as, "does the integer–valued function $\theta(n)$ equal 0 or 1?" which turn out to be undecidable in the general case. Analogous naïve–looking but intractable expressions for functions can also be found within more elaborate languages, as classical elementary analysis. With their help we can generate infinitely many undecidable sentences with a trivial appearance from arithmetic on and all the way up to the whole of mathematics.

Some of those intractable expressions represent the *halting function* $\theta(m, n)$, that tells us whether the Turing machine $M_m(n)$ stops over its input n. Once we have an expression for the halting function, we can obtain explicit expressions for all complete arithmetic degrees and even beyond. Therefore the associated undecidable predicates represent problems in all the corresponding degrees of unsolvability, both inside and outside the arithmetic hierarchy.

The Halting Function

We have seen that there is no computer program that can settle the halting problem, that is, we cannot devise a program that will tell us, for arbitrary programs, where they will fail to give some output (that is, where they will enter an infinite loop). However we can write down an explicit expression for a function that solves the halting problem (of course no computer program, that is, no mechanical procedure can calculate its values).

We'll exhibit its looks now.

We will need some notation. Let σ be the sign function, $\sigma(\pm x) = \pm 1$ and $\sigma(0) = 0$. The halting function $\theta(n, q)$ is explicitly given by[7]:

$$\theta(n, q) = \sigma(G_{n,q}),$$

$$G_{n,q} = \int_{-\infty}^{+\infty} C_{n,q}(x) e^{-x^2} dx,$$

---

[7] Among other expressions.

Let's explain the notation. First, consider θ:

- $\theta(n, q) = 1$ if and only if Turing machine of code (program) n stops over input q.
- $\theta(n, q) = 0$ if and only if Turing machine of code n enters an infinite loop over input q and never stops.

$C_{n,q}$ is obtained from the Feynmann–Richardson transform of polynomial $p_{n,q}$, which is the two–parameter universal Diophantine polynomial. $C_{n,q}(x)$ is different from zero over the reals if and only if machine of program n stops over q. Otherwise it is zero.

35.4.1 Chaos Is Undecidable

We used those undecidable predicates and noncomputable functions to settle several open decision problems, mainly in dynamical systems theory. One such problems dealt with a frequently handled question:

- *Is there a decision procedure for chaos?* Chaos theory has been a fast–growing research area since the early 70's, a decade after the discovery of (an apparent, but still unproved) chaotic behavior in a deterministic nonlinear dynamical system by E. Lorenz (for references see [8]). Chaos scientists usually proceed in one of two ways: whenever they wish to know if a given physical process is chaotic the usual starting point is to write down the equations that describe the process and out of them to check whether the process satisfies some of the established mathematical criteria for chaos and randomness.

However those equations are in most cases intractable nonlinear differential equations as they cannot in general be given explicit analytical solutions. Therefore, chaos theorists turn to computer simulations and for most nonlinear systems one sees a confusing, tangled pattern of trajectories on the screen. The system *looks* random, and there are statistical tests such as the Grassberger–Proccacia criterion that guarantee the existence of randomness in computer–simulated systems, modulo some error. Yet statistical tests furnish no *mathematical* proof of the existence of chaos in a dynamical system. There is always the chance that the system is undergoing a very long and complicated transient state, before it settles down to a nice and regular behavior. Therefore how can we prove that a dynamical system that *looks* chaotic is, in fact, chaotic?

This problem had been around since the discovery and early exploration of what is now called "deterministic chaos." In a 1983 conference (published in 1985) Morris Hirsch stated that time was ripe for a marriage between the "experimental" and "theoretical" sides of chaos research and posed the decision problem for chaotic systems [14].

Now, in a nutshell: let X be a dynamical system that describes a free particle; let Y be a (proved) chaotic system.[8] Then Z_n:

---

[8] There are infinitely many systems with a proved chaotic behavior.

$$Z_n = \theta(n)X + (1 - \theta(n))Y$$

is a family of dynamical systems which alternate between a free particle or a chaotic system, but such that we cannot algorithmically decide which is the case!
We can also concoct an example of Gödel incompleteness out of that one; we'll soon consider it.

The preceding discussion allows us to formulate a very general undecidability and incompleteness theorem for any mathematical theory like our S. Informally, let P be a property of the objects in S: if P isn't trivial (it applies to all objects in its domain, or to no object), then:

> P is undecidable — there is no algorithm to test for it — and entails incompleteness, that is, there is a sentence, for a x that we can explicitly construct, so that neither $P(x)$ nor $\neg P(x)$ can be proved in S.

This is our general Rice–like theorem.

35.4.2 Gödel Incompleteness in Physics

In order to construct an example of Gödel incompleteness in a physical theory, we had a blueprint to guide us: Post's example, which has just been sketched. However there was a major stumbling block in our way, as incompleteness requires an axiomatic framework, and the elaboration of such a framework is the challenge in Hilbert's 6th Problem. So, the construction of an example of incompleteness in physics through Post's technique involved a partial solution to a major problem.

Luckily for us, the domains of physics which interested us—classical mechanics, Schrödinger's quantum mechanics, some portions of quantum field theory without Feynmann's integrals—could be described with adequate mathematical rigor, and fit nicely within axiomatic set theory plus the axiom of choice. We then used the so–called Suppes predicates to present our axiomatic theory.

A Suppes predicate is the conjunction of two pieces (let's leave it vague). One piece describes the objects we require: manifolds, tensors, etc., for general relativity; the electromagnetic field, for electromagnetic theory, and so on. The second piece are the axioms that "physically" constrain the objects described in the first part; the second piece also includes the dynamics (or motion equations) for those objects—Einstein's equations for general relativity, Lagrangians submitted to a variational principle, in the case of ordinary classical mechanics, or d'Alembert's principle, for more general theories. The construction of a Suppes predicate can be quite cumbersome, but for our purposes it was enough to know that an axiomatic framework existed.

We then applied Post's trick and got an undecidable sentence in dynamical systems theory, which was seen as a portion of d'Alembert's mechanics. We obtained, for example, the following undecidable sentence:

> Z is a chaotic system.

Z was explicitly constructed; neither that sentence nor its negation could be proved from the axioms of our theory (supposed consistent). That sentence had a clear 'concrete' (please allow our abuse of language) meaning, and due to its construction, it is undecidable within our theory [3]. (For details see the reference.)

35.5 Mathematical Ecology; Economics

The chief explorers of undecidability and incompleteness in the social sciences have been Alain Lewis and Vela Velupillai. Lewis [17] pointed out that results such as our general incompleteness theorem (our Rice–like theorem) entail the incompleteness of the theory of Hamiltonian models in economics. Such results also entail the incompleteness of the theory of Arrow–Debreu equilibria and (what may at first look surprising) the incompleteness of the theory of *finite* games with Nash equilibria.

(Assuredly when a finite game is given through an explicit table of outcomes, it is always decidable. However for weak extensions of formalized arithmetic we can exhibit a description for the table of outcomes so that undecidability and incompleteness follows.)

The main result is:

> There is no general algorithm to test whether a given set of prices is an equilibrium set of prices for a competitive Arrow–Debreu market.

We can add similar undecidability (and of course incompleteness) results to population dynamics: an example of an undecidable question in that area is, can we check whether a given population[9] will bifurcate into two separate populations. We can actually concoct similar results everywhere in theoretical domains which are described by similar sets of equations.

In fact, we can assert: undecidability and incompleteness do matter in the formalized sciences; they matter a lot. Metamathematical phenomena are commonplace stuff in the sciences which have calculus and extensions as their main language.

We must learn how to deal with it.

Acknowledgements This paper was supported in part by CNPq, Philosophy Section. It is part of the research efforts of the Advanced Studies Group, Production Engineering Program, at COPPE–UFRJ and of the Logic Group, HCTE–UFRJ. We thank Profs. A. V. Assumpção, R. Bartholo, C. A. Cosenza, S. Fuks (in memoriam), S. Jurkiewicz, R. Kubrusly, M. Gomes, and F. Zamberlan for support.

---

[9] Described by a reaction–diffusion system of equations.

Appendix

We have added at this point a rather technical appendix where we give rigorous statements of the theorems that either bear on, or which imply, the results described above. It is to be noted that here we even go beyond the arithmetic hierarchy [20] and show how nasty (or how interesting, it depends on one's viewpoint…) these mathematical phenomena can be.

A First General Undecidability and Incompleteness Theorem

Let S be our formal theory and let L_S be its underlying formal language. We state here (without proof) our main undecidability and incompleteness theorems.

Definition A.1 A predicate P in L_S is **nontrivial** if there are term expressions ξ, ζ in S such that $S \vdash P(\xi)$ and $S \vdash \neg P(\zeta)$. $\square$

If $\xi \in L_S$ is any expression in that language (adequately extended), we write $\|\xi\|$ for its complexity, as measured by the number of letters from S's alphabet in ξ. Also we define (according to [11]) the *complexity of a proof* $C_S(\xi)$ of ξ in L_S to be the minimum length that a deduction of ξ from the axioms of S can have, subject to the conditions in the Ehrenfeucht–Mycielski paper.

Let P be any nontrivial predicate, and let $\mathcal{B}$ be an algebra of functions used in our theory; $\mathcal{F}$ is a subalgebra.

Then:

Proposition A.2 *Within S:*

1. *There is an expression* $\xi \in \lceil \mathcal{B} \rceil$ *so that* $S \nvdash \neg P(\xi)$ *and* $S \nvdash P(\xi)$*, while there is a model for our theory with standard arithmetic such that* $\mathbf{M} \models P(\xi)$.
2. *There is a denumerable set of expressions for functions* $\xi_m(x) \in \lceil \mathcal{B} \rceil$, $m \in \omega$*, such that there is no general decision procedure to ascertain, for an arbitrary* m*, whether* $P(\xi_m)$ *or* $\neg P(\xi_m)$ *is provable in S.*
3. *Given an arbitrary total recursive function* $g : \omega \to \omega$*, there is an infinite number of values for* m *so that* $C_S(P(\xi_m)) > g(\|P(\xi_m)\|)$. $\square$

That result was our first general incompleteness theorem [8]; it can be derived from Rice's theorem in computer science [20], which is an equally general result, but the proof we first gave for Proposition A.2 is weaker than Rice's theorem, since it only leads to unsolvable problems of Turing–degree not higher than $\mathbf{0}'$. However:

Proposition A.3 *We can explicitly and algorithmically construct in* L_S *an expression for the characteristic function of a subset of* ω *of degree* $\mathbf{0}''$. $\square$

That expression depends on recursive functions defined on ω and on elementary real–defined and real–valued functions plus the absolute value function and an integration, as in the case of the θ function (the Halting Function).

We simply use Theorem 9-II in [20] (p. 132). Actually the degree of the set described by the characteristic function whose expression we are going to obtain will depend on the fixed oracle set A; so, our construction is a general one.

Let $A \subset \omega$ be a fixed infinite subset of the integers. An oracle Turing machine ϕ_x^A with oracle A can be visualized as a two–tape machine where tape 1 is the usual computation tape, while tape 2 contains a listing of A. When the machine enters the oracle state s_0, it searches tape 2 for an answer to a question of the form "does $w \in A$?" Only finitely many such questions are asked during a converging computation; we can separate the positive and negative answers into two disjoint finite sets $D_u(A)$ and $D_v^*(A)$ with (respectively) the positive and negative answers for those questions; notice that $D_u \subset A$, while $D_v^* \subset \omega - A$. We can view those sets as ordered k– and k^*–ples; u and v are recursive codings for them [20]. The $D_u(A)$ and $D_v^*(A)$ sets can be coded as follows: only finitely many elements of A are queried during an actual converging computation with input y; if k' is the highest integer queried during one such computation, and if $d_A \subset c_A$ is an initial segment of the characteristic function c_A, we take as a standby for D and D^* the initial segment d_A where the length $l(d_A) = k' + 1$.

We can effectively list all oracle machines with respect to a fixed A so that given a particular machine, we can compute its index (or Gödel number) x, and given x we can recover the corresponding machine.

Now let us write $p(n, q, x_1, \dots, x_n)$ for a 2–parameter universal Diophantine polynomial. We can define the jump of A as follows:

$$A' = \{\rho(z) : \exists x_1, \dots, x_n \in \omega\, p(\rho(z), \langle z, d_{z,A}\rangle, x_1, \dots, x_n) = 0\}.$$

(ρ is an adequate recursive 1–1 function.) With the help of the λ map [3], we can now form a function modelled after the θ function; it is the desired characteristic function:

$$c_{\emptyset''}(x) = \theta(\rho(x), \langle x, d_{x,\emptyset'}\rangle).$$

(Actually we have proved more; we have obtained

$$c_{A'}(x) = \theta(\rho(x), \langle x, d_{x,A}\rangle),$$

with reference to an arbitrary $A \subset \omega$.)

We write $\theta^{(2)}(x) = c_{\emptyset''}(x)$.

Let $\mathbf{0}^{(n)}$ be the n–th complete Turing degree in the arithmetical hierarchy.

Corollary A.4 (Complete Degrees.) *For all $p \in \omega$, expressions $\theta^{(p)}(m)$ can be explicitly constructed for characteristic functions in the complete degrees $\mathbf{0}^{(p)}$.* ☐

General Incompleteness Theorems

Therefore,

Proposition A.5 *For every $n \in \omega$ there is a sentence ξ in S such that a model with standard arithmetic $\mathbf{M} \models \xi$ while for no $k \leq n$ there is a Σ_k sentence in N demonstrably equivalent to ξ.* □

Let $m_0(\emptyset^{(m)}) = \langle \rho(z), \langle z, d_{y,\emptyset^{(m)}} \rangle \rangle$ (use of the pairing function τ is supposed) such that

$$\mathbf{M} \models \forall x_1, \ldots, x_n [p(m_0, x_1, \ldots, x_n)]^2 > 0,$$

for an universal polynomial p.

Let $q(m_0(\emptyset^{(m)}), x_1, \ldots) = p(m_0(\emptyset^{(m)}), x_1, \ldots))^2$ be as after Proposition A.5. Then:

Corollary A.6 *Within T, for:*

$$\beta^{(m+1)} = \sigma(G(m_0(\emptyset^{(n)})),$$

$$G(m_0(\emptyset^{(n)})) = \int_{-\infty}^{+\infty} C(m_0(\emptyset^{(n)}), x) e^{-x^2} dx,$$

$$C(m_0(\emptyset^{(n)}), x) = \lambda q(m_0(\emptyset^{(n)}), x_1, \ldots, x_r),$$

$\mathbf{M} \models \beta^{(m+1)} = 0$ *but for all* $n \leq m+1$, $S^{(n)} \nvdash \beta^{(m+1)} = 0$ *and* $S^{(n)} \nvdash \neg(\beta^{(m+1)} = 0)$. □

Then,

Corollary A.7 *If L_S contains expressions for the $\theta^{(m)}$ functions as given in Corollary A.4, then for any nontrivial predicate P in N there is a $\zeta \in L_S$ such that the assertion $P(\zeta)$ is S–demonstrably equivalent to and S–arithmetically expressible as a Π_{m+1} assertion, but not as any assertion with a lower rank in the arithmetic hierarchy.* □

An extension of the preceding result is:

Corollary A.8 *For any nontrivial property P there is a $\zeta \in L_S$ such that the assertion $P(\zeta)$ is arithmetically expressible and for a model with standard arithmetic $\mathbf{M} \models P(\zeta)$, but it is only demonstrably equivalent to a Π_{n+1} assertion and not to a lower one in the hierarchy.* □

Definition A.9 $\emptyset^{(\omega)} = \{\langle x, y \rangle : x \in \emptyset^{(y)}\}$, for $x, y \in \omega$. □

Then:

Definition A.10 $\theta^{(\omega)}(m) = c_{\emptyset^{(\omega)}}(m)$, where $c_{\emptyset^{(\omega)}}(m)$ is obtained as in Proposition A.3. □

We can assuredly obtain arithmetic expressions for the characteristic functions in those higher degrees. However we are here especially interested in the analytic expressions for those functions.

Still,

Definition A.11 $\emptyset^{(\omega+1)} = (\emptyset^{(\omega)})'$. ☐

Corollary A.12 $\mathbf{0}^{(\omega+1)}$ *is the degree of* $\emptyset^{(\omega+1)}$. ☐

Corollary A.13 $\theta^{(\omega+1)}(m)$ *is the characteristic function of a nonarithmetic subset of* ω *of degree* $\mathbf{0}^{(\omega+1)}$. ☐

In the next results, **M** is again a model with standard arithmetic for S:

Corollary A.14 *Within S:*

$$\beta^{(\omega+1)} = \sigma(G(m_0(\emptyset^{(\omega)})),$$

$$G(m_0(\emptyset^{(\omega)})) = \int_{-\infty}^{+\infty} C(m_0(\emptyset^{(\omega)}), x)e^{-x^2} dx,$$

$$C(m_0(\emptyset^{(\omega)}), x) = \lambda q(m_0(\emptyset^{(\omega)}), x_1, \dots, x_r),$$

$\mathbf{M} \models \beta^{(\omega+1)} = 0$ *but* $S \nvdash \beta^{(\omega+1)} = 0$ *and* $S \nvdash \neg(\beta^{(\omega+1)} = 0)^{(\omega+1)} = 0)$. ☐

Details are found in [3]. Let's go beyond it:

Definition A.15 $\emptyset^{(\omega+1)} = (\emptyset^{(\omega)})'$. ☐

Corollary A.16 $\mathbf{0}^{(\omega+1)}$ *is the degree of* $\emptyset^{(\omega+1)}$. ☐

Corollary A.17 $\theta^{(\omega+1)}(m)$ *is the characteristic function of a nonarithmetic subset of* ω *of degree* $\mathbf{0}^{(\omega+1)}$. ☐

In the next results, **M** is a model (with standard arithmetic) for S:

Corollary A.18 *Within S:*

$$\beta^{(\omega+1)} = \sigma(G(m_0(\emptyset^{(\omega)})),$$

$$G(m_0(\emptyset^{(\omega)})) = \int_{-\infty}^{+\infty} C(m_0(\emptyset^{(\omega)}), x)e^{-x^2} dx,$$

$$C(m_0(\emptyset^{(\omega)}), x) = \lambda q(m_0(\emptyset^{(\omega)}), x_1, \dots, x_r),$$

$\mathbf{M} \models \beta^{(\omega+1)} = 0$ *but* $S \nvdash \beta^{(\omega+1)} = 0$ *and* $S \nvdash \neg(\beta^{(\omega+1)} = 0)$. ☐

Proposition A.19 (Nonarithmetic intractability.) *Given any nontrivial* P *such that for different* ξ, χ, $S \vdash P(\xi)$ *and* $S \vdash \neg P(\chi)$*:*

1. *There is a family of expressions* $\zeta_m \in L_S$ *such that there is no general algorithm to check, for every* $m \in \omega$*, whether or not* $P(\zeta_m)$*.*
2. *There is an expression* $\zeta \in L_S$ *such that* $\mathbf{M} \models P(\zeta)$ *while* $S \nvdash P(\zeta)$ *and* $S \nvdash \neg P(\zeta)$*.*
3. *Neither* ζ_m *nor* ζ *are arithmetically expressible.* ☐

References

1. Ash, C. J., & Knight, J. (2000). *Computable structures and the hyperarithmetical hierarchy*. Elsevier.
2. Chaitin, G., da Costa, N. C. A., & Doria, F. A. (2011). *Gödel's Way*, CRC Press.
3. da Costa, N. C. A., & Doria, F. A. (2007). Janus–faced physics: On Hilbert's 6th Problem. In Calude, C. (Ed.) *Randomness and complexity: From Leibniz to Chaitin*. World Scientific.
4. da Costa, N. C. A., & Doria, F. A. (2016). On the O'Donnell algorithm for *NP*–complete problems. To appear in the *Review of Behavioral Economics*.
5. da Costa, N. C. A. (1974). α-models and the systems T and T^*. *Notre Dame Journal of Formal Logic*, *15*, 443.
6. da Costa, N. C. A., Doria, F. A., & Bir, E. (2007). On the metamathematics of the *Pvs.NP* question. *Applied Mathematics and Computation*, *189*, 1223–1240.
7. da Costa, N. C. A. (2011). Gödel's incompleteness theorems and physics. *Principia*, *15*, 453.
8. da Costa, N. C. A., & Doria, F. A. (1991). *International Journal of Theoretical Physics*, *30*, 1041.
9. Davis, M. (1982). *Computability and Unsolvability*. Dover.
10. Doria, F. A. (2016). A monster lurks in the belly of computer science. preprint.
11. Ehrenfeucht, A., & Mycielski, J. (1971). *Bulletin of the AMS*, *77*, 366.
12. Gödel, K. (1964). What is Cantor's continuum problem? In P. Benacerraf & H. Putnam (Eds.) *Philosophy of Mathematics*. Prentice–Hall.
13. Gödel, K. (1974). J. W. Gibbs Lecture," (1954), quoted by H. Wang. In *From Mathematics to Philosophy*. Humanities Press.
14. Hirsch, M. (1985). The chaos of dynamical systems. In P. Fischer & W. R. Smith (Eds) *Chaos, Fractals and Dynamics*, M. Dekker (1985).
15. Kleene, S. C. (1935/6). General recursive functions of natural numbers. In *Math. Annalen112*, 727–742.
16. Kleene, S. C. (1952). *Introduction to Metamathematics*. Van Nostrand.
17. Lewis, A. A., & Inagaki, Y. (1991). *On the effective content of theories*. School of Social Sciences: U. of California at Irvine. preprint.
18. Post, E. (1944). Recursively enumerable sets of positive integers and their decision problems. *Bulletin of the American Mathematical Society*, *50*, 284–316.
19. Radò, T. (1962). On non-computable functions. *Bell System Technical Journal*, *41*, 887.
20. Rogers, H. Jr., (1967). *Theory of recursive functions and effective computability*. McGraw–Hill.
21. Shoenfield, J. (1967). *Mathematical Logic*. Addison–Wesley.
22. Suppes, P. (1988). *Scientific Structures and their Representation*, preliminary version. Stanford University.
23. Tsuji, M., da Costa, N. C. A., & Doria, F. A. (1998). The incompleteness of theories of games. *Journal of Philosophical Logic*, *27*, 553.

Chapter 36
Gödel's Ontological Dreams

Gary Mar

> *… There are more things in heaven and earth, Horatio, than are dreamt of in your philosophy.*
> —*Hamlet*, Act I, Scene V, 167.

Keywords Gödel's second incompleteness theorem • Löb's axiom • Time travel in Einstein's general theory of relativity • Gödel's ontological argument • Brouwersche axiom • Modal logic • Tense logic

Gödel, arguably the greatest logician of the 20th century,[1] dreamed of establishing philosophical theses and ontological results with the rigour and precision of mathematics. His dream was to a remarkable extent fulfilled. Despite their technical sophistication, Gödel's *logical* theorems such as the Completeness Theorem [22] and his Incompleteness Theorems [23] have perennially managed to escape mere mathematics and shed light on larger philosophical issues. Gödel's *philosophical* arguments for the unreality of time in the Theory of Relativity [1949–1952] and in his *Ontologischer Beweis* [*1970] have modal logical structures that mirror his mathematical reasoning. Gödel's success has often been attributed to his *philosophy of mathematics*, but his success is as much a tribute to his "*mathematics of philosophy*", i.e., his ability to formulate *philosophical* problems in a manner that made them amenable to *mathematical* methods.[2] We shall set forth two modal principles that characterize properties of *proof*, *time*, and *God* that reveal logical interconnections among these theorems and philosophical conclusions.

---

[1] Among the great logicians of the 20th century one must also include Alonzo Church, his dissertation student Alan Turing, and Alfred Tarski, who claimed only to be the "greatest living *sane* logician." [12], 5.

[2] This felicitous turn of phrase is due to Odifreddi [40].

G. Mar (✉)
Department of Philosophy, Stony Brook University, New York, USA
e-mail: Gary.Mar@stonybrook.edu

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_36

36.1 Gödel's Second Incompleteness Theorem

> I have been again concerned with logic recently, using the methods that you so successfully applied for the proof of undecidable properties. Here I came to a result that seems remarkable to me. Namely, I was able to prove that the consistency of mathematics is unprovable.
>
> —John von Neumann's letter to Gödel (Nov. 30, 1930) ([17], 337)

Gödel's Incompleteness Theorems [23] are among the most profound results of 20th century mathematics. Popular accounts state that Gödel's Theorem shows there are truths of arithmetic that cannot be proved. This statement is inaccurate on two counts. First, Gödel's theorem does not deal with provability in any *absolute* sense but only *relative* to a class of precisely defined formal systems. Secondly, there are *two* incompleteness *theorems*.

On September 7th, 1930, during the roundtable discussion on the foundations of mathematics that closed the Königsberg conference on the *Epistemology of the Exact Sciences*, Gödel quietly announced, in a meticulously crafted sentence, an early version of his First Incompleteness Theorem:

> (Assuming the consistency of classical mathematics) one can even give examples of propositions (and in fact of those of the type of Goldbach or Fermat) that, while contentually true, are unprovable in the formal system of classical mathematics ([13], 203).

Although leading logical positivists and philosophers of science such as Rudolf Carnap, Herbert Feigl and Friedrich Waismann attended the Königsberg conference, Gödel's concise announcement fell on uncomprehending ears, with one exception. The single exception was John von Neumann, who had been chosen to represent Hilbert's formalist views on the philosophy of mathematics at the conference. Gödel's remarks had piqued von Neumann's curiosity, and von Neumann pressed for a private conversation with Gödel to find out more about his meta-mathematical results.

One of the great ironies in the history of mathematics was that at another conference in Königsberg on the very next day, David Hilbert gave his famous lecture triumphantly ending with the credo: for the mathematician there is "no *Ignorabimus…. Wir müssen wissen. Wir werden wissen!*" Hilbert not only professed his credo that all mathematical truths are knowable but also proposed to limit the realm of what is mathematically knowable to what could be verified by a "mechanical procedure."[3]

Several weeks after their discussion, von Neumann sent the letter (quoted above) to Gödel announcing his remarkable discovery that "the consistency of mathematics

---

[3] A mechanical procedure, in contemporary terms, is an *algorithm.* Today Gödel's incompleteness theorems are studied through the perspective of algorithmic complexity. In a letter written in 1956 (but not discovered until the 1990s), Gödel wrote to von Neumann, who was dying from cancer, hoping to give him something to contemplate other than his impending death. In this remarkable letter, Gödel discusses what is now known as the P = NP conjecture, one of the most famous unsolved problem in computer science.

is unprovable." It must have been a great disappointment to von Neumann when Gödel informed him that thirteen days earlier he had submitted for publication his *Über formal unentscheidbare Sätze der* Principia Mathematica *und verwandter Systeme, I*, which already contained *Satz* XI, now known as Gödel's Second Incompleteness Theorem. The irony was that this theorem shattered Hilbert's dream of giving the foundations of mathematics an absolutely reliable justification through a finitary proof of the consistency of mathematical reasoning.[4]

Gödel informally explained his First Incompleteness Theorem noting that its analogy to the antinomy of the Liar (or other epistemological antinomies) "leaps to the eye ([13], 149)." Whereas the Liar sentence asserts of itself that it is *untrue*, the Gödel sentence says of itself that it is *unprovable* in a precisely specified formal system such as *Principia Mathematica*:

$$\text{G: G is unprovable in system PM}$$

Supposing G to be provable in PM, then PM would be *inconsistent* insofar as it proves a sentence asserting its own unprovability. Supposing $\sim$ G to be provable in PM, then G is unprovable in PM, assuming PM to be consistent. So neither G nor its negation is provable in PM, i.e., G is *undecidable* in PM. Therefore, assuming PM to be *consistent*, PM is *incomplete*.

> Gödel's First Incompleteness Theorem: if a formal system is consistent and its axiom system has enough arithmetic so that its theorems can be listed by some mechanical procedure, then there exists an *undecidable* sentence in that formal system, which is therefore *incomplete*.[5]

Elegant proofs of Gödel's Second Incompleteness Theorem were discovered in *modal provability logics,* which emerged from the 1950s–1970s. These logics were anticipated by Gödel's [24] "*An interpretation of intuitionistic propositional calculus.*" Gödel's insight was that intuitionistic *truth* was characterized in terms of *proof*, which is a kind of *necessity*, and so modal axioms could be used to formalize the *properties of provability*:

(T)	$\Box P \to P$	*What is provable is true.*
(**K**)	$\Box(P \to Q) \to (\Box P \to \Box Q)$	*Whatever follows from what is provable is provable.*
(**4**)	$\Box P \to \Box\Box P$	*What is provable is provably provable.*

Henkin [35] posed the intriguing question whether the *positive* Gödelian sentence "I am *provable*" is provable. Löb [37] answered Henkin's question in the affirmative by showing that Peano Arithmetic proves a counterpart to Löb's Axiom:

---

[4] Non-finitary proofs of consistency are possible, see Gentzen [21].

[5] The formal system is also *essentially incomplete*, i.e., one can add the undecidable Gödel sentence as a new axiom and the resulting system will have a new undecidable sentence, which is also undecidable in the original system.

(L)	$\Box(\Box P \rightarrow P) \rightarrow \Box P$	*Löb's axiom restricts* (**T**) *to what is provable*

A *Gödel-Löb modal probability logic* (GL) results from adding (**L**) to (**K**) with the rule of *necessitation* (i.e., if $\mathrm{GL} \vdash P$ then $\mathrm{GL} \vdash \Box P$), *modus ponens*, and a rule for proving all *tautologies*. (Adding Axiom (**4**) turns out to be redundant: in 1975 de Jongh proved that Axiom (**4**) is derivable from Löb's axiom and (**K**) using the substitution of '$(\Box P \wedge P)$' for 'P').

We can sketch an elegant proof in modal provability logic of Gödel's Second Incompleteness Theorem. First, we have a modal counterpoint to the *fixed-point theorem* that yields the *Gödel sentence*:

$$\vdash G \leftrightarrow \sim\Box G.$$

In his letter von Neumann noted that the consistency of Peano Arithmetic (PA) can be expressed by the formula that (1 = 2) is not provable:

$$\mathrm{Cons(PA)}: = \sim\Box(1=2).$$

Now the gist of the First Incompleteness Theorem is the demonstration that:

$$\text{if} \vdash G \leftrightarrow \sim\Box G, \text{ then} \vdash \sim\Box(1=2) \rightarrow \sim\Box G.$$

By the fixed-point theorem, $\sim\Box G$ is logically equivalent to G, so we have $\vdash \sim \Box\ 1 = 2 \rightarrow G$:

$$\begin{tikzcd}
\sim\Box\, 1=2 \rightarrow \arrow[dr] & \sim\Box G \arrow[d, leftrightarrow, "\textit{fixed-point theorem}"] \\
& G
\end{tikzcd}$$

By the rule of necessitation, we may prefix a $\Box$ and then distribute, using (**K**), the $\Box$ over the conditional:

$$\vdash \Box \sim \Box 1 = 2 \rightarrow \Box G$$

According to the First Incompleteness Theorem, the provability of the Gödel sentence implies the inconsistency of the system, so:

$$\begin{tikzcd}
\vdash \Box\sim\Box\, 1=2 \rightarrow \arrow[dr] & \Box G \arrow[d, "\textit{First Incompleteness Theorem}"] \\
& \Box\, 1=2
\end{tikzcd}$$

In short, we have:

$$\vdash \Box \sim \Box\, 1 = 2 \rightarrow \Box\, 1 = 2,$$

which, by contraposition, yields:

$$\vdash \sim \Box 1 = 2 \rightarrow \sim \Box \sim \Box 1 = 2.$$

Since $\sim \Box\ 1 = 2$, by definition, is Cons(PA), we have:

Gödel's Second Incompleteness Theorem. $\vdash (\text{Cons(PA)} \rightarrow \sim \Box\ \text{Cons(PA)})$, i.e., if Peano Arithmetic is *consistent*, then it cannot prove its own *consistency*.

In a lecture to a joint meeting of the Mathematical Association of America and the American Mathematical Society, Gödel summarized the significance of his result for Hilbert's program: the hope of finding "…a proof for freedom from contradiction undertaken by Hilbert and his disciples" had "vanished entirely in view of some recently discovered facts. It can be shown quite generally that there can exist no proof of the freedom of contradiction of a formal system S which could be expressed in terms of the formal system S itself …." ([25] in [15], 52).

36.2 Gödel, Einstein, and the Unreality of Time

> Time is no specific character of being. In relativity theory the temporal relation is like far and near in space. I do not believe in the objectivity of time. The concept of *Now* never occurs in science itself, and science is supposed to be concerned with all that is objective.
>
> —Gödel in Conversation with Wang ([46], 320, 9.5.10.)

During the last fifteen years of his life at the Institute for Advanced Studies (IAS), Einstein sought the company of the younger Gödel (who was born the year after Einstein's *annus mirabilis* in 1905. Einstein told Morgenstern that he went to the Institute "*um das Privileg zu haben, mit Gödel zu Fuss nach Hause gehen zu dürfen.*" A famous photo taken by Morgenstern shows Einstein, causally dressed in a rumpled sweatshirt with his wildly flowing hair, standing next to Gödel, dressed in a button-down double-breasted suit and impeccably groomed with every hair in place. The photo is a portrait of diametrical opposites.[6]

Einstein and Gödel, May 13, 1947. *Photo*: Oskar Morgenstern

---

[6] Reprinted in GCW-III, [15], 260. Curiously, the juxtaposition of Einstein and Gödel in the photo creates an Escher-like illusion: the arm suspended between the two looks as if it could belong to either.

In order to coax Gödel to publish so he could be advanced to full professor, von Neumann, as director of IAS, had come up with the idea of asking Gödel to submit a paper to a *Festschrift* in honor of his friend Einstein's 70th birthday. This led to Gödel's remarkable contributions to cosmology [26–30].

Gödel's contribution to physics is perhaps surprising to those who know of him only as a logician. However, Gödel was a physics student at the university before switching to mathematics, and his interest in physics was maintained by conversing with Einstein and attending the physics colloquia at IAS. After Gödel had agreed to contribute a short philosophical contribution for Einstein, Gödel came across a note by Gamow [20] in *Nature* challenging physicists to find a rotating universe consistent with general relativity. This convergence of challenges enabled Gödel to combine his *philosophical* defense of Kant's view on the non-objectivity of time [27] with his *mathematical* explorations into the cosmology of rotating Gödel universes in general relativity. Gödel's philosophical paper contained no mathematical equations but his technical paper [29] that appeared at the same time in a special issue of the *Review of Modern Physics* did. Gödel discovered a model in which the entire universe was rotating, creating time-like loops and making it possible to travel to the past by travelling into the future.[7]

Newton in his *Principia* [1687] ([39]) had postulated the *metaphysis* of "absolute, true, mathematical time… [which] flows equably without relation to anything external," whereas Einstein in his Special Theory of Relativity [10] adopted a *Machian* point of view, according to which time is something we abstract from *measurements* of time. Beginning with two principles:

> [*Relativity*] The laws of physics are valid for all observers in all inertial frames of reference.
>
> [*Constant c*] The vacuum speed of light is constant for all observers in all inertial frames of reference.

Einstein deduced that whether an observer measures two events to be happening "at the same time" depends on the observer's position and state of motion. This *Principle of the Relativity of Simultaneity*, Einstein [10] noted, undermines the privileged status of the present: "*The four-dimensional continuum is now no longer resolvable objectively into sections, all of which contain simultaneous events; 'now' loses for the spatially extended world its objective meaning.*"

Gödel noted, however, that Einstein's General Theory of Relativity [11] allows for the reintroduction of a *universal cosmic time*. If the universe were non-expanding, Gödel noted, there are distinguished frames of reference which "follow the mean motion of matter" of the cosmos. Gödel explained in his lecture at IAS that his search for rotating solutions, or *Gödel universes*, was prompted by a desire to counter this objection in order to prove that time is *not objective* and so *unreal*:

---

[7] Gödel casually remarks (GCW-III, [27], 271): "This contradicts Mach's principle but it does not contradict relativity theory."

> This incidentally also was the way in which I happened to arrive at the rotating solutions. I was working on the relationship… between Kant and relativistic physics insofar as in both theories the objective existence of a time in the Newtonian sense is denied. On this occasion one is led to observe that in the cosmological solutions known at present there does exist something like an absolute time…. So one is led to investigate whether or not this is a necessary property of all possible cosmological solutions.[8]

Gödel [27] linked relativity physics with the philosophical tradition of "Parmenides, Kant, and the modern idealists [i.e., McTaggart] in the common denial of the "objective existence of … time in the Newtonian sense" ([14], 202). McTaggart [42] in setting forth his famous philosophical argument in "The Unreality of Time" enunciated various philosophical *dicta*, including the following: "*If one of the determinations past, present and future can ever be applied to [an event] then one of them has always been and always will be applicable, though of course not always the same one.*"

There are various reasons for being interested in tense logic. Ordinary language is *tensed* whereas the language of physics is mathematical and *untensed*, and one can learn how to translate between the two types of expressions without confusing them. For example, the *connectedness* of time, the principle that "for any two distinct instants of time, one is earlier and other is later," can be expressed using modal temporal operators as "*whatever is going to have been the case either already has been or is now or will be.*"

Using the modal temporal operators:

$\Box$	It *will always* be the case that
$\Diamond$	It *will* be the case that
$\blacksquare$	It *has always been* the case that
$\blacklozenge$	It *was once* the case that

McTaggart's *dicta* above yield the following temporal principles, here listed with their corresponding properties of temporal ordering:

$\blacklozenge P \to \blacksquare(\Diamond P \vee P \vee \blacklozenge P)$	*Connectedness of the past*
$\Diamond P \to \Box(\Diamond P \vee P \vee \blacklozenge P)$	*Connectedness of the future*
$\blacklozenge P \to \Box \blacklozenge P$	*Transitivity of "earlier than"*
$\Diamond P \to \blacksquare \Diamond P$	*Transitivity of "later than"*
$P \to \blacksquare \Diamond P$	*The past and future are converses*
$P \to \Box \blacklozenge P$	*The future and are converses*[8]

The third and four principles are instances of Axiom (**5**) and the last two principles are instances of the *Brouwersche* axiom. Both Axiom (**5**) and the Brouwersche axioms play key roles in contemporary explications of the modal ontological argument.[9]

---

[8] Gödel's handwritten notes for his May 7, 1949 lecture at the IAS (GCW-III, [15], 274).

[9] One axiom can be derived from the other by reversing the conditional (taking its *converse*) and also reversing the temporal operators.

Consider the *tense logic*[10] axiomatized by:

Löb's Axiom: $\blacksquare(\blacksquare P \rightarrow P) \rightarrow \blacksquare P$
McKinsey's Axiom: $\Box\Diamond P \rightarrow \Diamond\Box P$

Löb's Axiom defines the *well-foundedness* of the past, i.e., there is no infinite regress of moments of time into the past and the *transitivity* of "*later than*." On transitive frames, the McKinsey's Axiom defines the *atomicity* of time for the future, i.e., for any moment of time, there is eventually a *maximal* moment of time. This tense logic, therefore, characterizes the intuitive notion of time proceeding in a linear sequence.

Löb's Axiom implies the transitivity of "*earlier than*" and so time in this tense logic is *isotropic*: it is transitive independent of temporal direction. However, it is precisely this *transitivity* of time that leads to inconsistency with intuitive principles about time in these Gödel universes. The problem with Special Relativity is that there is *no unique objective way* of identifying the present moment, but the problem with Gödel universes in the General Theory of Relativity that allows for closed time-like loops through every point, which is *incompatible* with a tensed theory of time.[11] Hence, given the possibility of Gödel universes, time as characterized by the above tense logic does not exist. Moreover, this modal characterization of temporality yields other intriguing results: (1) the corresponding relation for the McKinsey Axiom is *not first-order definable* and (2) the above logic is *modally incomplete* (i.e., the tense logic holds in no first-order frame and yet is not inconsistent.)[12]

---

[10] A linguistic motivation would be to combine modal tense logic with other features such as *counterfactuals*, *modality*, and Reichenbach's [1947] *analysis of tense using three points in time*. Given 'E' (the event), 'R' (a point of reference) and 'S' (point of speech) and two ordering relations, we can give analyses of sentences such as these: "If I *had asked* Rosemary to marry me when we first met, she would have accepted. If we *had married* back then, we *would have now been* married 28 years. Unfortunately, you can't change the past." Another reason concerns computer science and dynamic logic. "Temporal operators have been used to express such properties of programs as *termination*, *correctness*, *safety*, *deadlock freedom*, *clean behavior*, *data integrity*, *accessibility*, *responsiveness*, and *fair scheduling*" (see, Burgess [7], 95).

[11] Consider a tensed theory of time in which the *present* and *past* are *real* but in which the future is *not yet real*. For any two temporal points A and B on a closed loop such A is present and therefore real, B exists in both A's past and future and so is both real and not-real. According to *presentism* only the present is real. Assuming that A is the present and therefore real, in a closed time-like loop A would also be in its own past and future and hence unreal.

[12] See van Bentham [44], 198, 223. Löb's axiom requires *well-foundedness*, i.e., that there are no infinitely descending chains. The McKinsey axiom requires reflexive endpoints. But these two properties are inconsistent on transitive frames since any reflexive point implies the existence of an infinitely descending chain, which is forbidden by well-foundedness. However, the tense logic is consistent. Consider the collection of all finite and co-finite (i.e., a set whose complement is finite) subsets of natural numbers. This collection will have all finite sets and their complements, but not subsets such as the set of all even numbers, or the set of prime numbers. Let the accessibility relation on the set of finite and co-finite sets be the proper subset relation. The Löb Axiom holds because the frame is transitive and well-founded. The McKinsey Axiom also holds since its antecedent says that if any formula denotes a co-finite set, and this set will have a future stage that *stabilizes*, i.e., a set that contains all greater natural numbers. The empty set and the set of all natural numbers constitute non-reflexive "endpoints." Thomason [43], 153) speculates that the

On March 15, 1951, Einstein handed out the First Einstein Award awarded jointly to Julian Schwinger and Gödel, saying to Schwinger "you deserve it"[13] and to Gödel "you don't need it." On this occasion, it was fitting that von Neumann [45], the first mathematician to grasp the revolutionary significance of Gödel's Incompleteness Theorems twenty-one years earlier, delivered a tribute to Gödel's work hailing it as "a landmark which will remain visible far in space and time."

Einstein called Gödel's curious cosmological gift "an important contribution to the general theory of relativity" but sought physical constraints for ruling out such universes.[14] Gödel himself drew a different conclusion: if time travel in terms of closed time-like loops is possible, then time itself is unreal. Time, like God, is either necessary or nothing; if it disappears in one possible universe, it is undermined in every possible universe, including our own. In other words, if it is *possible* that the existence of time is *impossible*, then time does *not* exist. This reasoning is an instance of the logically dual form of the Brouwersche Axiom, $(\mathbf{B}\Diamond)\ \Diamond\Box \sim \mathrm{T} \rightarrow \sim \mathrm{T}$, which, as we shall see, plays a key role in Gödel's *Ontologischer Beweis*.

Gödel's and Einstein's belief in the *unreality* of time seemed to be a defense against the *reality* of death. After failing to get his doctorate in physics, Einstein had worked in obscurity in the patent office in Bern with Michele Besso, a friend since their years together at the Swiss Federal Polytechnic in Zurich. It was Besso who introduced Einstein to the positivism of Ernst Mach, which was formative in Einstein's discovery of relativity.[15] When Besso died, Einstein wrote in his letter to Besso's wife dated March 21, 1955: "*Now he has departed from this strange world a little ahead of me. That signifies nothing. For those of us who believe in physics, the distinction between past, present and future is only a stubbornly persistent illusion.*" When it came time for his own death less than two weeks later, Einstein simply said, "*It's time to go.*"

(Footnote 12 continued)

McKinsey Axiom has a "reasonable intuitive meaning" related to the Second Law of Thermodynamics since it says that "each proposition eventually ceases changing its truth value with time.".

[13] Schwinger went on to win the Nobel Prize in Physics in 1965 jointly with Richard Feynman and Shinichiro Tomonaga for work on Quantum Electrodynamics.

[14] Beginning in 1992, Stephen Hawking [34] postulated "Time Cops" (alluding to Isaac Asimov's *The End of Eternity*) to prevent the appearance of closed time-like curves but subsequently dropped the proposal as *ad hoc*. Also see Rindler [41]. The geometry of the multiverse of possible spacetimes in the Theory of Relativity is a subject of current research. For example, Chaitin et al. [8], 124: *The "typical" spacetime is exotic, without global time, and if properly axiomatized in set theory with Martin's axiom, it is set-theoretically generic.*

[15] Wang [46], 176, 5.4.19 reports Gödel's views: "Positivism is generally not fruitful in scientific research although it may have been valuable in the discovery of the special theory of relativity. Generally speaking, right ideas are fruitful. Positivism is pedagogically better for the special theory of relativity."

36.3 Gödel's God

> Visited Gödel at the Institute today. A true miracle: he has gained 18 lbs., is in best shape, sparkling conversation. Lots about politics but then about his ontological proof—he had the result several years ago, is not happy with it but hesitates over publishing it. It would be concluded that he really believes in God, whereas he is only undertaking a logical investigation (i.e., he shows that such a proof, appropriately axiomatized, is possible under classical assumptions—perfections, etc.). I joked that he should use a pseudonym—but he has already told Hempel and Scott about it and I said people would recognize 'the claw of the lion', as they did with Newton.
>
> —Oskar Morgenstern's diary entry August 29, 1970 ([9], 307).

Although in his younger days he attended the meetings of the Vienna Circle, Gödel distanced himself from the Circle's atheism and positivism. In his popularization of Logical Positivism, A. J. Ayer in *Language Truth and Logic* [1936] argued that "God exists" is not even false but *cognitively meaningless* because there could be no empirical experiences by which one could verify its truth or falsity. Gödel, however, regarded this atheism as a "prejudice of the times," and it was Gödel's *realism*—in contrast to the implicit *verificationism* of Brouwer's *intuitionism* and Hilbert's *formalism*—that led Gödel to distinguish between truth and proof and so discover his Completeness and Incompleteness Theorems.

Gödel, like Leibniz, worked on constructing an ontological modal proof for God's existence. Gödel's handwritten notes for his *Ontologischer Beweis* [31] occur in his notebooks c. 1941 preceding his preoccupation with Leibniz from 1943–6. In an unsent response to the famous Grandjean questionnaire,[16] Gödel [1975] wrote: "*My belief is theistic; not pantheistic, following Leibniz rather than Spinoza. Spinoza's God is less than a person. Mine is more than a person....*". [17]

However, it wasn't until February 1970 that Gödel, worried about dying, passed on his proof to Dana Scott.[18] Leibniz had criticized Descartes's ontological argument for failing to provide an argument for the possibility premise. Gödel bridges this gap by axiomatizing the notion of a positive property and then giving a *maximal consistency* argument to prove the possibility of the collection of all positive properties being instantiated. Gödel undoubtedly noticed the striking parallels between Leibniz's argument and his own non-constructive maximal consistency proof of the Completeness Theorem for first-order logic.[19]

---

[16] The Burke Grandjean questionnaire, the only first-person commentary on Gödel's early life, was first reported in Wang [47] and appears in GCW-IV, [16], 441ff.

[17] Wang [47], 16, 19, 21.

[18] When I asked Dana Scott at the 2014 *Vienna Summer of Logic* about how it felt to be entrusted with Gödel's ontological discovery, Scott said he regretted having to receive a stream of correspondence from religiously inspired admirers, and atheistic detractors, who had little appreciation for the logical properties of Gödel's ontological discovery.

[19] Fitting [18], 115: "Perhaps Gödel had in mind something like the notion of a maximal consistent set of formulas, familiar from the Lindenbaum/Henkin approach to proving classical completeness."

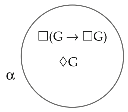
Here we can present the *modal core* of Gödel's ontological discovery.[20] The argument requires an Anselmian premise (the *converse* of the conditional in the antecedent of Löb's Axiom). If God exists, then God's existence would be a matter of *necessity*, or, in other words, God's existence, if God exists, could *not be accidental*. The first premise states that this conditional itself is a necessary truth. The second premise asserts that God's existence is *possible*, for which Gödel provided a demonstration. The surprising conclusion is that God *actually* exists:

Anselmian Axiom:	$\Box(G \rightarrow \Box G)$
Possibility Premise:	$\Diamond G$
Actualist Conclusion:	$\therefore G$

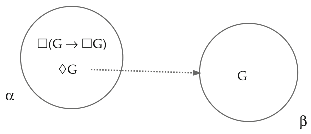
Using the Leibnizian idea that necessity is truth in *all* possible worlds and possibility is truth in *some* possible world, Kripke [36] showed that the proliferation of syntactical characterizations of modality elegantly correspond to natural semantic conditions that could be required of the accessibility relation of "relative possibility" among a set of possible worlds. Although Axiom (**5**) is used in many contemporary version of the modal ontological argument, the weaker Brouwersche Axiom is sufficient:

Brouwersche Axiom:	$G \rightarrow \Box\Diamond G$	*Symmetry*

We may sketch the modal core of Gödel's ontological argument using possible world diagrams. In the actual world α, the two modal premises are true:



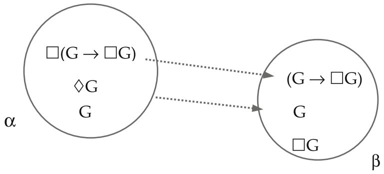
If $\Diamond G$ is true in the actual world α, then in *some* world β possible relative to α, G is true:



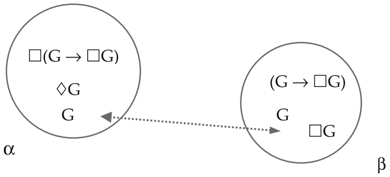
---

[20] See Hartshorne [33], which was praised by Gödel, and also Adams [1, 2] and Mar [38].

If $\Box(G \rightarrow \Box G)$ is true in α, then in *all* worlds possible relative to α, including β, $(G \rightarrow \Box G)$ is true. Hence, by *modus ponens* (since the laws of logic hold in all possible worlds) $\Box G$ is true in β:



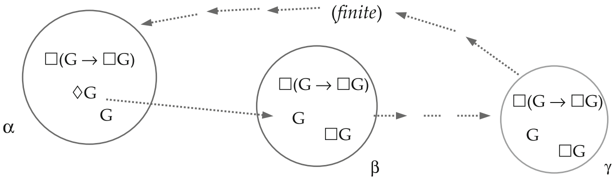
The Brouwersche Axiom makes the relative possibility relation *symmetric*, i.e., all accessibility arrows between worlds are *double* arrows. Since $\Box G$ is true in β, assuming symmetry, G is true in α.



In fact, the Brouwersche Axiom allows the deduction of the *stronger* conclusion $\Box G$, namely that God's existence is *necessary*. A *weaker* condition suffices for a proof of God's *actual* existence:

$$\Diamond G \wedge \Box(G \rightarrow \Box G) \rightarrow G$$

namely, the principle of "finite return",[21] i.e., from each successor world we can always return to α by a finite chain of successor worlds:



Some critics have accused Gödel of disingenuously distancing himself from his discovery as an *apologetic* argument, but we have taken Gödel at his word and engaged in a *logical* investigation.[22]

---

[21] See van Bentham [44], 176.

[22] Benzmüller and Paleo [5] programmed a computer to verify Gödel's ontological proof. For discussion of the logical properties of Gödel's ontological proof see Anderson [3] and Hájek [32].

In conclusion, Gödel had grand ontological dreams. Despite their technical sophistication, Gödel's results have managed to escape mere mathematics and shed light on larger philosophical issues. Why? One reason consists not merely in Gödel's judicious choice of problems but also in his virtuosity at *mathematizing* philosophical problems. Gödel in his conversations with Hao Wang remarked:

The significance of mathematical logic for philosophy lies in its power to make thoughts explicit by illustrating and providing a frame for the axiomatic method. Mathematical logic makes explicit the central place of predication in the philosophical foundation of rational thought.[23]

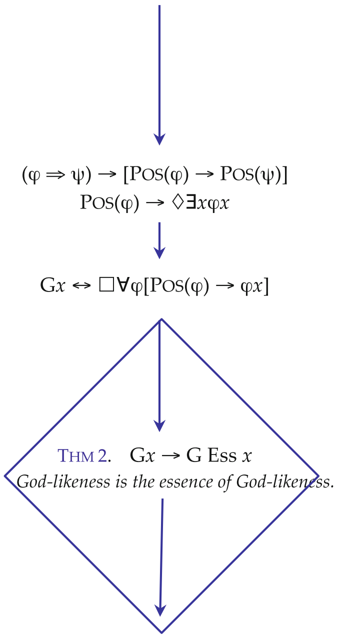
I have tried to show how deeply interconnected modal principles—such as variations of the Löb and the Brouwersche Axioms characterizing the properties of *proof*, *time*, and *God*—played a role in realizing Gödel's ontological dreams.[24]

Appendix: Gödel's *Ontologischer Beweis* (c. 1941)

---

[23] Wang [46], 293, 9.1.16.

[24] Gödel in his notebooks wrote: "*Philosophy*: The fundamental concept is cause. It involves: will, force, enjoyment, God, time, space" (GCW-III, [15], 433). I wish to thank Christian Stamos, John Foulks, Robert Pasternak, and Thomas Graf for enjoyable conversations, without which, this paper would not have become actual.



GÖDEL AX. 1 — $\mathrm{Pos}(\varphi) \lor \mathrm{Pos}(\sim\varphi)$ — Gödel Note: *"exclusive 'or'"*

SCOTT'S FORMULATION OF AX. 1 — $\mathrm{Pos}(\varphi) \leftrightarrow \sim\mathrm{Pos}(\sim\varphi)$ — *A property is positive if and only if its negation is not positive.*

ANDERSON'S EMENDATION OF AX. 1 — $\mathrm{Pos}(\varphi) \rightarrow \sim\mathrm{Pos}(\sim\varphi)$ — *The negation of a positive property is not positive.*

Gödel Note: *"The ontological proof must be grounded on the concept of* value *(p better than ~p) and on the* axioms*… It can be grounded only on axioms and not on definitions (= construction) of "positive," for a construction is compatible with an arbitrary relationship."* (GCW-III, 433)

Entailment $(\varphi \Rightarrow \psi) := \Box\forall x(\varphi x \rightarrow \psi x)$

AX. 2 — $(\varphi \Rightarrow \psi) \rightarrow [\mathrm{Pos}(\varphi) \rightarrow \mathrm{Pos}(\psi)]$ — *Entailment preserves positiveness*

THM. 1 — $\mathrm{Pos}(\varphi) \rightarrow \Diamond\exists x\varphi x$ — *Positive properties are possibly exemplified*

DEF. 1 — $Gx \leftrightarrow \Box\forall\varphi[\mathrm{Pos}(\varphi) \rightarrow \varphi x]$ — *A being is* God-like *if it has all positive properties.*

AX. 3 — Pos(G), i.e., *God-likeness is positive.*

AX. 4 — $\mathrm{Pos}(\varphi) \rightarrow \Box\mathrm{Pos}(\varphi)$, i.e., *Being positive is necessary, i.e., not accidental.*

Gödel Note: *"That the necessity of a positive property is positive is the essential presupposition of the ontological proof."* (GCW-III, 435)

AX. 5 — Pos(NE), *i.e., necessary existence is positive.*

DEF. 2 — $\varphi\ \mathrm{Ess}\ x \leftrightarrow \varphi x \land \forall\psi\,[\psi x \rightarrow (\varphi x \Rightarrow \psi x)]$, *i.e., an* essence *for x is a property that entails all other properties of x.*

THM 2. $Gx \rightarrow G\ \mathrm{Ess}\ x$

God-likeness is the essence of God-likeness.

DEF. 3 — $\mathrm{NE}\,x \leftrightarrow \forall\varphi\,[\varphi\ \mathrm{Ess}\ x \rightarrow \Box\exists x\varphi x]$, *"necessary existence"*

Gödel Note: *"For this it is required that all the properties of God are defined by a second-order-property."* (GCW-III, 431)

THM 3. $\exists xGx$ — *There exists a God-like being.*

COR. 3A. $\exists!xGx \land \Box\exists!xGx$ — *There exists a unique God-like being and it is necessary there is a unique God-like being.*

Gödel Remark: *"(Philosophy): If the ontological proof is correct, then one can obtain insight a priori into the existence (actuality) of a non-conceptual object."* (GCW-III, 431)

References

1. Adams, R. M. (1971). The logical structure of Anselm's arguments. *Philosophical Review, 80* (1), 28–54.
2. Adams, R. M. (1995). Introduction Note to *1970 [Ontological Proof]. In *GCW-III*, [15].
3. Anderson, C. A. (1990). Some emendations of Gödel's ontological proof. *Faith and Philosophy, 7*, 291–303.
4. Baaz, M., Papadimitriou, C., Putnam, H., Scott, D., & Harper, C., Jr. (Eds.). (2004). *Kurt Gödel and the foundations of mathematics: Horizons of truth*. New York: Cambridge University Press.
5. Benzmüller, C., & Paleo, B. W. (2003). *Formalization, mechanization and automation of Gödel's proof of God's existence*. http://arXiv.org/abs/1308.4526.
6. Bulloff, J., Holyoke, T., & Hahn, S. W. (Eds.). (1969). *Foundations of mathematics: Symposium papers commemorating the sixtieth birthday of Kurt Gödel*. New York: Springer-Verlag.
7. Burgess, J. (1984). Basic tense logic. In Gabbay and Guenthner [19], (pp. 89–113).
8. Chaitin, C., da Costa, N., & Doria, F. (2013). *Gödel's way: Exploits into an undecidable world*. New York: Taylor & Francis Group.
9. Dawson, J. W., Jr. (1997). *Logical Dilemmas: The life and work of Kurt Gödel*. Wellesley, MA: A. K. Peters Ltd.
10. Einstein, A. (1905). Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. *Annalen der Physik, 17*(6), 132–148.
11. Einstein, A. (1915). *Die Feldgleichungen der Gravitation, Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin*, (part 2, pp. 844–847).
12. Feferman, A. B., & Feferman, S. (2008). *Alfred Tarski: Life and logic*. Cambridge University Press.
13. Feferman, S., Dawson, J., Kleene, S., Moore, G., Solovay, R., van Heijenoort, J., et al. (Eds.) (1986). [GCW-I], *Kurt Gödel: Collected works, Vol. I, Publications 1929–1936*. New York: Oxford University Press.
14. Feferman, S., Dawson, J., Kleene, S., Moore, G., Solovay, R., van Heijenoort, J., et al. (Eds.) (1986). [GCW-II], *Kurt Gödel: Collected works, Vol. II, Publications 1938–1976*. New York: Oxford University Press.
15. Feferman, S., Dawson, J., Goldfarb, W., Parsons, C., & Solovay, R. [GCW-III]. (1995). *Kurt Gödel: Collected works, Vol. III, Unpublished essays and lectures*. New York: Oxford University Press.
16. Feferman, S., Dawson, J., Goldfarb, W., Parsons, C., & Sieg, W. (Eds.). [GCW-IV] (2003). *Kurt Gödel: Collected works, Vol. IV, Correspondence A–G*. New York: Oxford University Press.
17. Feferman, S., Dawson, J., Goldfarb, W., Parsons, C., & Sieg, W. (Eds.). [GCW-V], *Kurt Gödel: Collected Works, Vol. V, Correspondence H–Z*, (Oxford University Press: New York), 2003.
18. Fitting, M. (2002). *Types, Tableaus, and Gödel's God*. Kluwer Academic: North Holland.
19. Gabbay, D., & Guenthner, F. (Eds.) (1984). *Handbook of philosophical logic, Vol. II, Extensions of classical logic*. Dordrecht: D. Reidel.
20. Gamow, G. (1946). Rotating universe? *Nature, 158*, 549.
21. Gentzen, G. (1936). Die Widerspruchsfreiheit der reinen Zahlentheorie. *Mathematische Annalen, 112*, 493–565.
22. Gödel, K. (1930). Die Vollständigkeit der Axiome des logischen Funktionenkalküls. *Monatshefte für Mathematik und Physik, 37*, 349–60. In *GCW-I*, [13], pp. 102–23.
23. Gödel, K. (1931). Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme, I. *Monatshefte für Mathematik und Physik*, *38*. In *GCW-I*, [13], pp. 145–195.
24. Gödel, K. (1933). On Intuitionistic propositional calculus. In *GCW-I*, [13], pp. 222–225.

25. Gödel, K. (1933). The present situation in the foundations of mathematics. In *GCW-III*, [15], pp. 45–53.
26. Gödel, K. (1946/9-B2). Some observations about the relationship between theory of relativity and Kantian philosophy. In *GCW-III*, [15], pp. 230–259.
27. Gödel, K. (1949). A remark about the relationship between relativity theory and idealistic philosophy. In P. A. Schilpp (Ed.) *Albert Einstein: Philosopher-Scientist* (pp. 555–562) (Open Court: La Salle, Illinois) reprinted in *GCW-II*, [14], 202-7.
28. Gödel, K. (1949). Lecture on rotating universes. In *GCW-III*, [15], pp. 269–287.
29. Gödel, K. (1949). An example of a new type of cosmological solution of Einstein's field equations of gravitation. *Review of Modern Physics*, *21*, 447–450 (reprinted in *GCG-II*, [14], 190–98).
30. Gödel, K. (1952). Rotating universes in general relativity theory. In Graves et al. (Eds.), *Proceedings of the International Congress of Mathematicians* (Vol. 1, pp. 175–181) (reprinted in *GCW-II*, [14], 208–216).
31. Gödel, K. (*1970). Ontological proof. In *GCW-III*, [15], pp. 403–414.
32. Hájek, P. (2006). Gödel's ontological proof and its variants. In Baaz et al. [4], pp. 307–320.
33. Hartshorne, C. (1962). *The logic of perfection.* Open Court: LaSalle, Il.
34. Hawking, S. (2002). Chronological protection. In S. W. Hawking, et al. *The future of spacetime* (pp. 87–10). Norton: New York.
35. Henkin, L. (1952). A problem concerning provability. *Journal of Symbolic Logic, 17*, 160.
36. Kripke, S. (1959). A completeness theorem for modal logic. *Journal of Symbolic Logic, 24*, 1–14.
37. Löb, M. H. (1955). Solution of a problem of Leon Henkin. *Journal of Symbolic Logic, 20*, 115–118.
38. Mar, G. (1996). The modal unity of Anselm's *Proslogion*. *Faith and Philosophy, 13*, 50–67.
39. Newton, I. (1689). *Philosophia Naturalis principia mathematica, Book I* (transl. by A. Motte, rev. by F. Cajori). Berkeley: University of California Press.
40. Odifreddi, P. (2006). Gödel's mathematics of philosophy. In Baaz, et al. [2006], [4], pp. 299–304.
41. Rindler, W. (2006). Gödel, Einstein, Mach, Gamow, and Lanczos: Gödel's remarkable excursion into cosmology. In Baaz et al. [4], pp. 185–211.
42. McTaggart, J. M. E. (1908). The unreality of time. *Mind, 17*(68), 457–474.
43. Thomason, S. K. (1972). Semantic analysis of tense logic. *Journal of Symbolic Logic, 37*, 150–158.
44. van Bentham, J. (1984). II.4. Correspondence theory. In Gabbay and Gunethner [19], pp. 167–247.
45. von Neumann, J. (1951). Tribute to Dr. Gödel. In Bulloff, et al. [6], pp. 299–303.
46. Wang, H. (1997). *A logical journey: From Gödel to philosophy*. Cambridge, MA: MIT Press.
47. Wang, H. (1987). *Reflections on Kurt Gödel*. Cambridge, MA: MIT Press.

Part VI
Miscellaneous

Chapter 37
The Novel and the Map: Spatiotemporal Form and Discourse in Literary Cartography

Robert T. Tally Jr.

There is a moment in *Don Quixote* where the hero and his squire board an enchanted ship, in reality a small rowboat lacking oars, and set forth to “such longinquous ways and regions” as it may carry them. After floating a few yards downstream, the knight feels certain that they must have traveled at least two thousand miles. “If I only had an astrolabe here with which to take the height of the pole,” he says, “I would tell you how far we have gone; though if I know anything, we have passed, or soon shall pass, the equinoctial line which divides and cuts the opposing poles at equal distance.” In response to Sancho Panza’s question about this “noxious line,” Don Quixote cites Ptolemy, and observes that Sancho knows nothing of “colures, lines, parallels, zodiacs, ecliptics, poles, solstices, equinoxes, planets, signs of the zodiac and points, which are the measures of which the celestial and terrestrial spheres are composed.” In lieu of this scientific body of knowledge, the knight proposes another sure-fire test: “that according to the Spaniards and those who embark at Cadiz to go to the East Indies, one of the signs by which they know that they have passed the equinoctial line I mentioned is that the lice die on everyone about the ship.” Don Quixote entreats Sancho to check his person for lice, and the squire determines with absolutely certainty that they must not have yet crossed the equator, “not by many a long mile.”[1]

As so often occurs in this novel, the humor of the scene lies in the sometimes violent disjunction between reality and appearance, where rowboats can become enchanted ships, roadside inns take the form of grand palaces, or windmills in the shape of giants menace wayfarers with their mighty arms. But the comedy is heightened in this instance by the dual systems of knowledge by which to perceive and analyze the putative “reality” in question. That is, Don Quixote’s reference to Ptolemaic geography and cosmography, complete with an entire vocabulary of

---

[1] Miguel de Cervantes, *Don Quixote*, trans. J.M. Cohen (New York: Penguin, 1950), 657–659.

R.T. Tally Jr. (✉)
Department of English, Texas State University, 601 University Drive, San Marcos,
TX 78666, USA
e-mail: robert.tally@txstate.edu

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_37

scientific terms and concepts, is ultimately supported by what would appear to be a folkloric myth or sailor's fantasy about the disposition of vermin at a certain latitude. The grand abstractions of geometric figures and astrological signs yield to the visceral, earthly experience of lice on a peasant's thigh.

These two modes, the abstract and the experiential, could be said to reflect the narrative modes of the novel itself, which projects in its totality a vast map of the world it simultaneously presents and represents, while also carefully tracing the trajectories of its peregrinating protagonists, whose adventures give flavor—indeed, meaning—to the places and spaces laid out in this literary cartography. This is not just the case in *Don Quixote*, which has long served historians and theorists of the novel as an exemplary model of the form; arguably all novels, if not all narratives, are engaged in one type of mapping project or another. The map-like projection and meticulous description that so frequently characterize the form of the novel interact with the narrative exploration and movement of the plot, making for a spatiotemporal novelistic discourse that serves as a form of knowledge, but also as a form that troubles systems of knowledge, insofar as its imaginative and figurative language can, at times, serve to delegitimize or corrupt official discourses. In its heteroglossia and multiplicity of styles or forms, which Mikhail Bakhtin famously identified as the determining features of novelistic discourse,[2] the novel directly addresses basic concerns of epistemology, while also undermining its own findings.

In this chapter, I examine the novel as a form of spatiotemporal mapping of a world system it both mirrors and constructs. The novel, like the map, is a form a knowledge, registering accumulated information and experience, archiving and categorizing their relatives significance, framing the data and their interpretations, shaping it all into an intelligible array, and projecting potential future formations. The novel thus makes possible a visualization of the world system, as with such classic atlases as Abraham Ortelius's *Theatrum Orbis Terrarum*, except that this "theater" stages historical as well as geographical knowledge, which in turn must include the social, political, and economic forces that give form to the world system as it discloses itself in the novel. Formal techniques and narrative conventions help to define the world's spaces. Correspondingly, from the reader's perspective, matters of scale affect the perception and the interpretation of space and place, while the subject's position within these different scalar diagrams affects her or his ability to recognize their significance. In this sense, the novel's own *theatrum geographicum* sets the stage for a broader consideration of literary cartography of the world system, of which world literature is a protean counterpart.

Ortelius's 1570 atlas, titled *Theatrum Orbis Terrarum* (literally "the theater of the earthly orb," but more simply a "world map"), is among the most influential works of Renaissance art and science. One of the first atlases, collecting and binding in one volume some 70 maps, later expanded to include 167, the *Theatrum* quite literally defined the space of the world for generations. Along with Gerardus

---

[2] See Mikhail Bakhtin, "The Discourse of the Novel," in *The Dialogic Imagination: Four Essays*, trans. Caryl Emerson and Michael Holquist (Austin: University of Texas Press, 1981), 368.

Mercator's 1569 world map, which had employed Mercator's innovative projection, the Ortelius map world map gave form to the continents and seas in a new way, exaggerating the spaces furthest from the equator, while condensing those spaces closest to the line. The resulting misrepresentation of space has had notorious ideological uses and abuses, as the Global South could be actually diminished while the northern territories swell in size and purported value. (Mark Monmonier, in *How to Lie with Maps*,[3] observed that the cartographers and navigators of the British Empire embraced the "flattering" Mercator projection, with its use of Greenwich as the center and its enlargement of Canada, especially.) Ortelius was also the first popular world map to give the continents of the western hemisphere the label *America*, thus solidifying the legacy of Vespucci in this name over and against such rival toponyms as Columbia, New India, and so on. Above all, the new world map depicted the world as a political and geographical system, one that could be synoptically presented by the mapmaker and taken in by the user. The whole world, brought before one's eyes, in a single, theatrical moment.

When Ortelius's *Theatrum Orbis Terrarum* was first published, Miguel de Cervantes Saavedra was a soldier, freshly inducted into the *Infantería de Marina* (the Spanish Naval Infantry or Marines, although the phrase is quixotically suggestive of walking on water). It is not clear whether he was able to test the louse-at-the-equator theory personally, but he would take part in the Battle of Lapanto in Greece, pitting a "Holy League" against the expanding Ottoman Empire. Cervantes was wounded in battle and spent time convalescing in Italy, then he continued fighting elsewhere in the Mediterranean over the next few years. Sometime afterwards, famously, he was taken hostage by pirates and enslaved for five years in Algiers, before returning to Spain. Even if it limited itself to this period, Cervantes's biography already makes for the stuff of adventure novels or romances. His own trajectory from Spain to Italy, thence to the Greek isles and northern Africa, traced a personal itinerary through a key part of the emergent world system, that *Mediterranean* of Fernand Braudel's "geohistory" and Immanuel Wallerstein's sixteenth-century European "core."[4] Even before he began writing his own works, Cervantes's adventures placed him squarely on the map, while indubitably highlighting the crucial differences in the specific places, languages, and cultures of the various stops along his journeys. The "big picture" vision of the world figured on Mercator's and Ortelius's maps undoubtedly influenced the novelist's perspective on the new world into which the heroes of his own novels would move, but his own peripatetic movements, characterized by a good deal of peripety, certainly colored his understanding of those spaces.

The experience of place, as the geographer Yi-Fu Tuan has repeatedly observed, comes down to this fluctuating mixture of movement and rest. Tuan has that "Space

---

[3] Mark Monmonier, *How to Lie with Maps* (Chicago: University of Chicago Press, 1991), 94–99.

[4] See Fernand Braudel, *The Mediterranean and the Mediterranean World in the Age of Phillip II*, trans. Siân Reynolds (New York: Harper & Row, 1972); Immanuel Wallerstein, *The Modern World-System*, 3 volumes (New York: Academic Press, 1974).

is transformed into place when it has acquired definition and meaning," at which point it becomes the subject of interpretation, the traditional purview of literature.[5] Yet the more abstract conception of space, as a largely undifferentiated zone in which the subject moves without awareness or identification of discrete places is also crucial to literary discourses, since the distinction Tuan makes requires a sort of symbolic or representational activity whereby the individual subject connects his or her direct experience to a broader system or structure that, in various ways, gives form to or makes sense of that experience. As Fredric Jameson has argued persuasively, narrative is itself a socially symbolic act by which the writer coordinates the subjective or existential experience with the broader social totality, a national allegory or world system, that makes possible the "truth" of that experience.[6]

The grand world maps and atlases of Ortelius and Mercator were both representative and productive of the age of exploration that witnessed the rise of cartography. One tends to think of mapmaking as an innate, universal, and even "natural" aspect of human understanding of the world, and undoubtedly certain forms of primitive geographical sketches, along with the portulans charts and the medieval T-and-O maps, existed long before the fifteenth century. However, the explosion of ever more and more elaborate maps and charts in this age indicated that a revolution, not only in geography, but in the arts, sciences, and culture for generally, was under way. As Tom Conley has pointed out, "at the beginning of the fifteenth century, maps were practically non-existent, whereas only two centuries later they were the bedrock of most professions and disciplines."[7] The advent of this new age of cartography literally transformed the way we see the world and ourselves in it.

Recent scholarship on the theory and history of the novel has troubled the ease with which critics formerly named *Don Quixote* the first modern "novel" or identified the "rise of the novel" with sixteenth- and seventeenth-century literature,[8] but one may still observe that the novel form rapidly became a dominant genre in both European and world literature during this epoch. Philosophers as diverse as Georg Lukács and Michel Foucault have identified *Don Quixote* as the turning point, and the emergence of the novel as the aesthetic form expressing mankind's

---

[5] Yi-Fu Tuan, *Space and Place: The Perspective of Experience* (Minneapolis: University of Minnesota Press, 1977), 136.

[6] See Fredric Jameson, *The Political Unconscious: Narrative as a Socially Symbolic Act* (Ithaca: Cornell University Press, 1981); see also, Jameson, *Postmodernism, or, the Cultural Logic of Late Capitalism* (Durham: Duke University Press, 1990), especially 410–418.

[7] Tom Conley, *The Self-Made Map: Cartographic Writing in Early Modern France* (Minneapolis: University of Minnesota Press, 1996), 1.

[8] See, e.g., Alexander Beecroft, *An Ecology of World Literature: From Antiquity to the Present Day* (London: Verso, 2015); see also Franco Moretti's enormous editorial project, *Il Romanzo*, a five-volume collection of essays reconstellating the theory and history of the novel in a global context. It appears in English in two volumes as *The Novel, Volume 1: History, Geography, and Culture* and *The Novel, Volume 2: Forms and Themes* (Princeton: Princeton University Press, 2007).

"transcendental homelessness" brings to the fore the fundamentally literary cartographic project of the novel. For Lukács, "The novel is the epic of a world that has been abandoned by God."[9] Whereas the ancient and medieval epic had somehow assumed a clear connection between human experience and the world at large, a metaphysical unity essentially guaranteed by divine providence, the modern condition demands a form that attempts, and likely fails, to make those connections, to project that "archetypal map," as Lukács calls it.[10] The rise of the novel, not surprisingly, corresponds to the rise of cartography.

In *The Order of Things*, Foucault asserts that "*Don Quixote* is the first modern work of literature,"[11] which he explains by distinguishes the novel's epistemology from a Renaissance *episteme* characterizes by similitude. In the Renaissance world not yet abandoned by God, resemblances in nature could disclose the Almighty's signature; thus the natural world could be read like any other text. As Foucault explains, however,

> *Don Quixote* is the negative of the Renaissance world; writing had ceased to be the prose of the world; resemblances and signs had dissolved their former alliance; similitudes have become deceptive and verge upon the visionary or madness; things still remain stubbornly within their ironic identity: they are no longer anything but what they are; words wander off on their own, without content, without resemblance to fill their emptiness; they are no longer the marks of things; they lies sleeping between the pages of books and covered with dust. Magic, which permitted the decipherment of the world by revealing the secret resemblances beneath its signs, is no longer of any use except as an explanation, in terms of madness, of why analogies are always proved false. The erudition that once read nature and books alike as parts of a single text has been relegated to the same category as its own chimeras: lodged in the yellowed pages of books, the signs of language no longer have any value apart from the slender fiction which they represent. The written word and things no longer resemble one another. And between them, Don Quixote wanders off on his own.[12]

After this moment, the novel will have as its vocation the attempt to give some sort of reasonable shape to a world no longer guaranteed of its recognizable contours by a transcendent reality. Like the modern map, which uses figuration, exaggeration, and distortion in attempting to "realistically" represent the spaces on its surface, the novel cannot simply hold up a mirror to reality, but shapes and molds the images, characters, events, and places it represents.

The novel is what Lukács calls "a form-giving form," which also suggests its epistemological role, since the natural or social world it presents cannot simply be known objectively. Knowledge had to become the province of the knower, and the writer cannot be expected merely to reveal the truth, but like the reader must interpret the world. Hence the novel is an essentially epistemological form, and like

---

[9] Georg Lukács, *Theory of the Novel*, trans. Anna Bostock (Cambridge, MA: The MIT Press, 1971), 88.

[10] Ibid., 31.

[11] Michel Foucault, *The Order of Things*, trans. anon. (New York: Vintage, 1973), 48.

[12] Ibid., 47–48.

the map, it is a form of knowledge as well as an attempt to know. Bakhtin makes this very point in contrasting the epic and the novel. Whereas in the epic or ancient literature in general "it is memory, and not knowledge, that serves as the source and power for the creative impulse," writes Bakhtin, "[t]he novel, by contrast, is determined by experience, knowledge and practice (the future)." Bakhtin concludes by saying that, when "the novel became the dominant genre, epistemology became the dominant discipline."[13]

The epistemological or scientific impulse underlying cartographic and novelistic practice should not be taken in a strictly empirical sense. The will to knowledge in such work confronts a persistence of ambiguity that ultimately frustrates, but at the same time sustains, the project. As it becomes apparent that there can be no "true maps," as Jameson noted in a "digression on cartography," since there can be no perfectly mimetic representation of the spaces depicted on them, "it also becomes clear that there can be scientific progress, or better still, a dialectical advance, in the various historical moments of mapmaking."[14] The knowledge to be gained or advanced through these practices thus will remain provisional, tentative, incomplete, and therefore ultimately erroneous, but this means that the epistemic efforts can be directed at producing better maps or narratives, with it also understand that what counts as "better" may vary from time to time and place to place. As I have put it elsewhere, "[i]f failure is inevitable, the goal must be to fail in interesting ways."[15]

A more recent novel, Daniel Kehlmann's *Measuring the World* (2005), explicitly takes up the epistemological and cartographic projects of the modern novel. *Measuring the World* is not exactly a historical novel, but by interweaving the fictional and real lives of mathematician Carl Friedrich Gauss and geographer Alexander van Humboldt, Kehlmann evokes the intellectual fervor of the Goethezeit and its aftermath in Germany and elsewhere. The plot involves the crucial distinction between the scientific methods of these giants as they go about their revolutionary work, each measuring the world, thereby changing it forever, but it vastly different ways. Gauss rarely left his home in Göttingen, conducting the occasional experiment and consulting his telescope, but for the most part his labors involve speculation and deduction. Humboldt, famously, travelled to the Americas, scaling mountains and descending into volcanoes, exploring the Amazon, interviewing indigenous peoples, and always, taking special care to measure everything and record his finding. The abstract mathematical speculation is thus contrasted with the physically intensive empirical exploration. (Somewhat lesser characters in

---

[13] Bakhtin, "Epic and Novel," in *The Dialogic Imagination: Four Essays*, trans. Caryl Emerson and Michael Holquist (Austin: University of Texas Press, 1981), 15.

[14] Jameson, *Postmodernism*, 52.

[15] See my "Translator's Preface: The Timely Emergence of Geocriticism," in Bertrand Westphal, *Geocriticism: Real and Fictional Spaces*, trans. Robert T. Tally Jr. (New York: Palgrave Macmillan, 2011), xi.

the novel, like Gauss's son Eugen who wishes to study languages or Humboldt's gifted brother Friedrich, the philologist and philosopher, provide the barest glimpse of other forms by which we measure our world.) The two approaches, so different yet entirely complementary to the impossible project of the novel's title, also represent the two discursive modes of literary cartography, as the speculative or totalizing abstraction of the map provides the necessary framework for an experiential perambulations of the itinerary, which in turn gives shape, texture, color, and other characteristics to the places figured on the map.

In the end, this mapping affects the territory, which itself conditions the possible way in which its maps can be imagined, and so on. Late in *Measuring the World*, Kehlmann's Gauss thinks about this very thing, only he does so while he is engaged in his own wanderings, and he realizes that the mapping project forever alters the landscapes.

> In the afternoon he [Gauss] took long walks through the woods. Over time he'd ceased to get lost, he knew this area better than anyone, he'd fixed every detail of it on the map. Sometimes it was as if he hadn't just measured the region, but invented it, as if it had only achieved its reality through him. Where once there had been nothing but trees, peat bogs, stones, and grassy mounds, there now was a net of grades, angles, and numbers. Nothing someone had ever measured was now or could ever be the same as before. Gauss wondered if Humboldt would understand that. It began to rain, and he took shelter under a tree. The grass shivered, it smelled of fresh earth, and there was nowhere else he could ever want to be but here.[16]

Gauss's surmises, punctuated by the sensual pleasure and homeliness of the sylvan scene, brings the abstract and the experiential back into amenable relationship to one another.

The realism of a novel, as with the basic practicality of a map, can lead one to miss the intensively figurative, imaginative function of the form. For all their epistemological value, the novel and the map are far better at reminding us of the artificiality of representation, of the trickiness associated with languages and images, and of the potential for these forms to create radically alternatives visions of the world. "While the map is never the reality," the great geographer J. B. Harley once observed, "it helps us to create a different reality."[17] The various anecdotes and examples of this essay suggest the ways in which novels and maps, two exceedingly powerful forms of knowledge, give form to or make sense of the world system presented in and by them. Together, they disclose a *theatrum geographicum* in which the places and spaces of our world are made meaningful.

---

[16] Daniel Kehlmann, *Measuring the World*, trans. Carol Brown Janeway (New York: Vintage, 2006), 229.

[17] J. B. Harley, *The New Nature of Maps: Essays in the History of Cartography*, ed. Paul Laxton (Baltimore: Johns Hopkins University Press, 2001), 168.

Chapter 38
Time, Space, and the Human Geographies of Opportunity

Donald G. Janelle

Abstract Through space-adjusting technologies and social practices, innovation has progressively transformed the space and time constraints on the geographical range and nature of human activities, and on the interactivity of people at local through global scales. This essay explores explicit measurements and representations of the time-space convergence and divergence of places and their impacts on individuals, families, communities, regions, and nations. At such scales, the nature of time-space is differentially experienced, based on gradations and cleavages in human conditions that result from socio-economic systems that differentially allocate wealth, education, infrastructure, and other affordances for breaking down the rigidity and restrictions on mobility and opportunity. The essay addresses questions regarding the measurement of time-space convergence at different geographical scales, considers the implications of convergence processes for human well-being, and explores the use of time-geography concepts to help reduce constraints on opportunities for people to engage successfully and equitably in a world where the significance of distance can change rapidly and with uneven impact on places, regions, and nations.

Keywords Human extensibility • Information-communication technologies (ICT) • Space-adjusting technologies • Time-space convergence • Transportation

38.1 Introduction

The synchronization of interactions among human beings is essential to human survival. Yet, the sequencing of events and processes to secure livelihood are complexly entwined with how individuals and societies have negotiated the

---

D.G. Janelle (✉)
Western University, London, Canada
e-mail: janelle@geog.ucsb.edu

D.G. Janelle
Department of Geography, Center for Spatial Studies, University of California,
Santa Barbara, CA 93106-4060, USA

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_38

meaning of and control over time and space. As dimensions of human behavioral spaces (within homes, workplaces, neighborhoods, cities, regions, and the world), time and space pose fundamental constraints on mobility, location, and accessibility to the necessities of life. This chapter focuses on the role of space-adjusting technologies [2] and on how they interactively align with human spatial behavior and the physical organization of human built environments.

Combinations of *space-adjusting technologies* (i.e., information and communication technologies (ICT), and transportation) and strategic tradeoffs between space and time alter dramatically the potentials for human beings, firms, and other agents to expand opportunities for productive engagement over broader spatial domains (locally, regionally, and globally) and more intensively through time. Conversely, the affordances for socio-economic development offered by ICT create often conflicting challenges to individual autonomy, protection of privacy, national sovereignty, and the value of cultural identity. Indeed, the range of discrepancy in how different regions and social groups take advantage of these technologies and confront their contingent threats to identity is enormous and difficult to grasp. A key to recognizing and resolving such problems relates to fundamental conceptions of time and space—details and examples from a geographical perspective are reviewed in this chapter.

38.2 Time-Space Convergence—Measurements and Perspectives on a Shrinking World

Imagine Toronto and Montreal (500 km) colliding with one another, or Mumbai encroaching on Kolkata (1,168 km). Ridiculous? Yes/maybe, but it's also a matter of perspective. Distances are indeed real, but their measurement is a product of human conception. Having agreed on units of measurement (meters and kilometers, hours and minutes, or some other depiction of the energy and effort required to move from one place to another), it has been apparent for millennia that human beings are not content with the state of their distances to other places. Through invention, investments in new technologies, and expansion of transport infrastructure, some places have successfully lessened the aggregate effort required to achieve levels of utility that far exceed other places. Places that seek economic and political clout have been, and are, on a mission to be as close as possible to all other places in their regions and countries, and in some cases, the world.

Time-space convergence (TSC), the rate at which the time required to move between places changes over historical time, provides a means to measure the results of innovations in transport and communication technologies. TSC is measured as TT1–TT2/(Y2–Y1), where TT1 and TT2 are travel times in different years, Y1 and Y2 [12, 13, 15]. The convergence of Boston–New York, averaging 20 min per year over 210 years, illustrates the significant changes in time-distance that occur in regions of substantial economic growth based in part on transitions in ground transportation technologies from horse-drawn stagecoaches to railway and improved road facilities for automotive and truck transportation (Fig. 38.1). The inflection in the Boston–New York convergence curve in the late 1800s accords with the asymptotic nature of the convergence process as travel-times approach technical, financial, and geographical limits to further gains from transport speed and logistic efficiencies.

Figure 38.2 illustrates general patterns of convergence by railways for pairs of major world cities: Beijing–Shanghai, Mumbai–Delhi, St. Petersburg–Moscow, London–Edinburgh, and Boston–New York. Thus, technological improvements on the St. Petersburg–Moscow line produced an average annual convergence of 6.8 min per year from the introduction of passenger service in 1851 to the inauguration of new high-speed trains by 2010. This corresponds with about 3.0 min per year for rail transportation convergence in similar time frames for London–-Edinburgh and Boston–New York.

For China, regular railway passenger service for Beijing–Shanghai commenced in 1913, requiring about 44 h for the journey. With improved ferry crossings over the Yangtze River, this was reduced to about 36 h in 1933. By the mid-1950s, the trip was down to 28 h. The bridging of the Yangtze in 1968 saw this decline to about 22 h. Engine and track improvements and logistical innovations reduced it further to about 17 h by 1986; but, by 2010, dedicated infrastructure for high-speed trains traveling 300 km/hr reduced the journey to only 4.8 h. The average annual convergence through investment in rail passenger services for 1913–2010 equaled about 24 min per year. Similar investments connecting key economic centers by

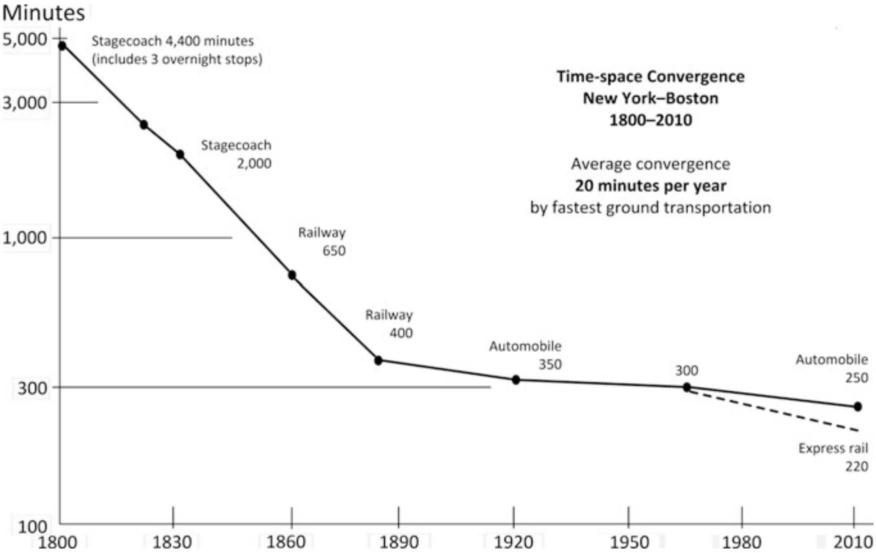


Fig. 38.1 Time-space convergence for Boston–New York, 1800–2010. Calculations and graphic by author from data in multiple public domain sources

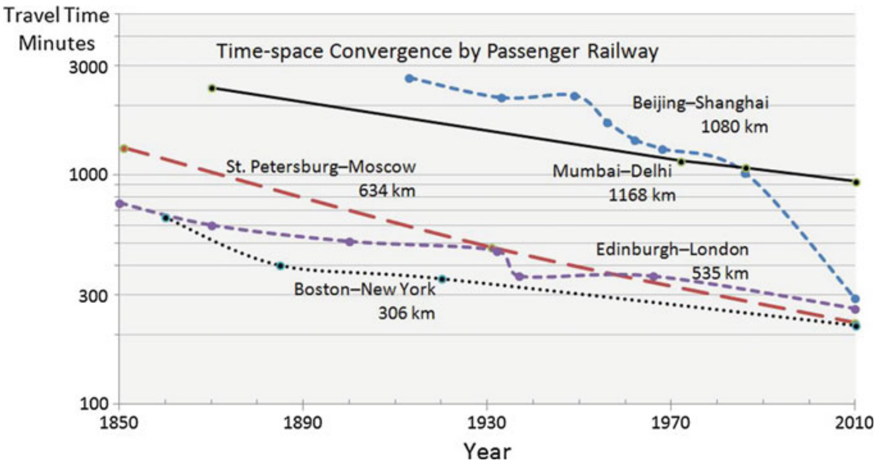


Fig. 38.2 Time-space convergence between major national centers by passenger railway. Calculations and graphic by author from data in multiple public domain sources

highway, rail, and air transportation networks have placed China on a favorable path to competitiveness in the global economy. For rail passenger service, the convergence between Delhi and Mumbai equaled about 10.5 min per year from the early 1870s to 2010. As one of the world's most railway intensive countries, India is currently seeking high-speed rail connectivity along major inter-urban corridors, an initiative to enhance both internal economic integration and access to the global community.

Some of the general properties of the convergence process are described briefly in reference to Fig. 38.3. Curve A is a smoothing out of data points based on known travel-times required for given years, whereas Curve C is the average level of convergence or divergence over a designated time period.

Curve B represents a more literal description of actual behavior and deserves more attention. While some processes of improvement occur gradually through time (road widening, shorter routes, changes in speed limits, better logistical control over resources, engine efficiencies, etc.), convergence tends to occur abruptly based on investments to enhance facilities through minor improvements and major new innovations. Such changes can result in a sudden collapse in distance that irrevocably alters the patterns of opportunities for places joined to, or excluded from, access to the new transportation alternatives. Conversely, failure to maintain facilities or to meet market demands can lead to congestion and time-space divergence. Generally, short of major disasters that sever routes temporarily or permanently, divergence occurs gradually—a situation that is all too common in large congested urban areas.

There is one other feature of the convergence process that is not illustrated in Fig. 38.3, namely that places more distant from one another inevitably converge more rapidly upon each other over time, even if they share the same levels of

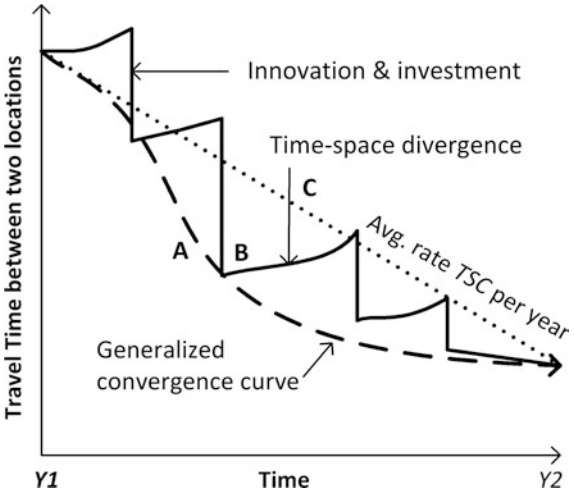


Fig. 38.3 Alternative models for representing time-space convergence. **a** generalized asymptotic convergence curve; **b** step-pattern of convergence to reflect abrupt changes at points of innovation implementation and time-space divergence in periods of neglected maintenance of transportation infrastructure; and **c** average annual rate of convergence over a designated span of years. Graphic by author

technological improvement. The same level of improvement over a longer distance will result in greater levels of convergence than over shorter distances. This inevitability works to the advantage of a country's (and the world's) largest metropolitan and regional centers, conferring upon them favorable cost and access benefits for economic development. For example, given the same levels of technology for their transportation linkages, San Francisco and New York (on the coastal margins of North America—4,135 km) will converge more rapidly than will Houston and Chicago (1,518 km). If travel speeds are increased uniformly between the centers over a 20-year period from 100 km/h to 130 km/h, Houston–Chicago will converge on average 9.0 min per year, whereas the San Francisco–New York convergence will average 28.6 min per year.

Removing the distance bias in the convergence measure is possible. Table 38.1 provides alternative measures based on the examples shown in Fig. 38.2. Thus, the dominant convergence of Beijing–Shanghai according to minutes per year is stripped of the distance bias by measuring the decline as minutes per 100 km or as minutes per 100 km per year. These measures elevate the convergence posture of St. Petersburg–Moscow and illustrate a lag in investment for rail infrastructure serving Edinburgh–London.

Regardless of the measure selected, the examples shown in Table 38.1 reveal uneven patterns of convergence across space and time that pose substantive issues for scientific interpretation and for evaluation of policy responses to implementations of space-adjusting technologies [16]. Uneven convergence among different parts of a city, among places in a region, or cities within a country can have significant implications for the unfolding patterns of economic development, political cohesion, and opportunity.

An unfortunate attribute of many advanced transportation developments is that unevenness in accessibility increases as investment requirements for higher speeds

Table 38.1 Alternative measures of time-space convergence

	Mumbai–Delhi 1870–2010	Beijing–Shanghai 1913–2010	Moscow–St. Petersburg 1851–2010	London–Edinburgh 1850–2010	New York–Boston 1860–2010
Distance (km) shortest path[a]	1,168	1,080	634	535	306
Travel time (minutes) year 1	2,400	2,640	1,305	750	650
Travel time (minutes) year 2	935	288	225	260	219
Reduction in travel time	1,465	2,352	1,080	490	431
Convergence: minutes per year	10.5	24.2	6.8	3.1	2.9
Convergence: minutes per 100 km	125	218	170	92	141
Convergence: minutes per 100 km/year	0.90	2.25	1.07	0.58	0.94

[a] Shortest path is calculated as the great-circle distance between city centers. Actual land-route distances can be significantly longer; the physical shortening of these routes over time would be captured in the measures of convergence

increase. For instance, modern limited-access highways (i.e., freeways and motorways) have very few points of entry when compared to ordinary streets and rural roads. Small communities may have no local entry points and must continue using older systems or travel long distances to enter at freeway interchanges. Similarly, high-speed rail networks benefit large places, while potential users in smaller intervening bypassed communities get to watch trains whiz by at 200 km per hour or more. This pattern also plays out with commercial shipping—large capital-intensive terminals, such as container ports, are few in number even at continental scales, further concentrating the spatial patterns of opportunities for regional economic growth.

Users of maps have traditionally seen the distribution of towns, cities, metropolitan centers, roads, and railways as fixed geometrical patterns that abide by Euclidean axioms. However, when the distance metric of the map shifts to functional measures of effort required to move between places (e.g., travel time or cost, usually over multiple exclusive networks instead of uniform isotropic surfaces), Euclidean rules prove inadequate and the resulting complex and multiple geometries become difficult to visualize. Cartographers and spatial theorists have grappled with how to represent such complexity graphically. It is not possible to review such attempts in this chapter but readers interested in this problem can explore the innovative work of Spiekermann and Wegener [23], Tobler [24], Axhausen, et al. [5], and L'Hostis [18]. A broader social contextualization of time-space convergence is provided by Warf [25].

38.3 Human Extensibility in a Shrinking World

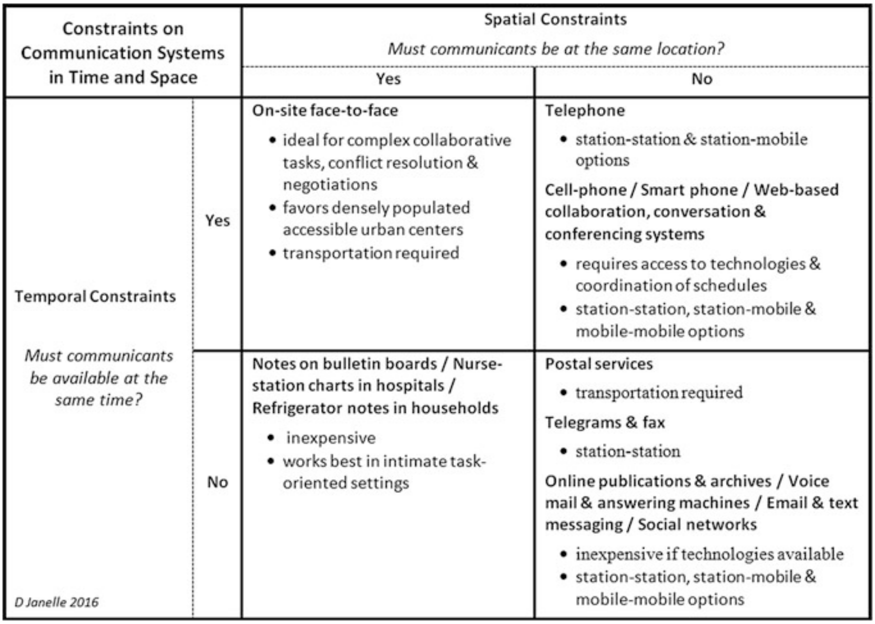
The previous section has focused on characteristics and implications of time-space convergence at levels of cities, regions, and countries. But, it is also important to address explicitly how convergence and divergence processes relate to human well-being and to the daily time-geographies of individuals, families, and communities. At such scales, the nature of time-space is differentially experienced, based on gradations and cleavages in human conditions that result from socioeconomic systems that differentially allocate wealth, education, infrastructure, and other affordances for breaking down the rigidity and restrictions on mobility and opportunity. This shift in scale allows for a focused look at how the time-space shrinkage of functional distances between places can be exploited by individuals, firms, and other institutions to reach beyond their actual physical locations—a concept referred to as *human extensibility* [14].

For much of human history, communication between individuals across space required that at least one of the communicants transport himself or herself to the location of the other person to carry out face-to-face exchanges or that messages be carried and delivered to recipients. Public presentations in town squares and halls have long provided for localized mass communication, postal delivery services have attended to private communications, and newspaper deliveries have addressed the needs for broad dissemination. Transportation is a significant means of such extensibility, but emphasis in recent decades has shifted increasingly toward the importance of ICT in structuring economic and social life, and to the integration of resources that facilitate the production, analysis, evaluation, packaging, distribution, and storage of information for diverse uses.

Cell phones, smart phones, and Internet services create opportunities for engagement with distant locations for family, professional, political, and entertainment purposes. Increasingly, the global Internet and cyber-infrastructure are revolutionizing pathways to communication and creating potentials to greatly democratize opportunities for human space-time extensibility. Online communication offers a conduit for human extension and influence to any geographical scale, doing so without leaving the comforts of home, school, or office——selectively through email and social networks or, less discriminately, with the broadcasting capabilities of blog, wiki, and video archive technologies. Nonetheless, there remain enormous barriers to accessing these tools, limiting use to only those with resources to cover the costs of equipment, licenses, subscriptions, and user fees, and with the technical training, time, and commitment for developing effective applications. Perspectives on both the technical and social issues related to human extensibility and ICT are addressed in publications by Adams [3, 4], and Kwan [17].

Given a plethora of traditional and ICT-based means for communication in the current era, it is useful to consider how individuals and collective social agents select from alternatives. In part, decisions often invoke evaluation of tradeoffs between the needs to be at the same place at the same time against the cost and efficacy associated with different means of communication. To help distinguish the tradeoffs between space and time, Fig. 38.4 provides a categorization of communication tools according to their suitability for overcoming spatial and temporal constraints based on requirements for geographical and/or temporal coincidence among communicants.

Changes in human behavior as a consequence of communication systems becoming more adaptable to human needs in space and time (e.g., the transition from station-based landlines to person-based mobile phones) remain difficult to assess at this early stage. However, scholars have been speculating about the prospects for some time, even before widespread availability of mobile devices and online services. Abler's [1] *polarization conjecture* suggested that an increase in the availability of information from multiple sources coupled with enhanced communication options could result in like-minded people coalescing on the basis of specialized interests and biases, contrary to expectations that human extensibility might promote a more uniform world view. In 2004, Couclelis advanced her hypothesis of *activity fragmentation*, suggesting that as ICTs become increasingly widespread, activities will become more fragmented in space and time and more person-based than space-based, resulting in situations where the time, duration, and synchronization of human activities in space will become less predictable and more autonomous. In combination with time-space convergence, it is also likely that increased access to



Constraints on Communication Systems in Time and Space		Spatial Constraints *Must communicants be at the same location?*	
		Yes	No
Temporal Constraints *Must communicants be available at the same time?*	Yes	**On-site face-to-face** • ideal for complex collaborative tasks, conflict resolution & negotiations • favors densely populated accessible urban centers • transportation required	**Telephone** • station-station & station-mobile options **Cell-phone / Smart phone / Web-based collaboration, conversation & conferencing systems** • requires access to technologies & coordination of schedules • station-station, station-mobile & mobile-mobile options
	No	**Notes on bulletin boards / Nurse-station charts in hospitals / Refrigerator notes in households** • inexpensive • works best in intimate task-oriented settings	**Postal services** • transportation required **Telegrams & fax** • station-station **Online publications & archives / Voice mail & answering machines / Email & text messaging / Social networks** • inexpensive if technologies available • station-station, station-mobile & mobile-mobile options

D Janelle 2016

Fig. 38.4 Communication technologies and their tradeoffs in meeting the temporal and spatial constraints of communicants. *Source* author

information resources and virtual connectivity through online services could build awareness to opportunities elsewhere in the world that increase the desire for, and the likelihood of, travel and commerce among distant and dispersed locations, continuing the growth in per capita global consumption of distance.

38.4 A Time-Geography Model of Society

Swedish geographer Torsten Hägerstrand (1916–2004), a pioneer in the study of innovation diffusion and the originator of *time geography* [11], saw clearly how the human structuring of time and space can impose constraints on the abilities of people to engage successfully and equitably in a world where new technologies and social practices open up possibilities or pose threats to the sustainability of local and global environments. Time geography provides a set of concepts and a graphical notation system that integrates time and space and that contextualizes the situations that people encounter in their interdependent activities of normal daily life and in projects over extended periods, even a lifetime. Individual human beings are the basic building blocks of social systems and individual activity behavior is a process, graphically encoded as a *time-space path* (Fig. 38.5a). These paths encompass the set of activity and travel episodes of one or more people over a 24-hour cycle, but

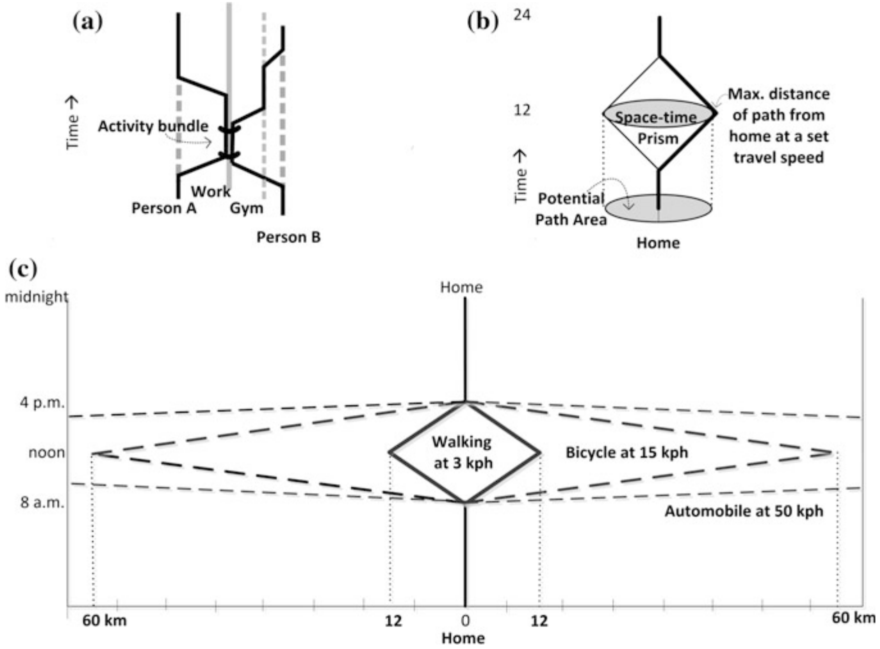


Fig. 38.5 Basic elements of time geography. **a** Daily space-time paths and activity bundles; **b** space-time prism and the maximum potential path area as constrained by the need to be at home during designated periods for personal maintenance (e.g., sleep and household obligations) and the speed of movement along one's time-space path; and **c** the maximum extent of space-time prisms and potential activity areas associated with different levels of personal mobility and travel mode. Graphic by author

are extendable as *lifelines* to cover a lifetime. In addition, they are extendable across space to encompass larger regions and broader social systems. Although this discussion emphasizes the paths of human beings, time geography also embraces the movement and ecological realms of material objects and physical resources that are essential for fulfilling many human endeavors.

Time-space paths are constrained by what Hägerstrand referred to as coupling, authority, and capability constraints.

- *Coupling constraints* refer primarily to social ties (e.g., the need to join with others to carry out activities (e.g., a game of basketball, a work task that requires the sharing of expertise, and essential family obligations, such as caring for a child, etc.). An *activity bundle* designates a coupling event involving the deliberate engagement of two or more individuals for carrying out a joint task for a prescribed time at the same location or *station* (e.g., workplace, gym, home).
- *Authority constraints* emphasize controls over access to space (e.g., laws regarding ownership of property, the designated service hours of banks and

other facilities, and the design of built environments that create physical barriers or restricted gateways to activity spaces).

- *Capability constraints* concern human biological needs for daily maintenance (food, sleep, hygiene, etc.), access to means (and speed) of movement, and the social capital, personal income, and talents to take advantage of opportunities.

The *space-time prism* (Fig. 38.5b) delimits the outer-bounds of movement possibilities and activity choices over a specified time period, usually constrained by speed of movement and by coupling constraints (e.g., the obligation to be at home during certain hours). The prism determines the *potential path area*, a geographical area that expands with faster means of travel, opening up a broader range of opportunities in time and space (Fig. 38.5c). However, participation in activities that consume time at stations limits the future geographical extent of activities within the specified time period. More complete time-geography representations can model the tradeoffs that individuals, households, businesses, and institutions make in allocating time for on-site activities versus travel through space.

Time geography constitutes a way of thinking about problems and potential solutions. Applications in the early 1970s were mostly conceptual but, increasingly, computational capabilities, new data sources (especially from geo-spatially referenced data generated through online social networks), and visualization capabilities (e.g., geographic information systems (GIS) and visual animation tools) have supported expanded research applications in many fields. Examples are numerous in the areas of transportation and land use planning—e.g., evaluating policy options according to how they might enhance or inhibit the range of choices and opportunities for households in different urban neighborhoods or rural communities, assessing how social benefits might accrue to households from implementation of flexible work schedules or telecommuting options; considerations of how land use patterns in cities encourage or inhibit pedestrian activities and healthy exercise; and search and surveillance strategies for police protection of neighborhoods and apprehension of criminals.

All of these examples invoke an understanding of, or a manipulation of, the temporal and spatial dimensions of society. As such, they constitute negotiations over the meaning of space and time with regard to possible conflicting values of different entities, such as households, private firms, public institutions, and nations, and interests (e.g., rich and poor, labor and management, rural and urban, majorities and minorities, and other possible categorizations) that divide the human community.

Many areas of application of time geography extend spatially beyond local regions to national and global scales and temporally to document and reason about human behavioral processes in the distant past and plausible futures. Thus, human extensibility and time-space convergence processes weigh prominently in how problems are recognized and addressed. Examples include predicting locations of archaeological sites and the patterns of movement within ancient settlement systems; studying economic- and social-benefit tradeoffs between travel events and virtual substitutes via online services; exploring disease pathways through social networks in epidemiology; the formation and role of invisible colleges in the history of science; investigating regional and global migration patterns and diasporas; and the extension of markets and governance systems across space. Time-geography also lends itself as a supplement to other established research methodologies in the social sciences and humanities, such as event analysis in historical studies, externalities in economics, and life-cycle and cohort analyses in demography and related branches of sociology.

Readers interested in learning more about the theoretical foundations and extensions of time geography are encouraged to consult research by Burns [7], Miller [19, 20], Yu and Shaw [26], Shaw and Yu [22], Miller and Bridwell [21], Buliung [6], and Ellegård and Svedin [9].

38.5 Conclusions

This chapter has presented geographical perspectives on time and space as basic dimensions of the human experience, focusing on the role of space-adjusting technologies in expanding access to resources in support of livelihood and opportunities for socio-economic development. The basic utility of a place on Earth is a function, in part, of the ability of its residents to access and communicate with others elsewhere—locally, regionally, and globally. In an ideal world, we might hope that such abilities would be distributed equitably and that access and communication would be achievable through minimal expenditures of time and financial resources. As seen in interrelated discussions about processes of time-space convergence and human extensibility, transportation, information, and communication technologies have enabled substantial mediation of the spatial and temporal constraints on mobility and human interactions. But, on the other hand, these processes promote spatial concentrations of development that can aggravate levels of disparity at local scales within cities and regions, and at national and global scales. It is argued that greater documentation and understanding of convergence and extensibility processes are required in order to better direct their consequences toward more balanced distributions of opportunity.

Negotiations over the meaning of time and space are central to modifying the basic capability, coupling, and authority constraints over time-space paths, space-time prisms, and the projects that people and institutions engage in over their lifespans. Society is still in an early phase of integrating ICT into the functioning of built environments and local communities, yet this technology is critical to human extensibility processes that can be channeled into breaking down constraints that currently limit opportunities for jobs and fulfilling lifestyles for many subpopulations (e.g., those with physical disabilities and impaired mobility). An emphasis in education on critically informed uses of ICT can alert populations to the opportunities and security threats of integrating online resources and ICT tools into daily activities and to extending professional and other interactions beyond local communities through virtual contacts in cyberspace.

ICT and transportation are continually evolving new forms and capabilities that will alter the meaning of space and time in the context of time-geography. A look at current and emergent practices is instructive. Global positioning systems (GPS) and tracking technologies are currently common features of transportation vehicles and mobile phones, and geo-spatial information is increasingly guiding ICT and transportation operations. Location-based services (LBS) exploit information on the locations of users in real-time for marketing purposes. These technologies can also alert one to the locations of friends in the immediate vicinity and they can provide critical information to the public during emergencies. They have also enabled the creation of new specialized information resources (e.g., details on bird sightings for bird watchers) and databases through *volunteered geographic information* (VGI) [10]. In essence, dense networks of mobile sensors, currently in place, provide potentials for documenting in real-time the locations and movements of vehicles and humans at local through global scales. Adding to the mix, autonomous self-driving vehicles are on the horizon and air-borne drones are in use for surveillance, parcel deliveries, and agricultural applications. The time-geography landscape is unfolding in ways to enable new human activity patterns, restructuring of spatial economies, and continued processes of time-space convergence and human extensibility. Yet, humankind remains in search of ways to share opportunities for people to engage successfully and equitably in a world where the significance of distance can change rapidly and with uneven impact on places, regions, and nations.

References

1. Abler, R. (1973). Monoculture or miniculture? The impact of communications media on culture in space. In D. Lanegran & R. Palm (Eds.), *An invitation to geography* (pp. 186–195). New York: McGraw-Hill.
2. Abler, R. F. (1975). Effects of space-adjusting technologies on the human geography of the future. In R. Abler, D. Janelle, A. Philbrick, & J. Sommer (Eds.), *Human geography in a shrinking world* (pp. 35–56). North Scituate, MA, USA: Duxbury Press.
3. Adams, P. C. (1995). A reconsideration of personal boundaries in space-time. *Annals of the Association of American Geographers, 85*(2), 267–285.
4. Adams, P. C. (2000). Application of a CAD-based accessibility model. In D. G. Janelle & D. Hodge (Eds.), *Information, place, and cyberspace: issues in accessibility* (pp. 217–239). Berlin: Springer.
5. Axhausen, K. W., Dolci, C., Fröhlich, Ph., Scherer, M., & Carosio, A. (2008). Constructing time scaled maps: Switzerland 1950–2000. *Transport Reviews, 28*(3), 391–413.
6. Buliung, R. N. (2011). Wired people in wired places: stories about machines and the geography of activity. *Annals of the Association of American Geographers, 101*(6), 1365–1381.
7. Burns, L. (1979). *Transportation, temporal and spatial components of accessibility*. Lexington, MA, USA: Lexington Books.
8. Couclelis, H. (2004). Pizza over the internet: E-commerce, the fragmentation of activity, and the tyranny of the region. *Entrepreneurship & Regional Development, 16*(1), 41–54.
9. Ellegård, K., & Svedin, U. (2012). Torsten Hägerstrand's time-geography as the cradle of the activity approach in transport geography. *Journal of Transport Geography, 23*, 17–25.

10. Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal, 69*(4), 211–221.
11. Hägerstrand, T. (1970). What about people in regional science? *Papers of the Regional Science Association, 24*, 1–12.
12. Janelle, D. G. (1968). Central place development in a time-space framework. *The Professional Geographer, 20*, 5–10.
13. Janelle, D. G. (1969). Spatial reorganization: A model and concept. *Annals of the Association of American Geographers, 59*, 348–364.
14. Janelle, D. G. (1973). Measuring human extensibility in a shrinking world. *Journal of Geography, 72*, 8–15.
15. Janelle, D. G. (2014). Time-space convergence. In R. J. Stimson (Ed.), *Handbook of research methods and applications in spatially integrated social science* (pp. 43–60). Cheltenham, UK: Edward Elgar.
16. Knowles, R. D. (2006). Transport shaping space: Differential collapse in time-space. *Journal of Transport Geography, 14*, 407–425.
17. Kwan, M.-P. (2000). Human extensibility and individual hybrid-accessibility in space-time: A multi-scale representation using GIS. In D. G. Janelle & D. Hodge (Eds.), *Information, place, and cyberspace: Issues in accessibility* (pp. 241–256). Berlin: Springer.
18. L'Hostis, A. (2009). The shriveled USA: Representing time–space in the context of metropolitanization and the development of high-speed transport. *Journal of Transport Geography, 17*(6), 433–439.
19. Miller, H. J. (1991). Modeling accessibility using space-time prism concepts within geographical information systems. *International Journal of Geographical Information Systems, 5*, 287–301.
20. Miller, H. J. (2005). A measurement theory for time geography. *Geographical Analysis, 37*(1), 17–45.
21. Miller, H. J., & Bridwell, S. A. (2009). A field-based theory for time geography. *Annals of the Association of American Geographers, 99*, 49–75.
22. Shaw, S.-L., & Yu, H. (2009). A GIS-based time-geographic approach of studying individual activities and interactions in a hybrid physical-virtual space. *Journal of Transport Geography, 17*(2), 141–149.
23. Spiekermann, K., & Wegener, M. (1994). The shrinking continent: New time-space maps of Europe. *Environment and Planning B: Planning and Design, 21*(6), 653–673.
24. Tobler, W. (2001). The world is shriveling as it shrinks, presentation at Texas A&M University, 26 October, http://www.geog.ucsb.edu/~tobler/presentations/. Accessed 2 May 2016.
25. Warf, B. (2008). *Time-space compression: Historical geographies*. New York: Routledge.
26. Yu, H., & Shaw, S.-L. (2007). Revisiting Hägerstrand's time geographic framework for individual activities in the age of instant access. In H. Miller (Ed.), *Societies and cities in the age of instant access* (pp. 103–118). Dordrecht, NL: Springer Science.

Author Biography

Donald G. Janelle's career as a geographer and academic administrator spans five decades, with appointments at the U.S. Air Force Academy, the University of Western Ontario, and the University of California Santa Barbara. He is currently Professor Emeritus (Western University) and Researcher Emeritus (UCSB). Janelle's research and publications are based broadly within geography and affiliated social and behavioral sciences. Primary themes include space-time analyses of individual behavior, the time-geography of cities, the temporal-spatial ordering of social systems, locational conflict analysis, social issues in transportation, and the role of space-adjusting technologies in structuring new patterns of social and economic organization. He is a recipient of the *Edward L. Ullman Award for Career Contributions to Transportation Geography* and the *Ronald F. Abler Honors Award for Distinguished Service* from the Association of American Geographers. Janelle has authored more than 130 journal articles and book chapters, and has co-edited six books. Recent titles include: *Information, Place, and Cyberspace: Issues in Accessibility (2000, Springer-Verlag); Spatially Integrated Social Science (2004, Oxford University Press); and Space in Mind. Concepts for Spatial Learning and Education (2014, MIT Press).*

Chapter 39
Losing Time and Space: Experiencing Immersion

Diana J. Reichenbach

Abstract The sense of immersion—the feeling of losing track of time and space—is something all of us have experienced. Whether it is through a film, an interactive game, or hiking in a landscape, these immersive experiences have been sought after since the beginning of humankind to open our minds to new possibilities and expand perceptions of our world. This article explores the role of art in crafting immersive experiences and how a variety of mediums activate our senses in very different ways.

39.1 Introduction

> The picture that located the observer in a virtual environment can arouse a powerful experience. There are metaphors to describe it: one is taken out of oneself, one is transported… It is not an illusion of reality that is induced in these pictures, but an awareness of being in the world. This is no illusion. [5, p. 232].

In his article, "The Ecological Approach to the Visual Perception of Pictures," James T. Gibson studies how virtual depictions of an environment influence the way we see them. What he describes is a moment of immersion for the viewer of an image; a moment that transports them from the current time and space and results in an awareness that previously was not perceived. Two elements of this statement: (1) the feeling of being "transported," and (2) the resulting "awareness of the world," will be recurring themes of this article. These are the elements that define the immersive experience of losing time and space and the resulting change in the way one thinks.

Immersive experiences can be long or short form. *Flow*, a term coined by positive psychologist Mihály Csíkszentmihályi, is a longer form immersive experience in which a person is engaged in, "a state of heightened focus in activities

D.J. Reichenbach (✉)
Savannah College of Art and Design,
3515 Montgomery Street, Savannah, GA 31405-3219, USA
e-mail: diana.reichenbach@gmail.com

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5_39

such as art, play and work." Csíkszentmihályi's theory asserts a person has a limited amount of information they can process per second, and that to achieve flow requires a temporary suspension of existence [4]. Flow is also commonly referred to as "being in the zone."

A shorter form immersive experience, and perhaps more common, can be described as an *insight*, or a, "sudden awareness of a complex process of interaction." [8, p. 90]. In *Interactive Storytelling,* game designer Chris Crawford explains these insights require a reshaping of our information network, resulting in a radical change of thought [3, p. 15].

In this article I would like to explore: (1) why immersive experiences are important, (2) how these experiences relate to art and technology, and (3) ways an immersive experience can be crafted to reach others.

39.2 Why Are Immersive Experiences Important?

Recently I visited Gettysburg Battlefield, the site of the legendary battle between the Union and the Confederacy and a turning point in the American Civil War. Like many children in America, I grew up reading about this battle in school and had seen it depicted in Hollywood films and documentary reenactments. It wasn't until visiting the site in person, however, that I had an experience that really changed the way I looked at the battle.

I was standing on the field that day looking at a panoramic painting. This image was depicting the same scenery in front of me, but with the rage of war inflicted upon it. I glanced at the painting and then the landscape behind it, identifying the same hills and slant of the horizon line. Then I felt something. It was as if in that moment the frame of the image faded away and the wounded soldiers, the fallen horses, and the exploding cannons had taken a presence, once again, in that space. At that moment, I could feel the sorrow and destruction that had visited this now very peaceful place. The previously abstract event of "Gettysburg" now had a very personal and emotional familiarity. I could see that day as clearly as any other in my own life, and although we were separated by centuries, these soldiers now had a tangible humanity.

It can be difficult to place yourself in a space and time that feels so disconnected and foreign from your own; that moment needs translation into a personal form to take relevance. It is during moments of immersion that these connections take place. This is the importance of an immersive experience: each new connection made provides increased empathy and broadens our personal understanding.

So why aren't we constantly in a state of immersion? Our nervous system is only capable of processing approximately 110 bits of information per second (as a point of reference, being in a conversation requires 60 bits per second) and immediate concerns related to work, family, and basic needs like eating take up most or all of our available processing capacity [4]. Although this explanation is very calculated, we can all recognize that our ability to focus has a limit, which is why we shift our attention from one thing to another. In order to break us out of these routines and expand our thinking a trigger is needed, and this trigger can take many forms. One of these forms is through exploiting the limits of our physical senses.

39.3 Modes of Physical Perception

Before we look at how immersive experiences are created, it's important to understand how time and space are interpreted through our physical senses. Each innovation in art and technology is the result of a new discovery of an ability or, just as importantly, a limitation which affects how we perceive our surroundings. In this section, I will focus specifically on how depth, color and light, and time are perceived and how those elements are conveyed through art and technology.

39.3.1 Depth

A few years ago, I was teaching a University level image-making course to a mixture of art and science students. One exercise in the course required students to manipulate color values and scale in Photoshop to create a landscape scene. After a few frustrating attempts, a computer programming major reached out for help and we reviewed his work during a meeting. His image depicted an apple tree and a house, but there was no depth to the image. For starters, the tree and the house were similar sizes. Furthermore, all of the colors in his scene had the same value: the grass and leaves were bright green, the apples were bright red, and the house was a bright yellow. Lastly, every color was solid with no gradation in value from edge to edge.

When I asked the student how the different objects in his scene—a tree, a house, and a field—related to each other in terms of depth, he explained that the house was supposed to be far off in the distance with the tree in the foreground. "Let's work with the house first," I said, "and see how making it bigger or smaller affects its depth in the scene." As I started to manipulate the house to make it smaller, I could sense that he was starting to understand the concept. When the house was significantly smaller than the tree, he stopped me.

"Right there! Wow, that really makes a difference."

"Next," I said, "let's change the color of the grass so that some of it appears closer and some appears farther away." I created a gradient for the ground, making the grass near the tree in the foreground a dark green and with the color growing increasingly brighter as it got closer to the house. "Finally," I said, "there is no sky in this image. Let's create a horizon line so we can orient where the ground meets the sky." His eyes widened.

"That looks amazing!"

I could sense that at that moment the lesson really clicked. His work improved over the course of the semester and he was able to critically evaluate visual cues in an image and apply them to his own work. That had been his “Ah-ha!” moment.

I bring up this example because being able to recognize or create a three-dimensional space on a two-dimensional surface is not something we can inherently do. In fact, it is physically impossible for us to see true three-dimensional space; the image captured by the retina in each eye is two-dimensional, with the third dimension being created by the stereoscopic effect of our binocular vision [1, p. 88].

Recreating this sense of depth had long challenged artists. It wasn’t until the fifteenth century, when fixed-point perspective was introduced, that we were able to communicate three-dimensional attributes successfully [1, p. 88]. Fixed-point perspective and other visual techniques use cues from our own perception to communicate spatial relationships in a scene. For example, in animation, depth is often created in a scene by designing the foreground, middle ground, and background as separate elements. In the foreground, elements appear larger, have a darker value, and will move faster as we pass by them. Elements in the background tend to be lighter, smaller, have less contrast, and will move more slowly in the frame as you travel through a scene.

Stereoscopic filmmaking, on the other hand, creates depth not from a perceptual cue but by recreating our binocular vision. In stereoscopic filmmaking, two cameras, each imitating what the left and right eyes of a person would see, capture the on-screen subject. Depth is created by the relationship of an object between the left image and the right image; the greater the difference in its position between the left and right image, the greater depth the object will appear to have in the stereoscopic frame. Like fixed-point perspective, however, this method does not create a true depth, but rather relies on the brain to stitch the two separate images together into a cohesive whole. Conclusively, if depth is a function of our mind putting together the two images received from our eyes, do different individuals perceive different depths? Is there a “true depth”?

An additional nuance is that not everyone interprets information from images in the same way. In the example with my student, although he could identify that the image lacked depth he could not identify why. Donald Hoffmann, in his book, *Visual Intelligence*, cites a case study about a boy who, after surgery, is able to see after being blind his whole life from congenital cataracts.

During the study, the boy had problems distinguishing between a cat and a dog until he was able to identify the animal through touch. In another instance, it was observed that it took months for him to recognize two-dimensional images as representing solid structures, considering them to be only abstract, flat planes comprised of colors. Then after realizing their representation, he was surprised that although they looked “round and uneven” that they were still flat to the touch [6, pp. 17–19].

These examples illustrate that visual depictions are not complete replications of reality, but rather that they simply carry information to communicate attributes of it. The ability to recognize visual cues and relate them to spatial depth required a moment of insight. Through moments of immersion, starting at a young age, we gain the ability to receive these new discoveries and expand our visual perception.

39.3.2 Light and Color

During my time in graduate school I experienced a transformative moment relating to light and color. I was taking a video course and we were discussing the additive color system: the mode of color used in light-based technologies such as projectors and television monitors. To demonstrate the additive color system, the instructor held color gels—thin layers of laminate that change the color of a light when positioned in front of it—up to a spot light projecting onto a nearby wall. He had three gels on hand: one red, one blue, and one green. "These are the primary colors of the additive light system," he announced.

"Green?" I thought. "Aren't the primary colors red, blue, and *yellow*?"

After holding each gel up in front of the light, reproducing the red, blue, and green colors as expected, he started to mix the colors. "Secondary colors," he said, "are created when the primary colors are mixed. When red and blue mix, it creates magenta, and when blue and green mix we get cyan."

"This makes sense," I thought, "magenta has some red and blue in it, just as cyan has a bit of green and blue."

"Finally," he continued, "red and green light combine to make yellow." He held the red and green gels up against the light and, to my surprise, in the space where the gels overlapped the light was indeed yellow.

This moment inspired further research into how light is processed by the human eye. When light waves hit the back of your retina, it activates one of two types of receptor cells, rods or cones. Rods are active in low light situations and consist of one type only. Cones become active in well-lit environments and come in three types that correspond to the colors red, green, and blue [7].

So what about yellow? Actually, we cannot perceive true yellow, as the color we see is the result of both our red and green cones being stimulated to create the yellow color we perceive. This is true of all non-red, green, and blue colors; those three cones, stimulated in different proportions, create the millions of colors we perceive. This phenomenon is completely perceptual; if our sight were engineered differently (having instead blue, green, and ultraviolet cones) how might our perception change?

Just as the discovery of depth perception cues inspired fixed-point perspective and stereoscopic imaging, discoveries in how we perceive light and color have influenced our technology. All light-based technology (such as video, computer monitors, and digital cameras) operate using the same RGB color system as our eyes, creating all of the colors you see using varied mixtures of red, green, and blue. By extending our physical sense of light and color into our technology we continue to create images replicating our own reality.

39.3.3 Time

Unlike depth and color, time is a much more difficult concept to perceive in a physical sense. One can argue that we can detect time through vision, such as when the changing direction and color of sunlight can signal its passage and duration [2, p. 312]. In this way, the change in the light is perceived through the contrast of the images we observe over time; in a sense we say that we *move* from day to night. Is time perceived through movement?

I will not delve into the concept of motion perception as an entire field of study; however, I will touch upon the topic of *persistence of vision*, the phenomenon that enables us to perceive motion in video and animation. To illustrate persistence of vision for yourself, hold a pencil or pen in your hand and use it to tap on the table at a moderate pace. Focus on the tip of the pen and you'll notice a streaking effect, which, in video, we call *motion blur*. What causes this blurring effect? If the tip of the pen is only in one position at any given time, how do we see it in more than one place?

Similar to depth and color perception, our eyes and brain have limits in terms of deciphering visual information related to movement and time. When an image reaches the back of our retina it takes a moment to fade as the new image is processed. Motion pictures exploit this latency to create a sense of movement, as it is presently impossible to display the billions of images required to fill every nanosecond of a film. In fact, the current frame rate for most motion pictures is much lower, at only twenty-four images per second. It is through persistence of vision that we are able to stitch these separate images together into one cohesive whole. Furthermore, this is another example of creating technology that imitates and leverages our perceptual attributes.

Some of the first motion pictures documented everyday events of the human experience, for example the famous Lumière film, *L'arrivée d'un train en gare de La Ciotat (The Arrival of a Train at La Ciotat Station).* This short film depicted a train arriving at La Ciotat Station in France, and was shot from a human perspective: the lens provided a field of vision similar to that of the naked eye, the camera was set at eye level, and the camera was static—no movement—during the entirety of the shot.

In the century to follow however, experimentation in camera techniques and editing expanded the ability of the moving image to communicate. We saw the rise of the montage, a technique for condensing the passage of time, and the extreme close-up, which brought us closer to a subject than we could achieve in our physical lives. Through varying the playback rate, time could move forward, backward, and at unnatural speeds (like slow motion), and spatial relations could be communicated by sequencing a series shots from different angles as opposed to showing the entire scene in one image.

Animation pushed this experimentation further still by enabling abstract characters to exude human characteristics and for stories to be told through movement and rhythm of color, line, and texture. The addition of sound unified the visual and acoustic spaces, an attribute that distinguished it from previous art forms such as painting, literature, music, and others.

By leveraging the physical anomaly of seeing sequential images as movement, we opened a new door of possibilities to how we perceive the world. As with depth and light and color, we used newfound knowledge of the human perception of time to develop re-creations of our physical experience. Through expanded technological advancements in the manipulation of time and space, new mediums of communication have been developed to enhance our perception and trigger immersive experiences.

39.4 How Are Immersive Experiences Achieved?

39.4.1 Breaking the Frame

Thinking "outside the box" is a metaphor that is commonly used to describe the ability to see something from a new perspective. In relation to this essay, it represents the transformative moment of awareness experienced through immersion. While these moments of immersion grant us opportunities to see outside the box, there are barriers in our daily lives that keep us from this expanded perspective, many of which are literally box-like in shape.

Our workspaces, our modes of transportation, and many of our mediums of expression have embedded the presence of "the box." Have a look around. As I type this I am looking at my box-shaped screen, with a "window" open to a box-shaped text document, and I am sitting in a box-shaped room filled with rectangular books, shelves, tables, and chairs. In fact, the most prominent non-box-shaped object in the room is a circular clock; what does that say of how we perceive space versus how we perceive time?

In the theory that, "The medium is the message," Marshall McLuhan suggests that it's not necessarily the content, but rather the form that influences how a message is perceived, and ultimately, that each new form of communication shifts the way in which a culture perceives its surroundings. He also suggests that in each new situation and technology is the "ghost of the old one," in that embedded into these new methods are artifacts of how we previously perceived our experiences [8 p. 83].

In alignment with McLuhan's research, I believe we are in an age where we have achieved the medium, namely virtual and augmented reality, to break outside of the box perceptually. However this medium, as with all other mediums, began by reflecting the past. For example, when cyberspace first became accessible to the public, this *virtual* space integrated many attributes of our *physical* space. When you opened a portal to the Internet, you opened a new "window." When you entered a space to connect with others you entered a "chat room." Although cyberspace has expanded beyond these initial boundaries, it is still restrained by the spatial qualities of the screen you access it through. The current shift toward virtual and augmented reality has the power to break the frame, immersing a viewer inside the medium itself. Of course, content will always play a role in communicating a message and experience to the viewer, but how will this new medium change the way in which we perceive? How will this change our awareness of our environment?

39.4.2 Immersive Mediums

Much of my personal artwork explores breaking the frame, and in the past several years I have focused on creating films for the fulldome, or planetarium, theater. A predecessor to virtual reality technologies, this venue has offered a space for experiencing 360-degree immersive films, breaking the traditional frame of filmmaking.

The planetarium was originally designed to mimic the experience of stargazing, situating the audience in an immersive physical situation. Through its hemispherical design, once the show begins the walls of the theater seem to disappear, simulating a nighttime sky. Much like the early Lumière film at the train station, with many planetarium shows the viewer's perspective is static and "natural" as it simply recreates the act of being outside on a perfect night for stargazing. As advancements were made in projection technology, however, this convex surface could host films that moved the viewer through other environments, both abstract and physical.

In his paper, "The Language of Immersive Cinema," Daniel Neafus, co-founder of IMERSA, an organization promoting the advancement of immersive technologies, explains what this new medium means for viewer experience:

> When our audiences are immersed in the seamless enveloping sensations of fulldome, they can become involved and truly engrossed in the experience. This intellectually stimulating, virtual reality can fully occupy the attention of participants both physically and emotionally. Audiences today appear to be moving beyond their passive observation of framed media and are discovering new ways to participate in immersive storytelling [9, p. 2].

Immersive mediums, with their capacity to engage the full attention of their audiences, have potential to trigger moments of immersion for viewers. These mediums also have the capacity to expand our perceptual boundaries. In his article, "Towards an Immersive Intelligence," author Joseph Nechvatal suggests that extending the field of view to the peripheral not only affects how we view the frame, but also has power to focus our point of view *beyond* the edges of our current concept of time, space, and perception [10, 418]. If this is true, can a heightened awareness of one's own perception also lead to an ability to control it?

39.4.3 Abstract, but not Too Abstract

One last topic I'd like to touch upon is abstraction, another powerful tool that can trigger an immersive experience or moment of insight. In sharing my own films with audiences, I have noticed that when the content is too abstract they have a difficult time connecting to the film. As it is a natural tendency for people to want to connect their present experience to a past one, many viewers lose interest when what they see and hear has no point of reference to their own personal history.

After creating a handful of films that strictly explored combinations of color, light, and texture in very abstract ways, I decided to put a tangible image in my next film. That film, *A Moment of Silence*, starts as a journey through a forest from the perspective of a traveller. In the beginning of the film, trees move through the frame, utilizing depth, light, and color cues, to recreate the experience in a natural way. Slowly, the forest is infiltrated by abstract imagery until suddenly it bursts into complete abstraction of light, color, and movement. Towards the end of the film, the abstract imagery gradually fades away and the forest returns, restoring the natural viewpoint of the traveller.

The reaction to this film was overwhelmingly positive, and many people commented that they felt transported during the screening. What made this film more effective than the others was that it provided viewers with a familiar point of reference; in this case, the trees and how they moved naturally through the frame. This grounding in real-world experience gave the abstract section that followed more prominence. It was only after viewers could place themselves in the film that they were willing to be challenged with something outside of their viewpoint. Through my experience, I've found that as long as this jump from reality to abstraction is not too far, most viewers are willing to go along with it.

39.5 Conclusion

Immersive experiences challenge our perceptions. These experiences shape how we engage with one another and our surrounding environment. Shifts in perception craft new ways of communicating and bring us closer to understanding ourselves beyond our physical sense.

Examples in this article have demonstrated that by exposing the abilities and limitations of human perception we are able to conceptualize beyond them. With each new discovery we advance the methods used to replicate this perception: from communicating three dimensions on a two-dimensional surface, to manipulating time through a series of images.

Ultimately we are creating ways to communicate, to both ourselves and to others, our perception of what is real. This is important because as a collective we determine the course the future will take and how we find our place in it. Through innovations in communication mediums such as virtual reality, we are on the brink of a new development in consciousness that will alter what we are able to perceive and, as a result, what we understand as reality.

References

1. Bertol, D. (1996). “Architecture of Images”: An investigation of architectural representations and the visual perception of three-dimensional space. *Leonardo*, *29*(2), 87–94. http://doi.org/10.2307/1576337.
2. Beveridge, P. (2000). Color perception and the art of James Turrell. *Leonardo*, *33*(4), 305–313. Retrieved from http://www.jstor.org/stable/1576905.
3. Crawford, C. (2013). *Chris crawford on interactive storytelling*. Berkeley, CA: New Riders Games.
4. Csikszentmihalyi, M. (2004, February). *Mihaly Csikszentmihalyi: Flow, the secret to happiness* [Video file]. Retrieved from https://www.ted.com/talks/mihaly_csikszentmihalyi_on_flow.
5. Gibson, J. J. (1978). The ecological approach to the visual perception of pictures. *Leonardo*, *11*(3), 227–235. http://doi.org/10.2307/1574154.
6. Hoffman, D. D. (1998). *Visual intelligence: How we create what we see*. New York: W.W. Norton.
7. Kelleher, C. (2013, January 8). *How we see color—Colm Kelleher* [Video file]. Retrieved from http://ed.ted.com/lessons/how-we-see-color-colm-kelleher.
8. McLuhan, M., McLuhan, E., & Zingrone, F. (1995). *Essential McLuhan*. New York, NY: Basic Books.
9. Neafus, D. (2014). *The language of immersive cinema*. Publisher: Author.
10. Nechvatal, J. (2001). Towards an Immersive Intelligence. *Leonardo*, *34*(5), 417–422. Retrieved from http://www.jstor.org/stable/1577234

Science, Mind, and Limits of Understanding

One of the most profound insights into language and mind, I think, was Descartes's recognition of what we may call "the creative aspect of language use": the ordinary use of language is typically innovative without bounds, appropriate to circumstances but not caused by them—a crucial distinction—and can engender thoughts in others that they recognize they could have expressed themselves. Given the intimate relation of language and thought, these are properties of human thought as well. This insight is the primary basis for Descartes's scientific theory of mind and body. There is no sound reason to question its validity, as far as I am aware. Its implications, if valid, are far-reaching, among them what it suggests about the limits of human understanding, as becomes more clear when we consider the place of these reflections in the development of modern science from the earliest days.

It is important to bear in mind that insofar as it was grounded in these terms, Cartesian dualism was a respectable scientific theory, proven wrong (in ways that are often misunderstood), but that is the common fate of respectable theories.

The background is the so-called "mechanical philosophy"—mechanical science in modern terminology. This doctrine, originating with Galileo and his contemporaries, held that the world is a machine, operating by mechanical principles, much like the remarkable devices that were being constructed by skilled artisans of the day and that stimulated the scientific imagination much as computers do today; devices with gears, levers, and other mechanical components, interacting through direct contact with no mysterious forces relating them. The doctrine held that the entire world is similar: it could in principle be constructed by a skilled artisan, and was in fact created by a super-skilled artisan. The doctrine was intended to replace the resort to "occult properties" on the part of the neoscholastics: their appeal to mysterious sympathies and antipathies, to forms flitting through the air as the means of perception, the idea that rocks fall and steam rises because they are moving to their natural place, and similar notions that were mocked by the new science.

The mechanical philosophy provided the very criterion for intelligibility in the sciences. Galileo insisted that theories are intelligible, in his words, only if we can "duplicate [their posits] by means of appropriate artificial devices." The same conception, which became the reigning orthodoxy, was maintained and developed

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5

by the other leading figures of the scientific revolution: Descartes, Leibniz, Huygens, Newton, and others.

Today Descartes is remembered mainly for his philosophical reflections, but he was primarily a working scientist and presumably thought of himself that way, as his contemporaries did. His great achievement, he believed, was to have firmly established the mechanical philosophy, to have shown that the world is indeed a machine, that the phenomena of nature could be accounted for in mechanical terms in the sense of the science of the day. But he discovered phenomena that appeared to escape the reach of mechanical science. Primary among them, for Descartes, was the creative aspect of language use, a capacity unique to humans that cannot be duplicated by machines and does not exist among animals, which in fact were a variety of machines, in his conception.

As a serious and honest scientist, Descartes therefore invoked a new principle to accommodate these non-mechanical phenomena, a kind of creative principle. In the substance philosophy of the day, this was a new substance, *res cogitans*, which stood alongside of *res extensa.* This dichotomy constitutes the mind-body theory in its scientific version. Then followed further tasks: to explain how the two substances interact and to devise experimental tests to determine whether some other creature has a mind like ours. These tasks were undertaken by Descartes and his followers, notably Géraud de Cordemoy; and in the domain of language, by the logician-grammarians of Port Royal and the tradition of rational and philosophical grammar that succeeded them, not strictly Cartesian but influenced by Cartesian ideas.

All of this is normal science, and like much normal science, it was soon shown to be incorrect. Newton demonstrated that one of the two substances does not exist: *res extensa*. The properties of matter, Newton showed, escape the bounds of the mechanical philosophy. To account for them it is necessary to resort to interaction without contact. Not surprisingly, Newton was condemned by the great physicists of the day for invoking the despised occult properties of the neo-scholastics. Newton largely agreed. He regarded action at a distance, in his words, as "so great an Absurdity, that I believe no Man who has in philosophical matters a competent Faculty of thinking, can ever fall into it." Newton however argued that these ideas, though absurd, were not "occult" in the traditional despised sense. Nevertheless, by invoking this absurdity, we concede that we do not understand the phenomena of the material world. To quote one standard scholarly source, "By 'understand' Newton still meant what his critics meant: 'understand in mechanical terms of contact action'."

It is commonly believed that Newton showed that the world is a machine, following mechanical principles, and that we can therefore dismiss "the ghost in the machine," the mind, with appropriate ridicule. The facts are the opposite: Newton exorcised the machine, leaving the ghost intact. The mind-body problem in its scientific form did indeed vanish as unformulable, because one of its terms, body, does not exist in any intelligible form. Newton knew this very well, and so did his great contemporaries.

John Locke wrote that we remain in “incurable ignorance of what we desire to know” about matter and its effects, and no “science of bodies [that provides true explanations is] within our reach.” Nevertheless, he continued, he was “convinced by the judicious Mr. Newton’s incomparable book, that it is too bold a presumption to limit God’s power, in this point, by my narrow conceptions.” Though gravitation of matter to matter is “inconceivable to me”, nevertheless, as Newton demonstrated, we must recognize that it is within God’s power “to put into bodies, powers and ways of operations, above what can be derived from our idea of body, or can be explained by what we know of matter.” And thanks to Newton’s work, we know that God “has done so.” The properties of the material world are “inconceivable to us,” but real nevertheless. Newton understood the quandary. For the rest of his life, he sought some way to overcome the absurdity, suggesting various possibilities, but not committing himself to any of them because he could not show how they might work and, as he always insisted, he would not “feign hypotheses” beyond what can be experimentally established.

Replacing the theological with a cognitive framework, David Hume agreed with these conclusions. In his history of England, Hume describes Newton as “the greatest and rarest genius that ever arose for the ornament and instruction of the species.” His most spectacular achievement was that while he “seemed to draw the veil from some of the mysteries of nature, he shewed at the same time the imperfections of the mechanical philosophy; and thereby restored [Nature’s] ultimate secrets to that obscurity, in which they ever did and ever will remain.”

Modern commentators observe that Einstein’s relativity theory provides a local interpretation of gravitational attraction, overcoming the non-locality that was an absurdity to Newton and others. But while the observation is correct, it does not bear on the recognition that the world is unintelligible to us. Such concepts as curved space-time are no less remote than action-at-a distance from the mechanical philosophy that provided the very criterion of intelligibility and understanding for the great founders of modern science—and that also seems to be basically our common sense conception of physical reality and the material world.

As the import of Newton’s discoveries was gradually assimilated in the sciences, the “absurdity’ recognized by Newton and his great contemporaries became scientific common sense. The properties of the natural world are inconceivable to us, but that does not matter. The goals of scientific inquiry were implicitly restricted: from the kind of conceivability that was a criterion for true understanding in early modern science from Galileo through Newton and beyond, to something much more limited: intelligibility of theories about the world. This seems to me a step of considerable significance in the history of human thought and inquiry, more so than is generally recognized, though it has been understood by historians of science.

Friedrich Lange, in his classic 19th century history of materialism, observed that we have “so accustomed ourselves to the abstract notion of forces, or rather to a notion hovering in a mystic obscurity between abstraction and concrete comprehension, that we no longer find any difficulty in making one particle of matter act upon another without immediate contact,… through void space without any material link. From such ideas the great mathematicians and physicists of the seventeenth century were far removed. They were all in so far genuine Materialists in the sense of ancient Materialism that they made immediate contact a condition of influence." This transition over time is "one of the most important turning-points in the whole history of Materialism," he continued, depriving the doctrine of much significance, if any at all. "What Newton held to be so great an absurdity that no philosophic thinker could light upon it, is prized by posterity as Newton's great discovery of the harmony of the universe!"

Similar conclusions are commonplace in the history of science. In the mid-twentieth century, Alexander Koyré observed that Newton demonstrated that "a purely materialistic pattern of nature is utterly impossible (and a purely materialistic or mechanistic physics, such as that of Lucretius or of Descartes, is utterly impossible, too)"; his mathematical physics required the "admission into the body of science of incomprehensible and inexplicable 'facts' imposed up on us by empiricism," by what is observed and our conclusions from these observations.

With the disappearance of the scientific concept of body (material, physical, etc.), what happens to the "second substance," res cogitans/mind, which was left untouched by Newton's startling discoveries? A plausible answer was suggested by John Locke, also within the reigning theological framework. He wrote that just as God added to matter such inconceivable properties as gravitational attraction, he might also have "superadded" to matter the capacity of thought. In the years that followed, Locke's "God" was reinterpreted as "nature," a move that opened the topic to inquiry. That path was pursued extensively in the years that followed, leading to the conclusion that mental processes are properties of certain kinds of organized matter. Restating the fairly common understanding of the time, Charles Darwin, in his early notebooks, wrote that there is no need to regard thought, "a secretion of the brain," as "more wonderful than gravity, a property of matter"—all inconceivable to us, but that is not a fact about the external world; rather, about our cognitive limitations.

It is of some interest that all of this has been forgotten, and is now being rediscovered. Nobel laureate Francis Crick, famous for the discovery of DNA, formulated what he called the "astonishing hypothesis" that our mental and emotional states are "in fact no more than the behavior of a vast assembly of nerve cells and their associated molecules." In the philosophical literature, this rediscovery has sometimes been regarded as a radical new idea in the study of mind. To cite one prominent source, the radical new idea is "the bold assertion that mental phenomena are entirely natural and caused by the neurophysiological activities of the brain." In fact, the many proposals of this sort reiterate, in virtually the same words, formulations of centuries ago, after the traditional mind-body problem became unformulable with Newton's demolition of the only coherent notion of body (or physical, material, etc.). For example, 18th century chemist/philosopher Joseph Priestley's conclusion that properties "termed mental" reduce to "the organical structure of the brain," stated in different words by Locke, Hume, Darwin, and many others, and almost inescapable, it would seem, after the collapse of the mechanical philosophy that provided the foundations for early modern science, and its criteria of intelligibility.

The last decade of the twentieth century was designated “the Decade of the Brain.” In introducing a collection of essays reviewing its results, neuroscientist Vernon Mountcastle formulated the guiding theme of the volume as the thesis of the new biology that “Things mental, indeed minds, are emergent properties of brains, [though] these emergences are…produced by principles that… we do not yet understand”—again reiterating eighteenth century insights in virtually the same words.

The phrase “we do not yet understand,” however, should strike a note of caution. We might recall Bertrand Russell’s observation in 1927 that chemical laws “cannot at present be reduced to physical laws.” That was true, leading eminent scientists, including Nobel laureates, to regard chemistry as no more than a mode of computation that could predict experimental results, but not real science. Soon after Russell wrote, it was discovered that his observation, though correct, was understated. Chemical laws never would be reducible to physical laws, as physics was then understood. After physics underwent radical changes, with the quantum-theoretic revolution, the new physics was unified with a virtually unchanged chemistry, but there was never reduction in the anticipated sense.

There may be some lessons here for neuroscience and philosophy of mind. Contemporary neuroscience is hardly as well-established as physics was a century ago. There are what seem to me to be cogent critiques of its foundational assumptions, notably recent work by cognitive neuroscientists C.R. Gallistel and Adam Philip King. The common slogan that study of mind is neuroscience at an abstract level might turn out to be just as misleading as comparable statements about chemistry and physics 90 years ago. Unification may take place, but that might require radical rethinking of the neurosciences, perhaps guided by computational theories of cognitive processes, as Gallistel and King suggest.

The development of chemistry after Newton also has lessons for neuroscience and cognitive science. The 18th century chemist Joseph Black recommended that “chemical affinity be received as a first principle, which we cannot explain any more than Newton could explain gravitation, and let us defer accounting for the laws of affinity, till we have established such a body of doctrine as he has established concerning the laws of gravitation.” The course Black outlined is the one that was actually followed as chemistry proceeded to establish a rich body of doctrine. Historian of chemistry Arnold Thackray observes that the “triumphs” of chemistry were “built on no reductionist foundation but rather achieved in isolation from the newly emerging science of physics.” Interestingly, Thackray continues, Newton and his followers did attempt to “pursue the thoroughly Newtonian and reductionist task of uncovering the general mathematical laws which govern all chemical behavior” and to develop a principled science of chemical mechanisms based on physics and its concepts of interactions among “the ultimate permanent particles of matter.” But the Newtonian program was undercut by Dalton’s “astonishingly successful weight-quantification of chemical units,” Thackray continues, shifting “the whole area of philosophical debate among chemists from that of chemical mechanisms (the why? of reaction) to that of chemical units (the what? and how much?),” a theory that “was profoundly antiphysicalist and anti-Newtonian in its rejection of the unity of matter, and its dismissal of short-range forces." Continuing, Thackray writes that "Dalton's ideas were chemically successful. Hence they have enjoyed the homage of history, unlike the philosophically more coherent, if less successful, reductionist schemes of the Newtonians."

Adopting contemporary terminology, we might say that Dalton disregarded the "explanatory gap" between chemistry and physics by ignoring the underlying physics, much as post-Newtonian physicists disregarded the explanatory gap between Newtonian dynamics and the mechanical philosophy by rejecting the latter, and thereby tacitly lowering the goals of science in a highly significant way, as I mentioned.

Contemporary studies of mind are deeply troubled by the "explanatory gap" between the science of mind and neuroscience—in particular, between computational theories of cognition, including language, and neuroscience. I think they would be well-advised to take seriously the history of chemistry. Today's task is to develop a "body of doctrine" to explain what appear to be the critically significant phenomena of language and mind, much as chemists did. It is of course wise to keep the explanatory gap in mind, to seek ultimate unification, and to pursue what seem to be promising steps towards unification, while nevertheless recognizing that as often in the past, unification may not be reduction, but rather revision of what is regarded as the "fundamental discipline," the reduction basis, the brain sciences in this case.

Locke and Hume, and many less-remembered figures of the day, understood that much of the nature of the world is "inconceivable" to us. There were actually two different kinds of reasons for this. For Locke and Hume, the reasons were primarily epistemological. Hume in particular developed the idea that we can only be confident of immediate impressions, of "appearances." Everything else is a mental construction. In particular, and of crucial significance, that is true of identity through time, problems that trace back to the pre-Socratics: the identity of a river or a tree or most importantly a person as they change through time. These are mental constructions; we cannot know whether they are properties of the world, a metaphysical reality. As Hume put the matter, we must maintain "a modest skepticism to a certain degree, and a fair confession of ignorance in subjects, that exceed all human capacity"—which for Hume includes virtually everything beyond appearances. We must "refrain from disquisitions concerning their real nature and operations." It is the imagination that leads us to believe that we experience external continuing objects, including a mind or self. The imagination, furthermore, is "a kind of magical faculty in the soul, which…is inexplicable by the utmost efforts of human understanding," so Hume argued.

A different kind of reason why the nature of the world is inconceivable to us was provided by "the judicious Mr. Newton," who apparently was not interested in the epistemological problems that vexed Locke and Hume. Newton scholar Andrew Janiak concludes that Newton regarded such global skepticism as "irrelevant—he takes the possibility of our knowledge of nature for granted." For Newton, "the primary epistemic questions confronting us are raised by physical theory itself." Locke and Hume, as I mentioned, took quite seriously the new science-based skepticism that resulted from Newton's demolition of the mechanical philosophy, which had provided the very criterion of intelligibility for the scientific revolution. That is why Hume lauded Newton for having "restored [Nature's] ultimate secrets to that obscurity, in which they ever did and ever will remain."

For these quite different kinds of reasons, the great figures of the scientific revolution and the Enlightenment believed that there are phenomena that fall beyond human understanding. Their reasoning seems to me substantial, and not easily dismissed. But contemporary doctrine is quite different. The conclusions are regarded as a dangerous heresy. They are derided as "the new mysterianism," a term coined by philosopher Owen Flanagan, who defined it as "a postmodern position designed to drive a railroad spike through the heart of scientism." Flanagan is referring specifically to explanation of consciousness, but the same concerns hold of mental processes in general.

The "new mysterianism" is compared today with the "old mysterianism," Cartesian dualism, its fate typically misunderstood. To repeat, Cartesian dualism was a perfectly respectable scientific doctrine, disproven by Newton, who exorcised the machine, leaving the ghost intact, contrary to what is commonly believed.

The "new mysterianism," I believe, is misnamed. It should be called "truism"—at least, for anyone who accepts the major findings of modern biology, which regards humans as part of the organic world. If so, then they will be like all other organisms in having a genetic endowment that enables them to grow and develop to their mature form. By simple logic, the endowment that makes this possible also excludes other paths of development. The endowment that yields scope also establishes limits. What enables us to grow legs and arms, and a mammalian visual system, prevents us from growing wings and having an insect visual system.

All of this is indeed truism, and for non-mystics, the same should be expected to hold for cognitive capacities. We understand this well for other organisms. Thus we are not surprised to discover that rats are unable to run prime number mazes no matter how much training they receive; they simply lack the relevant concept in their cognitive repertoire. By the same token, we are not surprised that humans are incapable of the remarkable navigational feats of ants and bees; we simply lack the cognitive capacities, though we can sometimes duplicate their feats with sophisticated instruments. The truisms extend to higher mental faculties. For such reasons, we should, I think, be prepared to join the distinguished company of Newton, Locke, Hume and other dedicated mysterians.

For accuracy, we should qualify the concept of "mysteries" by relativizing it to organisms. Thus what is a mystery for rats might not be a mystery for humans, and what is a mystery for humans is instinctive for ants and bees.

Dismissal of mysterianism seems to me one illustration of a widespread form of dualism, a kind of epistemological and methodological dualism, which tacitly adopts the principle that study of mental aspects of the world should proceed in some fundamentally different way from study of what are considered physical aspects of the world, rejecting what are regarded as truisms outside the domain of mental processes. This new dualism seems to me truly pernicious, unlike Cartesian dualism, which was respectable science. The new methodological dualism, in contrast, seems to me to have nothing to recommend it.

Far from bewailing the existence of mysteries-for-humans, we should be extremely grateful for it. With no limits to growth and development, our cognitive capacities would also have no scope. Similarly, if the genetic endowment imposed no constraints on growth and development of an organism it could become only a shapeless amoeboid creature, reflecting accidents of an unanalyzed environment, each quite unlike the next. Classical aesthetic theory recognized the same relation between scope and limits. Without rules, there can be no genuinely creative activity, even when creative work challenges and revises prevailing rules.

Contemporary rejection of mysterianism—that is, truism—is quite widespread. One recent example that has received considerable attention is an interesting and informative book by physicist David Deutsch. He writes that potential progress is "unbounded" as a result of the achievements of the Enlightenment and early modern science, which directed science to the search for best explanations. As philosopher/physicist David Albert expounds his thesis, "with the introduction of that particular habit of concocting and evaluating new hypotheses, there was a sense in which we could do anything. The capacities of a community that has mastered that method to survive, and to learn, and to remake the world according to its inclinations, are (in the long run) literally, mathematically, infinite."

The quest for better explanations may well indeed be infinite, but infinite is of course not the same as limitless. English is infinite, but doesn't include Greek. The integers are an infinite set, but do not include the reals. I cannot discern any argument here that addresses the concerns and conclusions of the great mysterians of the scientific revolution and the Enlightenment.

We are left with a serious and challenging scientific inquiry: to determine the innate components of our cognitive nature in language, perception, concept formation, reflection, inference, theory construction, artistic creation, and all other domains of life, including the most ordinary ones. By pursuing this task we may hope to determine the scope and limits of human understanding, while recognizing that some differently structured intelligence might regard human mysteries as simple problems and wonder that we cannot find the answers, much as we can observe the inability of rats to run prime number mazes because of the very design of their cognitive nature.

There is no contradiction in supposing that we might be able to probe the limits of human understanding and try to sharpen the boundary between problems that fall within our cognitive range and mysteries that do not. There are possible experimental inquiries. Another approach would be to take seriously the concerns of the great figures of the early scientific revolution and the Enlightenment: to pay attention to what they found "inconceivable," and particularly their reasons. The "mechanical philosophy" itself has a claim to be an approximation to common sense understanding of the world, a suggestion that might be clarified by experimental inquiry. Despite much sophisticated commentary, it is also hard to escape the force of Descartes's conviction that free will is "the noblest thing" we have, that "there is nothing we comprehend more evidently and more perfectly" and that "it would be absurd" to doubt something that "we comprehend intimately, and experience within ourselves" merely because it is "by its nature incomprehensible to us," if indeed we do not "have intelligence enough" to understand the workings of mind, as he speculated. Concepts of determinacy and randomness fall within our intellectual grasp. But it might turn out that "free actions of men" cannot be accommodated in these terms, including the creative aspect of language and thought. If so, that might be a matter of cognitive limitations—which would not preclude an intelligible theory of such actions, far as this is from today's scientific understanding.

Honesty should lead us to concede, I think, that we understand little more today about these matters than the Spanish physician-philosopher Juan Huarte did 500 years ago when he distinguished the kind of intelligence humans shared with animals from the higher grade that humans alone possess and is illustrated in the creative use of language, and proceeding beyond that, from the still higher grade illustrated in true artistic and scientific creativity. Nor do we even know whether these are questions that lie within the scope of human understanding, or whether they fall among what Hume took to be Nature's ultimate secrets, consigned to "that obscurity in which they ever did and ever will remain."

Noam Chomsky, MIT[1]

---

[1] This artticle is based on a talk given at "The Science and Faith Foundation (STOQ), The Vatican, January 2014." Copyright © 2015 Noam Chomsky.

Authors Information—Affiliation/Mailing Address and Email Information

NB: Joint authors are marked with asterisk and are placed together.
Philosophy:

(1) *Krishnamurthy Ramasubramanian, Cell for Indian Science and Technology in Sanskrit, Department of Humanities and Social Sciences, IIT Bombay, Mumbai, India. E-mail: ram@hss.iitb.ac.in, mullaikramas@gmail.com
(2) *Jonathan Duquette, Newton International Fellow, University of Oxford, Oxford, United Kingdom. E-mail: jonathan.duquette@orinst.ox.ac.uk
(3) James Robert Brown, Department of Philosophy, University of Toronto, 170 St George St, Toronto, ON M5R 2M8, Canada. Email: jrbrown@chass.utoronto.ca
(4) Nathan Salmon, Philosophy Department, University of California, Santa Barbara, South Hall 3431/3432, Santa Barbara, CA 93106-3090. Email: nsalmon@philosophy.ucsb.edu
(5) Nicholas Maxwell, Science and Technology Studies, University College London, 13 Tavistock Terrace, London, N19-4BZ, UK. Email: nicholas.maxwell@ucl.ac.uk
(6) Francesca Biagioli, Address: University of Konstanz, Zukunftskolleg, P.O. Box 216, 78457 Konstanz, Germany. Email: francesca.biagioli@uni-konstanz.de
(7) Valia Allori, Department of Philosophy, Northern Illinois University, Zulauf Hall 915, Dekalb IL 60115 USA. E-mail: vallori@niu.edu
(8) *Michael Witzel, Department of South Asian Studies, Harvard University, 1 Bow Street, 3rd fl., Cambridge, MA 02138, USA. Email: witzel@fas.harvard.edu
(9) *Nataliya Yanchevskaya, Princeton University, Princeton, NJ, USA. Email: nay@princeton.edu
(10) Hans Herlof Grelland, University of Agder, Jon Lilletuns vei 9, N-4879 Grimstad. Email: hans.grelland@uia.no
(11) E. C. G. Sudarshan, Physics Department, University of Texas Austin, Texas, USA. Email: bhamathig@gmail.com

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5

Theoretical Physics:

1. Anthony Sudbery Department of Mathematics, University of York, Heslington, York, England Y010 5DD. Email: as2@york.ac.uk
2. Gerard ’t Hooft, Institute for Theoretical Physics, Utrecht University & Spinoza Institute, Postbus 80.089, 3508 TB Utrecht, The Netherlands Email: g.thooft@uu.nl
3. Joan A. Vaccaro, Center for Quantum Dynamics, School of Science, Griffith University, 170 Kessels Road, Nathan 4111, Australia. Email: J.A.Vaccaro@griffith.edu.au
4. Vesselin Petkov, Institute for Foundational Studies Hermann Minkowski, Montreal, Quebec, Canada. Email: vpetkov@minkowskiinstitute.org
5. Norbert Straumann, Physik-Institut, Universitat Zurich, CH–8057 Zurich, Switzerland. Email: norbert.straumann@gmail.com
6. *W. M. Stuckey, Department of Physics, Elizabethtown College, Elizabethtown, PA 17022. Email: stuckeym@etown.edu
7. *Michael Silberstein, Department of Philosophy and Foundations of Physics, Committee for Philosophy and the Sciences, University of Maryland, College Park, MD 20742. Email: msilbers@umd.edu
8. *Timothy McDevitt, Department of Mathematical Sciences, Elizabethtown College, Elizabethtown, PA 17022 Email: mcdevittt@etown.edu
9. C. S. Unnikrishnan, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai - 400 005, India. Email: unni@tifr.res.in
10. George Musser, Contributing editor, Scientific American, 80 Glen Ridge Ave, Glen Ridge, NJ 07028, USA. Email: gmusser@nasw.org

Mathematics:

1. Dr Mary Leng, Department of Philosophy, University of York, Heslington, N. Yorkshire, Y010 5DD Email: mary.leng@york.ac.uk
2. Raymond O. Wells, Jr., Jacobs University Bremen; Rice University; University of Colorado at Boulder, Email: row@raw.com
3. Goro C. Kato, Mathematics Department, California Polytechnic State University, San Luis Obispo, CA, USA 93407. Email: gkato@calpoly.edu
4. Paul Ernest, Graduate School of Education, Exeter University, Exeter, EX1 2LU, UK. Email: P.Ernest@exeter.ac.uk
5. James Isenberg, Department of Mathematics and Institute of Theoretical Science, University of Oregon. Email: isenberg@uoregon.edu
6. Reuben Hersh, Department of Mathematics and Statistics University of New Mexico Humanities 415 Albuquerque NM 87131. Email: rhersh@gmail.com
7. Julian Barbour, Visiting Professor in Physics at the University of Oxford, UK and College Farm, South Newington, Banbury, Oxon, OX15 4JG UK. Email: julian.barbour@physics.ox.ac.uk
8. Lorenzo Sadun, Department of Mathematics, University of Texas, Austin, TX 78712. Email: sadun@math.utexas.edu

Biology/Cognitive Science:

1. Rajesh Kasturirangan, Mind and Society Initiative, Azim Premji university, Bangalore. Email: rkasturi@gmail.com
2. *Vijay Kumar Sharma, Chronobiology Laboratory, Evolutionary and Organismal Biology Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore–560064, Karnataka, India. Email: vksharmas@gmail.com
3. *L Abhilash, Chronobiology Laboratory, Evolutionary and Organismal Biology Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore–560064, Karnataka, India. Email: abhilashl@jncasr.ac.in
4. * Pier Luigi Luisi, Dipartimento di Biologia, Università degli Studi di Roma Tre Viale Guglielmo Marconi, 446, 00146 ROMA. Email: luisi@mat.ethz.ch
5. * Mauro Bergonzi, Dipartimento di Scienze Umane e Sociali, Università degli Studi di Napoli "L'Orientale", Palazzo Giusso, Largo S. Giovanni Maggiore, 30-80134 Napoli, Italy. Email: bergma@libero.it
6. Norman Sieroka, Eidgenössische Technische Hochschule Zürich, Professur für Philosophie, CLW, Clausiusstrasse 49, CH-8092 Zürich. Email: sieroka@phil.gess.ethz.ch
7. Randall E. Auxier, Southern Illinois University, 1263 Lincoln Drive, Carbondale, IL 62901-6899. Email: personalist61@gmail.com

Logic/Computer Science:

1. *David H. Bailey, Lawrence Berkeley National Laboratory (retired), Berkeley, CA 94720, and University of California, Davis, CA 95616, USA. Email: david@davidhbailey.com
2. *Jonathan M. Borwein, CARMA, University of Newcastle, Callaghan, NSW 2308, Australia. Email: jon.borwein@gmail.com
3. Alexander Keewatin (A. K.) Dewdney, Dept. of Computer Science, University of Western Ontario, London, Ontario N6A 5B7. Email: akd@uwo.ca
4. Gary Mar, Department of Philosophy, Stony Brook University, New York, USA. Email: Gary.Mar@stonybrook.edu
5. Doron Swade, School of Computing, University of Portsmouth, Buckingham Building, Lion Terrace, Portsmouth, Hants, PO1 3HE. Email: doron.swade@blueyonder.co.uk
6. *N. C. A. da Costa, Department of Philosophy, Federal University of Santa Catarina, Brazil. Email: ncacosta@terra.com.br
7. *F. A. Doria, Advanced Studies Research Group; HCTE; Fuzzy Sets Laboratory; Mathematical Economics Group, Production Engineering Program COPPE, UFRJ P.O. Box 68507, 21945–972 Rio RJ Brazil. Email: fadoria63@gmail.com

Miscellaneous

1. Robert T. Tally Jr., Department of English, Texas State University, 601 University Drive, San Marcos, TX 78666, USA. Email: robert.tally@txstate.edu
2. Donald G. Janelle, Center for Spatially Integrated Social Science (CSISS), Department of Geography/NCGIA, University of California, Santa Barbara CA 93106-4060. Email: janelle@geog.ucsb.edu
3. Diana J. Reichenbach, Maryland Institute College of Art, 1300 W Mount Royal Avenue, Baltimore MD 21217-4134. Email: diana.reichenbach@gmail.com

Titles in this Series

Quantum Mechanics and Gravity
By Mendel Sachs

Quantum-Classical Correspondence
Dynamical Quantization and the Classical Limit
By Dr. A.O. Bolivar

Knowledge and the World: Challenges Beyond the Science Wars
Ed. by M. Carrier, J. Roggenhofer, G. Küppers and P. Blanchard

Quantum-Classical Analogies
By Daniela Dragoman and Mircea Dragoman

Life—As a Matter of Fat
The Emerging Science of Lipidomics
By Ole G. Mouritsen

Quo Vadis Quantum Mechanics?
Ed. by Avshalom C. Elitzur, Shahar Dolev and Nancy Kolenda

Information and Its Role in Nature
By Juan G. Roederer

Extreme Events in Nature and Society
Ed. by Sergio Albeverio, Volker Jentsch and Holger Kantz

The Thermodynamic Machinery of Life
By Michal Kurzynski

Weak Links
The Universal Key to the Stability of Networks and Complex Systems
By Csermely Peter

The Emerging Physics of Consciousness
Ed. by Jack A. Tuszynski

Quantum Mechanics at the Crossroads
New Perspectives from History, Philosophy and Physics
Ed. by James Evans and Alan S. Thorndike

How Should Humanity Steer the Future
Ed. by Anthony Aguirre, Brendan Foster and Zeeya Merali

S. Wuppuluri and G. Ghirardi (eds.), *Space, Time and the Limits of Human Understanding*, The Frontiers Collection, DOI 10.1007/978-3-319-44418-5

Mind, Matter and the Implicate Order
By Paavo T.I. Pylkkanen

Particle Metaphysics
A Critical Account of Subatomic Reality
By Brigitte Falkenburg

The Physical Basis of the Direction of Time
By H. Dieter Zeh

Asymmetry: The Foundation of Information
By Scott J. Muller

Decoherence and the Quantum-To-Classical Transition
By Maximilian A. Schlosshauer

The Nonlinear Universe
Chaos, Emergence, Life
By Alwyn C. Scott

Quantum Superposition
Counterintuitive Consequences of Coherence, Entanglement, and Interference
By Mark P. Silverman

Symmetry Rules
How Science and Nature are Founded on Symmetry
By Joseph Rosen

Mind, Matter and Quantum Mechanics
By Henry P. Stapp

Entanglement, Information, and the Interpretation of Quantum Mechanics
By Gregg Jaeger

Relativity and the Nature of Spacetime
By Vesselin Petkov

The Biological Evolution of Religious Mind and Behavior
Ed. by Eckart Voland and Wulf Schiefenhovel

Homo Novus-A Human without Illusions
Ed. by Ulrich J. Frey, Charlotte Stormer and Kai P. Willfiihr

Brain-Computer Interfaces
Revolutionizing Human-Computer Interaction
Ed. by Bernhard Graimann, Brendan Allison and Gert Pfurtscheller

Extreme States of Matter
On Earth and in the Cosmos
By Vladimir E. Fortov

Searching for Extraterrestrial Intelligence
SETI Past, Present, and Future
Ed. by H. Paul Shuch

Essential Building Blocks of Human Nature
Ed. by Ulrich J. Frey, Charlotte Störmer and Kai P. Willführ

Mindful Universe
Quantum Mechanics and the Participating Observer
By Henry P. Stapp

Principles of Evolution
From the Planck Epoch to Complex Multicellular Life
Ed. by Hildegard Meyer-Ortmanns and Stefan Thurner

The Second Law of Economics
Energy, Entropy, and the Origins of Wealth
By Reiner Köummel

States of Consciousness
Experimental Insights into Meditation, Waking, Sleep and Dreams
Ed. by Dean Cvetkovic and Irena Cosic

Elegance and Enigma
The Quantum Interviews The Quantum Interviews
Ed. by Maximilian Schlosshauer

Humans on Earth
From Origins to Possible Futures
By Filipe Duarte Santos

Evolution 2.0
Implications of Darwinism in Philosophy and the Social and Natural Sciences
Ed. by Martin Brinkworth and Friedel Weinert

Probability in Physics
Ed. by Yemima Ben-Menahem and Meir Hemmo

Chips 2020
A Guide to the Future of Nanoelectronics
Ed. by Bernd Hoefflinger

From the Web to the Grid and Beyond
Computing Paradigms Driven by High-Energy Physics
Ed. by Rene Brun, Federico Carminati and Giuliana Galli Carminati

The Language Phenomenon
Human Communication from Milliseconds to Millennia
Ed. by P.-M. Binder and K. Smith

The Dual Nature of Life
By Gennadiy Zhegunov

Natural Fabrications
By William Seager

Ultimate Horizons
By Helmut Satz

Physics, Nature and Society
By Joaquín Marro

Extraterrestrial Altruism
Ed. by Douglas A. Vakoch

The Beginning and the End
By Clément Vidal

A Brief History of String Theory
By Dean Rickles

Singularity Hypotheses
Ed. by Amnon H. Eden, James H. Moor, Johnny H. Søraker and Eric Steinhart

Why More Is Different
Philosophical Issues in Condensed Matter Physics and Complex Systems
Ed. by Brigitte Falkenburg and Margaret Morrison

Questioning the Foundations of Physics
Ed. by Anthony Aguirre, Brendan Foster and Zeeya Merali

It From Bit or Bit From It?
Ed. by Anthony Aguirre, Brendan Foster and Zeeya Merali

Trick or Truth?
Ed. by Anthony Aguirre, Brendan Foster and Zeeya Merali

The Challenge of Chance
Ed. By Klaas Landsman

Quantum (Un)Speakables II
Half a Century of Bell's Theorem
Ed. by Reinhold Bertlmann, Anton Zeilinger

Energy, Complexity and Wealth Maximization
Ed. by Robert Ayres

Ancestors, Territoriality And Gods
By Ina Wunn

Space,Time and the Limits of Human Understanding
By Sham Wuppuluri

Information and Interaction
Ed. by Durham

The Technological Singularity
Ed. by Victor Callaghan