ERNST ZERMELO

Collected Works
Gesammelte Werke

VOLUME II
BAND II

Calculus of Variations,
Applied Mathematics,
and Physics

Variationsrechnung,
Angewandte Mathematik
und Physik

HEIDELBERGER AKADEMIE
DER WISSENSCHAFTEN

Springer

Schriften der Mathematisch-naturwissenschaftlichen Klasse
der Heidelberger Akademie der Wissenschaften
Nr. 23 (2013)

ERNST ZERMELO

Collected Works
Gesammelte Werke

Zermelo around 1910

ERNST ZERMELO

Collected Works
Gesammelte Werke

ERNST ZERMELO

Collected Works
Gesammelte Werke

VOLUME I
BAND I

Set Theory,
Miscellanea

Mengenlehre,
Varia

VOLUME II
BAND II

Calculus of Variations,
Applied Mathematics,
and Physics

Variationsrechnung,
Angewandte Mathematik
und Physik

ERNST ZERMELO

Collected Works
Gesammelte Werke

VOLUME II
BAND II

Calculus of Variations, Applied Mathematics, and Physics
Variationsrechnung, Angewandte Mathematik und Physik

Edited by
Herausgegeben von

Heinz-Dieter Ebbinghaus,
Akihiro Kanamori

Springer

Editors
Herausgeber

Heinz-Dieter Ebbinghaus
University of Freiburg, Germany

Akihiro Kanamori
Boston University, MA, USA

ISBN 978-3-540-70855-1 e-ISBN 978-3-540-70856-8
DOI 10.1007/978-3-540-70856-8
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2007921876

Mathematics Subject Classification (2010): 01A55, 01A60, 01A75, 49-02, 49K05, 49K15, 49Q05, 49S05, 53C22, 62J15, 70B99, 76B47, 76M23, 76R50, 80A05, 82C03, 82C40

Cover design: WMXDesign GmbH, Heidelberg
Typesetting and production: le-tex publishing services oHG, Leipzig

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface to the Zermelo edition

This is a complete edition of the published works of Ernst Zermelo which moreover includes selected correspondence and unpublished manuscripts. Zermelo is generally acknowledged for his pioneering work in axiomatic set theory and for introducing the axiom of choice as a basic principle of mathematics. In contrast, his work in applied mathematics and physics, despite its originality, is hardly recognized or has even been attributed to others. This edition of Zermelo's collected works provides a picture of the entire mathematician. It appears in two volumes. The first volume comprises Zermelo's published papers in set theory and the foundation of mathematics together with isolated papers of an algebraic, analytic, or number-theoretic character. The second volume is dedicated to Zermelo's work in the calculus of variations, mathematical physics, and fluid dynamics. Both volumes are supplemented by selected notes and manuscripts, mainly from Zermelo's *Nachlass*, which throw additional light on his papers, reflect his point of view, or are unpublished continuations of published work. To the best judgment of the editors, the selected notes and manuscripts fully and faithfully represent the essential unpublished writings of Zermelo concerning mathematics. Nevertheless, a possible edition of a third volume comprising further unpublished notes and letters from the *Nachlass* has *expressis verbis* been left open.

Both volumes contain some writings by other authors which include contributions actually written by Zermelo or which react to criticism Zermelo had made. Details are given in the prefaces to the respective volumes.

In order to provide access to a wider audience, the original papers are printed face to face with English translations. As both versions use the same layout, it is easy to go from the translation to the original version and vice versa. The layout itself tries to preserve the appearance of the original papers. For details we refer to the editorial information below.

Each paper or coherent group of papers is preceded by an introductory note which comments on contents, motivation, aims, and influence of the paper(s) concerned. Written by an expert in the field, it came to its final form in discussions with the editors.

Each volume contains a full bibliography of Zermelo together with a schematic *curriculum vitae* which will enable the reader to become acquainted with the personal circumstances from which a paper arose. In addition, Volume I starts with a more detailed biographical sketch of Zermelo's life and work.

Many of these features found their inspiration in the exemplary edition of Kurt Gödel's collected works by Solomon Feferman, John W. Dawson, Jr., and others.

The edition of Zermelo's collected works has a prehistory. Already as early as 1912, at the age of 41 and faced with a serious recurrence of his tuberculosis,

Zermelo conceived plans for an edition of his collected papers, but did not pursue them when his health improved. In 1949, under likewise deplorable personal circumstances, he tried again, this time approaching several publishers, among them Springer-Verlag. But the difficult situation in post-war Germany precluded such an enterprise. Immediately after Zermelo's death, in 1953, the historian of mathematics Helmuth Gericke and the philosopher Gottfried Martin, who had gotten to know Zermelo in the 1930s in Freiburg, started work on a two-volume edition, in 1956 gaining Paul Bernays as a third editor. Support was provided by the Kant-Gesellschaft. However, the plans were not realized; in 1962 work on the edition came to a definite end.

When in early 2004 new plans for an edition of Zermelo's collected works became more concrete, they found the enthusiastic support of Martin Peters of Springer-Verlag. In discussions with him it became clear very quickly that the edition should provide English translations and detailed comments. As Zermelo had been a member of the Heidelberger Akademie der Wissenschaften, the editors turned to the academy for financial support. The application found warm backing of Hans Günter Dosch, then Sekretar of the class for mathematics and the sciences of the academy. The application was successful. Even more, besides providing generous funding, the academy offered to let the edition appear in its regular series of publications of the class for mathematics and the sciences published by Springer-Verlag.

The editors wish to express deep gratitude to the Heidelberg academy for their ideal, financial support and to Springer-Verlag for their open-minded cooperation. In particular, many thanks go to Hans Günter Dosch and Martin Peters.

Freiburg, Toronto, and Boston
September 2009

Heinz-Dieter Ebbinghaus
Craig G. Fraser
Akihiro Kanamori

Preface to volume II

This second volume concludes the edition of Ernst Zermelo's collected works. The volume focuses on his contributions mainly to analysis and physics. Except for an excursion into physical chemistry (*Riesenfeld and Zermelo 1909*), the papers come from the decade around 1900 when Zermelo was in Berlin and Göttingen and about two years around 1930 when he was in Freiburg. They are accompanied by three items found in Zermelo's and in David Hilbert's *Nachlass*. For orientation especially about the personal circumstances accompanying the genesis of the papers, the volume starts with Zermelo's *curriculum vitae*, the one given in volume I.

Zermelo's works of an applied character may hold pioneering ideas and insights, but they did not receive the attention they deserved. One reason may be the sheer diversity of topics he treated. Of course, one should also take into consideration that starting soon after the turn of the century his mathematical work shifted elsewhere for more than two decades, to set theory and mathematical logic, research in these disciplines leading him to his most influential scientific achievements.

The Berlin-Göttingen period comprises three topics: the calculus of variations, the kinetic theory of gases, and hydrodynamics.

The engagement with the calculus of variations started with Zermelo's Ph. D. thesis (*1894*), written at the University of Berlin under the guidance of Hermann Amandus Schwarz.

The engagement in the kinetic theory of gases started in 1896, also in Berlin, when Zermelo became an assistant to Max Planck. It lasted for about ten years. Its best-known part, a controversy with Ludwig Boltzmann, is described and analyzed here in full with the inclusion of *Boltzmann 1896, 1897*.

Zermelo's interest in meteorology led him to hydrodynamics, work that culminated in his 1899 *Habilitation* thesis (*1902a, s1902b, s1902c*) in Göttingen.

In the late 1920s, Zermelo came back to his "old, even though mostly unhappy love for the 'applications' ". The starting paper (*1928*) on the evaluation of chess tournaments, with its early use of the maximum likelihood method, was to remain unknown until several decades later other people rediscovered his methods and results. Motivated by the circumnavigation of the earth by the airship Graf Zeppelin in August 1929, Zermelo wrote two papers (*1930c, 1931a*) on optimal steering methods of airships. Soon, however, this return to mathematics of an applied character came to an end when Zermelo got involved in a serious foundational debate which fully occupied what strength was left him after a serious illness.

The introductory notes are a crucial part of the Zermelo edition. Those who agreed to comment on a paper or a group of papers in this volume generously

shared their experience and knowledge with us and the potential reader. We at times had involved discussions toward securing the most informative and accurate presentations, and we appreciate the professionalism that was brought to bear.

The translations of the original papers were carried out by Enzo de Pellegrin. We again admire his extraordinary care and his feeling for both languages when handling Zermelo's style with its richness in nuances and its involved sentential structures. The introductory notes of Rüdiger Thiele were translated by David Kramer who with diligence and care successfully mirrored the style of the original German.

We express our gratitude to all who have supported us during our work. In this connection we would like to mention Ruth Allewelt from Springer-Verlag, Andrea Köhler and Petra Möws of Le-Tex Publishing Services, and Marlies Würth, the librarian of the Freiburg Mathematical Institute.

Again, Martin Peters of Springer-Verlag was ready to offer valuable help and advice.

We appreciate that Craig Fraser, while not being able to continue with his participation in the edition, was ready to contribute two substantial introductory notes.

Freiburg and Boston
December 2012

Heinz-Dieter Ebbinghaus
Akihiro Kanamori

Editorial information

Layout. The layout of the texts as well as of the translations mirrors the layout of the originals. Emphasized words, i.e., words in italics or words spaced out or consisting of small capitals, are given in italics. Original pagebreaks are indicated in the texts by "|", and the number of the new original page beginning there is given on the margin.

Editorial annotations. These are set in double square brackets "⟦" and "⟧".

Misprints and errors. Small textual errors in the originals are tacitly corrected; larger ones are corrected with the corrections commented on in editorial annotations.

Wrong words or words missing in the originals have been replaced or added in double square brackets.

Misprints in mathematical expressions in the originals are not corrected in the texts. They are, however, corrected in the translations and noted by an editorial annotation.

References. In the texts Zermelo's references to the literature are not altered. Translations as well as introductory notes refer to the main bibliography at the end of the volume instead and have the form *author(s) year of appearance*, followed by an additional index *a*, *b*, *c*, ... if necessary. An example: *Hahn and Zermelo 1904*. If the authors are clear from the context, their names may be omitted; in such a case, *1904* may be short for *Hahn and Zermelo 1904*. References to page numbers are kept in both the texts and the translations; they can be traced via the original pagebreaks and the original page numbers provided in the texts.

Footnotes. Whereas the translations use natural numbers in ascending order as footnote marks, the texts preserve the original marks. It may thus happen that a page of the text may contain identical footnote marks. In such cases the original page numbers on the margin allow for quick correlation of mark and footnote.

Figures. Whenever possible, a figure is located at the same position as in the original. If this is not possible for a figure, say Fig. n, then its original position is indicated on the margin by "Fig. n" and the figure itself appears as close as possible, at worst on the top of the next page.

Copyright permissions

The editors express their gratitude to the following copyright holders for granting permission for the inclusion of the original texts and their English translations:

Universitätsbibliothek der Humboldt-Universität Berlin (*1894*);

Wiley-VCH (*1896a, Boltzmann 1896, 1896b, Boltzmann 1897, 1931a, 1933a*);

Akademie der Wissenschaften zu Göttingen (*1899a*);

Niedersächsische Staats- und Universitätsbibliothek Göttingen (*s1899b*);

S. Hirzel Verlag (*1900, Riesenfeld and Zermelo 1909*);

B. G. Teubner (*1902a, 1902d, Hahn and Zermelo 1904, 1906, 1930c*);

Universitätsarchiv Freiburg (*s1902b, s1902c*);

Springer Verlag (*1904a, 1928*).

Further thanks go to the archive of the Universität Zürich for granting permission for the inclusion of the title photo (Universitätsarchiv Zürich, sign. (UAZ) AB.1.1165).

Editors and contributors of introductory notes

Hans-Georg Bartel
Institut für Chemie
Humboldt-Universität zu Berlin
12489 Berlin
Germany
hg.bartel@yahoo.de

Alexey V. Borisov
Institute of Computer Science
Udmurt State University
Izhevsk 426034
Russia
borisov@rcd.ru

Rainer Bruggemann
Department of Ecohydrology
Leibniz-Institute of Freshwater Ecology and Inland Fisheries
Müggelseedamm 310
12587 Berlin
Germany
brg_home@web.de

Heinz-Dieter Ebbinghaus
Mathematisches Institut
Universität Freiburg
79104 Freiburg
Germany
hde@uni-freiburg.de

Craig G. Fraser
Institute for the History and Philosophy of Science and Technology
Victoria College
University of Toronto
Toronto, Ontario M5S 1K7
Canada
cfraser@chass.utoronto.ca

Larisa A. Gazizullina
Institute of Computer Science
Udmurt State University
Izhevsk 426034
Russia
lag@rcd.ru

Mark E. Glickman
Center for Health Quality, Outcomes & Economics Research
Edith Nourse Rogers Memorial Hospital
200 Springs Road, mail drop 152
Bedford, MA 01730
USA
mg@bu.edu

Akihiro Kanamori
Department of Mathematics
Boston University
Boston, MA 02215
USA
aki@math.bu.edu

Sergei M. Ramodanov
Institute of Computer Science
Udmurt State University
Izhevsk 426034
Russia
ramodanov@mail.ru

Rüdiger Thiele
Senefelderstraße 7
06114 Halle
Germany
ruediger.thiele3@freenet.de

Jos Uffink
Department of Philosophy
University of Minnesota
Minneapolis, MN 55455
USA
jbuffink@umn.edu

Contents of volume II

Contents of volume I

Corrections to volume I

pp. xxiii–xxiv: Some English titles have been slightly changed.

p. xxiv, line 12: Replace "*Zermelo 1903*" by "*Zermelo 1904a*".

p. 239, footnote 4: Replace "art. 64" by "art. 66" and delete the part in brackets.

p. 364, lines 21/22 and p. 365, lines 23/24: Instead of

> "indem sie sich nicht sowohl auf die einzelnen 'definiten' Sätze p als auf ihre *Gesamtheit* P bezieht"

Zermelo means

> "indem sie sich nicht auf die einzelnen 'definiten' Sätze p, sondern vielmehr auf ihre *Gesamtheit* P bezieht".

Hence, instead of

> "since it does not refer both to the individual 'definite' propositions p and to their *totality* P"

the translation should be changed to

> "since it does not refer to the individual 'definite' propositions p, but rather to their *totality* P".

p. 367, line 1: Delete "0".

p. 537, line -18: Replace "*Husserl 1928*" by "*Husserl 1922*".

p. 539, line -1: Replace "partial ground relation" by "partial justification relation".

p. 540, line 6: Add a right parenthesis after "Sect. 1".

p. 541, line -3: Replace "the opening paragraph of Sect. 1" by "the paragraph preceding Sect. 1".

p. 642, line 11: Replace "1903" by "1904a" and shift the newly named item *1904a* behind item *1904*.

Ernst Zermelo

Collected Works
Gesammelte Werke

Volume II
Band II

Calculus of Variations,
Applied Mathematics,
and Physics

Variationsrechnung,
Angewandte Mathematik
und Physik

Ernst Zermelo's *curriculum vitae*

Heinz-Dieter Ebbinghaus

1871

27 July: Zermelo is born in Berlin as the second child and the only son of the *Gymnasialprofessor* Theodor Zermelo and his wife Auguste neé Zieger.

1878

3 June: Death of Zermelo's mother.

1880

April: Zermelo enters the Luisenstädtisches Gymnasium in Berlin.

1889

24 January: Death of Zermelo's father.
1 March: Zermelo finishes school. Remarks in his leaving certificate show that he suffers from physical fatigue.
Summer semester – summer semester 1890: Zermelo studies mathematics and physics at the University of Berlin with, among others, Lazarus Fuchs and Johannes Knoblauch.

1890

Winter semester 1890/91: Zermelo studies at the University of Halle-Wittenberg with, among others, Georg Cantor and Edmund Husserl.

1891

Summer semester 1891: Zermelo studies at the University of Freiburg, with his subjects including, as in Berlin and Halle, philosophy and psychology.
Winter semester 1891/92 – summer semester 1897: Zermelo studies again at the University of Berlin with, among others, Ferdinand Frobenius, Max Planck, Hermann Amandus Schwarz, and Wilhelm Wien.

1894

23 March: Zermelo applies to begin the Ph. D. process.
6 October: Zermelo obtains his Ph. D. degree. His dissertation *Untersuchungen zur Variations-Rechnung* was supervised by Hermann Amandus Schwarz.
December – September 1897: Zermelo is an assistant to Max Planck at the Institute for Theoretical Physics of the University of Berlin.

1895

December: Zermelo completes his first paper, *1896a,* which sets out his opposition to Ludwig Boltzmann's statistical theory of heat.

1896

Summer: Zermelo applies for an assistantship at the Deutsche Seewarte in Hamburg, but then decides to pursue an academic career.

H.-D. Ebbinghaus, A. Kanamori (Eds.), *Ernst Zermelo - Collected Works/ Gesammelte Werke II*, DOI 10.1007/978-3-540-70856-8_1, Schriften der Mathematisch-naturwissenschaftlichen Klasse der Heidelberger Akademie der Wissenschaften 23,

15 September: Zermelo completes his second paper, *1896b,* opposing Boltzmann.

1897

German translation *Glazebrook 1897* of *Glazebrook 1894.*
2 February: Zermelo passes the state exam for *Gymnasiallehrer* (high school teachers) that allowed him to teach mathematics and physics as main subjects and chemistry, geography, and mineralogy as additional subjects. According to the reports Zermelo exhibited a broad knowledge in German literature, philosophy, and religion.
19 July: Zermelo asks Felix Klein in Göttingen for support for his *Habilitation.*
Winter semester 1897/98: Zermelo continues his studies at Göttingen with, among others, David Hilbert, Felix Klein, and Arthur Schoenflies.

1899

3 February: David Hilbert presents Zermelo's first paper in applied mathematics, *1899a,* to the Königliche Gesellschaft der Wissenschaften zu Göttingen; it treats differential equations with inequalities.
Zermelo initiates the *Habilitation* process with the *Habilitation* thesis "Hydrodynamische Untersuchungen über die Wirbelbewegungen in einer Kugelfläche" the first part of which is published as *1902a.* The second part (*1902b, 1902c*) remained unpublished; it contains a solution of the 3-vortex problem on the sphere.
4 March: Zermelo gives his *Habilitation* address, *1900,* which proposes an alternative probabilistic approach to Bolzmann's in the latter's work in statistical mechanics. He is granted the *venia legendi* for mathematics at the University of Göttingen.

Around 1900

Beginning of the cooperation with Hilbert on the foundations of mathematics. Zermelo formulates the Zermelo-Russell paradox.

1900

Winter semester 1900/01: Zermelo gives his first course on set theory, the main topic being the Cantorian theory of cardinals.

1901

9 March: David Hilbert presents Zermelo's result on the addition of cardinals, *1901,* to the Königliche Gesellschaft der Wissenschaften zu Göttingen. The proof uses the axiom of choice.

1902

12 May: Zermelo gives a talk on Frege's foundation of arithmetic before the Göttingen Mathematical Society.
Summer semester – winter semester 1906/07: Zermelo receives a *Privatdozenten* grant.

Publication of *1902d,* the first paper on the calculus of variations after the Ph.D. dissertation. It treats shortest lines of bounded steepness with or without bounded torsion.

1903

June: Zermelo is under consideration for an extraordinary professorship of mathematics at the University of Breslau. He is shortlisted in the second position after Gerhard Kowalewski, Franz London, and Josef Wellstein who are shortlisted *aequo loco* in the first position.
1 December: Zermelo completes his second paper on the calculus of variations, *1904a.* It gives two simple proofs of a result of Paul du Bois-Reymond on the range of the method of Lagrange.

1904

Beginning of a life-long friendship with Constantin Carathéodory.
Together with Hans Hahn, Zermelo writes a contribution on the calculus of variations, *Hahn and Zermelo 1904,* for the *Encyklopädie der mathematischen Wissenschaften.*
August: Third International Congress of Mathematicians at Heidelberg. Julius König gives a flawed refutation of Cantor's continuum hypothesis. The error is detected by both Zermelo and Felix Hausdorff.
24 September: Zermelo informs Hilbert about his proof of the well-ordering theorem and the essential role of the axiom of choice. The letter is published as *1904.*
15 November: During a meeting of the Göttingen Mathematical Society, Zermelo defends his well-ordering proof against criticism by Julius König, Felix Bernstein, and Arthur Schoenflies.

1905

January: Zermelo falls seriously ill. In order to recover, he spends spring and early summer in Italy.
German translation *Gibbs 1905* of *Gibbs 1902.*
Spring: Zermelo works on the theory of finite sets which finally results in *1909a* and *1909b.*
21 December: Zermelo receives the title "Professor". The application had been filed by Hilbert in December 1904.

1906

Early that year: Zermelo catches pleurisy.
Zermelo works on a book on the calculus of variations together with Carathéodory.
Zermelo publishes a final criticism of Boltzmann's statistical interpretation of the second law of thermodynamics in the review *1906* of *Gibbs 1902.*
Summer semester: Zermelo lectures on "Mengenlehre und Zahlbegriff". He formulates an axiom system of set theory which comes close to the one published by him in 1908.
June: Medical doctors diagnose tuberculosis of the lungs.

Summer: Zermelo spends a longer time at the seaside.
Autumn: Zermelo is under discussion for a full professorship of mathematics at the University of Würzburg. The professorship is given to the extraordinary Würzburg professor Georg Rost. According to Hermann Minkowski Zermelo's difficulties in obtaining a professorship are rooted in his "nervous haste".
Winter 1906/07 – winter 1907/08: Several extended stays in Swiss health resorts for lung diseases.

1907

March: Zermelo applies for a professorship at the Academy of Agriculture in Poppelsdorf without success.
May: During a stay in Montreux Zermelo finishes his paper *1909a.*
14 July and 30 July: During a stay in the Swiss alps Zermelo completes his papers on a new proof of the well-ordering theorem and on the axiomatization of set theory, *1908a* and *1908b,* respectively.
20 August: Following an application by the Göttingen Seminar of Mathematics and Physics, the ministry commissions Zermelo to give lecture courses in mathematical logic and related matters, thus installing the first official lectureship for mathematical logic in Germany.

1908

April: Fourth International Congress of Mathematicians in Rome. Zermelo presents his work on finite sets, *1909b.* He becomes acquainted with Bertrand Russell. Together with Gerhard Hessenberg and Hugo Dingler he conceives plans for establishing a quarterly journal for the foundations of mathematics. The project fails because of diverging views between the group and the Teubner publishing house.
Summer semester: Zermelo gives a course on mathematical logic in fulfilment of his lectureship for mathematical logic and related topics.

1909

July: Zermelo is under consideration for an extraordinary professorship of mathematics at the University of Würzburg. He is shortlisted in the first position. The professorship is given to Emil Hilb shortlisted in the second position.
September: Completion of *Riesenfeld and Zermelo 1909.*

1910

24 January: The board of directors of the Göttingen Seminar of Mathematics and Physics applies to the minister to appoint Zermelo an extraordinary professor.
21 January: Zermelo, being under consideration for a full professorship of mathematics at the University of Zurich, is shortlisted in the first position.
24 February: The *Regierungsrat* of the Canton Zurich approves the choice of Zermelo.
15 April: Zermelo is appointed a full professor at the University of Zurich for an initial period of six years.

1911

28 January: Zermelo applies for leave for the coming summer semester because of a worsening of his tuberculosis.
February and March: Together with a partner, Zermelo applies for several patents concerning, for example, a regulator for controlling the revolutions of a machine.
Zermelo is awarded the interest from the Wolfskehl prize, Hilbert being chairman of the Wolfskehl committee of the Gesellschaft der Wissenschaften zu Göttingen.
Summer semester – winter semester 1911/12: Leave for a cure because of tuberculosis.

1912

January: Serious worsening of tuberculosis diagnosed.
Beginning of the cooperation with Paul Bernays who completes his *Habilitation* with Zermelo in 1913 and stays at the University of Zurich as an assistant to Zermelo and later as a *Privatdozent* until 1919.
August: Fifth International Congress of Mathematicians in Cambridge. Following an invitation by Bertrand Russell, Zermelo gives two talks, one on axiomatic and genetic methods in the foundation of mathematical disciplines and one on the game of chess. The second one results in the paper *1913* which may be considered the first paper in game theory.
Faced with the seriousness of his illness, Zermelo conceives plans for an edition of his collected papers.

1913

Spring: Zermelo is considered for a full professorship in mathematics at the Technical University of Breslau. He is shortlisted in the first position. The professorship is given to Max Dehn, shortlisted in the second position together with Issai Schur.
December: Zermelo completes his paper *1914* on subrings of whole transcendental numbers of the field of the real numbers and the complex numbers, respectively; it makes essential use of the axiom of choice.

1914

Early that year: Zermelo has regular discussions with Albert Einstein.
March: Operation of the thorax by Ferdinand Sauerbruch, the pioneer of thorax surgery.

Around 1915

Zermelo develops a theory of ordinal numbers where the ordinals are defined as by John von Neumann in 1923.

1915

Spring: A new serious outbreak of tuberculosis forces Zermelo to take a one-year leave.
July: Waldemar Alexandrow completes his Ph. D. thesis *Alexandrow 1915.* It is the only thesis guided by Zermelo alone. Kurt Grelling's thesis *Grelling*

1910, which extends Zermelo's theory of finite sets, was officially supervised by David Hilbert, but guided by Zermelo.
Autumn: Several surgical treatments of a tuberculosis of the vocal chords.

1916

21 March: As his illness is expected to extend into the summer semester, Zermelo is urged to agree to retire.
5 April: Zermelo agrees to retire.
15 April: Zermelo retires from his professorship.
31 October: Zermelo is awarded the annual Alfred Ackermann-Teubner prize of the University of Leipzig for the promotion of the mathematical sciences. Later prize winners include, for example, Emil Artin and Emmy Noether.
1 November – February 1917: Zermelo stays in Göttingen.
7 November 1916: Zermelo presents his theory of ordinal numbers to the Göttingen Mathematical Society.

1917

March – October 1919: Zermelo stays in various health resorts in the Swiss alps.

1919

July: First draft of the paper *1928* wherein Zermelo develops a procedure for evaluating the result of a tournament by using a maximum likelihood method.
November – March 1921: Zermelo stays at Locarno, Switzerland.

1921

Spring: Zermelo stays in Southern Tyrol and has correspondence with Abraham Fraenkel.
6 May 1921: Fraenkel informs Zermelo about a gap he has discovered in Zermelo's 1908 axiom system of set theory.
10 May: In his answer to Fraenkel, Zermelo proposes a second-order version of the axiom of replacement in order to close the gap, at the same time criticizing it because of its non-definite character.
17 July (?): Zermelo formulates his "infinity theses" (*s1921*) where he describes the aims of his research in infinitary languages and infinitary logic as carried out in the early 1930s.
22 September: Fraenkel announces his axiom of replacement in a talk delivered at the annual meeting of the Deutsche Mathematiker-Vereinigung. Zermelo agrees in principle, but maintains a critical attitude because of a deficiency of definiteness.
1 October: Zermelo settles in Freiburg, Germany.

1923

Winter semester 1923/24: Zermelo attends Edmund Husserl's course "Erste Philosophie".
– 1929: Cooperation with Marvin Farber on the development of a semantically based logic system, in 1927 leading to plans for a monograph on logic.

1924

Summer: Zermelo loses interest in Husserl's phenomenology. Discussions with Marvin Farber on the possibility of obtaining a professorship in the USA.

1926

or earlier: Zermelo starts a translation of Homer's *Odyssey*, one that aims at "liveliness as immediate as possible".
22 April: Zermelo is appointed an *ordentlicher Honorarprofessor* at the Mathematical Institute of the University of Freiburg.
Winter semester – winter semester 1934/35: Zermelo gives regular courses in various fields of mathematics.
– 1932: Zermelo works on the edition of Cantor's collected papers, *Cantor 1932*. He is supported by the mathematicians Reinhold Baer and Arnold Scholz and the philosopher Oskar Becker. The participation of Abraham Fraenkel leads to a mutual estrangement.

1927

12 June: Zermelo completes his paper *1927* on measurability, where he presents results which he had obtained around 1914 and which had first been presented in Alexandrow's thesis *Alexandrow 1915*.

1928

3 August: Zermelo completes his paper *1928* on the evaluation of tournaments.

1929

– 1931: Zermelo receives a grant from the Notgemeinschaft der Deutschen Wissenschaft (Deutsche Forschungsgemeinschaft) for a project on the nature and the foundations of pure and applied mathematics and the significance of the infinite in mathematics.
May and June: Zermelo spends several weeks in Poland, giving talks in Cracow and Lvov and a series of talks in Warsaw. In the latter he presents his view of the nature of mathematics, arguing strongly against intuitionism and formalism.
11 July: Zermelo completes his paper *1929a* wherein he responds to criticism of his notion of definiteness as put forward by Abraham Fraenkel, Thoralf Skolem, Hermann Weyl, and others.
18 September: At the annual meeting of the Deutsche Mathematiker-Vereinigung in Prague Zermelo gives a talk, *1930c*, on the solution of what is now called the "Zermelo navigation problem". An extension of the result is published as *1931a*.
Arnold Scholz becomes an assistant at the Mathematical Institute of the University of Freiburg, staying there for five years. Until his death on 1 February 1942, he will be Zermelo's closest friend and scientific partner.

1930

13 April: Zermelo completes the paper *1930a* wherein he formulates the second-order Zermelo-Fraenkel axiom system and presents an incisive picture of the cumulative hierarchy.
– 1932: Zermelo's controversy with Skolem and Gödel about finitary mathematics, in particular about Skolem's first-order approach to set theory and Gödel's first incompleteness theorem.

1931

Zermelo develops infinitary languages and an infinitary logic as a response to Skolem and Gödel.
15 September: Zermelo presents his work on infinitary languages and infinitary logic at the annual meeting of the Deutsche Mathematiker-Vereinigung. The talk results in the polemical paper *1932a* and the straightforward *1932b*.
September/October: Correspondence with Gödel about the proof of Gödel's first incompleteness theorem and Zermelo's infinitistic point of view.
18 December: Zermelo is elected a corresponding member of the Gesellschaft der Wissenschaften zu Göttingen on a proposal of Richard Courant.
– 1935: Zermelo continues his research on infinitary languages and infinitary logic which results in the paper *1935*. He works on large cardinals and on a monograph on set theory.

1932

Spring: Zermelo goes on a cruise that visits ancient sites in Greece, Italy, and Northern Africa.
June: Zermelo devises an electrodynamic clutch for motorcars.
July: Zermelo is invited to contribute a paper to a special issue of *Zeitschrift für angewandte Mathematik und Mechanik* in honor of its founder and editor Richard von Mises. He accepts the invitation and contributes the paper *1933a*.

1933

16 February: Zermelo is elected an extraordinary member of the Heidelberger Akademie der Wissenschaften on a proposal of Heinrich Liebmann and Artur Rosenthal.

1934

Zermelo moves to Bernshof, a remote country house in the hilly outskirts of Freiburg where he lives until his death.
2 November: Zermelo presents his paper *1934* on elementary number theory to the Göttingen academy.

1935

2 March: Zermelo resigns his honorary professorhip when denounced for his unwillingness to give the Hitler salute.
– 1940: Smaller scientific projects in various fields of mathematics, further work on a book on set theory and work on a collection of mathematical "miniatures" representing several of his own results.

1937

4 October: Zermelo gives a flawed refutation, *s1937,* of the existence of countable models of set theory.

1941

19 July: Arnold Scholz organizes a colloquium in Göttingen on the occasion of Zermelo's 70th birthday; Zermelo gives three talks that correspond to three items of his collection of mathematical "miniatures". Other speakers include Konrad Knopp and Bartel van der Waerden.

1944

14 October: Marriage to Gertrud Seekamp.
Zermelo suffers from a glaucoma that can no longer be treated adequately and will eventually lead to total blindness.

1946

23 January: Zermelo turns to the Rector's office of the University of Freiburg to request his reappointment as a full honorary professor.
23 July: Zermelo is reappointed an honorary professor. Because of age and increasing blindness, he is unable to lecture.

1947

– 1948: Zermelo stays several times in Switzerland. He can live there on his pension access to which is barred from Germany, leaving the Zermelos in a difficult financial situation.
– 1950: In order to escape financial need, Zermelo tries to move back to Switzerland. His applications fail; the Swiss authorities argue that his Swiss pension does not suffice to provide for his wife as well.

1949

Spring: Zermelo tries to arrange an edition of his collected works. He fails.

1953

21 May: Zermelo, nearly 82 years old, dies at Bernshof in Freiburg.
23 May: Zermelo is buried.
– 1962: Helmuth Gericke and Gottfried Martin work on the edition of Zermelo's collected works. In 1956 Paul Bernays agrees to take part in the edition. The project was not realized.
The Zermelo *Nachlass* is acquired by the University of Freiburg.

2003

15 December: Gertrud Zermelo, 101 years old, dies at Bernshof in Freiburg.

Introductory note to 1894

Craig G. Fraser

In 1894 Zermelo published his doctoral dissertation from the University of Berlin, written under the direction of Hermann Amandus Schwarz and devoted to a study of Karl Weierstrass's methods in the calculus of variations. In an introductory note Zermelo stated that he had become familiar with the contents of Weierstrass's lectures in 1892 from the copy in the "Mathematischer Verein" in Berlin, as well as from a lecture given by Schwarz. Weierstrass had investigated the simplest case in which only the first derivatives of the variables appear in the variational integrand function. Zermelo's main goal was to extend Weierstrass's results on necessary and sufficient conditions involving the "excess" or E function (the material in the twentieth to the twenty-third lectures of Weierstrass's lectures as they were eventually published (*1927*)) to variational problems in parametric form in which the integrand contains derivatives of order higher than one.

1. Sufficient conditions before Weierstrass

A major goal of the calculus of variations in the nineteenth century was to identify conditions that ensure that a proposed solution to a given variational problem is a maximum or a minimum. Any such solution will have to satisfy the Euler differential equation and will also have to satisfy Legendre's condition. It was noticed that a function that satisfied these conditions turned out in certain instances not to be a genuine extremum. It was required to assemble a set of conditions that taken together are sufficient to ensure a maximum or a minimum. In researches of the late 1830s Carl Gustav Jacobi (*1837, 1838*) introduced some new ideas in this direction that became the basis for a very active program of research. Jacobi formulated a certain condition, known in the later subject as Jacobi's criterion, that must be satisfied by any solution to the problem. Jacobi's theory was also based on a new transformation of the second variation. The variational integrand was expressed in a form that enabled one to infer Legendre's condition for very general integrals. (The relevant history is presented in *Todhunter 1861, Goldstine 1980* and *Fraser 2003*.)

The primary object of interest here is an integral involving a single independent variable of the form $\int_a^b f(x, y, y', ..., y^{(n)})\, dx$, where the integrand function f is a function of x, y and the derivatives of y with respect to x up to order n. It is necessary to find the particular function $y = y(x)$ that maximizes or minimizes this integral. In the elementary case where $n = 1$, researches succeeded in providing a completely satisfactory theory. In *1857*

H.-D. Ebbinghaus, A. Kanamori (Eds.), *Ernst Zermelo - Collected Works/ Gesammelte Werke II*, DOI 10.1007/978-3-540-70856-8_2, Schriften der Mathematisch-naturwissenschaftlichen Klasse der Heidelberger Akademie der Wissenschaften 23,

Ludwig Otto Hesse showed in this case that if the Euler, Legendre and Jacobi conditions are satisfied then the resulting curve is indeed a maximum or a minimum. The more interesting and much more difficult case occurred when $n \geq 2$. Here it was found that constants appearing in functions required in the transformation of the second variation must satisfy certain conditions. It was necessary to show that it was possible to find a set of constants that worked in the general case. Essentially the problem was one of existence, of finding suitable mathematical objects that allowed the transformation to take place. (A succinct statement of the point in question is given in *Lindelöf and Moigno 1861.*)

The central question was resolved by the Leipzig mathematician Adolph Mayer in his *Habilitation* thesis *1866*, a work whose core content was presented by Mayer two years later in an article in Crelle's journal (*1868*). Mayer showed that if Jacobi's criterion held, then it was possible to carry out the desired transformation of the second variation. Assuming the validity of Legendre's criterion, one may infer that the given function satisfying the Euler equation is indeed a maximum or a minimum. Mayer presented his result in a very general setting, using a formulation of the variational problem that had been developed by Alfred Clebsch. Mayer's investigation showed both technical sophistication and a deep understanding of the theoretical issues at the foundation of the theory.

2. Weierstrass

Weierstrass's contributions to the calculus of variations were a product of his middle and late years. Although he began lecturing on the subject at the University of Berlin as early as 1865, his most significant results were presented in the summer lectures of 1879, when he was sixty-three years old. The edition which was eventually published in 1927 is based on these as well as a second set of lectures given in 1883. Although this delay in publication somewhat limited the dissemination of his ideas, he exerted considerable influence on contemporary variational research. Copies of his notes circulated privately and his results began to be disseminated in published form by other researchers beginning in the middle 1890s. The appearance of Zermelo's dissertation in 1895 was among the first publications of Weierstrass's ideas, developed in a more general setting than the one adopted by Weierstrass.

More than any other researcher Weierstrass established the critical outlook of the calculus of variations as a modern mathematical subject. In his lectures the distinction between necessary and sufficient conditions appears clearly for the first time. He carefully specified the continuity properties that must be satisfied by functions and their variations. In problems of constrained optimization he used theorems on implicit functions to ensure that the optimizing arc was embedded in a suitable family of comparison curves.

Traditionally researchers in the calculus of variations did not identify at the outset of their investigation the precise class of comparison arcs in a given

variational problem. There was no prior logical conception concerning the nature of this class. However, the δ-process introduced by Lagrange required that both the comparison arc and its slope at each point differ by only a small amount from the value and slope of the solution curve. This condition was imposed by the nature of the variational process, which involved expanding the integrand function as a Taylor series and investigating the behavior of the second variation arising in this expansion. Isaac Todhunter (*1871*, 269) in an essay on what were known as "discontinuous" solutions seems to have been the first to explicitly call attention to this limitation on the class of comparison arcs:

> [. . .] if we assert that the relation [Euler equation] does give a minimum, we must bear in mind that this means a minimum with respect to admissible variations [. . .] our investigation is not applicable to such a variation as would be required in passing from the cycloid to the discontinuous figure: in such a passage p [$= \delta y'$] would not always be indefinitely small. Of course it might be possible to give some special investigation for such a case, but certainly the case is not included in the ordinary methods of the Calculus of Variations.

In his Berlin lectures Weierstrass developed an alternative to the traditional expansion methods that extended the variational theory to a larger class of comparison curves. The precise nature of these curves was still determined by the particular technical requirements of the new method, but the logical orientation of the subject had shifted. In earlier variational research the nature of the mathematical objects was determined implicitly by the variational process that was employed. By contrast, in Weierstrass there was a self-conscious and explicit focus on the objects being studied. His work involved a more intimate connection between the foundations of real analysis and the collection of concrete techniques and results that made up the variational theory.

It is necessary to call attention to one aspect of the style in which Weierstrass developed the theory. Traditionally researchers in the calculus of variations had adopted what is referred to as the ordinary or functional approach, in which the curve is expressed as $y = y(x)$ and the variational integrand (in the simplest case) takes the form $f(x, y, y')$. A distinctive aspect of Weierstrass's approach was his adoption of a parametric approach. The curve C is represented parametrically in the form $y = y(t)$ and $x = x(t)$. Here the variational integrand takes the form $I = \displaystyle\int_{t_0}^{t_1} f(x, y, x', y')\, dt$ where $x' = \dfrac{dx}{dt}$ and $y' = \dfrac{dy}{dt}$. In the parametric (or homogeneous) formulation of the variational problem it is necessary to impose conditions on the variational integrand in order to ensure that the problem is independent of the particular parameterization chosen. One must attend to these conditions in developing the theory.

Weierstrass used a parametric approach throughout his lectures on the calculus of variations. Researchers who adopted his new method tended to also use a parametric approach, and this was true in the case of Zermelo. However, not all researchers followed this practice. Although the parametric approach has certain advantages, particularly from a geometric viewpoint, its analytical development is less natural than the ordinary theory. During the years around 1900 when Weierstrass's ideas were becoming more widely known, researchers such as Oskar Bolza (*1909*), William Osgood (1901) and Emile Goursat (*1905*) went to some effort to reformulate his results in terms of the ordinary theory. In the large majority of the textbook literature of the past one hundred years the ordinary approach is taken as the standard formulation of the variational problem while the parametric theory is presented as a special subject.

For the sake of exposition we will adopt the ordinary theory in explaining some of the basic ideas that underlie Weierstrass's theory. At the conclusion we indicate by way of comparison the parametric form in which Weierstrass originally presented his result. (In our account we have adopted the sensible notation used by Adolf Kneser to denote extremal and comparison curves: $y = y(x)$ is the comparison curve, while $\overline{y} = \overline{y}(x)$ is the extremal curve; both Weierstrass and Zermelo use the opposite convention.)

Suppose $y = y_0(x)$ is an arc C_0 on which the variational integral $I = \int_a^b f(x, y, y')\, dx$ is a minimum. (The case of a maximum is similar, with the inequalities reversed.) In the traditional formulation of the theory involving expansion methods and the second variation, the conditions that must be satisfied are all specified in terms of the function $y = y_0(x)$ and this function alone. Thus $y = y_0(x)$ will be a solution to the Euler equation and we must have $\dfrac{\partial^2 f(x, y_0(x), y_0'(x))}{\partial y'^2} \geq 0$ (Legendre's condition) and the Jacobi criterion must hold. One typically supposes that the family of comparison curves is of the form $C : y = y_0(x) + \epsilon\zeta(x)$. Here $\zeta(x)$ is any function subject to the usual continuity restrictions with $\zeta(a) = \zeta(b) = 0$. More generally we may have a family of comparison curves of the form $C : y = y(x, \epsilon)$ where $y(x, 0) = y_0(x)$ and $y(a, \epsilon) = y(b, \epsilon) = 0$. It is evident that the class of comparison curves is very extensive. Nevertheless, as Todhunter observed in 1870, it is also clear that there are some restrictions on this class. For example, if $y = y_0(x) + \epsilon\zeta(x)$ (ϵ small) then $y'(x) - y_0'(x) = \epsilon\zeta'(x)$ with similar relations for higher derivatives of x. It follows that the neighboring curve $y = y(x)$ differs by only a small amount from the optimizing curve $y = y_0(x)$ not just for corresponding values of y but for derivatives of y of all orders.

It turns out that it is possible for a variational integral to be a minimum for the function $y = y_0(x)$, considered with respect to a class of comparison curves of the type $y = y(x, \epsilon)$, but not be a minimum if we allow comparison

curves whose slope differs by a finite amount from $y = y_0(x)$. The following simple example illustrates this situation. It is taken from Bolza's *Lectures on the calculus of variations* (*1904*, 74), one of the earliest expositions of the new theory developed by Weierstrass and extended by Zermelo. (The term "die neue Variationsrechnung" was sometimes used in reference to the Weierstrass theory.) We have $f(x, y, y') = y'^2 + y'^3$ defined on the interval $[0, 1]$.The variational integral is $I = \int_0^1 (y'^2 + y'^3)\, dx$. It follows that a solution to the Euler equation will be $y' = constant$. Since the solution must pass through the endpoints it follows that the hypothetical minimizing curve is simply $y = 0$. We have $\dfrac{\partial^2 f(x, y, y')}{\partial y'^2} = 2 \geq 0$, so the Legendre condition is satisfied. It is also the case that Jacobi's condition holds in this example. Hence the curve $y = 0$ minimizes the integral with respect to comparison arcs of the form $y = y(x, \epsilon)$. Consider now the comparison curve C consisting of two straight lines, the first joining the origin $(0, 0)$ to the point $(1 - p, q)$ and the second joining $(1 - p, q)$ to $(1, 0)$ (see Fig. 1). Here p is a number with $0 < p < 1$ and q is a small positive quantity. For this comparison curve we have

$$\Delta I = \frac{q^2}{p(1-p)}\left(1 + \frac{q}{1-p} - \frac{q}{p}\right).$$

We can make C lie within any neighborhood of $y = 0$ by making q sufficiently small. With q specified, it is clear that $\Delta I < 0$ for $p \ll q$. Hence I is not a minimum for the larger class of curves that includes the curve C.

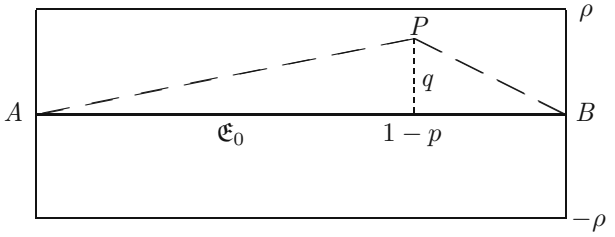


Fig. 1 (p. 74 of *Bolza 1904*)

In the terminology that was introduced by Kneser in *1900* and became standard, the traditional variational theory yields sufficient conditions for a *weak* extremum. Here each comparison curve is close to the minimizing curve at y and at all derivatives of y. By contrast, a solution will be a *strong* extremum if it is a minimum for the wider class of curves which are close to the solution curve but may have a slope that differs by a finite amount from the solution curve.

Consider again the problem of finding the curve $C_0 : y = y_0(x)$ that maximizes or minimizes $I = \int_a^b f(x, y, y')\, dx$. Suppose that the Euler and

Jacobi conditions hold for the arc C_0. We now enlarge the class of possible comparison curves to include ones whose slope differs by a finite amount from that of C. In order to establish that C_0 is a minimum with respect to this enlarged class of comparison curves it is necessary to formulate a condition that involves not just the function $y = y_0(x)$ but also the curves $C : y = y(x)$ in the comparison class. Perhaps the simplest approach would be simply to require that

$$f(x, y_0, y_0') \leq f(x, y, y') \ (a \leq x \leq b)$$

for all comparison curves $y = y(x)$. Imposing this condition is evidently not very informative, and indeed is only a restatement of the problem. Weierstrass succeeded in formulating a more meaningful condition involving a function E called the excess function. We are given the proposed minimizing arc $C_0 : y = y_0(x)$. We consider any comparison curve $C : y = y(x)$ and take any point on this curve with coordinates x and y. By assumption the point (x, y) on the comparison curve C is close to the point (x, y_0) on C_0. Let $y'(x)$ be the slope of the comparison curve at the given point. This quantity may vary by an arbitrary amount from $y_0'(x)$.

A solution to the Euler equation is called an extremal. We typically require that such a solution passes through the initial point. A key idea introduced by Weierstrass was to introduce a function $p(x, y)$—known as the slope function—defined as the slope of the extremal passing through the point (x, y) at this point. We now introduce the excess function $E(x, y, y', p)$ defined as

$$E(x, y, y', p) = f(x, y, y') - f(x, y, p) - (y' - p)\frac{\partial f}{\partial y'}(x, y, p)\,. \tag{1}$$

Consider a region or strip about the curve C_0 and suppose that for each point in this strip it is possible to determine an extremal joining the initial point and the given point. Such a set of solutions to the Euler equation is today called a field of extremals, a term introduced by Kneser. Consider the condition

$$E(x, y, y', p) \geq 0\,. \tag{2}$$

This condition is known as Weierstrass's condition. If this condition is satisfied for all comparison curves $y = y(x)$ on the interval then the value of I along C_0 is less than its value along any member of C of the comparison class. We may conclude in this case that $C_0 : y = y_0(x)$ is a strong minimum.

In later mathematics this result would be proved using something called the Hilbert invariant integral, introduced by David Hilbert in 1900 to simplify Weierstrass's method. Here we present Weierstrass's original idea, which was also adopted by Zermelo. The minimizing arc is given as $C_0 : (x, y_0(x))$ while

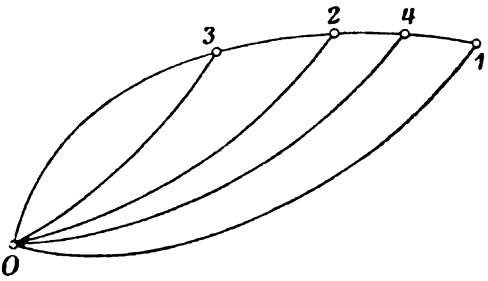


Fig. 2 (p. 219 of *Weierstrass 1927*)

the neighboring curve C is given as $C : (x, y(x))$. In Fig. 2 C_0 is the bottom arc **01**, while the neighboring curve C is the arc **0321**. We suppose that C lies in a narrow region about C_0. This region or "field" has the following property: for any point within it, there is a unique solution curve to the Euler equation—an extremal—passing through the initial point **0** and the given point. We designate extremal curves using the functional notation $\overline{y} = \overline{y}(x)$. Let (x, y) be any point on C; in Fig. 1 this point is labelled **2**. Consider the extremal curve $(x, \overline{y}(x))$ passing from **0** through **2**; in Fig. 1 it is the curve **02**. We introduce the function $\phi(x)$ defined as

$$\phi(x) = \int_{\mathbf{0}}^{\mathbf{2}} f(x, \overline{y}, \overline{y}')\, dx + \int_{\mathbf{2}}^{\mathbf{1}} f(x, y, y')\, dx\,. \tag{3}$$

This integral is taken along the extremal arc from **0** to **2** and then along the comparison arc from **2** to **1**. Evidently we have $\phi(0) = \displaystyle\int_{\mathbf{0}}^{\mathbf{1}} f(x, y, y')\, dx$ and $\phi(1) = \displaystyle\int_{\mathbf{0}}^{\mathbf{1}} f(x, y_0, y_0')\, dx$. The statement that $\displaystyle\int_{\mathbf{0}}^{\mathbf{1}} f(x, y_0, y_0')\, dx$ is a minimum is equivalent to the inequality $\phi(0) \geq \phi(1)$, which would follow if we are able to show that $\phi(x)$ is a decreasing function of x. To do this we calculate the derivative of $\phi(x)$ and show that it is negative.

Let the integrals $\overline{I}_{02}$ and I_{21} be defined as

$$\overline{I}_{02} = \int_{\mathbf{0}}^{\mathbf{2}} f(x, \overline{y}, \overline{y}')\, dx,\ \ I_{21} = \int_{\mathbf{2}}^{\mathbf{1}} f(x, y, y')\, dx\,. \tag{4}$$

From the standard formula for the variation of the integral when the endpoint is allowed to vary in both the x and y directions we have

$$\delta\overline{I}_{02} = \frac{\partial f}{\partial \overline{y}'}(x, \overline{y}, \overline{y}')\delta y + \big(f(x, \overline{y}, \overline{y}') - \overline{y}'\frac{\partial f}{\partial \overline{y}'}(x, \overline{y}, \overline{y}')\big)\delta x\,. \tag{5}$$

We now let $\delta x = dx$ and $\delta y = dy$. Note that at the point **2** we have $\overline{y} = y$ and $\overline{y}' = p(x, y)$, where p is the slope function for the given field. Hence (5)

becomes

$$\begin{aligned}\frac{d\overline{I}_{02}}{dx} &= \frac{\partial f}{\partial y'}(x,y,p)y' + f(x,y,p) - p\frac{\partial f}{\partial y'}(x,y,p) \\ &= \frac{\partial f}{\partial y'}(x,y,p)(y'-p) + f(x,y,p)\,.\end{aligned} \tag{6}$$

The derivative of I_{21} is given immediately as

$$\frac{dI_{21}}{dx} = -f(x,y,y')\,. \tag{7}$$

Thus we have

$$\begin{aligned}\phi'(x) &= \frac{d\overline{I}_{02}}{dx} + \frac{dI_{21}}{dx} \\ &= -\left(f(x,y,y') - f(x,y,p) - (y'-p)\frac{\partial f}{\partial y'}(x,y,p)\right),\end{aligned}$$

or

$$\phi'(x) = -E(x,y,y',p)\,. \tag{8}$$

If $E(x,y,y',p) \geq 0$ on the interval then $\phi'(x)$ is negative and $\phi(x)$ is a decreasing function, which is what was required to be proved.

It is apparent that $\phi(0) - \phi(1) = \int_{\mathbf{0}}^{\mathbf{1}} E(x,y,y',p)\,dx$. But $\phi(0) - \phi(1)$ is equal to ΔI, the variation of the integral with respect to the comparison arc. Hence we have

$$\Delta I = \int_{\mathbf{0}}^{\mathbf{1}} E(x,y,y',p)\,dx\,. \tag{9}$$

(9) is known in the modern literature as Weierstrass's theorem, although it does not appear explicitly in Weierstrass's lectures. From (9) it is clear why the function E is called the excess function, since the excess of the variational integral I in going from $y = y_0(x)$ to $y = y(x)$ is the integral of E over the given interval.

In Weierstrass's original parametric approach the variational integrand takes the form $\int_{t_0}^{t_1} F(x,y,x',y')\,dt$. The minimizing arc is given as C_0 : $(x_0(t), y_0(t))$ while the neighboring curve C is given as C : $(x(t), y(t))$. In Fig. 2 C_0 is the bottom arc **01**, while the neighboring curve C is the arc **0321**. We suppose that C lies in a narrow region about C_0. This region is supposed to be a "field" in the sense defined above: for any point within it there is a unique solution curve to the Euler equation passing through the initial point and the given point. Let $(x(t), y(t)$ be any point on C; in Fig. 2 this point is labelled **2**. Consider the extremal passing through this point; in Fig. 2 it is the curve **02**. Let $p(t), q(t)$ be the coordinate slope functions of

the extremal at (x, y). The excess function E in parametric form is defined as

$$E(x, y, p, q, x', y') = F(x, y, x', y') - x' \frac{\partial F}{\partial p}(x, y, p, q) - y' \frac{\partial F}{\partial q}(x, y, p, q)\,. \quad (10)$$

If for each such comparison curve C we have

$$E(x, y, p, q, x', y') \geq 0 \quad (11)$$

for all values of x, y, x', y' then we may conclude that C_0 minimizes the integral $\int_{t_0}^{t_1} F(x, y, x', y')\, dt$.

In the ordinary calculus a condition that $y = y(x)$ be a minimum at $x = a$ is that $\frac{dy}{dx}(x = a) = 0$. No reference is made in this condition to neighboring values of a. Similarly, in the case of weak extrema in the calculus of variations the conditions are formulated solely in terms of the curve $C : y = y_0(x)$ and do not involve any reference to neighboring curves or functions. By contrast, Weierstrass's condition involves the comparison curve as well as the field function $p(x, y)$ defined in a neighborhood of C. It should be noted that while it is true that Weierstrass has obtained a stronger result, this is possible because the condition that must be satisfied is more restrictive; the stronger result is achieved at a higher price.

3. Zermelo's dissertation

It was inevitable that Zermelo's readership would be restricted because he was extending a mathematical theory that itself had not been published and that would have been familiar only to a fairly small group of researchers either at German universities or who had studied there. Furthermore, Weierstrass's parametric approach was not widely used in the calculus of variations; even investigators such as Ludwig Scheeffer (*1885*), Georg Erdmann (*1877*) and Edmund Husserl (*1882*) who were in a general sense part of the Weierstrass "school" and were influenced by his ideas employed the ordinary formulation of the variational problem in their researches of the 1870s and 1880s.

Zermelo's dissertation was also written in a rather formal manner, with very limited exposition of basic ideas and principles, and was excessively concerned with procedural matters, detailed formulations required in the general theory, and points of rigor. In matters of style, in its propensity for strenuous formal development, his approach bore similarities with the work of such earlier researchers as Hesse. Zermelo's work displayed as well a new element in variational research, a tendency to want to develop the subject from a larger viewpoint and to present the results as an instance of some more general and yet to be precisely specified subject. This tendency was manifested in his study of homogeneity properties of functions in the first chapter, as well as in his classification of families of curves in the second chapter.

Zermelo's dissertation would have been of primary interest to a reader who was either already familiar with Weierstrass's lectures or who was very motivated to learn about his new methods. For such a reader, the work would have been a valuable piece of mathematical research and exposition. In a fully detailed and methodical manner Zermelo developed a general theory, showing fully the non-trivial considerations that are involved in extending Weierstrass's methods to the problem $n > 1$. The first chapter was devoted to the homogeneity relations that must be satisfied in the parametric theory; the second to necessary conditions; the third to the excess function; and the final fourth chapter to sufficiency conditions involving the excess function.

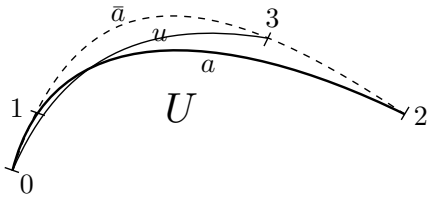


Fig. 3 (p. 90 of *Zermelo 1894*)

To indicate the basic idea behind Zermelo's development we describe how it plays out for the ordinary problem in the case where the variational integrand is a function of x, y, y' and y''. This was the setting in which Jacobi and so many other researchers set forth the theory. As before let $y = y_0(x)$ be the solution curve to the Euler equation joining **1** and **2**. In Fig. 3, from p. 90 of Zermelo's dissertation, this curve is denoted as a (note that the points are numbered slightly differently than in Weierstrass). We suppose there is a strip or region (what later was called a field) about a with the property that there is a unique extremal joining **0** (a point very close and to the left of **1** on a) and any given point **3** of the region. (In a technical refinement of Weierstrass's method, Zermelo takes the common starting point of the extremals to be **0** rather than **1** in order to simplify the analysis needed to establish the existence of the desired field. The value of the variational integral from **0** to **1** is taken to be negligible.) An arbitrary comparison curve 132: $y = y(x)$ is designated as $\overline{a}$. It is assumed that for each point **3** on $\overline{a}$ there is a unique extremal curve (a solution to the Euler equation) $\overline{y} = \overline{y}(x)$, joining **0** to **3**. This curve is designated as u in Fig. 3. Consider the function $\phi(x)$ given as

$$\phi(x) = \int_{\mathbf{0}}^{\mathbf{3}} f(x, \overline{y}, \overline{y}', \overline{y}'')\, dx + \int_{\mathbf{3}}^{\mathbf{2}} f(x, y, y', y'')\, dx\,. \tag{12}$$

Let the integrals in (12) be designated as $\overline{I}_{03}$ and I_{32}:

$$\overline{I}_{03} = \int_{\mathbf{0}}^{\mathbf{3}} f(x, \overline{y}, \overline{y}', \overline{y}'')\, dx,\ \ I_{32} = \int_{\mathbf{3}}^{\mathbf{2}} f(x, y, y', y'')\, dx\,. \tag{13}$$

$\overline{I}_{03}$ is evaluated along the extremal curve u from **0** to **3**, while I_{32} is evaluated along the comparison curve $\overline{a}$ from **3** to **2**. The key idea is to write $\delta\overline{I}_{03}$ using the variable endpoint formula applied to the case where there are second derivatives in the variational integrand. We have

$$\delta\overline{I}_{03} = \frac{\partial f(x,\overline{y},\overline{y}',\overline{y}'')}{\partial\overline{y}'}\delta y + \frac{\partial f(x,\overline{y},\overline{y}',\overline{y}'')}{\partial\overline{y}''}\delta y' + \big(f(x,\overline{y},\overline{y}',\overline{y}'') - \overline{y}'\big(\frac{\partial f(x,\overline{y},\overline{y}',\overline{y}'')}{\partial\overline{y}'} - \frac{d}{dx}\frac{\partial f(x,\overline{y},\overline{y}',\overline{y}'')}{\partial\overline{y}''}\big) - \overline{y}''\frac{\partial f(x,\overline{y},\overline{y}',\overline{y}'')}{\partial\overline{y}''}\big)\delta x\,. \tag{14}$$

We now let $\delta x = dx$ and $\delta y = dy$. Note that at the point **3** we have $\overline{y} = y$ and $\overline{y}' = p(x,y)$, $\overline{y}'' = q(x,y)$, where p and q are the field functions for the first and second derivatives of the extremal passing though **3**. With these designations (14) becomes

$$\frac{d\overline{I}_{03}}{dx} = \frac{\partial f(x,y,p,q)}{\partial p}y' + \frac{\partial f(x,y,p,q)}{\partial q}y'' + f(x,y,p,q) - p\big(\frac{\partial f(x,y,p,q)}{\partial p} - \frac{d}{dx}\frac{\partial f(x,y,p,q)}{\partial q}\big) - q\frac{\partial f(x,y,p,q)}{\partial q}\,. \tag{15}$$

We also have

$$\frac{dI_{32}}{dx} = -f(x,y,y',y'')\,. \tag{16}$$

Hence the derivative of $\phi(x)$ is

$$\phi'(x) = \int_{\mathbf{0}}^{\mathbf{3}} -E(x,y,y',y'',p,q)\,dx\,, \tag{17}$$

where

$$E(x,y,y',y'',p,q) = f(x,y,y',y'') - f(x,y,p,q) - (y'-p)\big(\frac{\partial f}{\partial p} - \frac{d}{dx}\frac{\partial f}{\partial q}\big) - (y''-q)\frac{\partial f}{\partial q}\,. \tag{18}$$

If $E(x,y,y',y'',p,q) \geq 0$ then it follows that

$$\phi'(x) \leq -1 \text{ and } \int_{\mathbf{0}}^{\mathbf{2}} f(x,y_0,y_0',y_0'')\,dx \leq \int_{\mathbf{0}}^{\mathbf{2}} f(x,y,y',y'')\,dx\,.$$

It is also apparent from (17) that $\phi(0) - \phi(1) = \Delta I$. Hence we have the following expression for the variation of the integral with respect to the comparison arc 032:

$$\Delta I = \int_{\mathbf{0}}^{\mathbf{2}} E(x,y,y',y'',p,q)\,dx\,. \tag{19}$$

(19) is Weierstrass's theorem and is the culminating result of Zermelo's treatise. It is stated on p. 79 of his dissertation in parametric form for the general case involving derivatives of order up to n. (It should be noted that the appellation "Weierstrass's theorem" was not used by Zermelo.) In the case $n = 1$ we have:

$$\Delta I = \int_{\sigma_1}^{\sigma_2} E(x, y, x', y', p, q)\, d\sigma\,, \tag{20}$$

where σ is the parameter and where the excess function in parametric form is given as

$$E = f(x, y, x', y') - f(x, y, p, q) - \frac{\partial f(x, y, p, q)}{\partial x'}(x' - p) - \frac{\partial f(x, y, p, q)}{\partial y'}(y' - q)\,. \tag{21}$$

In the traditional theory of sufficiency based on expansion methods it is necessary to ensure that there is no admissible δy which makes the second variation vanish. A point at which the second variation vanishes came to be called a conjugate point (a term coined by Weierstrass) and the problem has a solution only if there are no conjugate points on the interval. It is also necessary to show that there exist certain functions that allow one to transform the second variation to a suitable quadratic form. Mayer's achievement in his publications of the 1860s was to show in a very general setting that if there is no conjugate point on the interval then it is possible to produce the requisite functions needed in the transformation of the second variation. The basic problem here is one of mathematical existence. Zermelo following Weierstrass was confronted with a different kind of existence question. In order to carry out the derivation of equation (20) it is necessary to embed the extremal joining the endpoints in a field of extremals. Zermelo supplemented his presentation of (20) with an extended discussion of the existence of such a field and the conditions that are required for it. His approach was to write down an analytical condition stating that there is no conjugate point on the interval. From this condition it is shown that there is a strip or field about the given extremal joining the points A and B with the following property: for each point P in this region there is a unique extremal passing through it. If the variational integrand contains derivatives up to order n, then the extremal at P will have an nth order derivative at P that is a function of the values of $x, y', y'', ..., y^{(n-1)}$ there. Field-theoretic questions were an important part of Zermelo's theory and would become the focus of much further work in the calculus of variations.

4. Further discussion of Zermelo's theory

The variable endpoint formula (14) plays an essential role in the derivation of the condition $E(x, y, y', y'', p, q) \geq 0$. In (14) the increments δx, δy and $\delta y'$ are small increments in x, y and y'. It is immediately clear that the slope

of any comparison curve may only differ from the slope of the actual solution curve by a small amount. Thus in the Weierstrassian theory the case $n = 2$ is essentially different from the case $n = 1$, where the slope of the comparison curve may differ by any finite amount from the slope of the solution curve. In fact, the restriction on the slope of the comparison curve in the case $n = 2$ is the same as in the case of weak extrema for $n = 1$! In the general case in which the integrand function f contains derivatives of y up to order n, the comparison curve must differ by only a small amount at its derivatives up to order $n-1$. Of course, the derivatives of order n and higher may take on any value, so it is clear that the class of comparison curves is still larger than in the case of weak extrema.

There are several aspects of Zermelo's theory that somewhat limited its influence on the later development of the calculus of variations. His formulation using a parametric approach seems to have stemmed from a desire to remain faithful to Weierstrass's original exposition. However, the parametric formulation really constitutes a special topic, valuable from a certain geometric viewpoint but much too awkward to form the primary basis of the subject. Another important event was Hilbert's introduction (*1900a, 1905*) of the invariant integral, giving rise to an essential tool that transformed the theory. As we show below, the use of the invariant integral simplified the derivation of Weierstrass's theorem and provided a tool that could be applied to more general problems.

Mention should also be made of the central problem of concern to Zermelo. Although the variational problem with higher-order derivatives had been very prominent in the writings of Jacobi and his successors, it virtually disappeared from the textbook literature in the twentieth century. Instead one developed the theory for n dependent variables with variational integrands that contain only the first derivatives of the variables. The general variational integral is here $\int_a^b f(x, y_1, y_2, ..., y_n, y_1', y_2', ..., y_n')\, dx$. The investigation of sufficiency is carried out in this setting. The case of higher order derivatives is then treated as an optimization problem subject to constraint. The basic idea goes back to *Clebsch 1858a,b* and is illustrated by the problem of minimizing $\int_a^b f(x, y, y', y'')\, dx$. This problem can be reformulated as the problem of minimizing the integral $\int_a^b f(x, y_1, y_2, y_2')\, dx$ subject to the side constraint $y_1' - y_2 = 0$. Using the multiplier rule this problem is equivalent to minimizing the integral $\int_a^b \big(f(x, y_1, y_2, y_2') + \lambda(x)(y_1' - y_2)\big)\, dx$. The Euler equations for this problem are $\dfrac{\partial f}{\partial y_1} - \dfrac{d(\lambda(x))}{dx} = 0$ and $\dfrac{\partial f}{\partial y_2} - \lambda(x) - \dfrac{d(\frac{\partial f}{\partial y_2'})}{dx} = 0$. Noting that $y_1' - y_2 = 0$ we find that these two equations reduce to

$\dfrac{\partial f}{\partial y_1} - \dfrac{d}{dx}\dfrac{\partial f}{\partial y_1'} + \dfrac{d^2}{dx^2}\dfrac{\partial f}{\partial y_1''} = 0$, the Euler equation for $\displaystyle\int_a^b f(x, y, y', y'')\,dx$ with $y = y_1$. Sufficient conditions for the problem $\displaystyle\int_a^b f(x, y, y', y'')\,dx$ are in turn deduced from the general theory of sufficiency developed for the integral $\displaystyle\int_a^b f(x, y_1, y_2, ..., y_n, y_1', y_2', ..., y_n')\,dx$ and applied to the particular integral $\displaystyle\int_a^b \big(f(x, y_1, y_2, y_2') + \lambda(x)(y_1' - y_2)\big)\,dx$.

Despite these limitations, Zermelo's dissertation was important in bringing Weierstrass's ideas forward in published form and in developing the theory in new directions. It provided a source for the work of Kneser, Hilbert, Mayer (*1904*), Osgood (*1900/1901, 1901a*) and Bolza as well as the other researchers of the period. The work is cited no less than eight times in Bolza's *Lectures on the calculus of variations* (*1904*), on pp. 9, 35, 72, 76, 82, 119, 143, 174. Special note should be made of Bolza's discussion (p. 174) in Chapter V of transversals to sets of extremals, where attention is called to a result proved by Zermelo (p. 96 of his dissertation) concerning the envelope of a set of extremals.

In the history of the calculus of variations there are examples of researchers who began in this branch of mathematics and continued to make important contributions to it throughout their career. One might mention here such figures as Lagrange, Mayer and the American mathematician Gilbert Bliss. However, Zermelo belongs to another historical pattern of investigators who cut their teeth in the calculus of variations and then went on to prominence in very different fields of research. One could mention in addition to Zermelo (set theory) such figures as Charles Delaunay (celestial mechanics), Clebsch (algebraic geometry), Husserl (philosophy), and Herman Goldstine (computer science and numerical analysis).

5. Epilogue: Hilbert's invariant integral

Weierstrass's theorem in parametric form is given by (20). This statement is evidently relative to the particular parameterization chosen. Let us assume that we write the theorem in a form that is independent of any particular parameterization. One obvious way to do this would be to develop the theory in traditional ordinary form, using x as the independent variable and y as the dependent variable. In ordinary form Weierstrass's theorem is written:

$$\Delta I = \int_{x_1}^{x_2} E(x, y, y', p)\,dx \tag{22}$$

where

$$E = f(x, y, y') - f(x, y, p) - (y' - p)\frac{\partial f}{\partial y'}(x, y, p)\,. \tag{23}$$

We have

$$\Delta I = \int_{x_1}^{x_2} \big(f(x, y, y') - f(x, y_0, y_0')\big)\, dx\,, \tag{24}$$

where $y_0 = y_0(x)$ is the extremal joining the initial and final points. From (22), (23) and (24) it follows that

$$\begin{aligned} &\int_{x_1}^{x_2} \big(f(x, y, y') - f(x, y_0, y_0')\big)\, dx \\ &\qquad = \int_{x_1}^{x_2} \big(f(x, y, y') - f(x, y, p) - (y' - p)\frac{\partial f}{\partial p}(x, y, p)\big)\, dx\,, \end{aligned} \tag{25}$$

or

$$\int_{x_1}^{x_2} f(x, y_0, y_0')dx = \int_{x_1}^{x_2} \big(f(x, y, p) + (y' - p)\frac{\partial f}{\partial p}(x, y, p)\big)\, dx\,. \tag{26}$$

Because $y = y_0(x)$ is given, the quantity on the left side of (26) is constant. Hence from (26) we deduce that the integral

$$H = \int_{x_1}^{x_2} \big(f(x, y, p) + (y' - p)\frac{\partial f}{\partial p}(x, y, p)\big)\, dx$$

has the same value for all comparison curves $y = y(x)$: the integral H is invariant with respect to the path.

Hilbert did not discuss how he arrived at the idea of the invariant integral: in his account it is something that is introduced without any explanation. However, it is reasonable to suppose that he first came across the idea by simply writing down Weierstrass's theorem in ordinary form, and noticing as we did above that the integral H is invariant. It was then a simple matter to show directly that H is invariant. Using the Euler equation $\dfrac{\partial f}{\partial y} - \dfrac{d}{dx}\dfrac{\partial f(x, y, p)}{\partial p} = 0$ it is straightforward to prove that

$$\frac{\partial}{\partial y}\big(f(x, y, p) - p\frac{\partial f}{\partial p}(x, y, p)\big) = \frac{\partial}{\partial x}\big(\frac{\partial f}{\partial p}(x, y, p)\big)\,, \tag{27}$$

and so the condition for the integrability of the differential form $\big(f(x,y,p) - p\dfrac{\partial f}{\partial p}(x,y,p)\big)\,dx + \dfrac{\partial f}{\partial p}(x,y,p)\,dy$ is satisfied. Having established that H is invariant directly we can then use this fact to provide a new proof of Weierstrass's theorem, which is what Hilbert did. A significant advantage of Hilbert's approach is that a wider class of fields can be used in the sufficiency proof. In the theory of Weierstrass and Zermelo, the extremals of the field pass through a single point: in the case of Weierstrass this point is the initial point of the extremal, and in the case of Zermelo it is a point very close to the initial point (see *Bolza 1904*, 82, note 1). Such a field is said to be a central field. By contrast, the proof of Weierstrass's theorem using the invariant integral applies to any covering of a region surrounding the solution curve by a family of extremals in which only one extremal passes through each point of the region. The invariant integral can also be applied to more general variational problems, and is an important field-theoretic tool in the investigation of extrema. (For later literature related to this subject, see *Hadamard 1910, Bliss 1925* and *Bliss 1946*. The relevant history may be found in *Thiele 2007*.)

In the publication of his Paris address of 1900, where Hilbert first presented the idea of the invariant integral, he referred to Kneser's *Lehrbuch* but not to Zermelo. However, we know that he held Zermelo's work in the calculus of variations in high regard. In 1903 he recommended Zermelo for a position at the University of Breslau, writing[1] "Zermelo is a modern mathematician who combines versatility with depth in a rare way. He is an expert in the calculus of variations (and working on a comprehensive monograph about it). I regard the calculus of variations as a branch of mathematics which will belong to the most important ones in the future." Hilbert added that some years earlier, "Zermelo was my main mathematical company, and I have learnt a lot from him, for example, the Weierstrassian calculus of variations." Zermelo was not offered the position.

[1] In a letter to the hiring committee; cf. *Ebbinghaus 2007*, 35–36, 276–277.

Untersuchungen zur Variations-Rechnung

1894

Die Grundlage der nachstehenden Untersuchungen bilden die Vorlesungen des Herrn Prof. *Weierstrass* über „Variations-Rechnung“, die, während einer Reihe von Jahren an der Berliner Universität gehalten, mir zuerst im Sommer 1892 durch einige Ausarbeitungen im Besitz des „Mathematischen Vereins“, zunächst aber durch eine Vorlesung des Herrn Prof. *H. A. Schwarz* ihrem wesentlichen Inhalte nach zur Kenntnis gekommen sind. Meine Arbeit wird von dem Bestreben geleitet, einen Teil der von Herrn Weierstrass neu entwickelten strengen Methoden, die sich auf Maxima und Minima von Integralen der einfachsten Form beziehen, zunächst ohne Berücksichtigung von Nebenbedingungen und in wesentlich theoretischem Interesse, auf den allgemeineren Fall auszudehnen, wo die Function unter dem Integralzeichen Ableitungen beliebig hoher Ordnung enthält. Diese Verallgemeinerung bezieht sich auf die Darstellung der Curven durch einen Parameter, auf die Durchführung einer strengen Definition des Maximums oder Minimums und auf die Anwendung der von Herrn Weierstrass eingeführten Function E zur Auffindung notwendiger und hinreichender Bedingungen, während die auch in jenen Vorlesungen untersuchte „zweite Variation“ hier grundsätzlich ausser Betracht geblieben ist. Dagegen ist die wertvolle Arbeit *Ludwig Scheeffers* „Die Maxima und Minima der einfachen Integrale zwischen festen Grenzen“ (Math. Ann. XXV), die bei sehr ähnlicher Betrachtungsweise doch erhebliche Verschiedenheiten der Methode aufweist, hier ohne wesentlichen Einfluss gewesen.

Erster Abschnitt.

Über die Bedingungen, denen die Function unter dem Integralzeichen genügen muss.

Die Untersuchungen von Herrn *Weierstrass* beziehen sich auf Integrale der Form:

$$J = \int_{t_1}^{t_2} F(x, y; x', y')\, dt$$

$$\left(x' = \frac{dx}{dt}, \quad y' = \frac{dy}{dt} \right) ,$$

wo $x = \varphi(t)$, $y = \psi(t)$ als die laufenden Coordinaten eines Curvenstückes aufgefasst werden können, über welches die Integration zu erstrecken ist.

Investigations in the calculus of variations

1894

The following investigations are based on courses of lectures on the "calculus of variations" given by Prof. *Weierstrass* at the University of Berlin for a number of years. I originally became aware of their basic ideas in the summer of 1892 through several sets of notes in the possession of the "Mathematischer Verein", but first through a lecture by Prof. *H. A. Schwarz*. In my work, which is essentially guided by theoretical interest, I seek to extend some of the new, rigorous methods developed by Mr. Weierstrass concerning the maxima and minima of integrals of the simplest form to the more general case where the function under the integral sign contains derivatives of arbitrary order without, at first, taking into consideration ancillary conditions. This generalization concerns the representation of the curves by means of a parameter, the establishment of a rigorous definition of the maximum or the minimum, and the application of the function E introduced by Mr. Weierstrass in order to find necessary and sufficient conditions, whereas the "second variation" also investigated in those lectures will not, in general, be considered here. On the other hand, the valuable work "Die Maxima und Minima der einfachen Integrale zwischen festen Grenzen" by *Ludwig Scheeffer* (*Scheeffer 1885*), which, for all the similarities in perspective, uses a very different method, has had little influence on the present work.

First section.

On the conditions to be met by the function under the integral sign.

The investigations by Mr. *Weierstrass* are concerned with integrals of the form

$$J = \int_{t_1}^{t_2} F(x, y; x', y')\, dt$$

$$\left(x' = \frac{dx}{dt}, \quad y' = \frac{dy}{dt} \right) ,$$

where $x = \varphi(t)$, $y = \psi(t)$ may be considered the current coordinates of a curve segment along which the integration is to be taken.

Soll dieses Integral von der in gewisser Beziehung willkürlichen Art, wie die Coordinaten der Curve als Functionen des Parameters t dargestellt sind, *unabhängig* sein, so muss, wie gezeigt wird, F in Bezug auf seine beiden letzten Argumente x', y' *homogen* von erster Dimension sein, oder, was dasselbe sagt, der Bedingung genügen:

$$\frac{\partial F}{\partial x'}x' + \frac{\partial F}{\partial y'}y' = F \ ,$$

woraus weitere Formeln entspringen.

Zur analogen Untersuchung des allgemeineren Falles:

$$J = \int_{t_1}^{t_2} F\left(x, x', \ldots x^{(n)}; y, y', \ldots y^{(n)}\right) dt$$

haben wir zunächst nach den entsprechenden Bedingungsgleichungen für diese allgemeinere Function F zu fragen.

3 | Zur Herleitung derselben werde ich mich, da hier die direkte Substitution eines anderen Parameters nicht mehr zweckmässig scheint, in ähnlicher Weise der Variationsrechnung selbst bedienen, wie man es seit Euler zur Aufstellung der gewöhnlichen „Integrabilitäts-Bedingungen“ zu thun pflegt. Doch wird diese Anwendung eine rein formale sein und keine anderen Principien als die Elemente der Differential-Rechnung zu Grunde legen.

$$F \equiv F\left(x^{(\mu)}, y^{(\mu)}\right)$$

(in abgekürzter Schreibweise) werde als eine analytische Function ihrer sämtlichen Argumente vorausgesetzt, die in dem ganzen betrachteten Bereiche den Charakter einer ganzen Function, also auch partielle Ableitungen beliebig hoher Ordnung besitzt.

$$x = \varphi(\vartheta) \ , \quad y = \psi(\vartheta)$$

seien eindeutige und mit ihren r ersten Ableitungen stetige Functionen von ϑ im Intervall

$$\vartheta_1 \leqq \vartheta \leqq \vartheta_2 \ ,$$

wobei für den einen Teil der folgenden Betrachtungen die Annahme $r = n$, für den anderen erst $r = 2n$ genügt; ausserdem mögen $\varphi'(\vartheta)$ und $\psi'(\vartheta)$ an keiner Stelle des Intervalles gleichzeitig verschwinden.

Es wird dann der Punkt x, y ein zusammenhängendes Curvenstück 1 2 in einem Zuge und in bestimmtem Richtungssinn beschreiben, während ϑ immer wachsend das Intervall $\vartheta_1 \ldots \vartheta_2$ durchläuft; für eine beliebig gezeichnete rectificierbare Curve kann z. B. immer $\vartheta = s$ gesetzt, die Bogenlänge zur unabhängigen Variablen gewählt werden. Zu jedem Werte ϑ zwischen ϑ_1 und ϑ_2 gehört ein bestimmter Punkt $\varphi(\vartheta)$, $\psi(\vartheta)$ des Curvenstückes und umgekehrt

If this integral is to be *independent* of the, in certain respects arbitrary, manner in which the coordinates of the curve are represented as functions of the parameters t, then, as shall be shown, F must be *homogeneous* of first order in its last two arguments x', y', or, what amounts to the same, it must meet the condition

$$\frac{\partial F}{\partial x'}x' + \frac{\partial F}{\partial y'}y' = F \,,$$

from which further formulas arise.

To investigate the more general case

$$J = \int_{t_1}^{t_2} F\left(x, x', \ldots x^{(n)}; y, y', \ldots y^{(n)}\right) dt$$

along similar lines, we first must address the question of the corresponding constraint equations for this more general function F.

In order to deduce them I shall make use of the calculus of variations in a way similar to how it has been used to determine the usual "integrability conditions" ever since Euler, since in this case the direct substitution of another parameter no longer seems practical. However, this application will be a purely formal one and not require any other principles but the basic elements of the differential calculus.

$$F \equiv F\left(x^{(\mu)}, y^{(\mu)}\right)$$

(in abbreviated notation) is assumed to be an analytic function of all of its arguments that is of the character of an entire function over the whole domain under consideration, and hence also has partial derivatives of arbitrary order. Let

$$x = \varphi(\vartheta) \,, \quad y = \psi(\vartheta)$$

be single-valued functions of ϑ that, together with their first r derivatives, are continuous on the interval

$$\vartheta_1 \leqq \vartheta \leqq \vartheta_2 \,,$$

where, for one part of the following considerations, it suffices to assume that $r = n$, and for the other, only that $r = 2n$; furthermore, we assume that $\varphi'(\vartheta)$ and $\psi'(\vartheta)$ do not vanish simultaneously anywhere in the interval.

In this case, the point x, y traces a continuous curve segment 1 2 in one fell swoop in a specific direction as the ever-increasing ϑ passes through the interval $\vartheta_1 \ldots \vartheta_2$; given an arbitrarily drawn rectifiable curve, we can, e.g., always set $\vartheta = s$ and choose the arc length as the independent variable. To each value ϑ between ϑ_1 and ϑ_2 there belongs a particular point $\varphi(\vartheta)$, $\psi(\vartheta)$ of the curve segment, and, conversely, to each of these points a particular ϑ

zu jedem dieser Punkte ein gewisses ϑ des Intervalles. — Wir können nun $\vartheta = \vartheta(t)$ als eine beliebige im Intervall mit ihren ersten r Ableitungen stetige und beständig wachsende Function von t, oder, was dasselbe ist, t als eine eben solche Function von ϑ annehmen und dann die Endwerte t_1, t_2 bestimmt denken durch $\vartheta(t_1) = \vartheta_1$, $\vartheta(t_2) = \vartheta_2$. Wenn dann t das Intervall $t_1 \dots t_2$ zunehmend durchläuft, so muss auch ϑ immer wachsen von ϑ_1 bis ϑ_2 und das Curvenstück 1 2 erzeugen.

4 | $\vartheta = t$ ergiebt die ursprüngliche Darstellung der Curve:

$$x = \varphi(t)\ , \quad y = \psi(t)\ ,$$

$\vartheta = \overline{\vartheta}(t)$ aber eine andere:

$$x = \varphi(\overline{\vartheta}) = \overline{\varphi}(t)\,, \quad y = \psi(\overline{\vartheta}) = \overline{\psi}(t)\,;$$

und von dieser Form müssen auch alle den gemachten Voraussetzungen entsprechenden Darstellungen $\overline{\varphi}(t), \overline{\psi}(t)$ unseres Curvenstückes sein: immer giebt es eine solche Function $\vartheta = \overline{\vartheta}(t)$, welche die Überführung bewirkt.

Soll nun für ein beliebig vorgeschriebenes Curvenstück 1 2

$$x = \varphi(\vartheta)\ , \ \ y = \psi(\vartheta) \qquad (\vartheta_1 \leqq \vartheta \leqq \vartheta_2)$$

unser Integral

$$J = \int_{t_1}^{t_2} F\left(x^{(\mu)}, y^{(\mu)}\right) dt \qquad \left(x^{(\mu)} = \frac{d^\mu x}{dt^\mu}\right)$$

einen bestimmten, von der besonderen Form der Darstellung unabhängigen Wert besitzen, so darf es nach Fixierung der Functionen φ, ψ nur von den Endwerten ϑ_1 und ϑ_2 der Function $\vartheta = \vartheta(t)$ abhängen, also bei constantem Anfangspunkt ϑ_1 nur von dem variablen Endpunkt ϑ_2. Es muss also, wenn man ϑ_2, t_2 durch ϑ, t ersetzt:

$$\int_{t_1}^{t} F\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big)\, dt = J(\vartheta) \quad \left(D^\mu = \frac{d^\mu}{dt^\mu}\right)$$

eine blosse Function von ϑ sein, und durch Differentiation nach t:

$$F\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big) = D\,J(\vartheta) = J'(\vartheta)\vartheta' \tag{1}$$

für eine beliebige, nur den Stetigkeitsbedingungen und der Bedingung $\vartheta' > 0$ genügende Function $\vartheta = \vartheta(t)$, also auch für $\vartheta = t$, $\vartheta' = 1$,

$$D^\mu\varphi(\vartheta) = \varphi^{(\mu)}(\vartheta)\ , \ \ D^\mu\psi(\vartheta) = \psi^{(\mu)}(\vartheta)\ ,$$

5 | so dass wegen

$$F\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right) = J'(\vartheta)$$

of the interval.—We now may assume $\vartheta = \vartheta(t)$ to be an arbitrary function t that, together with its first r derivatives, is continuous and increasing on the interval. Or, what amounts to the same, we may assume t to be a function of ϑ of this sort, and the end values t_1, t_2 to be determined by $\vartheta(t_1) = \vartheta_1$, $\vartheta(t_2) = \vartheta_2$. If, then, t passes through the interval $t_1 \dots t_2$ as it increases, then ϑ, too, must increase from ϑ_1 to ϑ_2 and generate the curve segment 1 2.

$\vartheta = t$ yields the original representation of the curve:

$$x = \varphi(t)\,, \quad y = \psi(t)\,,$$

$\vartheta = \overline{\vartheta}(t)$, however, a different one:

$$x = \varphi(\overline{\vartheta}) = \overline{\varphi}(t)\,, \quad y = \psi(\overline{\vartheta}) = \overline{\psi}(t)\,;$$

and this is the form that all representations $\overline{\varphi}(t), \overline{\psi}(t)$ of our curve segment that satisfy the stated assumptions must have: there is always a function $\vartheta = \overline{\vartheta}(t)$ that achieves the transformation.

Now if, for an arbitrarily prescribed curve segment 1 2

$$x = \varphi(\vartheta)\,, \quad y = \psi(\vartheta) \qquad (\vartheta_1 \leqq \vartheta \leqq \vartheta_2)\,,$$

our integral

$$J = \int_{t_1}^{t_2} F\left(x^{(\mu)}, y^{(\mu)}\right)\, dt \qquad \left(x^{(\mu)} = \frac{d^\mu x}{dt^\mu}\right)$$

is supposed to have a particular value independent of the specific form of representation, then, once the functions φ, ψ have been fixed, it may only depend on the end values ϑ_1 and ϑ_2 of the function $\vartheta = \vartheta(t)$, and hence only on the variable endpoint ϑ_2, assuming a constant starting point ϑ_1. Thus, if ϑ_2, t_2 are replaced by ϑ, t,

$$\int_{t_1}^{t} F\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big)\, dt = J(\vartheta) \qquad \left(D^\mu = \frac{d^\mu}{dt^\mu}\right)$$

must only be a function of ϑ, and, by differentiation with respect to t,

$$F\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big) = D\,J(\vartheta) = J'(\vartheta)\vartheta' \tag{1}$$

for any function $\vartheta = \vartheta(t)$ satisfying only the continuity conditions and the condition $\vartheta' > 0$, and hence also for $\vartheta = t$, $\vartheta' = 1$,

$$D^\mu\varphi(\vartheta) = \varphi^{(\mu)}(\vartheta)\,, \quad D^\mu\psi(\vartheta) = \psi^{(\mu)}(\vartheta)\,,$$

so that on account of

$$F\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right) = J'(\vartheta)$$

nunmehr (1) geschrieben werden kann:

$$F\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big) = F\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right)\vartheta' \ . \tag{1a}$$

Die Differentiationen aber kann man ausführen:

$$x' = D\varphi(\vartheta) = \varphi'(\vartheta)\vartheta' \ , \quad x'' = \varphi''(\vartheta)\vartheta'^2 + \varphi'(\vartheta)\vartheta'' \ ,$$

u. s. w., allgemein

$$\left.\begin{aligned} x^{(\mu)} = D^\mu\varphi(\vartheta) &= R_\mu\left(\varphi(\vartheta), \varphi'(\vartheta), \ldots \varphi^{(\mu)}(\vartheta), \vartheta', \ldots \vartheta^{(\mu)}\right) \ , \\ \text{abgekürzt} &= R_\mu\left(\varphi^{(\nu)}(\vartheta), \vartheta^{(\nu)}\right) \ , \\ y^{(\mu)} = D^\mu\psi(\vartheta) &= R_\mu\left(\psi^{(\nu)}(\vartheta), \vartheta^{(\nu)}\right) \ . \end{aligned}\right\} \tag{2}$$

Hier ist R_μ eine ganze rationale Function ihrer sämtlichen Argumente mit ganzzahligen positiven Coefficienten, linear und homogen in Bezug auf die ersten μ, und enthält die höchste Ableitung $\vartheta^{(\mu)}$ nur in dem einen Gliede $\varphi'(\vartheta)\vartheta^{(\mu)}$.

Man kann aber immer $n+1$ willkürliche Grössen $\vartheta, \vartheta', \vartheta'', \ldots \vartheta^{(n)}$ ansehen als die Werte der successiven Ableitungen einer stetigen Function $\vartheta = \vartheta(t)$ für einen beliebigen Argumentwert $t = t'$; wir brauchen ja nur zu setzen

$$\vartheta(t) = \sum_{\mu=0}^{n} \frac{\vartheta^{(\mu)}}{\mu!}(t-t')^{(\mu)} \ .$$

Im Falle $\vartheta' > 0$ wird jede solche Function in einer gewissen Umgebung von $t = t'$ mit t immer nur zunehmen und somit allen an ϑ gestellten Forderungen genügen. Daher wird die Gleichung (1), die ja für beliebige Functionen $\vartheta(t)$ gelten sollte, nach Einsetzung der Ausdrücke (2) *identisch* bestehen für unbestimmte und unter einander unabhängige Grössen $\vartheta, \vartheta', \ldots \vartheta^{(n)}$ mit der
6 | einzigen Bedingung $\vartheta' > 0$, so dass

$$\Phi = \frac{F}{\vartheta'} = J'(\vartheta) = \Phi(\vartheta)$$

von $\vartheta', \vartheta'', \ldots \vartheta^{(n)}$ unabhängig ist, also

$$\frac{\partial\Phi}{\partial\vartheta^{(\mu)}} = 0 \ \ (\mu = 1, 2, \ldots n) \ , \ \frac{\partial\varphi}{\partial\vartheta} = \frac{d\Phi}{d\vartheta} = \frac{D\Phi}{\vartheta'} \ , \tag{3}$$

und weiter:

$$\delta_\vartheta\Phi \equiv \sum_{\mu=0}^{n} \frac{\partial\Phi}{\partial\vartheta^{(\mu)}}\delta\vartheta^{(\mu)} = \frac{\partial\Phi}{\partial\vartheta}\delta\vartheta = D\Phi\frac{\delta\vartheta}{\vartheta'} = D\Phi.\tau \ , \tag{4}$$

wo $\delta\vartheta$ eine *willkürliche*, mit ihren n ersten Ableitungen

$$D^\mu\delta\vartheta = \delta\vartheta^{(\mu)}$$

we now can write (1) as

$$F\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big) = F\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right)\vartheta' \,. \tag{1a}$$

But the differentiations can be carried out:

$$x' = D\varphi(\vartheta) = \varphi'(\vartheta)\vartheta' \,, \quad x'' = \varphi''(\vartheta)\vartheta'^2 + \varphi'(\vartheta)\vartheta'' \,,$$

e. t. c., generally,

$$\left.\begin{aligned} x^{(\mu)} = D^\mu\varphi(\vartheta) &= R_\mu\left(\varphi(\vartheta), \varphi'(\vartheta), \ldots \varphi^{(\mu)}(\vartheta), \vartheta', \ldots \vartheta^{(\mu)}\right) \,, \\ \text{abbreviated} &= R_\mu\left(\varphi^{(\nu)}(\vartheta), \vartheta^{(\nu)}\right) \,, \\ y^{(\mu)} = D^\mu\psi(\vartheta) &= R_\mu\left(\psi^{(\nu)}(\vartheta), \vartheta^{(\nu)}\right) \,. \end{aligned}\right\} \tag{2}$$

In this case, R_μ is an integral rational function of all its arguments with positive integral coefficients that is linear and homogeneous with respect to the first μ, and contains the highest-order derivative $\vartheta^{(\mu)}$ only in the term $\varphi'(\vartheta)\vartheta^{(\mu)}$.

But it is always possible to consider $n+1$ arbitrary quantities $\vartheta, \vartheta', \vartheta'', \ldots \vartheta^{(n)}$ as values of successive derivatives of a continuous function $\vartheta = \vartheta(t)$ for an arbitrary argument $t = t'$; for we only need to set

$$\vartheta(t) = \sum_{\mu=0}^{n} \frac{\vartheta^{(\mu)}}{\mu!}(t - t')^{(\mu)} \,.$$

When $\vartheta' > 0$, any function of this kind always only increases in the vicinity of a certain $t = t'$ along with t, thus meeting all requirements imposed on ϑ. Therefore, upon substitution of the expressions (2), the equation (1), which was supposed to hold for arbitrary functions $\vartheta(t)$, holds *identically* for indeterminate, mutually independent quantities $\vartheta, \vartheta', \ldots \vartheta^{(n)}$, under the sole condition that $\vartheta' > 0$, so that

$$\Phi = \frac{F}{\vartheta'} = J'(\vartheta) = \Phi(\vartheta)$$

is independent of $\vartheta', \vartheta'', \ldots \vartheta^{(n)}$, and hence

$$\frac{\partial\Phi}{\partial\vartheta^{(\mu)}} = 0 \;\; (\mu = 1, 2, \ldots n) \,, \; \frac{\partial\varphi}{\partial\vartheta} = \frac{d\Phi}{d\vartheta} = \frac{D\Phi}{\vartheta'} \,, \tag{3}$$

and furthermore

$$\delta_\vartheta\Phi \equiv \sum_{\mu=0}^{n} \frac{\partial\Phi}{\partial\vartheta^{(\mu)}}\delta\vartheta^{(\mu)} = \frac{\partial\Phi}{\partial\vartheta}\delta\vartheta = D\Phi\frac{\delta\vartheta}{\vartheta'} = D\Phi.\tau \,, \tag{4}$$

where $\delta\vartheta$ denotes an *arbitrary* function of t that, together with its first n derivatives

$$D^\mu\delta\vartheta = \delta\vartheta^{(\mu)} \,,$$

stetige Function von t bedeutet; auch

$$\tau = \frac{\delta\vartheta}{\vartheta'}$$

ist solch eine willkürliche Function.

Es ist aber allgemein:

$$\delta_\vartheta f\left(\vartheta, \vartheta', \ldots \vartheta^{(n)}\right) \equiv \sum_{\mu=0}^{n} \frac{\partial f}{\partial \vartheta^{(\mu)}} \delta\vartheta^{(\mu)} \tag{5}$$

$$= \left[\frac{\partial}{\partial\varepsilon} f\left(\vartheta + \varepsilon\delta\vartheta, \vartheta' + \varepsilon\delta\vartheta', \ldots\right)\right]_{\varepsilon=0}$$

„die *erste Variation* von f in Bezug auf ϑ“ und folgt den Gesetzen:

$$\delta H\left(f_1, f_2, \ldots\right) = \frac{\partial H}{\partial f_1}\delta f_1 + \frac{\partial H}{\partial f_2}\delta f_2 + \ldots$$

$$\delta D^\mu f = D^\mu \delta f\,.$$

Daher ist einerseits:

7 | $$\delta_\vartheta F \equiv \sum_{\mu=0}^{n} \frac{\partial F}{\partial \vartheta^{(\mu)}}\delta\vartheta^{(\mu)} = \sum_{\mu=0}^{n}\left\{\frac{\partial F}{\partial x^{(\mu)}}\delta x^{(\mu)} + \frac{\partial F}{\partial y^{(\mu)}}\delta y^{(\mu)}\right\}$$

$$= \sum_{\mu=0}^{n}\left\{\frac{\partial F}{\partial x^{(\mu)}}(x'\tau)^{(\mu)} + \frac{\partial F}{\partial y^{(\mu)}}(y'\tau)^{(\mu)}\right\}\,,$$

da

$$\delta x = \delta\varphi(\vartheta) = \varphi'(\vartheta)\delta\vartheta = \varphi'(\vartheta)\vartheta'\frac{\delta\vartheta}{\vartheta'}$$

$$= D\varphi(\vartheta).\tau = x'\tau$$

und ebenso:

$$\delta y = y'\tau$$

ist, andrerseits aber wegen (4)

$$\delta_\vartheta F = \delta_\vartheta(\Phi\vartheta') = \delta_\vartheta\Phi.\vartheta' + \Phi\delta\vartheta'$$

$$= D\Phi.\tau\vartheta' + \Phi D(\vartheta'\tau)$$

$$= D(\Phi\vartheta'\tau) = D(F\tau)\,,$$

also:

$$\delta_\vartheta F = \sum_{\mu=0}^{n}\left\{\frac{\partial F}{\partial x^{(\mu)}}(x'\tau)^{(\mu)} + \frac{\partial F}{\partial y^{(\mu)}}(y'\tau)^{(\mu)}\right\} = D(F\tau)\,. \tag{6}$$

is continuous; also

$$\tau = \frac{\delta\vartheta}{\vartheta'}$$

is an arbitrary function of this kind.

But, generally,

$$\delta_\vartheta f\left(\vartheta, \vartheta', \ldots \vartheta^{(n)}\right) \equiv \sum_{\mu=0}^{n} \frac{\partial f}{\partial \vartheta^{(\mu)}} \delta\vartheta^{(\mu)} \tag{5}$$

$$= \left[\frac{\partial}{\partial\varepsilon} f\left(\vartheta + \varepsilon\delta\vartheta, \vartheta' + \varepsilon\delta\vartheta', \ldots\right)\right]_{\varepsilon=0}$$

is "the *first variation* of f with respect to ϑ" and is subject to the laws

$$\delta H\left(f_1, f_2, \ldots\right) = \frac{\partial H}{\partial f_1}\delta f_1 + \frac{\partial H}{\partial f_2}\delta f_2 + \ldots$$

$$\delta D^\mu f = D^\mu \delta f\,.$$

Hence, on the one hand,

$$\delta_\vartheta F \equiv \sum_{\mu=0}^{n} \frac{\partial F}{\partial \vartheta^{(\mu)}}\delta\vartheta^{(\mu)} = \sum_{\mu=0}^{n}\left\{\frac{\partial F}{\partial x^{(\mu)}}\delta x^{(\mu)} + \frac{\partial F}{\partial y^{(\mu)}}\delta y^{(\mu)}\right\}$$

$$= \sum_{\mu=0}^{n}\left\{\frac{\partial F}{\partial x^{(\mu)}}(x'\tau)^{(\mu)} + \frac{\partial F}{\partial y^{(\mu)}}(y'\tau)^{(\mu)}\right\}\,,$$

since

$$\delta x = \delta\varphi(\vartheta) = \varphi'(\vartheta)\delta\vartheta = \varphi'(\vartheta)\vartheta'\frac{\delta\vartheta}{\vartheta'}$$

$$= D\varphi(\vartheta).\tau = x'\tau$$

and likewise

$$\delta y = y'\tau\,,$$

while, on the other hand,[1]

$$\delta_\vartheta F = \delta_\vartheta(\Phi\vartheta') = \delta_\vartheta\Phi.\vartheta' + \Phi\delta_\vartheta\vartheta'$$

$$= D\Phi.\tau\vartheta' + \Phi D(\vartheta'\tau)$$

$$= D(\Phi\vartheta'\tau) = D(F\tau)\,,$$

on account of (4), and hence

$$\delta_\vartheta F = \sum_{\mu=0}^{n}\left\{\frac{\partial F}{\partial x^{(\mu)}}(x'\tau)^{(\mu)} + \frac{\partial F}{\partial y^{(\mu)}}(y'\tau)^{(\mu)}\right\} = D(F\tau)\,. \tag{6}$$

[1] [[In the first line of the following formula, Zermelo erroneously writes "δ" for the fourth "δ_ϑ".]]

Diese Beziehung muss bestehen für beliebige Functionen x, y, τ von t, da ja auch die Functionen φ und ψ willkürlich sein sollen. Wir brauchen daher nur nach formaler Ausführung der Differentiationen die Coefficienten der $\tau, \tau', \ldots \tau^{(n)}$ beiderseits einander gleich zu setzen, um das System der von den $\tau^{(\mu)}$ unabhängigen Bedingungsgleichungen für F zu erhalten.

Diese Bedingungen sind zugleich auch hinreichend für das Bestehen der verlangten Eigenschaft. Denn sind sie erfüllt, so gelten auch (6), (4) für willkürliche $\varphi, \psi, \vartheta, \tau$ oder $\delta\vartheta$ und daher, durch Coefficientenvergleichung, auch (3). Dann lässt sich für irgend zwei Functionen φ, ψ auch immer $J'(\vartheta)$ der Gleichung (1) gemäss bestimmen, und durch Integration folgt schliesslich:

$$\int_{t_1}^{t_2} F\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big)\, dt = J(\vartheta_2) - J(\vartheta_1)\ ,$$

8 | in der That nur abhängig von der durch φ, ψ bestimmten Form der Curve und ihren durch ϑ_1 und ϑ_2 bestimmten Endpunkten für die Integration, aber unabhängig von der Beziehung zwischen t und ϑ, d. h. von der besonderen Darstellungsform.

Um nun die Bedingungsgleichungen noch in *andrer Gestalt* nebst den zwischen beiden bestehenden Beziehungen darzustellen, schicke ich einige allgemein gültige formale Entwicklungen voraus, in denen zur Abkürzung

$$\frac{\partial F}{\partial x^{(\mu)}} = X_\mu\ , \quad \frac{\partial F}{\partial y^{(\mu)}} = Y_\mu$$

geschrieben werde.

Nach einer in der Variationsrechnung allgemein gebräuchlichen Umformung durch partielle Integration ist für eine beliebige Function $F(x^{(\mu)}, y^{(\mu)})$ wenn x, y und τ irgendwelche Functionen von t bedeuten:

$$\sum_{\mu=0}^{n} X_\mu (x'\tau)^{(\mu)} = Px'\tau + D\sum_{\mu=1}^{n} P_\mu (x'\tau)^{(\mu-1)}\ , \tag{7}$$

wenn

$$X_\mu = P_\mu + D\,P_{\mu+1} \qquad (\mu = 0, 1, \ldots n) \qquad (P_{n+1} = 0\ , \quad \text{also} \quad P_n = X_n) \tag{8}$$

$$P_\mu = \sum_{\chi=0}^{n-\mu} (-1)^\chi D^\chi X_{\mu+\chi} \qquad (\mu = 0, 1, \ldots n,\ P_0 = P)\ .$$

Durch Entwicklung der $(x'\tau)^{(\mu)}$ und Zusammenfassung gleichnamiger Glieder geht (7) über in:

$$\sum_{\nu=0}^{n} \Xi_\nu \tau^{(\nu)} = \Pi\tau + D\sum_{\nu=1}^{n} \Pi_\nu \tau^{(\nu-1)}\ , \tag{9}$$

This relationship must hold for arbitrary functions x, y, τ of t, since the functions φ and ψ are supposed to be arbitrary as well. After having formally carried out the differentiation, we therefore only need to set the coefficients of the $\tau, \tau', \ldots \tau^{(n)}$ equal to one another on both sides in order to obtain for F the system of the constraint equations independent of the $\tau^{(\mu)}$.

At the same time, these conditions are also sufficient for the required property to obtain. For if they are satisfied, then (6), (4) also hold for arbitrary $\varphi, \psi, \vartheta, \tau$ or $\delta\vartheta$, and hence, by virtue of comparison of coefficients, (3) holds as well. For any two functions φ, ψ, it is then always possible to determine also $J'(\vartheta)$ in accordance with the equation (1), and by integration it eventually follows that

$$\int_{t_1}^{t_2} F\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big)\, dt = J(\vartheta_2) - J(\vartheta_1)\ ,$$

which really only depends on the shape of the curve, determined by φ, ψ, and on its endpoints, determined by ϑ_1 and ϑ_2, for the integration, but which is independent of the relation between t and ϑ, that is, of the particular form of representation.

In order to represent the constraint equations also *in another form* besides the relations obtaining between the two of them, we will first consider several generally valid expansions, using the abbreviations

$$\frac{\partial F}{\partial x^{(\mu)}} = X_\mu\ , \quad \frac{\partial F}{\partial y^{(\mu)}} = Y_\mu\ .$$

According to a transformation by partial integration commonly used in the calculus of variations, we have, for any function $F(x^{(\mu)}, y^{(\mu)})$, where x, y and τ denote arbitrary functions of t,

$$\sum_{\mu=0}^{n} X_\mu (x'\tau)^{(\mu)} = Px'\tau + D\sum_{\mu=1}^{n} P_\mu (x'\tau)^{(\mu-1)}\ , \tag{7}$$

if

$$X_\mu = P_\mu + D\,P_{\mu+1} \qquad (\mu = 0, 1, \ldots n) \tag{8}$$

$$(P_{n+1} = 0\ , \quad \text{hence} \quad P_n = X_n)$$

$$P_\mu = \sum_{\chi=0}^{n-\mu} (-1)^\chi D^\chi X_{\mu+\chi} \qquad (\mu = 0, 1, \ldots n,\ P_0 = P)\ .$$

By expanding the $(x'\tau)^{(\mu)}$ and collecting like terms, (7) is transformed into

$$\sum_{\nu=0}^{n} \Xi_\nu \tau^{(\nu)} = \Pi\tau + D\sum_{\nu=1}^{n} \Pi_\nu \tau^{(\nu-1)}\ , \tag{9}$$

wenn

$$\Xi_\nu = \sum_{\mu=\nu}^{n} \binom{\mu}{\nu} X_\mu x^{(\mu-\nu+1)} \tag{10}$$

9 |

$$\Pi_\nu = \sum_{\mu=\nu}^{n} \binom{\mu-1}{\nu-1} P_\mu x^{(\mu-\nu+1)} , \quad \Pi = \Pi_0 = Px' \tag{11}$$

$$(\nu = 1, 2, \ldots n) .$$

Aus (9) aber folgt durch Vergleichung der Coefficienten von $\tau^{(\nu)}$ links und rechts, was sich auch durch direkte Benutzung von (8), (10), (11) nachweisen liesse:

$$\Xi_\nu = \Pi_\nu + D\,\Pi_{\nu+1} \qquad (\nu = 0, 1, \ldots n) \tag{12}$$

$$(\Pi_{n+1} = 0, \text{ also } \Pi_\nu = \Xi_\nu) ,$$

daher:

$$\Pi_\nu = \sum_{\chi=0}^{n-\nu} (-1)^\chi D^\chi \Xi_{\nu+\chi} .$$

Die analogen Beziehungen bestehen, wenn man die Ausdrücke

$$x^{(\mu)}, \quad X_\mu, \quad P_\mu, \quad \Xi_\nu, \quad \Pi_\nu$$

der Reihe nach ersetzt durch:

$$y^{(\mu)}, \quad Y_\mu, \quad Q_\mu, \quad H_\nu, \quad P_\nu ,$$

und durch Zusammenziehung der entsprechenden Gleichungen erhält man aus (7), (9) und (12):

$$\delta_\vartheta F = \sum_{\mu=0}^{n} \left\{ X_\mu (x'\tau)^{(\mu)} + Y_\mu (y'\tau)^{(\mu)} \right\} \tag{7a}$$

$$= (Px' + Qy')\,\tau + D \sum_{\mu=1}^{n} \left\{ P_\mu (x'\tau)^{(\mu-1)} + Q_\mu (y'\tau)^{(\mu-1)} \right\}$$

$$= \sum_{\nu=0}^{n} (\Xi_\nu + H_\nu)\,\tau^{(\nu)} \tag{9a}$$

$$= (\Pi + P)\tau + D \sum_{\nu=1}^{n} (\Pi_\nu + P_\nu)\,\tau^{(\nu-1)}$$

10 |

$$\Xi_\nu + H_\nu = \Pi_\nu + P_\nu + D\,(\Pi_{\nu+1} + P_{\nu+1}) \tag{12a}$$

$$\Pi_\nu + P_\nu = \sum_{\chi=0}^{n-\nu} (-1)^\chi D^\chi\,(\Xi_{\nu+\chi} + H_{\nu+\chi})$$

$$(\nu = 0, 1, 2, \ldots n) .$$

if

$$\Xi_\nu = \sum_{\mu=\nu}^{n} \binom{\mu}{\nu} X_\mu x^{(\mu-\nu+1)} \tag{10}$$

$$\Pi_\nu = \sum_{\mu=\nu}^{n} \binom{\mu-1}{\nu-1} P_\mu x^{(\mu-\nu+1)}\,, \quad \Pi = \Pi_0 = Px' \tag{11}$$

$$(\nu = 1, 2, \ldots n)\,.$$

But from (9) we obtain by comparison of the coefficients of $\tau^{(\nu)}$ on the right and left sides what could also be verified by direct use of (8), (10), (11):

$$\Xi_\nu = \Pi_\nu + D\,\Pi_{\nu+1} \qquad (\nu = 0, 1, \ldots n) \tag{12}$$

$$(\Pi_{n+1} = 0, \text{ also } \Pi_\nu = \Xi_\nu)\,,$$

and therefore

$$\Pi_\nu = \sum_{\chi=0}^{n-\nu} (-1)^\chi D^\chi \Xi_{\nu+\chi}\,.$$

The analogous relations hold when the expressions

$$x^{(\mu)}, \quad X_\mu, \quad P_\mu, \quad \Xi_\nu, \quad \Pi_\nu$$

are replaced by, respectively,

$$y^{(\mu)}, \quad Y_\mu, \quad Q_\mu, \quad H_\nu, \quad P_\nu\,.$$

Collecting the corresponding equations, we obtain from (7), (9) and (12):

$$\delta_\vartheta F = \sum_{\mu=0}^{n} \left\{ X_\mu (x'\tau)^{(\mu)} + Y_\mu (y'\tau)^{(\mu)} \right\} \tag{7a}$$

$$= (Px' + Qy')\,\tau + D \sum_{\mu=1}^{n} \left\{ P_\mu (x'\tau)^{(\mu-1)} + Q_\mu (y'\tau)^{(\mu-1)} \right\}$$

$$= \sum_{\nu=0}^{n} (\Xi_\nu + H_\nu)\,\tau^{(\nu)} \tag{9a}$$

$$= (\Pi + P)\tau + D \sum_{\nu=1}^{n} (\Pi_\nu + P_\nu)\,\tau^{(\nu-1)}$$

$$\Xi_\nu + H_\nu = \Pi_\nu + P_\nu + D\,(\Pi_{\nu+1} + P_{\nu+1}) \tag{12a}$$

$$\Pi_\nu + P_\nu = \sum_{\chi=0}^{n-\nu} (-1)^\chi D^\chi\,(\Xi_{\nu+\chi} + H_{\nu+\chi})$$

$$(\nu = 0, 1, 2, \ldots n)\,.$$

Die bisherigen Beziehungen gelten, wenn X_μ, Y_μ beliebige Functionen von t sind. Ist aber:

$$F = F\left(x^{(\mu)}, y^{(\mu)}\right), \quad X_\mu = \frac{\partial F}{\partial x^{(\mu)}}, \quad Y_\mu = \frac{\partial F}{\partial y^{(\mu)}},$$

so ist nach (10):

$$\Xi_0 + H_0 = \Xi + H = \sum_{\mu=0}^{n} \left\{ X_\mu x^{(\mu+1)} + Y_\mu y^{(\mu+1)} \right\} = DF, \tag{13}$$

und weil nach (12a)

$$\begin{aligned} \Pi + P &= \Xi + H - D(\Pi_1 + P_1) \\ Px' + Qy' &= \Pi + P = D\left(F - \Pi_1 - P_1\right). \end{aligned} \tag{14}$$

Für solche F aber, welche unserer Forderung, d. h. der Gleichung (1a) genügen, ist

$$\begin{aligned} \delta_\vartheta F &= \sum_{\mu=0}^{n} \left\{ X_\mu (x'\tau)^{(\mu)} + Y_\mu (y'\tau)^{(\mu)} \right\} \\ &= \sum_{\nu=0}^{n} \left(\Xi_\nu + H_\nu\right) \tau^{(\nu)} = D(F\tau), \end{aligned} \tag{6}$$

oder wegen (7a) und (9a)

$$\begin{aligned} D(F\tau) &= (Px' + Qy')\,\tau + D \sum_{\mu=1}^{n} \left\{ P_\mu (x'\tau)^{(\mu-1)} + Q_\mu (y'\tau)^{(\mu-1)} \right\} \\ &= (\Pi + P)\tau + D \sum_{\nu=1}^{n} \left(\Pi_\nu + P_\nu\right) \tau^{(\nu-1)}. \end{aligned} \tag{6a}$$

11 | Es müsste also $(Px' + Qy')\tau = (\Pi + P)\tau$ für willkürliches $\tau = \tau(t)$ eine vollständige Ableitung sein, was nur möglich ist, wenn

$$\Pi + P = Px' + Qy' = 0, \tag{15}$$

The previous relations hold when X_μ, Y_μ are arbitrary functions of t. But if

$$F = F\left(x^{(\mu)}, y^{(\mu)}\right), \quad X_\mu = \frac{\partial F}{\partial x^{(\mu)}}, \quad Y_\mu = \frac{\partial F}{\partial y^{(\mu)}},$$

then, by (10),

$$\Xi_0 + H_0 = \Xi + H = \sum_{\mu=0}^{n} \left\{ X_\mu x^{(\mu+1)} + Y_\mu y^{(\mu+1)} \right\} = DF, \tag{13}$$

and since, by (12a),

$$\begin{aligned} \Pi + P &= \Xi + H - D(\Pi_1 + P_1) \\ Px' + Qy' &= \Pi + P = D\left(F - \Pi_1 - P_1\right). \end{aligned} \tag{14}$$

But for F satisfying our demand, i.e., the equation (1a),

$$\begin{aligned} \delta_\vartheta F &= \sum_{\mu=0}^{n} \left\{ X_\mu (x'\tau)^{(\mu)} + Y_\mu (y'\tau)^{(\mu)} \right\} \\ &= \sum_{\nu=0}^{n} \left(\Xi_\nu + H_\nu\right) \tau^{(\nu)} = D(F\tau), \end{aligned} \tag{6}$$

or, on account of (7a) and (9a),

$$\begin{aligned} D(F\tau) &= (Px' + Qy')\,\tau + D \sum_{\mu=1}^{n} \left\{ P_\mu (x'\tau)^{(\mu-1)} + Q_\mu (y'\tau)^{(\mu-1)} \right\} \qquad (6a) \\ &= (\Pi + P)\tau + D \sum_{\nu=1}^{n} \left(\Pi_\nu + P_\nu\right) \tau^{(\nu-1)}. \end{aligned}$$

Hence, $(Px' + Qy')\tau = (\Pi + P)\tau$ would have to be a total derivative for an arbitrary $\tau = \tau(t)$, which is possible only if

$$\Pi + P = Px' + Qy' = 0, \tag{15}$$

sodass die vorhergehende Gleichung durch Integration übergeht in:

$$\sum_{\mu=1}^{n} \left\{ P_\mu (x'\tau)^{(\mu-1)} + Q_\mu (y'\tau)^{(\mu-1)} \right\} \tag{16}$$
$$= \sum_{\nu=1}^{n} (\Pi_\nu + P_\nu)\, \tau^{(\nu-1)} = F\tau \ ,$$

wo die Integrationsconstante offenbar verschwindet.

Durch Coefficienten-Vergleichung erhält man daraus mit Hilfe von (10) und (11)

$$\Xi_\nu + H_\nu = \sum_{\mu=\nu}^{n} \binom{\mu}{\nu} \left\{ X_\mu x^{(\mu-\nu+1)} Y_\mu y^{(\mu-\nu+1)} \right\} = e_{\nu,1}\, F \ , \tag{17}$$

$$\Pi_\nu + P_\nu = \sum_{\mu=\nu}^{n} \binom{\mu-1}{\nu-1} \left\{ P_\mu x^{(\mu-\nu+1)} + Q_\mu y^{(\mu-\nu+1)} \right\} = e_{\nu,1}\, F \tag{18}$$
$$(\nu = 1, 2, \ldots n) \ ,$$

wo, wie auch im Folgenden,

$$e_{\mu,\nu} = 1 \ \ (\mu = \nu) \ , \qquad e_{\mu,\nu} = 0 \ \ (\mu \neq \nu)$$

die Bedeutung des Kroneckerschen Symbols ($\delta_{\mu,\nu}$) besitzt.

Diese beiden Systeme von n Gleichungen stellen, jedes für sich allein, die *vollständigen* Bedingungen für die Erfüllung unserer Forderung dar. Denn die Coefficienten von τ auf beiden Seiten von (6) stimmen nach (13) *identisch* überein, die von $\tau^{(\mu)}$ $(\mu > 0)$ aber werden durch (17) zur Übereinstimmung gebracht. Ferner entsteht (15) wegen (14) durch einfache Differentiation aus
12 | der zu (18) gehörigen Gleichung:

$$\Pi_1 + P_1 = F \ ,$$

sodass auch (18) gleichbedeutend ist mit (6a). Es folgt also aus (17) für den ursprünglichen, aus (18) für den transformierten Ausdruck jedesmal dieselbe Formel:

$$\delta_\vartheta F = D(F\tau) \ , \tag{6}$$

die auch als hinreichende Bedingung bereits nachgewiesen ist.

Der unmittelbare analytische Zusammenhang dieser beiden Systeme (17) und (18) wird durch die Formeln (12a) gegeben, durch deren zweite man unmittelbar (18) aus (17), durch deren erste aber umgekehrt (17) aus (18) ableiten kann.

Trotz dieser Äquivalenz wird man naturgemäss das System (17) als die einfachere, ursprünglichere Form der Bedingungen ansehen müssen, aus welcher die andere (18) erst durch Differentiation hervorgegangen ist. Doch wird

so that, by integration, the previous equation is transformed into

$$\sum_{\mu=1}^{n} \left\{ P_\mu (x'\tau)^{(\mu-1)} + Q_\mu (y'\tau)^{(\mu-1)} \right\} \tag{16}$$
$$= \sum_{\nu=1}^{n} (\Pi_\nu + P_\nu)\, \tau^{(\nu-1)} = F\tau \,,$$

where the constant of integration obviously vanishes.

From this we obtain, by comparison of coefficients and with the help of (10) and (11),

$$\Xi_\nu + H_\nu = \sum_{\mu=\nu}^{n} \binom{\mu}{\nu} \left\{ X_\mu x^{(\mu-\nu+1)} Y_\mu y^{(\mu-\nu+1)} \right\} = e_{\nu,1}\, F \,, \tag{17}$$

$$\Pi_\nu + P_\nu = \sum_{\mu=\nu}^{n} \binom{\mu-1}{\nu-1} \left\{ P_\mu x^{(\mu-\nu+1)} + Q_\mu y^{(\mu-\nu+1)} \right\} = e_{\nu,1}\, F \tag{18}$$
$$(\nu = 1, 2, \ldots n) \,,$$

where, as also in the following,

$$e_{\mu,\nu} = 1 \;\; (\mu = \nu) \,, \qquad e_{\mu,\nu} = 0 \;\; (\mu \neq \nu)$$

has the meaning of Kronecker's symbol ($\delta_{\mu,\nu}$).

Each of these two systems of n equations represents the *complete* requirements for the satisfaction of our demand. For the coefficients of τ on both sides of (6) agree *identically* according to (13), while those of $\tau^{(\mu)}$ ($\mu > 0$) are brought into agreement by means of (17). Furthermore, we obtain (15) on account of (14) by simple differentiation from the equation belonging to (18):

$$\Pi_1 + P_1 = F \,,$$

so that (18), too, has the same meaning as (6a). Hence, the same formula follows from (17) for the original expression and from (18) for the transformed expression, namely

$$\delta_\vartheta F = D(F\tau) \,, \tag{6}$$

which has already been shown to be a sufficient condition as well.

The direct analytic connection between the two systems (17) and (18) is given by the formulas (12a), the second of which allows for the immediate derivation of (18) from (17) and, conversely, the first of which, that of (17) from (18).

This equivalence notwithstanding, we will of course have to consider the system (17) the more primitive and more natural representation of the conditions, from which (18) arises only by differentiation. But it is just the latter

gerade die zweite in Gestalt der Gleichungen (16) und

$$\Pi_1 + P_1 = \sum_{\mu=1}^{n} \left(P_\mu x^{(\mu)} + Q_\mu y^{(\mu)} \right) = F \qquad (18)_1$$

bei den späteren Untersuchungen angewandt werden.

Beide Systeme haben eine einzige Gleichung gemeinsam, nämlich ($\nu = n$):

$$\begin{aligned} X_n x' + Y_n y' &= P_n x' + Q_n y' \qquad (17)_n, (18)_n \\ &= \frac{\partial F}{\partial x^{(n)}} x' + \frac{\partial F}{\partial y^{(n)}} y' = e_{n,1} F \ , \end{aligned}$$

welche für $n = 1$ in der Form:

$$\frac{\partial F}{\partial x'} x' + \frac{\partial F}{\partial y'} y' = F$$

als die in diesem Falle *einzige* Bedingung die Homogeneität der Function $F = F(x, y; x', y')$ in Bezug auf x' und y' ausdrückt.

Durch partielle Differentiation der allgemeinen Formel $(17)_n$, $(18)_n$ nach $x^{(n)}$ und $y^{(n)}$ ergiebt sich:

13 |
$$\begin{aligned} \frac{\partial^2 F}{\partial x^{(n)} \partial x^{(n)}} x' + \frac{\partial^2 F}{\partial x^{(n)} \partial y^{(n)}} y' &= 0 \\ \frac{\partial^2 F}{\partial x^{(n)} \partial y^{(n)}} x' + \frac{\partial^2 F}{\partial y^{(n)} \partial y^{(n)}} y' &= 0 \ , \end{aligned}$$

sodass man, wie es später geschehen wird, setzen kann:

$$\begin{aligned} \frac{\partial^2 F}{\partial x^{(n)} \partial x^{(n)}} = y'^2 F_1 \ , \quad \frac{\partial^2 F}{\partial y^{(n)} \partial y^{(n)}} = x'^2 F_1 \ , \qquad (19) \\ \frac{\partial^2 F}{\partial x^{(n)} \partial y^{(n)}} = -x' y' F_1 \ , \end{aligned}$$

wo

$$F_1 = \left(\frac{\partial^2 F}{\partial x^{(n)} \partial x^{(n)}} + \frac{\partial^2 F}{\partial y^{(n)} \partial y^{(n)}} \right) : \left(x'^2 + y'^2 \right)$$

eine von $x^{(\mu)}, y^{(\mu)}$ ($\mu = 0, 1, \ldots n$) abhängige Function ist, welche endlich, eindeutig und stetig bleibt, so lange die partiellen Ableitungen von F es sind und x', y' nicht beide gleichzeitig verschwinden.

Der ursprünglichen Bedingungsgleichung (1) oder (1a) kann man noch eine sehr gebräuchliche Form geben, wenn man in der Umgebung einer Stelle $t = \vartheta$, wo $\varphi'(\vartheta) \gtrless 0$ ist, eine Function $\vartheta(t)$ bestimmt durch die Gleichung: $t = \varphi(\vartheta) = x$, also

$$D^\mu \varphi(\vartheta) = e_{\mu,1} \quad (\mu > 0) \ ; \qquad D^\mu \psi(\vartheta) = \frac{d^\mu y}{dx^\mu} \ ,$$

which will be used in the form of the equations (16) and

$$\Pi_1 + P_1 = \sum_{\mu=1}^{n} \left(P_\mu x^{(\mu)} + Q_\mu y^{(\mu)} \right) = F \qquad (18)_1$$

later in our investigations.

Both systems have a single equation in common, namely ($\nu = n$):

$$\begin{aligned} X_n x' + Y_n y' &= P_n x' + Q_n y' \qquad (17)_n,\ (18)_n \\ &= \frac{\partial F}{\partial x^{(n)}} x' + \frac{\partial F}{\partial y^{(n)}} y' = e_{n,1}\, F\ , \end{aligned}$$

which, for $n = 1$, in the form

$$\frac{\partial F}{\partial x'} x' + \frac{\partial F}{\partial y'} y' = F$$

expresses the homogeneity of the function $F = F(x, y; x', y')$ with respect to x' and y', being the *sole* condition in this case.

Partial differentiation of the general formula $(17)_n$, $(18)_n$ with respect to $x^{(n)}$ and $y^{(n)}$ yields

$$\begin{aligned} \frac{\partial^2 F}{\partial x^{(n)} \partial x^{(n)}} x' + \frac{\partial^2 F}{\partial x^{(n)} \partial y^{(n)}} y' &= 0 \\ \frac{\partial^2 F}{\partial x^{(n)} \partial y^{(n)}} x' + \frac{\partial^2 F}{\partial y^{(n)} \partial y^{(n)}} y' &= 0\ , \end{aligned}$$

so that we can put, as will be done later,

$$\begin{aligned} \frac{\partial^2 F}{\partial x^{(n)} \partial x^{(n)}} = y'^2\, F_1\ , \quad \frac{\partial^2 F}{\partial y^{(n)} \partial y^{(n)}} = x'^2\, F_1\ , \qquad (19) \\ \frac{\partial^2 F}{\partial x^{(n)} \partial y^{(n)}} = -x' y'\, F_1\ , \end{aligned}$$

where

$$F_1 = \left(\frac{\partial^2 F}{\partial x^{(n)} \partial x^{(n)}} + \frac{\partial^2 F}{\partial y^{(n)} \partial y^{(n)}} \right) : \left(x'^2 + y'^2 \right)$$

is a function dependent on $x^{(\mu)}, y^{(\mu)}$ $(\mu = 0, 1, \ldots n)$, which remains finite, single-valued, and continuous as long as the partial derivatives of F do and as long as x', y' do not both simultaneously vanish.

We can also represent the original constraint equation (1) or (1a) in a very common form by determining a function $\vartheta(t)$ in the vicinity of a position $t = \vartheta$ where $\varphi'(\vartheta) \gtrless 0$ by means of the equation: $t = \varphi(\vartheta) = x$, and hence

$$D^\mu \varphi(\vartheta) = e_{\mu,1} \quad (\mu > 0)\ ; \qquad D^\mu \psi(\vartheta) = \frac{d^\mu y}{dx^\mu}\ ,$$

wo

$$\frac{dy}{dx} = \frac{y'}{x'} , \qquad \frac{d^2y}{dx^2} = \frac{x'y'' - y'x''}{x'^3}$$

u. s. w., allgemein:

$$\frac{d^\mu y}{dx^\mu} = \frac{S_\mu \left(x', x'', \ldots x^{(\mu)}; y', y'', \ldots y^{(\mu)}\right)}{x'^{2\mu-1}} = \frac{S_\mu \left(x^{(\nu)}, y^{(\nu)}\right)}{x'^{2\mu-1}}$$

gesetzt werden kann, wenn $x^{(\mu)} = D^\mu\varphi(\vartheta)$, $y^{(\mu)} = D^\mu\psi(\vartheta)$ bedeutet für beliebige Beziehungen zwischen t und ϑ, also auch für $\vartheta = t$, $x^{(\mu)} = \varphi^{(\mu)}(\vartheta)$, $y^{(\mu)} = \psi^{(\mu)}(\vartheta)$. Setzt man nun die oben gefundenen Ausdrücke in (1a) ein, so erhält man:

14 |
$$F\left(x, 1, 0, \ldots 0; y, \frac{dy}{dx}, \ldots \frac{d^n y}{dx^n}\right) = F\left(x, \frac{d^\mu y}{dx^\mu}\right) = F\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right) \frac{d\vartheta}{dx} .$$

Nun ist, wieder nach (1a), für beliebiges $\vartheta = \vartheta(t)$

$$F\left(x^{(\mu)}, y^{(\mu)}\right) = F\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big) = F\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right)\vartheta'$$
$$= F\left(x, \frac{d^\mu y}{dx^\mu}\right)\frac{dx}{d\vartheta}\cdot\frac{d\vartheta}{dt} = F\left(x, \frac{d^\mu y}{dx^\mu}\right)x' ,$$

d. h.

$$F\left(x^{(\mu)}, y^{(\mu)}\right) = F\left(x, y, \frac{dy}{dx}, \ldots \frac{d^n y}{dx^n}\right)x' , \tag{20}$$

eine identische Beziehung, wenn man für $\frac{d^\mu y}{dx^\mu}$ ihre eben gefundenen Ausdrücke durch die $x^{(\mu)}, y^{(\mu)}$ einsetzt.

Das angewandte Verfahren und die gefundenen Bedingungsgleichungen lassen sich unmittelbar auf den allgemeineren Fall einer *grösseren Anzahl von Variablen* $x, y, z, \ldots$ übertragen. So wird z. B. ein Integral:

$$J = \int_{t_1}^{t_2} F\left(x^{(\mu)}, y^{(\mu)}, z^{(\mu)}\right) dt$$

über ein Stück 1 2 einer *Raumcurve*: $x = \varphi(t)$, $y = \psi(t)$, $z = \chi(t)$ erstreckt, einen von der besonderen Form der Darstellung, d. h. von der Wahl des Parameters t unabhängigen Wert besitzen, wenn:

$$\Xi_\nu + H_\nu + Z_\nu = \Pi_\nu + P_\nu + T_\nu = e_{\nu,1}F \qquad (17), (18)$$
$$(\nu = 1, 2, \ldots n) ,$$

where it is possible to put

$$\frac{dy}{dx} = \frac{y'}{x'} , \qquad \frac{d^2y}{dx^2} = \frac{x'y'' - y'x''}{x'^3}$$

e. t. c., generally,

$$\frac{d^\mu y}{dx^\mu} = \frac{S_\mu \left(x', x'', \ldots x^{(\mu)}; y', y'', \ldots y^{(\mu)}\right)}{x'^{2\mu-1}} = \frac{S_\mu \left(x^{(\nu)}, y^{(\nu)}\right)}{x'^{2\mu-1}} ,$$

if $x^{(\mu)}$ is $D^\mu\varphi(\vartheta)$ and $y^{(\mu)}$ is $D^\mu\psi(\vartheta)$ for any relation between t and ϑ, and hence also for $\vartheta = t$, $x^{(\mu)} = \varphi^{(\mu)}(\vartheta)$, $y^{(\mu)} = \psi^{(\mu)}(\vartheta)$. Substituting the expressions found above into (1a), one obtains

$$F\left(x, 1, 0, \ldots 0; y, \frac{dy}{dx}, \ldots \frac{d^n y}{dx^n}\right) = F\left(x, \frac{d^\mu y}{dx^\mu}\right) = F\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right) \frac{d\vartheta}{dx} .$$

We now have, again on account of (1a), for arbitrary $\vartheta = \vartheta(t)$,

$$F\left(x^{(\mu)}, y^{(\mu)}\right) = F\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big) = F\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right)\vartheta' = F\left(x, \frac{d^\mu y}{dx^\mu}\right)\frac{dx}{d\vartheta}\cdot\frac{d\vartheta}{dt} = F\left(x, \frac{d^\mu y}{dx^\mu}\right)x' ,$$

i. e.,

$$F\left(x^{(\mu)}, y^{(\mu)}\right) = F\left(x, y, \frac{dy}{dx}, \ldots \frac{d^n y}{dx^n}\right)x' , \tag{20}$$

an identical relation if the $\dfrac{d^\mu y}{dx^\mu}$ are replaced by the expressions in $x^{(\mu)}, y^{(\mu)}$ just found for them.

The method employed and the constraint equations found can be immediately applied to the more general case of a *greater number of variables* $x, y, z, \ldots$. For instance, the integral

$$J = \int_{t_1}^{t_2} F\left(x^{(\mu)}, y^{(\mu)}, z^{(\mu)}\right) dt$$

taken along a segment 1 2 of a *space curve*: $x = \varphi(t)$, $y = \psi(t)$, $z = \chi(t)$, possesses a value independent of the particular form of representation, i. e., the choice of the parameter t, if

$$\Xi_\nu + H_\nu + Z_\nu = \Pi_\nu + P_\nu + T_\nu = e_{\nu,1}F \qquad (17), (18)$$
$$(\nu = 1, 2, \ldots n) ,$$

wo die Ausdrücke Z_ν, T_ν ebenso nach z gebildet sind wie $\Xi_\nu, \Pi_\nu; H_\nu, P_\nu$ nach x und y.

Die Untersuchung ist hier mit grösserer Vollständigkeit geführt worden, als es für die unmittelbare Anwendung auf das Problem der Variationsrechnung notwendig gewesen wäre; die Frage ist als eine Aufgabe von selbständi-
15 gem Interesse aufgefasst | worden, die bisher eine ausreichende Beantwortung noch nicht gefunden zu haben scheint.

Zuerst, soweit ich in Erfahrung bringen konnte, stellt sie *Lagrange* in seinen „Leçons sur le Calcul des Fonctions“ in dem der Variationsrechnung gewidmeten Abschnitte und zwar in der hier durch (20) gegebenen Form, und er beweist hier die *Notwendigkeit* der einen Gleichung:

$$Px' + Qy' = 0 \tag{15}$$

(in meiner Bezeichnung), ohne sich über die hinreichenden Bedingungen zu äussern. Nach ihm gelangt *Poisson* in seinem „Mémoire sur le Calcul des Variations“ auf ähnlichem Wege zu demselben Ergebnis, hält aber, wie er ausdrücklich erklärt, diese eine Bedingung zugleich für *hinreichend*, ein Irrtum, den auch *Todhunter* (History on the Progress of the Calculus of Variations) in seiner Besprechung der Poissonschen Schrift wiederholt hat.

Dass (15) in Wirklichkeit für *keinen* Wert von n eine hinreichende Bedingung ist, geht schon daraus hervor, dass

$$F = C = \text{const}$$

vermöge $P = 0$, $Q = 0$ diese Bedingung befriedigt, während doch

$$C = \frac{C}{\vartheta'}\vartheta' = \frac{C}{x'}x'$$

von keiner der Formen (1a) oder (20) ist und auch das Integral

$$\int_{t_1}^{t_2} C\, dt = C(t_2 - t_1)$$

von der Wahl des Parameters abhängig ist. Eine beliebige Function F von der in Frage stehenden Eigenschaft muss daher nach Hinzufügung einer beliebigen, nicht verschwindenden Constanten diese Eigenschaft verlieren, während $Px' + Qy' = 0$ unverändert bestehen bleibt. Thatsächlich drückt die Gleichung (15) nur die Bedingung dafür aus, dass $F(x^{(\mu)}, y^{(\mu)})$, wenn für y irgend eine Function von x, oder allgemeiner, wenn für x, y irgend welche Functionen von ϑ eingesetzt werden, *„integrabel“* sei, dass nämlich:

16 | $$\int_{t_1}^{t_2} F\left(x^{(\mu)}, y^{(\mu)}\right) dt = \int_{t_1}^{t_2} F_0\left(\vartheta, \vartheta', \ldots \vartheta^{(n)}\right) dt = \left[J\left(\vartheta, \vartheta', \ldots \vartheta^{(n)}\right)\right]_{t_1}^{t_2}$$

where the expressions Z_ν, T_ν are formed after z, just as $\Xi_\nu, \Pi_\nu; H_\nu, P_\nu$ are after x and y.

This case has been investigated more thoroughly than necessary for the immediate application to the problem of the calculus of variations; the question has been considered a problem of independent interest and does not seem to have been completely answered yet.

As far as I was able to learn, it was *Lagrange* who posed this question for the first time, namely in the form it takes here for (20), in the part of his *1806* that is devoted to the calculus of variations. There, he proves the *necessity* of the one equation

$$Px' + Qy' = 0 \tag{15}$$

(in my terminology), without saying anything about the sufficient conditions. Subsequently, *Poisson* obtains the same result in a similar fashion in his *1823* but, as he expressly states, takes this one condition to be also *sufficient*, a mistake that *Todhunter* (*1861*) duplicates in his review of Poisson's treatise.

That (15) is, in fact, *not* a sufficient condition for *any* value of n can already be seen when we consider that

$$F = C = \text{const}$$

satisfies this condition by dint of $P = 0$, $Q = 0$, whereas

$$C = \frac{C}{\vartheta'}\vartheta' = \frac{C}{x'}x'$$

is neither of the form (1a) nor (20), and also the integral

$$\int_{t_1}^{t_2} C\,dt = C(t_2 - t_1)$$

depends on the choice of the parameter. Adding an arbitrary nonvanishing constant to an arbitrary function F of the property under consideration must therefore deprive the function of this property, whereas $Px' + Qy' = 0$ continues to hold. In fact, the equation (15) only expresses the condition that $F(x^{(\mu)}, y^{(\mu)})$ is "*integrable*" whenever some arbitrary function of x is substituted for y or, more generally, some arbitrary functions of ϑ for x, y, namely that

$$\int_{t_1}^{t_2} F\left(x^{(\mu)}, y^{(\mu)}\right) dt = \int_{t_1}^{t_2} F_0\left(\vartheta, \vartheta', \ldots \vartheta^{(n)}\right) dt = \left[J\left(\vartheta, \vartheta', \ldots \vartheta^{(n)}\right)\right]_{t_1}^{t_2}$$

nur von den Endwerten:

$$\vartheta_1, \vartheta_1', \ldots \vartheta_1^{(n)};\ \vartheta_2, \vartheta_2', \ldots \vartheta_2^{(n-1)}\ ,$$

nicht aber von dem ganzen Verlauf, von der Form der Function ϑ abhängt.

Die bekannte Eulersche „Integrabilitätsbedingung“

$$\sum_{\mu=0}^{n} (-1)^\mu D^\mu \frac{\partial F}{\partial \vartheta^{(\mu)}} = 0$$

ist nämlich äquivalent der Eigenschaft, dass der Ausdruck:

$$\delta_\vartheta F = \frac{\partial F}{\partial \vartheta}\delta\vartheta + \frac{\partial F}{\partial \vartheta'}\delta\vartheta' + \ldots + \frac{\partial F}{\partial \vartheta^{(n)}}\delta\vartheta^{(n)}$$

seinerseits „integrabel“ ist für eine willkürliche Function $\delta\vartheta = \vartheta'\tau$.

Es ist aber nach (7a)

$$\delta_\vartheta F = (Px' + Qy')\,\tau + D\sum_{\mu=1}^{n}\left\{P_\mu (x'\tau)^{(\mu-1)} + Q_\mu (y'\tau)^{(\mu-1)}\right\}\ ,$$

also die Integrabilitätsbedingung, wie behauptet:

$$Px' + Qy' = 0\ . \tag{15}$$

So ist z. B.:

$$F = xy'' - yx'' + xyx' = \frac{d}{dt}(xy' - yx') + xyx'$$

nach t integrabel, sobald für y irgend eine Function von x eingesetzt wird, das Integral aber hängt dann immer noch von x' oder von $y' = \frac{dy}{dx}x'$ ab, ändert sich also auch mit der Darstellung der Curve, über welche die Integration erstreckt wird.

17 | In der That ist hier auch:

$$P = \frac{\partial F}{\partial x} - \frac{d}{dt}\frac{\partial F}{\partial x'} + \frac{d^2}{dt^2}\frac{\partial F}{\partial x''} = y'' + yx' - \frac{d}{dt}(xy) - \frac{d^2}{dt^2}y = -xy'$$

$$Q = \frac{\partial F}{\partial y} - \frac{d}{dt}\frac{\partial F}{\partial y'} + \frac{d^2}{dt^2}\frac{\partial F}{\partial y''} = -x'' + xx' + \frac{d^2}{dt^2}x = xx'\ ,$$

also

$$Px' + Qy' = 0\ , \tag{15}$$

während doch:

$$\frac{\partial F}{\partial x''}x' + \frac{\partial F}{\partial y''}y' = -yx' + xy'$$

nicht identisch verschwindet, $(17)_2$ also nicht erfüllt ist.

only depends on the end values

$$\vartheta_1, \vartheta_1', \dots \vartheta_1^{(n)};\ \vartheta_2, \vartheta_2', \dots \vartheta_2^{(n-1)}\ ,$$

but not on the entire course, on the form of the function ϑ.

For Euler's well-known "integrability condition"

$$\sum_{\mu=0}^{n} (-1)^{\mu} D^{\mu} \frac{\partial F}{\partial \vartheta^{(\mu)}} = 0$$

is equivalent to the property that the expression

$$\delta_{\vartheta} F = \frac{\partial F}{\partial \vartheta}\delta\vartheta + \frac{\partial F}{\partial \vartheta'}\delta\vartheta' + \dots + \frac{\partial F}{\partial \vartheta^{(n)}}\delta\vartheta^{(n)}$$

is, in turn, "integrable" for an arbitrary function $\delta\vartheta = \vartheta'\tau$.

But, by (7a),

$$\delta_{\vartheta} F = (Px' + Qy')\,\tau + D\sum_{\mu=1}^{n} \left\{ P_{\mu}(x'\tau)^{(\mu-1)} + Q_{\mu}(y'\tau)^{(\mu-1)} \right\}\ ,$$

and hence the integrability condition, as asserted:

$$Px' + Qy' = 0\ . \tag{15}$$

Thus, for instance,

$$F = xy'' - yx'' + xyx' = \frac{d}{dt}(xy' - yx') + xyx'$$

is integrable with respect to t, as soon as an arbitrary function of x is substituted for y. But then the integral still depends on x' or on $y' = \dfrac{dy}{dx}x'$, and hence varies along with the representation of the curve along which the integration is carried out.

In fact, we also have

$$P = \frac{\partial F}{\partial x} - \frac{d}{dt}\frac{\partial F}{\partial x'} + \frac{d^2}{dt^2}\frac{\partial F}{\partial x''} = y'' + yx' - \frac{d}{dt}(xy) - \frac{d^2}{dt^2}y = -xy'$$

$$Q = \frac{\partial F}{\partial y} - \frac{d}{dt}\frac{\partial F}{\partial y'} + \frac{d^2}{dt^2}\frac{\partial F}{\partial y''} = -x'' + xx' + \frac{d^2}{dt^2}x = xx'\ ,$$

thus

$$Px' + Qy' = 0\ , \tag{15}$$

while

$$\frac{\partial F}{\partial x''}x' + \frac{\partial F}{\partial y''}y' = -yx' + xy'$$

does not vanish identically, and hence $(17)_2$ is not satisfied.

Das Verfahren, dessen wir uns zur Ableitung der Bedingungsgleichungen für F bedienten, lässt sich leicht auf *verwandte Aufgaben* übertragen.

Eine Function

$$\Phi\left(x^{(\mu)}, y^{(\mu)}\right) = \Phi\left(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\right)$$
$$(\mu = 0, 1, \ldots n, \vartheta = \vartheta(t))$$

soll dieselbe Eigenschaft besitzen wie das Integral von F, d. h. sie soll von der besonderen Darstellung der Curve $x = \varphi(\vartheta)$, $y = \psi(\vartheta)$, also von $\vartheta', \vartheta'', \ldots \vartheta^{(n)}$ unabhängig sein und nach Wahl von φ, ψ nur noch von ϑ abhängen, so dass (*vergl.* (1a)):

$$\Phi\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big) = \Phi\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right) . \tag{21}$$

Dann muss Φ wie vorher $J'(\vartheta)$ der Gleichung (4) genügen:

$$\begin{aligned}\delta_\vartheta\Phi &= \sum_{\mu=0}^{n}\left\{\frac{\partial\Phi}{\partial x^{(\mu)}}\delta x^{(\mu)} + \frac{\partial\Phi}{\partial y^{(\mu)}}\delta y^{(\mu)}\right\} \\ &= \sum_{\mu=0}^{n}\left\{X_\mu(x'\tau)^{(\mu)} + Y_\mu(y'\tau)^{(\mu)}\right\} = D\Phi\cdot\tau ,\end{aligned} \tag{22}$$

18 | wenn jetzt:

$$X_\mu = \frac{\partial\Phi}{\partial x^{(\mu)}} , \quad Y_\mu = \frac{\partial\Phi}{\partial y^{(\mu)}}$$

und wieder $\delta\vartheta = \vartheta'\tau$ gesetzt wird, wo τ eine willkürliche Function von t ist.

Nun gelten für $\delta_\vartheta\Phi$ dieselben Umformungen wie für $\delta_\vartheta F$, so dass man schliesslich mit Anwendung derselben Bezeichnungen die verlangten Bedingungen in einer der Formen schreiben kann:

$$\begin{aligned}&\Xi_\nu + H_\nu = 0 \\ &\qquad\qquad (\nu = 1, 2, \ldots n) . \\ &\Pi_\nu + P_\nu = 0\end{aligned} \tag{23}$$

Wenn man in (21) wieder $t = \varphi(\vartheta) = x$ einsetzt, so ergiebt sich:

$$\begin{aligned}\Phi\left(x^{(\mu)}, y^{(\mu)}\right) &= \Phi\left(x, 1, 0, \ldots 0;\ y, \frac{dy}{dx}, \ldots \frac{d^n y}{dx^n}\right) \\ &= \Phi\left(x, \frac{d^\mu y}{dx^\mu}\right) ,\end{aligned} \tag{24}$$

wo wie in (20) wieder

$$\frac{d^\mu y}{dx^\mu} = \frac{S_\mu\left(x^{(\nu)}, y^{(\nu)}\right)}{x'^{2\mu-1}}$$

gesetzt werden kann.

The procedure we used to derive the constraint equations for F can easily be applied to *related problems.*

Suppose that a function

$$\Phi\left(x^{(\mu)}, y^{(\mu)}\right) = \Phi\left(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\right)$$
$$(\mu = 0, 1, \ldots n, \vartheta = \vartheta(t))$$

has the same property as the integral of F, i. e., it is independent of the particular representation of the curve $x = \varphi(\vartheta)$, $y = \psi(\vartheta)$, and hence of $\vartheta', \vartheta'', \ldots \vartheta^{(n)}$, and only depends on ϑ, after φ, ψ have been chosen, so that (*cf.* (1a)):

$$\Phi\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big) = \Phi\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right) . \tag{21}$$

Then Φ, just like $J'(\vartheta)$ above, must satisfy the equation (4):

$$\begin{aligned}\delta_\vartheta\Phi &= \sum_{\mu=0}^{n}\left\{\frac{\partial\Phi}{\partial x^{(\mu)}}\delta x^{(\mu)} + \frac{\partial\Phi}{\partial y^{(\mu)}}\delta y^{(\mu)}\right\} \\ &= \sum_{\mu=0}^{n}\left\{X_\mu(x'\tau)^{(\mu)} + Y_\mu(y'\tau)^{(\mu)}\right\} = D\Phi\cdot\tau ,\end{aligned} \tag{22}$$

provided that we now set

$$X_\mu = \frac{\partial\Phi}{\partial x^{(\mu)}} , \quad Y_\mu = \frac{\partial\Phi}{\partial y^{(\mu)}}$$

and $\delta\vartheta = \vartheta'\tau$ again, where τ is any function of t.

In this case, the same transformations hold for $\delta_\vartheta\Phi$ as for $\delta_\vartheta F$ so that, using the same denotations, we can finally write the required conditions in one of the following forms:

$$\begin{aligned}&\Xi_\nu + H_\nu = 0 \\ &\qquad\qquad (\nu = 1, 2, \ldots n) . \\ &\Pi_\nu + P_\nu = 0\end{aligned} \tag{23}$$

If we again substitute $t = \varphi(\vartheta) = x$ into (21), we obtain

$$\begin{aligned}\Phi\left(x^{(\mu)}, y^{(\mu)}\right) &= \Phi\left(x, 1, 0, \ldots 0;\ y, \frac{dy}{dx}, \ldots \frac{d^n y}{dx^n}\right) \\ &= \Phi\left(x, \frac{d^\mu y}{dx^\mu}\right) ,\end{aligned} \tag{24}$$

where, as in (20),

$$\frac{d^\mu y}{dx^\mu} = \frac{S_\mu\left(x^{(\nu)}, y^{(\nu)}\right)}{x'^{2\mu-1}} .$$

Ebenso kann man auch die *Bogenlänge* einführen:

$$t = s = \int \sqrt{x'^2 + y'^2}\, dt \ ,$$

also:

$$\Phi\left(x^{(\mu)}, y^{(\mu)}\right) = \Phi\left(\frac{d^\mu x}{ds^\mu}, \frac{d^\mu y}{ds^\mu}\right)$$

setzen. Es ist aber:

$$\frac{dx}{ds} = \cos\alpha_1 \ , \quad \frac{dy}{ds} = \sin\alpha_1 \ ,$$

wo α_1 den von Tangente und x-Richtung gebildeten Winkel bezeichnet, und es können die

19 | $$\frac{d^\mu x}{ds^\mu} = \frac{d^{\mu-1}\cos\alpha_1}{ds^{\mu-1}} \ , \qquad \frac{d^\mu y}{ds^\mu} = \frac{d^{\mu-1}\sin\alpha_1}{ds^{\mu-1}}$$

ausgedrückt werden als ganze rationale Functionen von:

$$\cos\alpha_1 \ , \ \ \sin\alpha_1 \ , \ \ \alpha_2 = \frac{d\alpha_1}{ds} \ , \ \ \ldots \ \ \alpha_\mu = \frac{d^{\mu-1}\alpha_1}{ds^{\mu-1}} \ ,$$

so dass schliesslich:

$$\Phi\left(x^{(\mu)}, y^{(\mu)}\right) = \Phi_1\left(x, y, \alpha_1, \alpha_2, \ldots \alpha_n\right) = \Phi\left(x, y, a_\mu\right) \ . \tag{25}$$

Hier ist aber:

$$\alpha_1 = \operatorname{arctg}\frac{y'}{x'} \ , \qquad \alpha_2 = \frac{x'y'' - y'x''}{\left(x'^2 + y'^2\right)^{\frac{3}{2}}} \quad \text{(die Krümmung)}$$

und allgemein:

$$\alpha_\mu = \frac{d^{\mu-1}\alpha_1}{ds^{\mu-1}} = \frac{d^{\mu-2}\alpha_2}{ds^{\mu-2}} = \frac{T_\mu\left(x^{(\nu)}, y^{(\nu)}\right)}{\left(x'^2 + y'^2\right)^{\frac{3\mu-3}{2}}} \quad (\mu > 1) \, ,$$

wo $T_\mu(x^{(\nu)}, y^{(\nu)})$ eine ganze rationale Function von $x', x'', \ldots x^{(\mu)}$; y', y'', $\ldots$ $y^{(\mu)}$ bezeichnet.

Diese Ausdrücke der α_μ besitzen dieselbe Eigenschaft (21), unabhängig von der Differentiationsvariablen zu sein, wie Φ selbst. Nach (25) lässt sich also jede Function ψ von der betrachteten Eigenschaft, jede *„Osculations-Invariante“*, wie sie aus nachher anzugebendem Grunde genannt werden möge, durch die $n + 2$ besonderen $x, y, \alpha_1, \ldots \alpha_n$ ausdrücken, ebenso wie nach (24) durch die $x, y, \frac{d^\mu y}{dx^\mu}$. Doch hat die Darstellungsweise (25) den wesentlichen Vorzug vor der anderen, dass die α_μ immer *endlich* bleiben für alle

In the same way, we can introduce the *arc length*

$$t = s = \int \sqrt{x'^2 + y'^2}\, dt \ ,$$

and hence set

$$\Phi\left(x^{(\mu)}, y^{(\mu)}\right) = \Phi\left(\frac{d^\mu x}{ds^\mu}, \frac{d^\mu y}{ds^\mu}\right) \ .$$

But

$$\frac{dx}{ds} = \cos\alpha_1 \ , \quad \frac{dy}{ds} = \sin\alpha_1 \ ,$$

where α_1 denotes the angle between the tangent and the x direction, and we can express the

$$\frac{d^\mu x}{ds^\mu} = \frac{d^{\mu-1}\cos\alpha_1}{ds^{\mu-1}} \ , \qquad \frac{d^\mu y}{ds^\mu} = \frac{d^{\mu-1}\sin\alpha_1}{ds^{\mu-1}}$$

as integral rational functions of

$$\cos\alpha_1 \ , \quad \sin\alpha_1 \ , \quad \alpha_2 = \frac{d\alpha_1}{ds} \ , \quad \ldots \quad \alpha_\mu = \frac{d^{\mu-1}\alpha_1}{ds^{\mu-1}} \ ,$$

so that finally

$$\Phi\left(x^{(\mu)}, y^{(\mu)}\right) = \Phi_1\left(x, y, \alpha_1, \alpha_2, \ldots \alpha_n\right) = \Phi\left(x, y, a_\mu\right) \ . \tag{25}$$

But here,

$$\alpha_1 = \operatorname{arctg}\frac{y'}{x'} \ , \qquad \alpha_2 = \frac{x'y'' - y'x''}{\left(x'^2 + y'^2\right)^{\frac{3}{2}}} \quad \text{(the curvature)}$$

and, generally,

$$\alpha_\mu = \frac{d^{\mu-1}\alpha_1}{ds^{\mu-1}} = \frac{d^{\mu-2}\alpha_2}{ds^{\mu-2}} = \frac{T_\mu\left(x^{(\nu)}, y^{(\nu)}\right)}{\left(x'^2 + y'^2\right)^{\frac{3\mu-3}{2}}} \quad (\mu > 1) \, ,$$

where $T_\mu(x^{(\nu)}, y^{(\nu)})$ denotes an integral rational function of $x', x'', \ldots x^{(\mu)}$; y', $y'', \ldots y^{(\mu)}$.

These expressions for the α_μ possess the same property (21) of being independent of the variable of differentiation as Φ itself. By (25), we can therefore express every function ψ with the property under consideration, every "*osculation invariant*", as we will call it for reasons to be specified later, in terms of the $n+2$ special $x, y, \alpha_1, \ldots \alpha_n$ and, likewise, by (24), in terms of the $x, y, \dfrac{d^\mu y}{dx^\mu}$. But (25) has the essential advantage over the other mode of

„regulären“ Stellen der Curven, d. h. solche, wo die $x^{(\mu)}, y^{(\mu)}$ endlich sind und gleichzeitig $x'^2 + y'^2 > 0$, die x' und y' nicht beide verschwinden.

Auch für die Function F von der früher untersuchten Beschaffenheit ergiebt sich hieraus *eine neue Darstellungsweise*. Setzt man nämlich:

20 | $$F\left(x^{(\mu)}, y^{(\mu)}\right) = \Phi\frac{ds}{dt} = \Phi\left(x^{(\mu)}, y^{(\mu)}\right)\sqrt{x'^2 + y'^2}\,,$$

also für $x = \varphi(\vartheta)$, $y = \varphi(\vartheta)$, je nachdem man nach t oder ϑ differentiiert, einerseits

$$F\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big) = \Phi\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big).\sqrt{\varphi'^2(\vartheta) + \psi'^2(\vartheta)}\,\vartheta'\,,$$

andererseits aber

$$F\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right) = \Phi\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right)\sqrt{\varphi'^2(\vartheta) + \psi'^2(\vartheta)}\,,$$

so wird nach:

$$F\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big) = F\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right)\vartheta'\,, \tag{1a}$$

jetzt:

$$\Phi\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big) = \Phi\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right)\,,$$

Φ ist also eine „Osculations-Invariante“ und daher nach (25)

$$F\left(x^{(\mu)}, y^{(\mu)}\right) = \Phi(x, y, \alpha_\mu)\sqrt{x'^2 + y'^2} \tag{26}$$
$$F\,dt = \Phi(x, y, \alpha_\mu)\,ds\,.$$

Der für Functionen der hier betrachteten Art eingeführte Ausdruck gründet sich auf ihre Bedeutung für die *Berührung oder „Osculation“ zweier Curven*, über welche sich hier einige Bemerkungen anschliessen mögen, die in den folgenden Untersuchungen vielfache Anwendung finden werden.

Wenn zwei Curven $x = \varphi(u)$, $y = \psi(u)$; $x = \overline{\varphi}(v)$, $y = \overline{\psi}(v)$ in den Punkten $u = u_0$, $v = v_0$, in denen sie sich regulär verhalten, wo also auch $\varphi'^2(u) + \psi'^2(u)$ und $\overline{\varphi}'^2(v) + \overline{\psi}'^2(v) > 0$ sind, *gemeinsame Werte* besitzen für $x, y, \alpha_1, \alpha_2, \ldots \alpha_m$ und damit für alle Osculations-Invarianten $\Phi(x^{(\mu)}, y^{(\mu)})$ bis zur $m$ten Ordnung ($\mu = 0, 1, \ldots m$), so sagt man, sie *berühren einander von $m$ter Ordnung* in dem Punkte:

$$x_0 = \varphi(u_0) = \overline{\varphi}(v_0)\,, \qquad y_0 = \psi(u_0) = \overline{\psi}(v_0)\,.$$

21 | Genauer wäre dieses Verhältnis auszudrücken: sie berühren einander „von mindestens m ter Ordnung“, noch genauer „von einer Ordnung $> m - 1$“, im Sinne der strengeren von Herrn *Weierstrass* in seinen Vorlesungen gegebenen Definition, der zufolge die wahre Ordnungszahl der Berührung auch

representation that the α_μ always remain *finite* for all "regular" positions of the curves, i. e., those where the $x^{(\mu)}, y^{(\mu)}$ are finite and, at the same time, $x'^2 + y'^2 > 0$, the x' and y' do not both vanish.

This also yields *a new mode of representation* for the function F of the constitution investigated earlier. For if we set

$$F\left(x^{(\mu)}, y^{(\mu)}\right) = \Phi\frac{ds}{dt} = \Phi\left(x^{(\mu)}, y^{(\mu)}\right)\sqrt{x'^2 + y'^2}\,,$$

and hence, for $x = \varphi(\vartheta)$, $y = \varphi(\vartheta)$, depending on whether we differentiate with respect to t or ϑ, on the one hand

$$\begin{aligned} &F\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big) \\ &= \Phi\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big).\sqrt{\varphi'^2(\vartheta) + \psi'^2(\vartheta)}\,\vartheta'\,, \end{aligned}$$

but on the other hand

$$F\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right) = \Phi\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right)\sqrt{\varphi'^2(\vartheta) + \psi'^2(\vartheta)}\,,$$

then, by

$$F\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big) = F\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right)\vartheta'\,, \tag{1a}$$

now

$$\Phi\big(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\big) = \Phi\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right)\,,$$

and hence Φ is an "osculation invariant", and therefore, by (25),

$$\begin{aligned} F\left(x^{(\mu)}, y^{(\mu)}\right) &= \Phi(x, y, \alpha_\mu)\sqrt{x'^2 + y'^2} \\ F\,dt &= \Phi(x, y, \alpha_\mu)\,ds\,. \end{aligned} \tag{26}$$

The expression introduced here for functions of the kind under consideration is based on their significance for the *contact or "osculation" of two curves*, about which a few comments are in order here which will be invoked many times in the investigations to follow.

If two curves $x = \varphi(u)$, $y = \psi(u)$; $x = \overline{\varphi}(v)$, $y = \overline{\psi}(v)$ have *common values* for $x, y, \alpha_1, \alpha_2, \ldots \alpha_m$, and hence for all osculation invariants $\Phi(x^{(\mu)}, y^{(\mu)})$ up to the m th order ($\mu = 0, 1, \ldots m$), at the points $u = u_0$, $v = v_0$ where they are regular, and hence where also $\varphi'^2(u) + \psi'^2(u)$ and $\overline{\varphi}'^2(v) + \overline{\psi}'^2(v) > 0$, then they are said to *have contact of the m th order* at the point

$$x_0 = \varphi(u_0) = \overline{\varphi}(v_0)\,, \qquad y_0 = \psi(u_0) = \overline{\psi}(v_0)\,.$$

It would be more precise to express this relationship as follows: they have contact "of at least m th order", or even more precisely, "of an order $> m-1$", in the sense of the more rigorous definition given by Mr. *Weierstrass* in his

gebrochen sein kann. Doch will ich mich hier immer der abgekürzten Ausdrucksweise bedienen.

Damit der betrachtete Umstand eintritt, müssen sich, wie sich streng beweisen liesse, für geeignete unabhängige Variable die Ableitungen der Coordinaten bis zur m ten Ordnung zur Übereinstimmung bringen lassen, d. h. es müssen sich u und v so als Functionen eines beliebigen Parameters t darstellen lassen, dass für:

$$t = t_0 \,, \qquad D^\mu u = \frac{d^\mu u}{dt^\mu} = u_0^{(\mu)} \,, \qquad D^\mu v = \frac{d^\mu v}{dt^\mu} = v_0^{(\mu)}$$

die Beziehungen bestehen:

$$D^\mu \varphi(u) = D^\mu \overline{\varphi}(v) \,, \qquad D^\mu \psi(u) = D^\mu \overline{\psi}(v) \qquad (27)$$
$$(\mu = 0, 1, 2, \ldots m)$$

oder, was nach (2) auf dasselbe hinauskommt, dass die Gleichungen:

$$R_\mu \left(\varphi^{(\nu)}(u_0), u_0^{(\nu)} \right) = R_\mu \left(\overline{\varphi}^{(\nu)}(v_0), v_0^{(\nu)} \right)$$
$$R_\mu \left(\psi^{(\nu)}(u_0), u_0^{(\nu)} \right) = R_\mu \left(\overline{\psi}^{(\nu)}(v_0), v_0^{(\nu)} \right)$$
$$(\mu = 0, 1, 2, \ldots m)$$

durch geeignete Werte der $u_0^{(\mu)}, v_0^{(\nu)}$ $(\mu = 0, 1, \ldots m)$ befriedigt werden können.

Ist dies der Fall, so ist auch in der That an der betrachteten Stelle nach (21):

$$\Phi \left(\varphi^{(\mu)}(u), \psi^{(\mu)}(u) \right) = \Phi\big(D^\mu \varphi(u), D^\mu \psi(u) \big) \qquad (28)$$
$$= \Phi\big(D^\mu \overline{\varphi}(v), D^\mu \overline{\psi}(v) \big) = \Phi \left(\overline{\varphi}^{(\mu)}(v), \overline{\psi}^{(\mu)}(v) \right) .$$

Hierbei ist jedoch für $m \geqq 1$, d. h. wenn eine *wirkliche Berührung* stattfinden soll, immer noch $u_0', v_0' \gtrless 0$ vorauszusetzen, weil wegen

22 | $$\varphi'(u) u' = \overline{\varphi}'(v) v' \,, \qquad \psi'(u) u' = \overline{\psi}'(v) v'$$

auch

$$\sqrt{\varphi'^2(u) + \psi'^2(u)}\, u' = \pm \sqrt{\overline{\varphi}'^2(v) + \overline{\psi}'^2(v)}\, v'$$

sein und daher

$$\frac{u'}{v'} = \pm \frac{\sqrt{\overline{\varphi}'^2(v) + \overline{\psi}'^2(v)}}{\sqrt{\varphi'^2(u) + \psi'^2(u)}}$$

lectures according to which the true order number of the contact may also be rational. But I shall always use the shorthand form here.

For the situation under consideration to arise, we must, as could be proved rigorously, be able to bring into agreement the derivatives of the coordinates up to the m th order for suitable independent variable, i. e., it must be possible to represent u and v as functions of an arbitrary parameter t so that for

$$t = t_0 \,, \qquad D^\mu u = \frac{d^\mu u}{dt^\mu} = u_0^{(\mu)} \,, \qquad D^\mu v = \frac{d^\mu v}{dt^\mu} = v_0^{(\mu)}$$

the relations

$$D^\mu \varphi(u) = D^\mu \overline{\varphi}(v) \,, \qquad D^\mu \psi(u) = D^\mu \overline{\psi}(v) \qquad (27)$$
$$(\mu = 0, 1, 2, \ldots m)$$

hold, or, what, by (2), amounts to the same, that the equations

$$R_\mu \left(\varphi^{(\nu)}(u_0), u_0^{(\nu)} \right) = R_\mu \left(\overline{\varphi}^{(\nu)}(v_0), v_0^{(\nu)} \right)$$
$$R_\mu \left(\psi^{(\nu)}(u_0), u_0^{(\nu)} \right) = R_\mu \left(\overline{\psi}^{(\nu)}(v_0), v_0^{(\nu)} \right)$$
$$(\nu = 0, 1, 2, \ldots m)$$

can be satisfied by suitable values of the $u_0^{(\mu)}, v_0^{(\mu)}$ $(\mu = 0, 1, \ldots m)$.[2]

If this is the case, then, at the position under consideration, we indeed have, by (21),

$$\begin{aligned} \Phi \left(\varphi^{(\mu)}(u), \psi^{(\mu)}(u) \right) &= \Phi \big(D^\mu \varphi(u), D^\mu \psi(u) \big) \qquad (28) \\ &= \Phi \big(D^\mu \overline{\varphi}(v), D^\mu \overline{\psi}(v) \big) = \Phi \left(\overline{\varphi}^{(\mu)}(v), \overline{\psi}^{(\mu)}(v) \right) . \end{aligned}$$

In this case, however, we still have to assume for $m \geqq 1$, i. e., if a *genuine contact* is to occur, that $u_0', v_0' \gtrless 0$, for on account of

$$\varphi'(u) u' = \overline{\varphi}'(v) v' \,, \qquad \psi'(u) u' = \overline{\psi}'(v) v'$$

we also must have

$$\sqrt{\varphi'^2(u) + \psi'^2(u)}\, u' = \pm \sqrt{\overline{\varphi}'^2(v) + \overline{\psi}'^2(v)}\, v' \,,$$

and hence

$$\frac{u'}{v'} = \pm \frac{\sqrt{\overline{\varphi}'^2(v) + \overline{\psi}'^2(v)}}{\sqrt{\varphi'^2(u) + \psi'^2(u)}}$$

[2] [[Zermelo erroneously writes "$v_0^{(\nu)}$" instead of "$v_0^{(\mu)}$".]]

einen bestimmten von 0 und ∞ verschiedenen Wert annehmen muss. Auch ist dabei zu unterscheiden zwischen „*gleichgerichteter* Berührung“, bei welcher u' und v', mithin auch $\varphi'(u)$ und $\overline{\varphi}'(v)$, $\psi'(u)$ und $\overline{\psi}'(v)$ gleiches Vorzeichen haben, von der *„entgegengesetzten“* Berührung, wo alle diese Grössen-Paare entgegengesetzte Vorzeichen besitzen; im ersten Falle werden die Coordinaten der beiden Curven in der Nähe des Berührungspunktes in gleichem, im anderen Falle in entgegengesetztem Sinne sich ändern, wenn jede Curve im Sinne der wachsenden Variablen u oder v beschrieben gedacht wird. Bei den späteren Anwendungen wird es immer so eingerichtet werden, dass nur „gleichgerichtete“ Berührungen in Betracht kommen.

Da die Variable t in der zuerst gegebenen Definition der Berührung selbst keine Rolle spielt, so müssen auch die Bedingungen (27) bestehen bleiben, wenn man darin diese Variable durch eine beliebige Function derselben ersetzt.

Sei etwa:

$$D^\mu_\vartheta \varphi(u) = D^\mu_\vartheta \overline{\varphi}(v)\ , \qquad D^\mu_\vartheta \psi(u) = D^\mu_\vartheta \overline{\psi}(v) \tag{27}$$
$$\left(\mu = 0, 1, 2, \ldots m; \qquad D^\mu_\vartheta = \frac{d^\mu}{d\vartheta^\mu}\right),$$

so wird mit Hilfe von (2) in der That auch:

$$D^\mu_t \varphi(u) = R_\mu\left(D^\nu_\vartheta \varphi(u), \vartheta^{(\nu)}\right) = R_\mu\left(D^\nu_\vartheta \overline{\varphi}(v), \vartheta^{(\nu)}\right) = D^\mu_t \overline{\varphi}(v)$$
$$D^\mu_t \psi(u) = R_\mu\left(D^\nu_\vartheta \psi(u), \vartheta^{(\nu)}\right) = R_\mu\left(D^\nu_\vartheta \overline{\psi}(v), \vartheta^{(\nu)}\right) = D^\mu_t \overline{\psi}(v)$$
$$(\mu = 0, 1, \ldots m)\ ,$$

23 | wo die $\vartheta^{(\mu)} = \dfrac{d^\mu \vartheta}{dt^\mu}$ an der betrachteten Stelle bis auf $\vartheta' \gtrless 0$ willkürlich gewählt sein können. Nun kann man die Variable t immer so wählen, dass u eine *beliebige* Function von t wird, also beliebig vorgeschriebene Werte der $u^{(\mu)} = \dfrac{d^\mu u}{dt^\mu}$ im Berührungspunkte annimmt; dann aber sind die anderen Grössen $v^{(\mu)} = \dfrac{d^\mu v}{dt^\mu}$ durch die Gleichungen (27) vollständig bestimmt. Diese lassen sich nämlich nach (2) in der Form schreiben:

$$\begin{aligned} R_\mu\left(\overline{\varphi}^{(\nu)}(v), v^{(\nu)}\right) &= D^\mu \varphi(u) \\ R_\mu\left(\overline{\psi}^{(\nu)}(v), v^{(\nu)}\right) &= D^\mu \psi(u) \end{aligned} \qquad (\mu = 0, 1, \ldots m) \tag{29}$$

und gestatten eine successive eindeutige Auflösung nach $v', v'', \ldots v^{(n)}$ durch Benutzung der oberen oder der unteren Gleichungen, je nachdem $\overline{\varphi}'(v)$ oder $\overline{\psi}'(v)$ von Null verschieden ist.

must take a particular value different from 0 and ∞. Also we have to distinguish here the "contact *of equal direction*", where u' and v', and hence also $\varphi'(u)$ and $\overline{\varphi}'(v)$, $\psi'(u)$ and $\overline{\psi}'(v)$ have the same sign, from the contact "*of opposite direction*", where all these pairs of quantities have opposite signs; in the first case, the coordinates of the two curves vary in the same sense near the point of contact, and in the other case, they vary in the opposite sense, assuming that we describe the curve in the sense of the increasing variables u or v. In later applications, we will always set up matters so that we only have to consider contacts "of equal direction".

Since the variable t is itself irrelevant in the first definition of contact, the conditions (27) must continue to hold when this variable is replaced by an arbitrary function.

For instance, let

$$D_\vartheta^\mu \varphi(u) = D_\vartheta^\mu \overline{\varphi}(v)\,, \qquad D_\vartheta^\mu \psi(u) = D_\vartheta^\mu \overline{\psi}(v) \tag{27}$$
$$\left(\mu = 0, 1, 2, \ldots m; \qquad D_\vartheta^\mu = \frac{d^\mu}{d\vartheta^\mu}\right),$$

then, by means of (2), we indeed also have

$$D_t^\mu \varphi(u) = R_\mu\left(D_\vartheta^\nu \varphi(u), \vartheta^{(\nu)}\right) = R_\mu\left(D_\vartheta^\nu \overline{\varphi}(v), \vartheta^{(\nu)}\right) = D_t^\mu \overline{\varphi}(v)$$
$$D_t^\mu \psi(u) = R_\mu\left(D_\vartheta^\nu \psi(u), \vartheta^{(\nu)}\right) = R_\mu\left(D_\vartheta^\nu \overline{\psi}(v), \vartheta^{(\nu)}\right) = D_t^\mu \overline{\psi}(v)$$
$$(\mu = 0, 1, \ldots m)\,,$$

where the $\vartheta^{(\mu)} = \dfrac{d^\mu \vartheta}{dt^\mu}$ can be chosen arbitrarily at the position under consideration but for $\vartheta' \gtrless 0$. Now we can always choose the variable t so that u becomes an *arbitrary* function of t, and hence assumes arbitrarily prescribed values of the $u^{(\mu)} = \dfrac{d^\mu u}{dt^\mu}$ at the point of contact; but the other quantities $v^{(\mu)} = \dfrac{d^\mu v}{dt^\mu}$ are then completely determined by the equations (27). For, by (2), they can be written in the form

$$\begin{aligned} R_\mu\left(\overline{\varphi}^{(\nu)}(v), v^{(\nu)}\right) &= D^\mu \varphi(u) \\ R_\mu\left(\overline{\psi}^{(\nu)}(v), v^{(\nu)}\right) &= D^\mu \psi(u) \end{aligned} \qquad (\mu = 0, 1, \ldots m) \tag{29}$$

and allow for a successive unique solution for $v', v'', \ldots v^{(n)}$ by use of the upper or lower equations, depending on whether $\overline{\varphi}'(v)$ or $\overline{\psi}'(v)$ is different from zero.

So kann man auch $t = u$, $u^{(\mu)} = e_{\mu,1}$ $(\mu = 1, 2, \dots m)$ vorschreiben und erhält dann die Bedingungen der Osculation in der Form:

$$\begin{gathered}\varphi^{(\mu)}(u) = D^\mu \overline{\varphi}(v) = R_\mu \left(\overline{\varphi}^{(\nu)}(v), v^{(\nu)}\right) \\ \psi^{(\mu)}(u) = D^\mu \overline{\psi}(v) = R_\mu \left(\overline{\psi}^{(\nu)}(v), v^{(\nu)}\right) \\ (\mu = 0, 1, 2, \dots m)\ ,\end{gathered} \tag{30}$$

wo die u, $v^{(\mu)}$ für den Berührungspunkt zu nehmen sind und wo im Falle „gleichgerichteter“ Berührung immer $v' > 0$ angenommen werden muss.

24 | Zweiter Abschnitt.

Definition und erste notwendige Bedingungen des Minimums.

Indem wir ein Maximum oder Minimum eines Integrales

$$J = \int_{t_1}^{t_2} F\, dt$$

suchen, können wir durch die Substitution $F \,||\, -F$ den Fall des Maximums auf den des *Minimums* zurückführen, brauchen uns also ohne Beschränkung der Allgemeinheit nur mit dem letzteren zu beschäftigen. Unsere Aufgabe ist demnach die folgende:

Ist

$$F\left(x^{(\mu)}, y^{(\mu)}\right) = F\left(x, x', \dots x^{(n)}\ ; y, y', \dots y^{(n)}\right)$$

eine vorgeschriebene analytische Function ihrer sämtlichen Argumente, die im ganzen betrachteten Bereiche den Charakter einer ganzen Function besitzt und den im ersten Abschnitte entwickelten Integrabilitätsbedingungen genügt, so suchen wir unter der *Gesamtheit A* aller Curven

$$x = \varphi(t)\ , \quad y = \psi(t)\ ,$$

die gewissen vorgeschriebenen Bedingungen genügen, eine solche besondere Curve a, welche einen *kleineren* Wert des auf der Curve zwischen bestimmten Grenzen erstreckten Integrales

$$J = \int_{t_1}^{t_2} F\left(x^{(\mu)}, y^{(\mu)}\right)\, dt \qquad \left(x^{(\mu)} = \frac{d^\mu x}{dt^\mu},\ y^{(\mu)} = \frac{d^\mu y}{dt^\mu}\right)$$

Thus we can also set $t = u$, $u^{(\mu)} = e_{\mu,1}$ $(\mu = 1, 2, \dots m)$, thereby obtaining the conditions of the osculation in the form

$$\begin{gathered} \varphi^{(\mu)}(u) = D^{\mu}\overline{\varphi}(v) = R_{\mu}\left(\overline{\varphi}^{(\nu)}(v), v^{(\nu)}\right) \\ \psi^{(\mu)}(u) = D^{\mu}\overline{\psi}(v) = R_{\mu}\left(\overline{\psi}^{(\nu)}(v), v^{(\nu)}\right) \\ (\mu = 0, 1, 2, \dots m)\ , \end{gathered} \tag{30}$$

where we must take the u, $v^{(\mu)}$ for the point of contact and always assume $v' > 0$ in the case of contact "of equal direction".

Second section.

Definition and first necessary conditions of the minimum.

When seeking to determine the maximum or minimum of an integral

$$J = \int_{t_1}^{t_2} F\, dt$$

we may reduce the case of the maximum to that of the *minimum* by the substitution $F|| - F$, and hence only need to consider the latter without loss of generality. Our task is therefore as follows:

Suppose that

$$F\left(x^{(\mu)}, y^{(\mu)}\right) = F\left(x, x', \dots x^{(n)}\ ; y, y', \dots y^{(n)}\right)$$

is a prescribed analytic function of all of its arguments that in the whole domain under consideration possesses the character of an entire function and satisfies the integrability conditions set out in the first section. Then among the *totality* A of all curves

$$x = \varphi(t)\ , \quad y = \psi(t)$$

satisfying certain prescribed conditions we seek to determine a special curve a furnishing a value for the integral

$$J = \int_{t_1}^{t_2} F\left(x^{(\mu)}, y^{(\mu)}\right)\, dt \qquad \left(x^{(\mu)} = \frac{d^{\mu}x}{dt^{\mu}},\ y^{(\mu)} = \frac{d^{\mu}y}{dt^{\mu}}\right)$$

25 liefert als alle *benachbarten* Curven $\overline{a}$ derselben Gesamtheit A, | so dass immer:

$$\Delta J = J(\overline{a}) - J(a) > 0 \tag{31}$$

wird.

Die unsere *Gesamtheit A* definierenden *Bedingungen* mögen sich für die nachstehenden Untersuchungen auf die folgenden beschränken.

1. *Grenzbedingungen.* Alle Curven A sollen an beiden Grenzen der Integration $t = t_1$ und $t = t_2$ in gegebenen Punkten 1 und 2 *zwei gegebene analytische Curven* d_1 und d_2

$$x = k_1(t)\ , \quad y = l_1(t)\ ; \quad x = k_2(t)\ , \quad y = l_2(t)$$

von $n-1$*ter Ordnung berühren,* so dass nach (30) die Gleichungen:

$$\begin{gathered} D^\mu\varphi(\vartheta_1) = k_1^{(\mu)}(t_1') = k_1^{(\mu)}\ , \quad D^\mu\psi(\vartheta_1) = l_1^{(\mu)}(t_1') = l_1^{(\mu)} \\ D^\mu\varphi(\vartheta_2) = k_2^{(\mu)}(t_2') = k_2^{(\mu)}\ , \quad D^\mu\psi(\vartheta_2) = l_2^{(\mu)}(t_2') = l_2^{(\mu)} \\ (\mu = 0, 1, 2, \ldots n-1) \end{gathered} \tag{32}$$

durch passende Werte der $\vartheta_1^{(\mu)} = \vartheta_1^{(\mu)}(t_1')$ und $\vartheta_2^{(\mu)} = \vartheta_2^{(\mu)}(t_2')$ ($\vartheta_1 = t_1$, $\vartheta_2 = t_2$) an den Stellen $t = t_1'$ und $t = t_2'$ befriedigt werden können; oder, was dasselbe ist, alle diese Curven sollen an den Grenzen *vorgeschriebene Werte der Osculations-Invarianten* $x, y, \alpha_1, \ldots \alpha_{n-1}$ besitzen.

2. *Stetigkeitsbedingungen.*

α) Damit die über die Curven A erstreckten Integrale immer einen bestimmten Sinn haben, wollen wir vorläufig voraussetzen, dass sich das ganze Intervall $t_1 \ldots t_2$ jedesmal *in eine endliche Anzahl von Teilen zerlegen* lässt, in deren jedem

$$x^{(\mu)} = \varphi^{(\mu)}(t)\ , \quad y^{(\mu)} = \psi^{(\mu)}(t) \quad (\mu = 0, 1, \ldots n) \tag{33}$$

eindeutige und stetige Functionen von t sind. Allerdings ist die Convergenz der Integrale auch unter allgemeineren Voraussetzungen möglich, worauf aber hier zur Vereinfachung der Untersuchung keine Rücksicht genommen werden soll.

26 | β) Ausserdem sollen die Functionen $\varphi(t)$ und $\psi(t)$ der Curven A *mit ihren r ersten Ableitungen* im ganzen Integrations-Intervall (34) *ausnahmslos stetig verlaufen*, wo $r \geqq 0$ vorläufig unbestimmt bleibt, und für $r \geqq 1$ der Bedingung genügen:

$$\varphi'^2(t) + \psi'^2(t) \geqq \gamma^2 > 0\ , \tag{35}$$

sodass „singuläre Punkte“ ($\varphi'(t) = 0, \psi'(t) = 0$) ausgeschlossen sind. Es wird sich nämlich (vergl. Satz IV) zeigen, dass ohne eine solche Voraussetzung ein Minimum überhaupt unmöglich wäre.

taken along the curve between certain limits that is *smaller* than that of all *neighboring* curves $\overline{a}$ of the same totality A so that we always have

$$\Delta J = J(\overline{a}) - J(a) > 0 \,. \tag{31}$$

In the subsequent investigations, we will restrict the *conditions* defining our *totality* A as follows:

1. *Limit conditions.* All curves A shall have *contact* of $n-1$ *th order* with *two given analytic curves* d_1 and d_2

$$x = k_1(t)\,, \quad y = l_1(t)\,; \quad x = k_2(t)\,, \quad y = l_2(t)$$

at both limits of the integration $t = t_1$ and $t = t_2$ at given points 1 and 2 so that, by (30), the equations

$$\begin{aligned} D^\mu\varphi(\vartheta_1) = k_1^{(\mu)}(t_1') = k_1^{(\mu)}\,, \quad D^\mu\psi(\vartheta_1) = l_1^{(\mu)}(t_1') = l_1^{(\mu)} \\ D^\mu\varphi(\vartheta_2) = k_2^{(\mu)}(t_2') = k_2^{(\mu)}\,, \quad D^\mu\psi(\vartheta_2) = l_2^{(\mu)}(t_2') = l_2^{(\mu)} \\ (\mu = 0, 1, 2, \ldots n-1) \end{aligned} \tag{32}$$

can be satisfied by appropriate values of $\vartheta_1^{(\mu)} = \vartheta_1^{(\mu)}(t_1')$ and $\vartheta_2^{(\mu)} = \vartheta_2^{(\mu)}(t_2')$ $(\vartheta_1 = t_1,\ \vartheta_2 = t_2)$ at the positions $t = t_1'$ and $t = t_2'$; or, what amounts to the same, all these curves shall possess *prescribed values of the osculation invariants* $x, y, \alpha_1, \ldots \alpha_{n-1}$ at the limits.

2. *Continuity conditions.*

α) In order to ensure that the integrals taken along the curves A always have a definite meaning, we shall, for the present, assume that the entire interval $t_1 \ldots t_2$ is always capable of decomposition into *a finite number of parts* in each of which

$$x^{(\mu)} = \varphi^{(\mu)}(t)\,, \quad y^{(\mu)} = \psi^{(\mu)}(t) \quad (\mu = 0, 1, \ldots n) \tag{33}$$

are single-valued and continuous functions of t. But in order to simplify our present investigation we shall not consider the possibility of the convergence of the integrals also under more general assumptions.

β) Furthermore, the functions $\varphi(t)$ and $\psi(t)$ of the curves A, *together with their first* r *derivatives*, shall be continuous on the entire interval of integration *without exception*, where $r \geqq 0$ remains indefinite for the present, and, for $r \geqq 1$, satisfy the equation (34)

$$\varphi'^2(t) + \psi'^2(t) \geqq \gamma^2 > 0\,, \tag{35}$$

so that "singular points" $(\varphi'(t) = 0, \psi'(t) = 0)$ are excluded. For we shall see (cf. Theorem IV) that a minimum would be entirely impossible without an assumption of this kind.

Streng genommen brauchten die $x^{(\mu)}$, $y^{(\mu)}$ $(\mu \leqq r)$ nicht selbst immer stetig zu bleiben, sondern nur die „Osculations-Invarianten“ $\Phi\left(x^{(\mu)}, y^{(\mu)}\right)$ bis zur r ten Ordnung, also namentlich die $x, y, \alpha_1, \ldots \alpha_r$, die nach (25) alle übrigen bestimmen. Denn durch Veränderung der Differentiations-Variablen, oder, was dasselbe ist, der „Darstellung“ der Curve für einzelne Teile des Intervalls, wodurch etwa die $\varphi^{(\mu)}(\vartheta)$, $\psi^{(\mu)}(\vartheta)$ in $D^\mu\varphi(\vartheta)$, $D^\mu\psi(\vartheta)$ übergehen, würde die Stetigkeit der $x^{(\mu)}, y^{(\mu)}$, wenn sie einmal bestände, beliebig aufgehoben und wieder eingeführt werden können. Es genügte also, die Stetigkeit der $\varphi^{(\mu)}(t)$, $\psi^{(\mu)}(t)$ für die einzelnen Teile der Curve vorauszusetzen, wenn nur ausserdem an den Übergangsstellen immer Berührungen r ter Ordnung angenommen werden. Dieser Fall aber wird sich durch eine geeignete Darstellung der Curve immer vermeiden lassen, und so soll denn auch den folgenden Untersuchungen die Bedingung (34) in voller Strenge zu Grunde gelegt werden.

Alle diesen Bedingungen (32)–(35) genügenden Curven A wollen wir als „*erlaubte Curven*“, den Übergang von einer zur anderen als eine „erlaubte Variation“ bezeichnen. Anderweitige Beschränkungen der Curven A sollen hier nicht betrachtet, überall soll „freie Variation“ vorausgesetzt werden.

Einer sorgfältigen Untersuchung bedarf jetzt noch der Begriff der „*benachbarten*“ *Curven* $(\overline{a})$, dessen Auffassung für das ganze Problem von entscheidender Bedeutung ist.

Herr Prof. *Weierstrass* betrachtet als zu a „benachbarte“ Curven $\overline{a}$ alle diejenigen, welche ganz innerhalb eines gewissen, durch eine Begrenzung C
27 eingeschlossenen Flächenstreifens ver-|laufen, wenn C das Curvenstück a vollständig umgiebt und ihm nirgend unendlich nahe kommt. Ist δ der kürzeste Abstand zwischen C und a, so liegen im Innern des Streifens alle Punkte, die von a Entfernungen $< \delta$ besitzen, und umgekehrt wird eine zweite Begrenzungscurve C' derselben Art, für welche δ den *grössten* Abstand von a darstellt, nur solche Punkte einschliessen, die a näher kommen als auf die Entfernung δ. Dieses zweite, durch C' eingeschlossene Gebiet wird aber für die Frage nach dem Minimum dieselbe Rolle spielen wie das grössere C, so dass man auch alle die Curven A als „benachbarte“ betrachten kann, deren

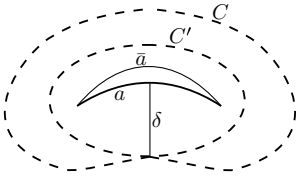


Fig. 1.

sämtliche Punkte von a Abstände $< \delta$ besitzen, oder, analytisch ausgedrückt, unter den A alle solchen Curven

$$x = \overline{\varphi}(\lambda)\ , \quad y = \overline{\varphi}(\lambda)\ ,$$

Strictly speaking, the $x^{(\mu)}$, $y^{(\mu)}$ $(\mu \leqq r)$ would not always have to be continuous themselves, but only the "osculation invariants" $\Phi\left(x^{(\mu)}, y^{(\mu)}\right)$ up to the r th order, and hence in particular $x, y, \alpha_1, \ldots \alpha_r$, which, by (25), determine all the others. For by changing the variable of differentiation, or, what amounts to the same, the "representation" of the curve for individual parts of the interval, whereby, say, the $\varphi^{(\mu)}(\vartheta)$, $\psi^{(\mu)}(\vartheta)$ are transformed into $D^\mu\varphi(\vartheta)$, $D^\mu\psi(\vartheta)$, we would be able to remove and reintroduce at will the continuity of the $x^{(\mu)}, y^{(\mu)}$ once it obtains. It is therefore sufficient to assume the continuity of the $\varphi^{(\mu)}(t)$, $\psi^{(\mu)}(t)$ for the individual parts of the curve, provided only that we also assume contacts of the r th order at the corners. But a suitable representation of the curve always helps avoid this scenario. Therefore, the rigorous application of the condition (34) shall underly the investigations to follow.

All curves A satisfying the conditions (32)–(35) will be called "*admissible curves*", and the transition from one to the other an "admissible variation". We will not consider any further restrictions on the curves A and always assume "free variation".

The concept of "*neighboring*" *curve* $(\overline{a})$ now still requires thorough investigation on account of its crucial significance for the entire problem.

Prof. *Weierstrass* considers all those curves $\overline{a}$ as curves "neighboring" a that run entirely within a certain strip of the plane enclosed by a boundary C, if C completely surrounds the curve segment a without getting infinitely close to it at any point. If δ is the shortest distance between C and a, then all those points lie in the interior of the strip whose distance from a is $< \delta$, and, conversely, a second boundary curve C' of the same kind for which δ is the *greatest* distance from a encloses only points that get closer to a than the distance δ. But this second region, which is enclosed by C', will play the same role in the question of the minimum as the greater one, C, so that we can consider all those curves A "neighboring" curves each of whose points lie

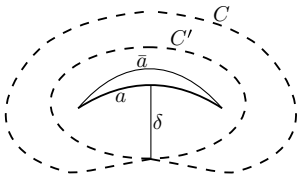


Fig. 1.

at a distance $< \delta$ from a, or, expressed in analytic terms, among the A all such curves

$$x = \overline{\varphi}(\lambda)\ , \quad y = \overline{\varphi}(\lambda)\ ,$$

für welche zu jedem Werte λ zwischen λ_1 und λ_2 stets ein $t = \varkappa$ zwischen t_1 und t_2 so bestimmt werden kann, dass

$$\sqrt{(\overline{\varphi}(\lambda) - \varphi(\varkappa))^2 + (\overline{\psi}(\lambda) - \psi(\varkappa))^2} < \delta \, ,$$

also auch

$$|\overline{\varphi}(\lambda) - \varphi(\varkappa)| < g \, , \quad \left|\overline{\psi}(\lambda) - \psi(\varkappa)\right| < g$$

sicher für $g \geqq \delta$. Bestehen aber umgekehrt diese letzten Ungleichheiten für $g \leqq \frac{\delta}{\sqrt{2}}$, so folgt daraus wieder die vorhergehende.

Ein Minimum wird daher für a dann und nur dann stattfinden, wenn alle diejenigen erlaubten Curven ein grösseres Intervall als a liefern, welche für irgend ein constantes positives g den beiden Ungleichheiten genügen.

Die so entwickelte Definition ist aber für unseren allgemeineren Fall nicht
28 ausreichend, da sie die Existenz eines Mini- | mums allzusehr beschränken, ja vielleicht überhaupt unmöglich machen würde. Wir müssen vielmehr voraussetzen, dass wenigstens in den Fällen $n > 1$ nicht nur die *Punkte* der Curven $\overline{a}$ solchen von a hinreichend nahe kommen, sondern auch die Tangentenrichtungen, die Krümmungen u. s. w., allgemein die Osculations-Invarianten bis zu einer gewissen Ordnung m. Es sollen also als „benachbart“ angesehen werden alle solchen Curven $x = \overline{\varphi}(\lambda)$, $y = \overline{\psi}(\lambda)$, für welche zu jedem λ des ganzen Intervalls immer ein $\varkappa$ zwischen t_1 und t_2 so bestimmt werden kann, dass eine Anzahl Ungleichheitsbedingungen befriedigt werden von der Form:

$$\left|\Phi\left(\overline{\varphi}^{(\mu)}(\lambda), \overline{\psi}^{(\mu)}(\lambda)\right) - \Phi\left(\varphi^{(\mu)}(\varkappa), \psi^{(\mu)}(\varkappa)\right)\right| < \delta \, , \tag{36}$$

wo

$$\Phi\left(x^{(\mu)}, y^{(\mu)}\right) = \Phi\left(x, x', \ldots x^{(m)} \, ; y, y', \ldots y^{(m)}\right)$$

eine Function von der Eigenschaft (21) und δ eine von λ unabhängige Grösse ist.

Diese Bedingungen lassen sich wieder auf die zwischen den *Coordinaten selbst* und ihren Ableitungen bestehenden Ungleichheiten zurückführen:

$$\left|D^\mu \overline{\varphi}(\lambda) - \varphi^{(\mu)}(\varkappa)\right| < g_\mu \, , \quad \left|D^\mu \overline{\psi}(\lambda) - \psi^{(\mu)}(\varkappa)\right| < g_\mu \tag{37}$$

$$(\mu = 0, 1, 2, \ldots m),$$

wo für ein jedes λ ausser $\varkappa$ noch die Grössen $\lambda' > 0, \lambda'', \ldots \lambda^{(m)}$, die Werte der ersten m Ableitungen einer Function $\lambda(t) = \lambda(t; \lambda)$ für $t = \varkappa$, *beliebig* angenommen werden können. Dabei brauchen die zu verschiedenen λ gehörigen $\varkappa, \lambda^{(\mu)}$ von vornherein in keinen Beziehungen zu einander zu stehen, während die positiven Grössen $g, g_1, \ldots g_m$ von λ unabhängige bestimmte Werte haben müssen.

Die Willkürlichkeit der Parameterdarstellung kommt für die Curven $\overline{\varphi}, \overline{\psi}$ hier nicht in Betracht, da jede Veränderung der Darstellung offenbar nur die

are considered "neighboring" curves for which it is always possible to determine for every value λ between λ_1 and λ_2 a $t = \varkappa$ between t_1 and t_2 so that

$$\sqrt{(\overline{\varphi}(\lambda) - \varphi(\varkappa))^2 + (\overline{\psi}(\lambda) - \psi(\varkappa))^2} < \delta \, ,$$

and hence also

$$|\overline{\varphi}(\lambda) - \varphi(\varkappa)| < g \, , \quad \left|\overline{\psi}(\lambda) - \psi(\varkappa)\right| < g$$

certainly for $g \geqq \delta$. But if, conversely, these latter inequalities hold for $g \leqq \frac{\delta}{\sqrt{2}}$, then, in turn, the former follows from this.

Hence a has a minimum if and only if all those admissible curves furnish a greater interval than a that satisfy both inequalities for some positive constant g.

But the definition thus developed is not sufficient for our more general case as it would greatly restrict the existence of a minimum, and perhaps even make it impossible. Rather, we have to assume that at least in the cases $n > 1$ not only the *points* of the curves $\overline{a}$ get sufficiently close to those of a but also the directions of the tangents, the curvatures e. t. c., generally, the osculation invariants up to a certain order m. Hence all those curves $x = \overline{\varphi}(\lambda)$, $y = \overline{\psi}(\lambda)$ are to be considered "neighboring" curves for which it is always possible to determine for every λ of the entire interval a $\varkappa$ between t_1 and t_2 so that several inequality conditions of the form

$$\left|\Phi\left(\overline{\varphi}^{(\mu)}(\lambda), \overline{\psi}^{(\mu)}(\lambda)\right) - \Phi\left(\varphi^{(\mu)}(\varkappa), \psi^{(\mu)}(\varkappa)\right)\right| < \delta \, , \tag{36}$$

are satisfied, where

$$\Phi\left(x^{(\mu)}, y^{(\mu)}\right) = \Phi\left(x, x', \ldots x^{(m)} \, ; y, y', \ldots y^{(m)}\right)$$

is a function with the property (21) and δ a quantity independent of λ.

These conditions can again be reduced to the inequalities obtaining between the *coordinates themselves* and their derivatives:

$$\left|D^\mu \overline{\varphi}(\lambda) - \varphi^{(\mu)}(\varkappa)\right| < g_\mu \, , \quad \left|D^\mu \overline{\psi}(\lambda) - \psi^{(\mu)}(\varkappa)\right| < g_\mu \tag{37}$$

$$(\mu = 0, 1, 2, \ldots m),$$

where for each λ but $\varkappa$ we may assume the quantities $\lambda' > 0, \lambda'', \ldots \lambda^{(m)}$, the values of the first m derivatives of a function $\lambda(t) = \lambda(t; \lambda)$ for $t = \varkappa$, *arbitrarily*. In this case, the $\varkappa, \lambda^{(\mu)}$ belonging to different λ do not need to stand in any relation to one another from the outset, whereas the positive quantities $g, g_1, \ldots g_m$ must have values independent of λ.

The arbitrariness of the parametric representation is out of the question here as far as the curves $\overline{\varphi}$, $\overline{\psi}$ are concerned since every change in the rep-

$\lambda', \lambda'', \ldots \lambda^{(m)}$ ändern würde. Dagegen muss immer eine *bestimmte* Darstellung $x = \varphi(t)$, $y = \psi(t)$ von a zu Grunde gelegt werden, ihre Veränderung müsste unter Umständen eine Veränderung der g_μ zur Folge haben, wenn die Beziehungen (37) fortbestehen sollen.

29 | Wie leicht zu zeigen ist, lassen sich die g_μ immer so klein angeben, dass für alle den Bedingungen (37) genügenden Curven $\overline{a}$ auch Bedingungen der Form (36) in beliebiger Anzahl und für beliebig kleine δ befriedigt werden können, und umgekehrt giebt es immer eine Anzahl von Ungleichheiten (36) mit so kleinen Werten der δ, dass auch die (37) gelten müssen für passende $\varkappa, \lambda^{(\mu)}$ und beliebig klein vorgeschriebene g_μ. Doch wird dabei vorausgesetzt, dass jedes $\Phi\left(\varphi^{(\mu)}(t), \psi^{(\mu)}(t)\right)$ auf dem ganzen Curvenstücke a überall endlich bleibt.

Von a soll also behauptet werden, es liefere ein Minimum in einer „Nachbarschaft m ter Ordnung", falls sich für irgend eine Darstellung: $x = \varphi(t)$, $y = \psi(t)$ von a positive Grössen $g, g_1, \ldots g_m$ so angeben lassen, dass alle „erlaubten" Curven $\overline{a}$

$$x = \overline{\varphi}(\lambda),\ y = \overline{\psi}(\lambda)\ ,$$

die für alle in Betracht kommenden λ und für jedesmal passend bestimmte $\varkappa, \lambda^{(\mu)}$ den Bedingungen (37) genügen, Inteqrale liefern

$$J(\overline{a}) > J(a).$$

Die „Ordnung der Nachbarschaft" m wird später auf die Fälle $m = n - 1$ und $m = n$ beschränkt werden, welche allein für unsere Untersuchung Interesse bieten.

Man erkennt, dass die hier gegebene Definition der „benachbarten Curven" sich ebenso wie die Weierstrass'sche nur auf ihr Verhalten in jedem einzelnen Punkte bezieht, nicht aber auf ihren allgemeinen Verlauf oder eine specielle Darstellungsform.

Um die Analogie noch deutlicher hervorzuheben, kann man *die Gesamtheit aller Eigenschaften, die eine Curve in einem bestimmten Punkte mit allen sie von m ter Ordnung berührenden Curven gemeinsam hat*, also die *Gesamtheit der Osculations-Invarianten $\Phi\left(x^{(\mu)}, y^{(\mu)}\right)$* bis zur m ten Ordnung, in einen einzigen Begriff, den eines „*Curvenelementes m ter Ordnung*", zusammenfassen, der sich für $m = 0$ auf den eines einfachen Punktes reduciert und somit als dessen Erweiterung aufgefasst werden kann. Bestimmt wird ein solches „Element" nach (24) und (25) durch die Werte der $x, y, \dfrac{dy}{dx}, \ldots \dfrac{d^m y}{dx^m}$ oder

30 der $x, y, \alpha_1, \ldots \alpha_m$, also durch | $m + 2$ unabhängige Variable, durch die $x^{(\mu)}$, $y^{(\mu)}$ $(\mu = 0, 1, \ldots m)$ selbst aber nur insofern, als dabei von allen nur durch die willkürliche Wahl der Variablen t entstehenden Verschiedenheiten abgesehen werden muss. Es müssen nämlich zwei Elemente $\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right)_m$ und $\left(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\right)_m$ für beliebige Werte der $\vartheta', \ldots \vartheta^{(m)}$ als *identisch* betrachtet werden, da sie nach (21) in Bezug auf sämtliche Osculations-Invarianten bis zur Ordnung m übereinstimmen. Somit haben zwei Curven $x = \varphi(t)$,

resentation would obviously only change the $\lambda', \lambda'', \ldots \lambda^{(m)}$. By contrast, a *particular* representation $x = \varphi(t)$, $y = \psi(t)$ of a must always be taken for granted, and its change would possibly have to lead to a change of the g_μ, assuming the relations (37) are to continue to obtain.

As is readily shown, it is always possible to choose sufficiently small g_μ so that for all curves $\overline{a}$ satisfying the conditions (37) an arbitrary number of conditions of the form (36) for arbitrarily small δ can be satisfied as well. Conversely, there is always a number of inequalities (36) whose values of the δ are so small that the (37), too, must hold for appropriate $\varkappa, \lambda^{(\mu)}$ and arbitrarily small prescribed g_μ. But it is taken for granted here that every $\Phi\left(\varphi^{(\mu)}(t), \psi^{(\mu)}(t)\right)$ remains finite everywhere on the entire curve segment a.

What we wish to assert of a is that it furnishes a minimum in a "neighborhood of m th order", if, for some representation: $x = \varphi(t)$, $y = \psi(t)$ of a, positive quantities $g, g_1, \ldots g_m$ can be specified so that all "admissible" curves $\overline{a}$

$$x = \overline{\varphi}(\lambda),\ y = \overline{\psi}(\lambda)\ ,$$

that satisfy the conditions (37) *for all relevant λ and for always suitably determined $\varkappa, \lambda^{(\mu)}$ furnish integrals*

$$J(\overline{a}) > J(a).$$

Later, we will restrict the "order of the neighborhood" m to the cases $m = n-1$ and $m = n$, which are the only ones of interest in our investigations.

We can see that, like Weierstrass's, the definition of the "neighboring curves" given here only refers to their behavior at each individual point but not to their general course or a special form of representation.

In order to bring out the analogy even more clearly we may capture *the totality of all properties shared by some curve at a certain point with all curves having m th order contact with it*, that is, the *totality of osculation invariants* $\Phi\left(x^{(\mu)}, y^{(\mu)}\right)$ up to the mth order, in a single concept, that of a "*curve element of m th order*", which is reduced to that of a simple point when $m = 0$, and hence may be considered as its extension. According to (24) and (25), such an "element" is determined by the values of the $x, y, \dfrac{dy}{dx}, \ldots \dfrac{d^m y}{dx^m}$ or of the $x, y, \alpha_1, \ldots \alpha_m$, and hence by $m + 2$ independent variables, by the $x^{(\mu)}$, $y^{(\mu)}$ $(\mu = 0, 1, \ldots m)$ themselves but only insofar as differences arising solely from the arbitrary choice of the variables t have to be left out of consideration. For two elements $\left(\varphi^{(\mu)}(\vartheta), \psi^{(\mu)}(\vartheta)\right)_m$ and $\left(D^\mu\varphi(\vartheta), D^\mu\psi(\vartheta)\right)_m$ for arbitrary values of the ϑ', $\ldots \vartheta^{(m)}$ must be considered *identical*, since, by (21), they agree with respect to all osculation invariants up to order m. Hence, two

$y = \psi(t)$ und $x = \overline{\varphi}(t)$, $y = \overline{\psi}(t)$ dann und nur dann ein „Element m ter Ordnung“ gemeinsam, wenn sie eine *Berührung* m ter Ordnung mit einander eingehen, wo nach (27)

$$D^\mu \varphi(u) = D^\mu \overline{\varphi}(v) \ , \quad D^\mu \psi(u) = D^\mu \overline{\psi}(v) \quad (\mu = 0, 1, \ldots m) \ .$$

Nun definiert (36) und in anderer Form (37) für jeden Wert von $\varkappa$ eine gewisse „*Umgebung*“ des Elementes $\left(\varphi^{(\mu)}(\varkappa), \psi^{(\mu)}(\varkappa)\right)_m$ und daher für variable $\varkappa$ zwischen t_1 und t_2 eine „*Nachbarschaft m ter Ordnung*“ des Curvenstückes a. Solche „Umgebungen“ und solche „Nachbarschaften“ sind, wie leicht zu zeigen ist, *continuierliche Bereiche* in der $m + 2$ fachen Mannigfaltigkeit der Curvenelemente, nämlich Gesamtheiten von der Beschaffenheit, dass jedes ihrer im Inneren gelegenen Elemente stets in eine solche (nicht unendlich kleine) Umgebung eingeschlossen werden kann, welche selbst wieder ganz der Gesamtheit angehört. Eine nach den Definitionen (36) und (37) zu a „benachbarte Curve“ $\overline{a}$ ist also eine solche, deren sämtliche Elemente einem solchen Nachbarschaftsbereiche von a angehören, jedes nämlich immer einer gewissen „Umgebung“ eines Elementes von a.

„Begrenzt“ werden solche „Bereiche“ oder „Gebiete m ter Ordnung“ nach (36) mit Hilfe von (24) und (25) durch Gebilde von höchstens m ter Ordnung, nämlich durch Gleichungen der Form:

$$\Phi\left(x^{(\mu)}, y^{(\mu)}\right)_m = \Phi_1(x, y; \alpha_\mu)_m = \Phi_2\left(x, y, \frac{d^\mu y}{dx^\mu}\right)_m = 0 \ ,$$

also durch Differentialgleichungen m ter Ordnung, deren Integrale in Gestalt von m fach unendlichen Curvenscharen eine geometrische Veranschaulichung
31 der „Gebiete“ ermöglichen und der im Anfang | erwähnten „Begrenzungscurve“ C eines Flächenstreifens entsprechen. Doch soll auf diese Betrachtungen vorläufig nicht weiter eingegangen, sondern zur eigentlichen Aufgabe der Variationsrechnung zurückgekehrt werden.

Es sei nur noch bemerkt, dass die hier gegebene Definition der Nachbarschaft am nächsten kommt der *Scheeffer*'schen („Über die Bedeutung der Begriffe: Maximum und Minimum in der Variationsrechnung“, Math. Ann. XXVI), in welche sie übergeht, wenn in (37) $\lambda = \varkappa = x = t$ und $m = n$ angenommen wird. Dann reducieren sich nämlich diese Bedingungen auf die folgenden:

$$\left|\overline{\psi}^{(\mu)}(x) - \psi^{(\mu)}(x)\right| < g_\mu \qquad (\mu = 0, 1, \ldots n) \ ,$$

wo durch $y = \psi(x)$ und $y = \overline{\psi}(x)$ die Curven a und $\overline{a}$ dargestellt werden.

curves $x = \varphi(t)$, $y = \psi(t)$ and $x = \overline{\varphi}(t)$, $y = \overline{\psi}(t)$ have a common "element of the m th order" if and only if they make m th order *contact* with one another, where, by (27),

$$D^\mu \varphi(u) = D^\mu \overline{\varphi}(v) \ , \quad D^\mu \psi(u) = D^\mu \overline{\psi}(v) \quad (\mu = 0, 1, \ldots m) \ .$$

Now (36), and also (37), albeit in a different form, defines for every value of $\varkappa$ a certain "*vicinity*" of the element $\left(\varphi^{(\mu)}(\varkappa), \psi^{(\mu)}(\varkappa)\right)_m$, and hence for variable $\varkappa$ between t_1 and t_2 a "*neighborhood of m th order*" of the curve segment a. Such "environments" and "neighborhoods" are, as is readily shown, *continuous domains* in the $(m+2)$-manifold of the curve elements, namely totalities constituted so that each of its interior elements can always be enclosed in a (not infinitely small) environment that, in turn, entirely belongs to the totality. Hence, a curve $\overline{a}$ that, according to the definitions (36) and (37), is a curve "neighboring" a is one whose elements all belong to such a neighborhood domain of a, that is, each of them always to a certain "vicinity" of an element of a.

Such "domains", or "regions of the m th order" are "delimited", according to (36) by means of (24) and (25), by structures of at most m th order, namely by equations of the form

$$\Phi\left(x^{(\mu)}, y^{(\mu)}\right)_m = \Phi_1(x, y; \alpha_\mu)_m = \Phi_2\left(x, y, \frac{d^\mu y}{dx^\mu}\right)_m = 0 \ ,$$

hence by differential equations of m th order whose integrals in the form of m-fold infinite families of curves admit a geometric realization of the "regions" and correspond to the "boundary curve" C of a strip of the plane mentioned at the beginning. But we will now return to the real task of the calculus of variations without further elaborating on these considerations for the present.

Let me add only that the definition of neighborhood given here comes closest to *Scheeffer*'s definition (*Scheefer 1886*), into which it is transformed if we set $\lambda = \varkappa = x = t$ and $m = n$ in (37). For these conditions are then reduced to the following ones:

$$\left|\overline{\psi}^{(\mu)}(x) - \psi^{(\mu)}(x)\right| < g_\mu \qquad (\mu = 0, 1, \ldots n) \ ,$$

where $y = \psi(x)$ and $y = \overline{\psi}(x)$ describe the curves a and $\overline{a}$.

Untersuchung der ersten Variation.

Zu den Curven $\overline{a}$, für welche im Fall, dass a ein Minimum liefert, $J(\overline{a}) > J(a)$ sein muss, gehören alle Curven der Form:

$$\begin{aligned} x &= \overline{\varphi}(t) = \varphi(t) + \varepsilon\xi(t) \\ y &= \overline{\psi}(t) = \psi(t) + \varepsilon\eta(t) \end{aligned} \tag{38}$$

mit den folgenden Eigenschaften:

1. An den Grenzen soll

$$\xi^{(\mu)}(t_1) = 0\,, \quad \eta^{(\mu)}(t_1) = 0\,; \quad \xi^{(\mu)}(t_2) = 0\,, \quad \eta^{(\mu)}(t_2) = 0 \qquad (\mu = 0, 1, \ldots n-1)\,, \tag{39}$$

wodurch auch für $\overline{\varphi}(t), \overline{\psi}(t)$ die Bedingungen (32) befriedigt werden.

2. Im ganzen Intervall $t_1 \leqq t \leqq t_2$ sollen $\xi(t)$ und $\eta(t)$, mithin auch $\overline{\varphi}(t)$ und $\overline{\psi}(t)$ den Stetigkeitsbedingungen (33) und (34) genügen.

Ausserdem soll $|\varepsilon|$ so klein gewählt werden, dass für $t_1 \leqq t \leqq t_2$ gemäss (35)

3. $(\varphi'(t) + \varepsilon\xi'(t))^2 + (\psi'(t) + \varepsilon\eta'(t))^2 > \gamma' > 0$,

32 | was, da schon a (35) genügt:

$$\varphi'^2(t) + \psi'^2(t) > \gamma^2 > 0$$

und immer $\varepsilon^2(\xi'^2(t) + \eta'^2(t)) \geqq 0$ ist, sicher erreicht werden kann durch:

$$\begin{aligned} &\left(\varphi'^2(t) + \psi'^2(t)\right)^2 + 2\varepsilon\left(\varphi'(t)\xi'(t) + \psi'(t)\eta'(t)\right) \\ &\quad > \gamma^2 - 2|\varepsilon|h = \gamma' > 0\,, \quad \text{also} \quad |\varepsilon| < \frac{\gamma^2}{2h}\,, \end{aligned}$$

wenn $|\varphi'(t)\xi'(t) + \psi'(t)\eta'(t)| \leqq h$ angenommen wird.

Endlich soll, damit $\overline{a}$ auch zu den „benachbarten Curven“ von a gehört, gleichfalls für das ganze Intervall $t_1 \leqq t \leqq t_2$:

4. $$\left|\varepsilon\xi^{(\mu)}(t)\right| < g_\mu\,, \qquad \left|\varepsilon\eta^{(\mu)}(t)\right| < g_\mu \qquad (\mu = 0, 1, 2, \ldots m)\,,$$

wodurch (37) sicher befriedigt wird mit

$$\lambda = \varkappa = t, \qquad \lambda^{(\mu)} = e_{\mu,1} \qquad (\mu = 1, 2, \ldots)\,.$$

Dazu aber braucht nur:

$$|\varepsilon| \leqq \frac{g_\mu}{h_\mu} \qquad (\mu = 0, 1, \ldots m)$$

angenommen zu werden, wenn

$$\left|\xi^{(\mu)}(t)\right| < h_\mu\,, \quad \left|\eta^{(\mu)}(t)\right| < h_\mu \quad (t_1 \leqq t \leqq t_2\,; \quad \mu = 0, 1, \ldots m)$$

Investigation of the first variation.

Among the curves $\overline{a}$ for which it must be the case that $J(\overline{a}) > J(a)$ whenever a furnishes a minimum are all curves of the form

$$\begin{aligned} x &= \overline{\varphi}(t) = \varphi(t) + \varepsilon\xi(t) \\ y &= \overline{\psi}(t) = \psi(t) + \varepsilon\eta(t) \end{aligned} \tag{38}$$

with the following properties:

1. At the limits we shall have

$$\xi^{(\mu)}(t_1) = 0\,, \quad \eta^{(\mu)}(t_1) = 0\,; \quad \xi^{(\mu)}(t_2) = 0\,, \quad \eta^{(\mu)}(t_2) = 0 \qquad (\mu = 0, 1, \ldots n-1)\,, \tag{39}$$

whereby the conditions (32) are satisfied also for $\overline{\varphi}(t), \overline{\psi}(t)$.

2. Both $\xi(t)$ and $\eta(t)$, and hence also $\overline{\varphi}(t)$ and $\overline{\psi}(t)$, shall satisfy the continuity conditions (33) and (34) in the entire interval $t_1 \leqq t \leqq t_2$.

Furthermore, $|\varepsilon|$ shall be chosen so small that, for $t_1 \leqq t \leqq t_2$, in accordance with (35),

3. $(\varphi'(t) + \varepsilon\xi'(t))^2 + (\psi'(t) + \varepsilon\eta'(t))^2 > \gamma' > 0$,

which, since a already satisfies (35):

$$\varphi'^2(t) + \psi'^2(t) > \gamma^2 > 0$$

and always $\varepsilon^2(\xi'^2(t) + \eta'^2(t)) \geqq 0$, can certainly be attained by means of

$$\begin{aligned} &\left(\varphi'^2(t) + \psi'^2(t)\right)^2 + 2\varepsilon\left(\varphi'(t)\xi'(t) + \psi'(t)\eta'(t)\right) \\ &\quad > \gamma^2 - 2|\varepsilon|h = \gamma' > 0\,, \quad \text{hence} \quad |\varepsilon| < \frac{\gamma^2}{2h}\,, \end{aligned}$$

provided we assume that $|\varphi'(t)\xi'(t) + \psi'(t)\eta'(t)| \leqq h$.

Finally, in order to ensure that $\overline{a}$ also belongs to the "neighboring curves" of a, we also assume that, for the entire interval $t_1 \leqq t \leqq t_2$,

4. $$\left|\varepsilon\xi^{(\mu)}(t)\right| < g_\mu\,, \qquad \left|\varepsilon\eta^{(\mu)}(t)\right| < g_\mu \qquad (\mu = 0, 1, 2, \ldots m)\,,$$

whereby (37) is certainly satisfied when

$$\lambda = \varkappa = t, \qquad \lambda^{(\mu)} = e_{\mu,1} \qquad (\mu = 1, 2, \ldots)\,.$$

For this to be the case, it suffices to assume that

$$|\varepsilon| \leqq \frac{g_\mu}{h_\mu} \qquad (\mu = 0, 1, \ldots m)\,,$$

when

$$\left|\xi^{(\mu)}(t)\right| < h_\mu\,, \quad \left|\eta^{(\mu)}(t)\right| < h_\mu \quad (t_1 \leqq t \leqq t_2\,; \quad \mu = 0, 1, \ldots m)$$

ihrer Endlichkeit wegen vorausgesetzt werden, während gleichzeitig die g_μ immer beliebig klein vorgeschrieben sein dürfen.

Setzt man nun zur Abkürzung:

$$F\left(\varphi^{(\mu)}(t), \psi^{(\mu)}(t)\right) = F\left(x^{(\mu)}, y^{(\mu)}\right) = F$$

und

$$F\left(\varphi^{(\mu)}(t) + \varepsilon\xi^{(\mu)}(t), \psi^{(\mu)}(t) + \varepsilon\eta^{(\mu)}(t)\right)$$
$$= F\left(x^{(\mu)} + \varepsilon\xi^{(\mu)}, y^{(\mu)} + \varepsilon\eta^{(\mu)}\right) = F_\varepsilon \,,$$

33 | $$\int_{t_1}^{t_2} F\,dt = J \,, \qquad \int_{t_1}^{t_2} F_\varepsilon\,dt = J_\varepsilon \,,$$

so kann man nach den über F gemachten Voraussetzungen für hinreichend kleine Werte von $|\varepsilon|$ nach Potenzen von ε entwickeln:

$$F_\varepsilon = F + \varepsilon\delta F + (\varepsilon)_2 \,, \quad J_\varepsilon = J + \varepsilon\delta J + (\varepsilon)_2 \tag{40}$$

und für $|\delta J| > 0$

$$J_\varepsilon - J = \Delta_\varepsilon J = \varepsilon\delta J\left(1 + \frac{(\varepsilon)}{\delta J}\right) \,,$$

wo der letzte Factor für hinreichend kleine $|\varepsilon|$ immer positiv, das Produkt also durch Wahl des Vorzeichens von ε beliebig auch negativ gemacht werden kann, sodass hier ein Minimum für a sicherlich *nicht* eintritt. Also:

Satz I. *Wenn einem der Gesamtheit A angehörenden Curvenstück a ein Minimum des Integrales J entsprechen soll, so muss vor allen Dingen die „erste Variation“ δJ für a immer verschwinden für beliebige den Bedingungen 1. und 2. genügende Functionen $\xi(t), \eta(t)$.*

Es ist aber:

$$\delta J = \int_{t_1}^{t_2} \delta F\,dt = \int_{t_1}^{t_2} \sum_{\mu=0}^{n} \left(X_\mu\xi^{(\mu)} + Y_\mu\eta^{(\mu)}\right) dt \,, \tag{41}$$

wenn, ebenso wie im ersten Abschnitte,

$$\frac{\partial F}{\partial x^{(\mu)}} = X_\mu \,, \quad \frac{\partial F}{\partial y^{(\mu)}} = Y_\mu \,, \quad \varphi^{(\mu)}(t) = x^{(\mu)} \,, \quad \psi^{(\mu)}(t) = y^{(\mu)}$$

gesetzt wird. Nun lässt sich die schon in (7a), dort freilich für $\xi = x'\tau, \eta = y'\tau$, angegebene Umformung:

$$\delta F = P\xi + Q\eta + D\sum_{\mu=1}^{n}\left(P_\mu\xi^{(\mu-1)} + Q_\mu\eta^{(\mu-1)}\right) \tag{42}$$

$$\left(P_\mu = \sum_{i=0}^{n-\mu}(-1)^i D^i x_{\mu+i} \,, \quad Q_\mu = \sum_{i=0}^{n-\mu}(-1)^i D^i Y_{\mu+i}\right)$$

are presupposed on account of their finiteness, while, at the same time, it is always possible to make the g_μ arbitrarily small.

If we now use the abbreviations

$$F\left(\varphi^{(\mu)}(t), \psi^{(\mu)}(t)\right) = F\left(x^{(\mu)}, y^{(\mu)}\right) = F$$

and

$$F\left(\varphi^{(\mu)}(t) + \varepsilon\xi^{(\mu)}(t), \psi^{(\mu)}(t) + \varepsilon\eta^{(\mu)}(t)\right)$$
$$= F\left(x^{(\mu)} + \varepsilon\xi^{(\mu)}, y^{(\mu)} + \varepsilon\eta^{(\mu)}\right) = F_\varepsilon \ ,$$
$$\int_{t_1}^{t_2} F\, dt = J \ , \qquad \int_{t_1}^{t_2} F_\varepsilon\, dt = J_\varepsilon \ ,$$

then, according to our assumptions about F, it is possible to expand in powers of ε for sufficiently small values of $|\varepsilon|$

$$F_\varepsilon = F + \varepsilon\delta F + (\varepsilon)_2 \ , \quad J_\varepsilon = J + \varepsilon\delta J + (\varepsilon)_2 \tag{40}$$

and for $|\delta J| > 0$

$$J_\varepsilon - J = \Delta_\varepsilon J = \varepsilon\delta J\left(1 + \frac{(\varepsilon)}{\delta J}\right) \ ,$$

where it is always possible to make the last factor positive for sufficiently small $|\varepsilon|$, and hence also make the product negative by appropriate choice of the sign of ε, so that there certainly exists *no* minimum for a in this case. Therefore,

Theorem I. *If a minimum of the integral J is supposed to correspond to a curve segment a belonging to the totality A, then, first of all, the "first variation" δJ for a must always vanish for all functions $\xi(t), \eta(t)$ satisfying conditions 1. and 2.*

But

$$\delta J = \int_{t_1}^{t_2} \delta F\, dt = \int_{t_1}^{t_2} \sum_{\mu=0}^{n} \left(X_\mu \xi^{(\mu)} + Y_\mu \eta^{(\mu)}\right) dt \tag{41}$$

if, as in the first section, one sets

$$\frac{\partial F}{\partial x^{(\mu)}} = X_\mu \ , \quad \frac{\partial F}{\partial y^{(\mu)}} = Y_\mu \ , \quad \varphi^{(\mu)}(t) = x^{(\mu)} \ , \quad \psi^{(\mu)}(t) = y^{(\mu)} \ .$$

Now it is only possible to carry out the transformation already stated in (7a), albeit for $\xi = x'\tau$, $\eta = y'\tau$,

$$\delta F = P\xi + Q\eta + D\sum_{\mu=1}^{n}\left(P_\mu \xi^{(\mu-1)} + Q_\mu \eta^{(\mu-1)}\right) \tag{42}$$

$$\left(P_\mu = \sum_{i=0}^{n-\mu} (-1)^i D^i x_{\mu+i} \ , \quad Q_\mu = \sum_{i=0}^{n-\mu} (-1)^i D^i Y_{\mu+i}\right) \ ,$$

34 | nur ausführen für solche Werte von t, für welche die P_μ, Q_μ, also die Ableitungen $x^{(n+1)}, \dots x^{(2n)}$; $y^{(n+1)}, \dots y^{(2n)}$ sämtlich *existieren*, was für $r < 2n$ in (33) und (34) noch nicht enthalten ist; ja, wir müssen sie als *stetige* Functionen von t in einem gewissen Intervalle voraussetzen, um den bekannten Schluss der Variationsrechnung machen zu können, dass P und Q überall verschwinden müssen. Dann aber lässt sich unter den allgemeinsten Voraussetzungen streng beweisen:

Satz II. *Existieren P und Q in irgend einem Teil-Intervall $t' \dots t''$ als stetige Functionen von t, so müssen sie für den Fall eines Minimums im ganzen Intervall überall verschwinden.*

Ist also $t_1 \leqq t' < t_0 < t'' \leqq t_2$, so wird behauptet, dass unter den gemachten Voraussetzungen $P(t_0) = 0$ sein muss.

Angenommen, es wäre z. B. $P(t_0) > p > 0$ positiv, so liesse sich der Stetigkeit von P wegen eine positive Grösse α so klein annehmen, dass

$$|P(t) - P(t_0)| < P(t_0) - p \qquad \text{für } |t - t_0| < \alpha \,,$$

so dass:

$$\begin{aligned} P(t) = P(t_0) + P(t) - P(t_0) &\geqq P(t_0) - |P(t) - P(t_0)| \\ &> P(t_0) - (P(t_0) - p) = p \,, \end{aligned}$$

also:

$$P(t) > p > 0 \tag{43}$$

ist im ganzen Intervall $t_0 - \alpha \dots t_0 + \alpha$.

In diesem wie in dem grösseren $t' \dots t''$ sind aber mit P, Q auch die übrigen P_μ, Q_μ stetige Functionen von t und daher die Umformung (42) gestattet, so dass

$$\begin{aligned} \delta J = \int_{t_1}^{t_2} \delta F \, dt &= \left(\int_{t_1}^{t_0 - \alpha} \delta F \, dt + \int_{t_0 + \alpha}^{t_2} \delta F \, dt \right) \\ &+ \left[\sum_{\mu=1}^{n} \left(P_\mu \xi^{(\mu-1)} + Q_\mu \eta^{(\mu-1)} \right) \right]_{t_0 - \alpha}^{t_0 + \alpha} \\ &+ \int_{t_0 - \alpha}^{t_0 + \alpha} (P\xi + Q\eta) \, dt = \delta J_1 + \delta J_2 + \delta J_0 \,. \end{aligned} \tag{44}$$

35 | Jetzt brauchen wir nur zu setzen:

$$\begin{aligned} \xi(t) &= 0 \qquad (t_1 \leqq t < t_0 - \alpha \,, \quad t_0 + \alpha < t \leqq t_2) \\ &= (t - t_0 + \alpha)^{r'+1} (t_0 + \alpha - t)^{r'+1} > 0 \\ &\qquad (t_0 - \alpha < t < t_0 + \alpha) \\ \eta(t) &= 0 \qquad (t_1 \leqq t \leqq t_2) \,, \end{aligned}$$

for those values of t for which the P_μ, Q_μ, that is, the derivatives $x^{(n+1)}, \dots x^{(2n)}$; $y^{(n+1)}, \dots y^{(2n)}$, all *exist*, which is not yet included in (33) and (34) when $r < 2n$; in fact, we must assume that they are *continuous* functions of t on a certain interval in order to draw the well-known conclusion of the calculus of variations, namely that P and Q must vanish everywhere. But, under the most general assumptions, it is then possible to furnish a rigorous proof of

Theorem II. *If P and Q exist in some partial interval $t' \dots t''$ as continuous functions of t, then they must vanish everywhere in the entire interval in case of a minimum.*

Hence, if $t_1 \leqq t' < t_0 < t'' \leqq t_2$, then what is being asserted is that, under the assumptions made, it must be the case that $P(t_0) = 0$.

Assuming, say, that $P(t_0) > p > 0$ were positive, it would then be possible, on account of the continuity of P, to choose a positive quantity α sufficiently small so that

$$|P(t) - P(t_0)| < P(t_0) - p \qquad \text{for } |t - t_0| < \alpha\,,$$

so that

$$\begin{aligned} P(t) = P(t_0) + P(t) - P(t_0) &\geqq P(t_0) - |P(t) - P(t_0)| \\ &> P(t_0) - (P(t_0) - p) = p\,, \end{aligned}$$

and hence

$$P(t) > p > 0 \tag{43}$$

in the entire interval $t_0 - \alpha \dots t_0 + \alpha$.

But, along with P, Q, the remaining P_μ, Q_μ are continuous functions of t on this interval as well as on the greater $t' \dots t''$. Hence, the transformation (42) is permissible, so that

$$\begin{aligned} \delta J = \int_{t_1}^{t_2} \delta F\, dt &= \left(\int_{t_1}^{t_0 - \alpha} \delta F\, dt + \int_{t_0 + \alpha}^{t_2} \delta F\, dt \right) \\ &+ \left[\sum_{\mu=1}^{n} \left(P_\mu \xi^{(\mu-1)} + Q_\mu \eta^{(\mu-1)} \right) \right]_{t_0 - \alpha}^{t_0 + \alpha} \\ &+ \int_{t_0 - \alpha}^{t_0 + \alpha} (P\xi + Q\eta)\, dt = \delta J_1 + \delta J_2 + \delta J_0\,. \end{aligned} \tag{44}$$

Now we only need to set

$$\begin{aligned} \xi(t) &= 0 \qquad (t_1 \leqq t < t_0 - \alpha\,, \quad t_0 + \alpha < t \leqq t_2) \\ &= (t - t_0 + \alpha)^{r'+1} (t_0 + \alpha - t)^{r'+1} > 0 \\ &\qquad (t_0 - \alpha < t < t_0 + \alpha) \\ \eta(t) &= 0 \qquad (t_1 \leqq t \leqq t_2)\,, \end{aligned}$$

wo $r' \geqq r$ und gleichzeitig $r' \geqq n-1$ eine beliebig grosse positive ganze Zahl sein kann. Dann sind wegen:

$$\xi^{(\mu)}(t_0 \pm \alpha) = 0 \qquad (0 \leqq \mu \leqq r')$$

ξ und η mit ihren Ableitungen bis zur $r$ten Ordnung im ganzen Intervall stetige Functionen von t (gemäss (33) und (34)) und genügen den Grenzbedingungen (39), müssen also nach Satz I die erste Variation zum Verschwinden bringen, während doch in (44)

$$\delta J_1 = 0\,, \quad \delta J_2 = 0\,,$$

also nach (43)

$$\delta J = \delta J_0 = \int_{t_0-\alpha}^{t_0+\alpha} P\xi\, dt > 0$$

wird, so dass die Annahme $P(t_0) > 0$ und ebenso, wie analog gezeigt wird, die anderen $P(t_0) < 0$ und $Q(t_0) \gtrless 0$ für ein Minimum unmöglich sind.

Wir können den Beweis auch führen, wenn wir nur solche Variationen zulassen, bei denen ξ, η im ganzen Intervall $t_1 \ldots t_2$ mit *allen* Ableitungen stetig und durch eine einzige analytische Formel ausgedrückt sind.

Setzen wir nämlich, einer zuerst von Herrn Prof. *Weierstrass* gegebenen Anregung folgend, wenn ϱ eine von t unabhängige, später zu bestimmende positive Grösse bedeutet,

$$\xi(t) = (t-t_1)^n (t_2-t)^n e^{-\varrho^2 (t-t_0)^2} \geqq 0$$
$$\eta(t) = 0$$

für

$$t_1 \leqq t \leqq t_2\,,$$

36 | wodurch allen Forderungen genügt wird, da auch

$$\xi^{(\mu)}(t_1) = \xi^{(\mu)}(t_2) = 0 \qquad (\mu = 0, 1, \ldots n-1)\,, \tag{39}$$

so sind die successiven Ableitungen von der Form:

$$\xi^{(\mu)}(t) = \xi_\mu(t; \varrho^2) e^{-\varrho^2 (t-t_0)^2}\,,$$

wo die ξ_μ ganze rationale Functionen ihrer beiden Argumente bedeuten.

In den Intervallen $t_1 \ldots t_0 - \alpha$ und $t_0 + \alpha \ldots t_2$ ist nun überall:

$$e^{-\varrho^2 (t-t_0)^2} \leqq e^{-\alpha^2 \varrho^2} \quad \text{wegen} \quad |t-t_0| \geqq \alpha\,,$$

also:

$$|\delta F| \leqq e^{-\alpha^2 \varrho^2} \sum_{\mu=0}^{n} \left| X_\mu(t) \xi_\mu(t, \varrho^2) \right| \leqq g_1(\varrho^2) e^{-\alpha^2 \varrho^2}\,,$$

where $r' \geqq r$, and, at the same time, $r' \geqq n-1$ can be an arbitrarily large positive integer. Then, on account of

$$\xi^{(\mu)}(t_0 \pm \alpha) = 0 \qquad (0 \leqq \mu \leqq r') ,$$

ξ and η, together with their derivatives up to the r th order, are continuous functions of t on the entire interval (according to (33) and (34)) and satisfy the limit conditions (39), and hence, by Theorem I, must make the first variation vanish, whereas in (44)

$$\delta J_1 = 0 , \quad \delta J_2 = 0 ,$$

and hence, by (43),

$$\delta J = \delta J_0 = \int_{t_0-\alpha}^{t_0+\alpha} P\xi \, dt > 0 ,$$

so that the assumption $P(t_0) > 0$, and also the other ones $P(t_0) < 0$ and $Q(t_0) \gtrless 0$, as is shown along similar lines, are not possible for a minimum.

We are also able to carry out the proof if we only permit variations for which ξ, η, together with *all* derivatives, are continuous on the entire interval $t_1 \ldots t_2$ and expressed by a single analytic formula.

For, if we, following a suggestion first made by Prof. *Weierstrass*, set

$$\xi(t) = (t-t_1)^n (t_2-t)^n e^{-\varrho^2 (t-t_0)^2} \geqq 0$$
$$\eta(t) = 0 ,$$

where ϱ denotes a positive quantity that is independent of t and to be specified later, for

$$t_1 \leqq t \leqq t_2 ,$$

whereby all requirements are met, since also

$$\xi^{(\mu)}(t_1) = \xi^{(\mu)}(t_2) = 0 \qquad (\mu = 0, 1, \ldots n-1) , \tag{39}$$

then the successive derivatives are of the form

$$\xi^{(\mu)}(t) = \xi_\mu(t; \varrho^2) e^{-\varrho^2 (t-t_0)^2} ,$$

where the ξ_μ denote integral rational functions of both of their arguments.

Now, in the intervals $t_1 \ldots t_0 - \alpha$ and $t_0 + \alpha \ldots t_2$, we everywhere have

$$e^{-\varrho^2 (t-t_0)^2} \leqq e^{-\alpha^2 \varrho^2} \quad \text{on account of} \quad |t-t_0| \geqq \alpha ,$$

and hence

$$|\delta F| \leqq e^{-\alpha^2 \varrho^2} \sum_{\mu=0}^{n} \left| X_\mu(t) \xi_\mu(t, \varrho^2) \right| \leqq g_1(\varrho^2) e^{-\alpha^2 \varrho^2} ,$$

wo $g_1(\varrho^2)$ die ganze Function von ϱ^2 mit positiven Coefficienten bezeichnet, die man erhält, wenn man in der Summe alle Coefficienten der Potenzen von ϱ^2 durch ihre grössten Beträge im Intervall $t_1 \leqq t \leqq t_2$ ersetzt.

Es wird demnach in (44)

$$|\delta J_1| \leqq \int_{t_1}^{t_0-\alpha} |\delta F|\, dt + \int_{t_0+\alpha}^{t_2} |\delta F|\, dt < (t_2 - t_1) g_1(\varrho^2) e^{-\alpha^2\varrho^2}$$

und ganz ebenso:

$$|\delta J_2| = e^{-\alpha^2\varrho^2} \left| \left[\sum_{\mu=1}^{n} P_\mu \xi_{\mu-1}(t, \varrho^2) \right]_{t_0-\alpha}^{t_0+\alpha} \right| \leqq g_2(\varrho^2) e^{-\alpha^2\varrho^2}$$

und

$$|\delta J_1 + \delta J_2| < \left((t_2 - t_1) g_1(\varrho^2) + g_2(\varrho^2) \right) e^{-\alpha^2\varrho^2} = g(\varrho^2) e^{-\alpha^2\varrho^2} \,.$$

Im Intervall $t_0 - \alpha \ldots t_0 + \alpha$ aber besitzt:

$$\xi_0(t) = (t - t_1)^n (t_2 - t)^n$$

eine positive untere Grenze q, so dass wegen (43) $(P > p)$

37 |
$$\delta J_0 = \int_{t_0-\alpha}^{t_0+\alpha} P\xi\, dt > pq \int_{t_0-\alpha}^{t_0+\alpha} e^{-\varrho^2 (t-t_0)^2}\, dt$$
$$> \frac{pq}{\varrho} \sum_{-\alpha\varrho}^{\alpha\varrho} e^{-z^2}\, dz = \frac{pq}{\varrho} (\sqrt{\pi} - \varkappa) > 0 \,,$$

wo wegen:

$$\int_{-\infty}^{+\infty} e^{-z^2}\, dz = \sqrt{\pi}$$

$\varkappa$ für jedes vorgeschriebene α durch Wahl eines hinreichend grossen ϱ beliebig klein gemacht werden kann.

Jetzt kann endlich die Beziehung:

$$\left| \frac{\delta J_1 + \delta J_2}{\delta J_0} \right| < \frac{\varrho g(\varrho^2) e^{-\alpha^2\varrho^2}}{pq(\sqrt{\pi} - x)} < 1$$

durch Vergrösserung von ϱ sicher erreicht werden, da

$$\lim_{\varrho=\infty} \varrho g(\varrho^2) e^{-\varrho^2\alpha^2} = 0$$

where $g_1(\varrho^2)$ denotes the entire function of ϱ^2 with positive coefficients which is obtained by replacing in the sum all coefficients of the powers of ϱ^2 by their greatest amounts in the interval $t_1 \leqq t \leqq t_2$.

In (44), we therefore have

$$|\delta J_1| \leqq \int_{t_1}^{t_0-\alpha} |\delta F|\, dt + \int_{t_0+\alpha}^{t_2} |\delta F|\, dt < (t_2 - t_1) g_1(\varrho^2) e^{-\alpha^2\varrho^2}$$

and likewise

$$|\delta J_2| = e^{-\alpha^2\varrho^2} \left| \left[\sum_{\mu=1}^{n} P_\mu \xi_{\mu-1}(t, \varrho^2) \right]_{t_0-\alpha}^{t_0+\alpha} \right| \leqq g_2(\varrho^2) e^{-\alpha^2\varrho^2}$$

and

$$|\delta J_1 + \delta J_2| < \left((t_2 - t_1) g_1(\varrho^2) + g_2(\varrho^2)\right) e^{-\alpha^2\varrho^2} = g(\varrho^2) e^{-\alpha^2\varrho^2} \,.$$

In the interval $t_0 - \alpha \ldots t_0 + \alpha$, however,

$$\xi_0(t) = (t - t_1)^n (t_2 - t)^n$$

possesses a positive lower limit q so that, on account of (43), $(P > p)$

$$\delta J_0 = \int_{t_0-\alpha}^{t_0+\alpha} P\xi\, dt > pq \int_{t_0-\alpha}^{t_0+\alpha} e^{-\varrho^2 (t-t_0)^2}\, dt$$

$$> \frac{pq}{\varrho} \sum_{-\alpha\varrho}^{\alpha\varrho} e^{-z^2}\, dz = \frac{pq}{\varrho}(\sqrt{\pi} - \varkappa) > 0 \,,$$

where $\varkappa$, on account of

$$\int_{-\infty}^{+\infty} e^{-z^2}\, dz = \sqrt{\pi} \,,$$

can be made arbitrarily small for every prescribed α by choice of a sufficiently large ϱ.

Finally, we are now able to easily arrive at the relation

$$\left| \frac{\delta J_1 + \delta J_2}{\delta J_0} \right| < \frac{\varrho g(\varrho^2) e^{-\alpha^2\varrho^2}}{pq(\sqrt{\pi} - x)} < 1$$

by increasing ϱ, since

$$\lim_{\varrho=\infty} \varrho g(\varrho^2) e^{-\varrho^2\alpha^2} = 0 \,,$$

ist, und es wird schliesslich nach (44)

$$\delta J = \delta J_0 \left(1 + \frac{\delta J_1 + \delta J_2}{\delta J_0}\right) > 0 \,,$$

während doch $\delta J = 0$ sein sollte, womit der Beweis des Satzes II vollendet ist.

Die beiden Gleichungen $P = 0$ und $Q = 0$ hängen zusammen durch (15):

$$Px' + Qy' = 0.$$

Setzt man daher $\qquad Py' - Qx' = G, \qquad$ so folgt:

$$P(x'^2 + y'^2) = y'G \,, \qquad Q(x'^2 + y'^2) = -x'G \,; \tag{45}$$

es werden also, da x' und y' nach (35) nicht gleichzeitig verschwinden sollen, die beiden Differentialgleichungen der Ordnung $2n$ ersetzt durch eine einzige: $G = 0$, die wir als „*die Differentialgleichung des Problems*“ bezeichnen wollen.

38 | Setzt man nun:

$$\left.\begin{aligned} x'\xi + y'\eta &= (x'^2 + y'^2)v \,, \quad \xi = x'v + y'w \,, \\ y'\xi + x'\eta &= (x'^2 + y'^2)w \,, \quad \eta = y'v - x'w \,, \end{aligned}\right\} \tag{46}$$

so wird

$$P\xi + Q\eta = (Px' + Qy')v + (Py' - Qx')w = G\,w \,, \tag{47}$$

und daher, wenn in (42) der integrierte Teil nur angedeutet wird:

$$\delta J = \Big[\quad\Big]_{t'}^{t''} + \int_{t'}^{t''} Gw\,dt$$

für jedes Intervall $t' \ldots t''$, für welches die Umformung (42) gestattet ist.

Stetigkeits-Bedingungen.

Was Herr *Weierstrass* für das Integral:

$$J = \int_{t_1}^{t_2} F(x, y; x', y')\,dt$$

von den Functionen $\dfrac{\partial F}{\partial x'}$ und $\dfrac{\partial F}{\partial y'}$ bewiesen hat, ähnlich wie *P. du Bois-Reymond* (Math. Annalen XV) und *Todhunter* (Researches on the Calculus of Variations) von analogen Functionen in der gewöhnlichen Darstellung, lässt sich folgendermassen verallgemeinern:

and, by (44), eventually

$$\delta J = \delta J_0 \left(1 + \frac{\delta J_1 + \delta J_2}{\delta J_0}\right) > 0 \,,$$

while we ought to have $\delta J = 0$, which completes the proof of theorem II.

The two equations $P = 0$ and $Q = 0$ are connected with one another via (15):

$$Px' + Qy' = 0.$$

If we therefore set $Py' - Qx' = G,$ it then follows that

$$P(x'^2 + y'^2) = y'G \,, \qquad Q(x'^2 + y'^2) = -x'G \,; \tag{45}$$

thus, since x' and y' are not supposed to simultaneously vanish according to (35), the two differential equations of order $2n$ are replaced by a single one: $G = 0$, which we shall call "*the differential equation of the problem*".

If one now sets

$$\begin{cases} x'\xi + y'\eta = (x'^2 + y'^2)v \,, & \xi = x'v + y'w \,, \\ y'\xi + x'\eta = (x'^2 + y'^2)w \,, & \eta = y'v - x'w \,, \end{cases} \tag{46}$$

then

$$P\xi + Q\eta = (Px' + Qy')v + (Py' - Qx')w = G\,w \,, \tag{47}$$

and hence, if the integrated part in (42) is merely indicated,

$$\delta J = \Big[\quad\Big]_{t'}^{t''} + \int_{t'}^{t''} Gw\,dt$$

for every interval $t' \ldots t''$ for which the transformation (42) is permissible.

Continuity conditions.

What *Weierstrass* has shown for the integral

$$J = \int_{t_1}^{t_2} F(x, y; x', y')\,dt$$

of the functions $\dfrac{\partial F}{\partial x'}$ and $\dfrac{\partial F}{\partial y'}$, like *P. du Bois-Reymond* (*1879a,b*) and *Todhunter* (*1871*) in the case of analogous functions in the usual representation, can be generalized as follows:

Satz III. *Für ein Curvenstück a, das den Bedingungen* (32) *bis* (35) *der A genügt und das ein Minimum des Integrales liefern, oder allgemeiner, dessen erste Variation in der oben angegebenen Weise verschwinden soll, dürfen die Grössen* $P_1, P_2 \dots P_n = X_n; Q_1, Q_2, \dots Q_n = Y_n$, *als Functionen von* t *betrachtet, an einer Stelle* t_0 *keine endlichen Sprünge erleiden, wenn zu beiden Seiten von* t_0 *in einer beliebig kleinen Umgebung* $t' \dots t''$ *die Functionen* $x, x', \dots x^{(2n)}$; $y, y', \dots y^{(2n)}$ *und damit auch* P, Q *stetig sind.*

Um diesen Satz für eine dieser Grössen P_λ zu erweisen, nehmen wir eine Variation ξ, η von der Beschaffenheit, dass

39 |
$$\begin{gathered}
\eta(t) = 0 \qquad (t_1 \leqq t \leqq t_2) \\
\xi(t) = 0 \qquad (t_1 \leqq t \leqq t',\ t'' \leqq t \leqq t_2), \\
\xi^{(\mu-1)}(t') = \xi^{(\mu-1)}(t'') = 0 \quad (\mu = 1, 2, \dots r' + 1 \geqq n) \qquad (48) \\
\xi^{(\mu-1)}(t_0) = e_{\mu,\lambda} = \begin{cases} 1 \ (\mu = \lambda) \\ 0 \ (\mu \gtrless \lambda) \end{cases} \quad (\mu = 1, 2, \dots n)
\end{gathered}$$

und $\xi(t)$ auch im Intervall $t' \dots t''$ mit seinen Ableitungen bis zur r' ten Ordnung stetig ist.

Dann ist:

$$\begin{aligned}
\delta J = \int_{t_1}^{t_2} \delta F \, dt &= \int_{t'}^{t_0} + \int_{t_0}^{t''} \\
&= \left[\sum_{\mu=1}^{n} P_\mu \xi^{(\mu-1)} \right]_{t'}^{t_0 - 0} + \int_{t'}^{t_0 - 0} Gw \, dt \\
&\quad + \left[\sum_{\mu=1}^{n} P_\mu \xi^{(\mu-1)} \right]_{t_0 + 0}^{t''} + \int_{t_0 + 0}^{t''} Gw \, dt \,,
\end{aligned}$$

also wegen $G = 0$ in den beiden Intervallen $t' \dots t_0 - 0$ und $t_0 + 0 \dots t''$ (Satz II) und wegen (48):

$$\delta J = \left[\sum_{\mu=1}^{n} P_\mu \xi^{(\mu-1)} \right]_{t_0 + 0}^{t_0 - 0} = P_\lambda(t_0 - 0) - P_\lambda(t_0 + 0) \,,$$

und da nach Satz I $\delta J = 0$ sein muss:

$$P_\lambda(t_0 - 0) = P_\lambda(t_0 + 0) \qquad (\lambda = 1, 2, \dots n) \,,$$

q. e. d.

Doch wurde bei diesem Beweise vorausgesetzt, dass P_λ auf beiden Seiten von t_0 nach bestimmten Grenzwerten $P_\lambda(t_0 \mp 0)$ convergiert, was z. B. für

Theorem III. *Given a curve segment a that satisfies the conditions* (32) – (35) *on the A and furnishes a minimum of the integral, or more generally, whose first variation shall vanish in the manner specified above, the quantities* $P_1, P_2 \dots P_n = X_n$; $Q_1, Q_2, \dots Q_n = Y_n$ *considered as functions of* t *may not suffer any finite jump discontinuities at a position* t_0 *if, on either side of* t_0, *the functions* $x, x', \dots x^{(2n)}$; $y, y', \dots y^{(2n)}$, *and hence also* P, Q, *are continuous on an arbitrarily small vicinity* $t' \dots t''$.

In order to prove this theorem for one of these quantities P_λ, we assume a variation ξ, η constituted so that

$$\begin{aligned} \eta(t) &= 0 \qquad (t_1 \leqq t \leqq t_2) \\ \xi(t) &= 0 \qquad (t_1 \leqq t \leqq t',\ t'' \leqq t \leqq t_2), \\ \xi^{(\mu-1)}(t') = \xi^{(\mu-1)}(t'') &= 0 \quad (\mu = 1, 2, \dots r' + 1 \geqq n) \\ \xi^{(\mu-1)}(t_0) = e_{\mu,\lambda} &= \begin{cases} 1\ (\mu = \lambda) \\ 0\ (\mu \gtrless \lambda) \end{cases} \quad (\mu = 1, 2, \dots n) \end{aligned} \tag{48}$$

and $\xi(t)$, together with its derivatives up the r' th order, is continuous on the interval $t' \dots t''$.

Then

$$\begin{aligned} \delta J = \int_{t_1}^{t_2} \delta F\, dt &= \int_{t'}^{t_0} + \int_{t_0}^{t''} \\ &= \left[\sum_{\mu=1}^{n} P_\mu \xi^{(\mu-1)} \right]_{t'}^{t_0 - 0} + \int_{t'}^{t_0 - 0} Gw\, dt \\ &\quad + \left[\sum_{\mu=1}^{n} P_\mu \xi^{(\mu-1)} \right]_{t_0 + 0}^{t''} + \int_{t_0 + 0}^{t''} Gw\, dt \end{aligned}$$

and hence, on account of $G = 0$, in the two intervals $t' \dots t_0 - 0$ and $t_0 + 0 \dots t''$ (Theorem II) and, on account of (48),

$$\delta J = \left[\sum_{\mu=1}^{n} P_\mu \xi^{(\mu-1)} \right]_{t_0 + 0}^{t_0 - 0} = P_\lambda(t_0 - 0) - P_\lambda(t_0 + 0)\ ,$$

and since, by Theorem I, necessarily $\delta J = 0$:

$$P_\lambda(t_0 - 0) = P_\lambda(t_0 + 0) \qquad (\lambda = 1, 2, \dots n)\ ,$$

q. e. d.

In this proof, however, we took for granted that, on either side of t_0, P_λ converges to certain limits $P_\lambda(t_0 \mp 0)$, which, e. g., is not the case for the

die Function $\sin\frac{1}{t}$ an ihrer einzigen Unstetigkeitsstelle $t = 0$ nicht stattfin-
40 det; ferner wurde eine Variation | $\xi(t)$ benutzt, deren höhere Ableitungen an den Stellen t', t'' unstetig sind, und es fehlte schliesslich noch der Beweis für die Existenz einer Function $\xi(t)$ mit den angenommenen Eigenschaften im Intervall $t' \leqq t \leqq t''$. Jetzt soll der Beweis von allen diesen Bedenken befreit werden durch Benutzung ganz ähnlicher Hilfsmittel wie der zum Beweise des vorigen Satzes II angewendeten.

Es seien α, β beliebig kleine positive Grössen,

$$t_1 \leqq t' < t_0 - \alpha < t_0 < t_0 + \beta < t'' \leqq t_2$$

und es werde vorausgesetzt, dass in den beiden Intervallen $t' \ldots t_0 - \alpha$ und $t_0 + \beta \ldots t''$, wie klein auch α und β genommen werden, P und Q stetig verlaufen, also nach Satz II beständig verschwinden, dann kann, so wird behauptet, für α, β eine obere Grenze α_0 so angegeben werden, dass

$$|P_\lambda(t_0 + \beta) - P_\lambda(t_0 - \alpha)| < \varepsilon \tag{49}$$

für ein beliebig klein vorgeschriebenes ε unter der Voraussetzung des Verschwindens der ersten Variation.

Wir betrachten dazu die folgende jedenfalls erlaubte Variation:

$$\begin{aligned} &\eta = 0 \\ &\xi = \xi(t;\varrho) = \xi_0(t,\varrho)e^{-\varrho(t-t_0)^2} \qquad (\varrho > 0)\ , \end{aligned}$$

wo $\xi_0(t,\varrho)$ eine ganze rationale Function von t und ϱ sein soll und so beschaffen, dass

$$\begin{aligned} \xi^{(\mu-1)}(t_1) &= \xi^{(\mu-1)}(t_2) = 0\ , \\ \xi^{(\mu-1)}(t') &= \xi^{(\mu-1)}(t'') = 0\ , \\ \xi^{(\mu-1)}(t_0) &= e_{\mu,\lambda} \qquad (\mu = 1, 2, \ldots n)\ . \end{aligned} \tag{50}$$

Den ersten Bedingungen wird genügt durch den Ansatz:

$$\begin{aligned} \xi_0(t,\varrho) &= (t-t_1)^n(t_2-t)^n(t-t')^n(t''-t)^n\varphi(t,\varrho) \\ &= h(t)\varphi(t,\varrho)\ , \end{aligned}$$

und es bleibt nur noch die ganze Function $\varphi(t,\varrho)$ so zu bestimmen, dass auch die letzte Bedingung befriedigt wird, nämlich

41 |
$$\begin{aligned} e_{\mu,\lambda} &= D^{\mu-1}\left[h(t)\varphi(t,\varrho)e^{-\varrho(t-t_0)^2}\right]\ , \qquad (t = t_0) \\ &= \sum_{i=0}^{\mu-1}\binom{\mu-1}{i}\varphi^{(i)}(t,\varrho)D^{\mu-i-1}\left[h(t)e^{-\varrho(t-t_0)^2}\right] \\ &= \sum_{i=0}^{\mu-1}\binom{\mu-1}{i}\varphi^{(i)}(t_0,\varrho)h_{\mu-i-1}(t_0,\varrho)\ , \end{aligned}$$

function $\sin\frac{1}{t}$ at its only discontinuity at $t=0$; furthermore, we used a variation $\xi(t)$ whose higher derivatives are discontinuous at the positions t', t''. Finally, what was missing was a proof of the existence of some function $\xi(t)$ with the asserted properties in the interval $t' \leqq t \leqq t''$. In what follows, the proof shall be rid of all these concerns by use of expedients very similar to those used for the proof of Theorem II.

Consider arbitrarily small positive quantities α, β such that

$$t_1 \leqq t' < t_0 - \alpha < t_0 < t_0 + \beta < t'' \leqq t_2$$

and assume that P and Q are continuous on both intervals $t' \ldots t_0 - \alpha$ and $t_0 + \beta \ldots t''$, however small α and β may be, and hence, by Theorem II, always vanish. Then, so it is asserted, it is possible to specify an upper limit α_0 for α, β so that

$$|P_\lambda(t_0+\beta) - P_\lambda(t_0-\alpha)| < \varepsilon \tag{49}$$

for an arbitrarily small prescribed ε, provided that the first variation vanishes.

For this purpose let us consider the following variation, which is certainly admissible:

$$\begin{aligned} &\eta = 0 \\ &\xi = \xi(t;\varrho) = \xi_0(t,\varrho)e^{-\varrho(t-t_0)^2} \qquad (\varrho > 0)\ , \end{aligned}$$

where $\xi_0(t,\varrho)$ is supposed to be an integral rational function of t and ϱ so constituted that

$$\begin{aligned} &\xi^{(\mu-1)}(t_1) = \xi^{(\mu-1)}(t_2) = 0\ , \\ &\xi^{(\mu-1)}(t') = \xi^{(\mu-1)}(t'') = 0\ , \\ &\xi^{(\mu-1)}(t_0) = e_{\mu,\lambda} \qquad (\mu = 1,2,\ldots n)\ . \end{aligned} \tag{50}$$

The first conditions are met by the ansatz

$$\begin{aligned} \xi_0(t,\varrho) &= (t-t_1)^n(t_2-t)^n(t-t')^n(t''-t)^n\varphi(t,\varrho) \\ &= h(t)\varphi(t,\varrho)\ , \end{aligned}$$

and the only task remaining is to determine the entire function $\varphi(t,\varrho)$ so that the last condition, too, is satisfied, namely

$$\begin{aligned} e_{\mu,\lambda} &= D^{\mu-1}\left[h(t)\varphi(t,\varrho)e^{-\varrho(t-t_0)^2}\right]\ , \qquad (t=t_0) \\ &= \sum_{i=0}^{\mu-1}\binom{\mu-1}{i}\varphi^{(i)}(t,\varrho)D^{\mu-i-1}\left[h(t)e^{-\varrho(t-t_0)^2}\right] \\ &= \sum_{i=0}^{\mu-1}\binom{\mu-1}{i}\varphi^{(i)}(t_0,\varrho)h_{\mu-i-1}(t_0,\varrho)\ , \end{aligned}$$

wenn

$$D^i\left[h(t)e^{-\varrho(t-t_0)^2}\right] = h_i(t,\varrho)e^{-\varrho(t-t_0)^2}$$

gesetzt wird, wo auch h_i eine ganze Function ist und der Exponentialfactor für $t = t_0$ sich auf 1 reduciert.

Nun lassen sich die vorliegenden Gleichungen für $\mu = 1, 2 \ldots n$ successive auflösen nach den Unbekannten

$$\varphi(t_0,\varrho)\ , \quad \varphi'(t_0,\varrho) = \left[\frac{\partial\varphi(t,\varrho)}{\partial t}\right]_{t=t_0} , \ldots \varphi^{(n-1)}(t_0,\varrho)\ ,$$

da immer die höchste Ableitung $\varphi^{(\mu-1)}(t_0,\varrho)$ den Coefficienten $h_0(t,\varrho) = h(t_0) > 0$ besitzt. So ergeben sich diese Grössen $\varphi^{(\mu-1)}(t_0,\varrho)$ sämtlich als ganze Functionen von ϱ, und man braucht nur noch

$$\varphi(t,\varrho) = \sum_{i=0}^{n=1} \frac{1}{i!}\varphi^{(i)}(t_0,\varrho)(t-t_0)^i$$

zu setzen, um eine Function

$$\xi(t) = h(t)\varphi(t,\varrho)e^{-\varrho(t-t_0)^2} = \xi_0(t,\varrho)e^{-\varrho(t-t_0)^2}$$

zu erhalten, die in der That allen Bedingungen (50) genügt und deren Ableitungen von der Form sind:

$$\xi^{(\mu)}(t) = \xi_\mu(t,\varrho)e^{-\varrho(t-t_0)^2}\ .$$

Ganz dieselbe Methode unter der vereinfachenden Annahme $\varrho = 0$ liesse sich anwenden zur Bildung der vorhin benutzten Function $\xi(t)$ mit den Eigenschaften (48).

42 | Für unsere Variation $\delta x = \xi(t,\varrho), \delta y = 0$ wird nun:

$$\left.\begin{aligned} \delta J &= \int_{t_1}^{t_2} \delta F\, dt \\ &= \left(\int_{t_1}^{t'} + \int_{t''}^{t^2}\right) + \int_{t_0-\alpha}^{t_0+\beta} + \left(\int_{t'}^{t_0-\alpha} + \int_{t_0+\beta}^{t''}\right) \\ &= \quad \delta J_1 \quad + \ \delta J_2 \ + \quad \delta J_3\ . \end{aligned}\right\} \tag{51}$$

Ähnlich wie beim Beweise von II lässt sich jetzt eine ganze rationale Function $g(\varrho)$ so angeben, dass für das ganze Intervall $t_1 \leqq t \leqq t_2$

$$\begin{aligned} |\delta F| = \left|\sum_{\mu=0}^{n} X_\mu \xi^{(\mu)}\right| &= e^{-\varrho(t-t_0)^2}\left|\sum_{\mu=0}^{n} X_\mu(t)\xi_\mu(t,\varrho)\right| \\ &< e^{-\varrho(t-t_0)^2} g(\varrho) \end{aligned} \tag{52}$$

if we set

$$D^i\left[h(t)e^{-\varrho(t-t_0)^2}\right] = h_i(t,\varrho)e^{-\varrho(t-t_0)^2}\ ,$$

where h_i, too, is an entire function and the exponential factor for $t = t_0$ is reduced to 1.

Now it is possible to solve the equations at hand for $\mu = 1, 2 \ldots n$ successively for the unknowns

$$\varphi(t_0,\varrho)\ , \quad \varphi'(t_0,\varrho) = \left[\frac{\partial\varphi(t,\varrho)}{\partial t}\right]_{t=t_0}\ , \ldots \varphi^{(n-1)}(t_0,\varrho)\ ,$$

since the highest derivative $\varphi^{(\mu-1)}(t_0,\varrho)$ always possesses the coefficient $h_0(t,\varrho) = h(t_0) > 0$. Thus arise all these quantities $\varphi^{(\mu-1)}(t_0,\varrho)$ as entire functions of ϱ, and we only need to set

$$\varphi(t,\varrho) = \sum_{i=0}^{n=1}\frac{1}{i!}\varphi^{(i)}(t_0,\varrho)(t-t_0)^i$$

in order to obtain a function

$$\xi(t) = h(t)\varphi(t,\varrho)e^{-\varrho(t-t_0)^2} = \xi_0(t,\varrho)e^{-\varrho(t-t_0)^2}$$

which indeed satisfies all conditions (50) and whose derivatives are of the form

$$\xi^{(\mu)}(t) = \xi_\mu(t,\varrho)e^{-\varrho(t-t_0)^2}\ .$$

It is possible to use precisely the same method under the simplifying assumption $\varrho = 0$ in order to form the function $\xi(t)$ used above with the properties (48).

In the case of our variation $\delta x = \xi(t,\varrho), \delta y = 0$ we now have

$$\left.\begin{aligned}\delta J &= \int_{t_1}^{t_2}\delta F\,dt \\ &= \left(\int_{t_1}^{t'} + \int_{t''}^{t^2}\right) + \int_{t_0-\alpha}^{t_0+\beta} + \left(\int_{t'}^{t_0-\alpha} + \int_{t_0+\beta}^{t''}\right) \\ &= \quad \delta J_1 \quad + \ \delta J_2 \ + \quad \delta J_3\ .\end{aligned}\right\} \tag{51}$$

Similar to the proof of II, it is now possible to specify an entire rational function $g(\varrho)$ so that, for the entire interval $t_1 \leqq t \leqq t_2$,

$$\begin{aligned}|\delta F| = \left|\sum_{\mu=0}^{n} X_\mu \xi^{(\mu)}\right| &= e^{-\varrho(t-t_0)^2}\left|\sum_{\mu=0}^{n} X_\mu(t)\xi_\mu(t,\varrho)\right| \\ &< e^{-\varrho(t-t_0)^2}g(\varrho)\end{aligned} \tag{52}$$

wird, da nach unserer Annahme die X_μ auch an der Unstetigkeitsstelle endlich bleiben sollen. Für $(t-t_0)^2 \geqq \tau^2$, also in den Intervallen $t_1 \ldots t'$ und $t'' \ldots t_2$, wenn τ die kleinere der Differenzen $t_0 - t'$ und $t'' - t_0$ bezeichnet, wird daher:

$$|\delta F| < e^{-\varrho\tau^2 g(\varrho)}$$

und

$$|\delta J_1| = \left| \int_{t_1}^{t'} + \int_{t''}^{t_2} \right| < (t_2 - t_1) g(\varrho) e^{-\varrho\tau^2} < \frac{\varepsilon}{3} \tag{53}$$

für ein beliebig kleines ε, wenn bei vorgeschriebenem τ der Parameter ϱ hinreichend gross genommen wird.

Ferner wird für $t_0 - \alpha < t < t_0 + \beta$:

$$|\delta F| < e^{-\varrho(t-t_0)^2} g(\varrho) \leqq g(\varrho)$$

wegen $\varrho > 0$, und daher kann auch für das eben bestimmte ϱ noch der zweite Teil in (51)

$$|\delta J_2| = \left| \int_{t_0-\alpha}^{t_0+\beta} \delta F \, dt \right| < g(\varrho)(\alpha+\beta) < 2\alpha_0 g(\varrho) < \frac{\varepsilon}{3} \tag{54}$$

43 | gemacht werden durch hinreichend klein gewähltes α_0 und

$$\alpha < \alpha_0 \,, \ \beta < \alpha_0 \,.$$

In den Intervallen $t' \ldots t_0 - \alpha$ und $t_0 + \beta \ldots t''$ endlich sind nach der Annahme alle P_μ stetige Functionen von t und daher ist die Formel (42) anwendbar, so dass

$$\begin{aligned} \delta J_3 = &\left[\sum_{\mu=1}^{n} P_\mu \xi^{(\mu-1)} \right]_{t'}^{t_0-\alpha} + \int_{t'}^{t_0-\alpha} P\xi \, dt \\ &+ \left[\sum_{\mu=1}^{n} P_\mu \xi^{(\mu-1)} \right]_{t_0+\beta}^{t''} + \int_{t_0+\beta}^{t''} P\xi \, dt \,, \end{aligned}$$

oder, wegen $P = 0$, da sonst nach II gewiss kein Minimum stattfinden könnte, mit Hilfe von (50)

$$\delta J_3 = \left[\sum_{\mu=1}^{n} P_\mu \xi^{(\mu-1)} \right]_{t_0+\beta}^{t_0-\alpha}$$

since, by our assumption, the X_μ are supposed to remain finite at the discontinuity as well. For $(t-t_0)^2 \geqq \tau^2$, and hence in the intervals $t_1 \ldots t'$ and $t'' \ldots t_2$, where τ denotes the smaller of the difference $t_0 - t'$ and $t'' - t_0$, therefore

$$|\delta F| < e^{-\varrho\tau^2 g(\varrho)}$$

and

$$|\delta J_1| = \left| \int_{t_1}^{t'} + \int_{t''}^{t_2} \right| < (t_2 - t_1) g(\varrho) e^{-\varrho\tau^2} < \frac{\varepsilon}{3} , \tag{53}$$

given an arbitrarily small ε, if we take the parameter ϱ sufficiently large for prescribed τ.

Furthermore, for $t_0 - \alpha < t < t_0 + \beta$,

$$|\delta F| < e^{-\varrho(t-t_0)^2} g(\varrho) \leqq g(\varrho) ,$$

on account of $\varrho > 0$, and therefore it is also possible to make the second part in (51)

$$|\delta J_2| = \left| \int_{t_0-\alpha}^{t_0+\beta} \delta F \, dt \right| < g(\varrho)(\alpha+\beta) < 2\alpha_0 g(\varrho) < \frac{\varepsilon}{3} \tag{54}$$

for the ϱ just determined by choice of a sufficiently small α_0 and

$$\alpha < \alpha_0, \ \beta < \alpha_0 .$$

Finally, all P_μ are, by assumption, continuous functions of t in the intervals $t' \ldots t_0 - \alpha$ and $t_0 + \beta \ldots t''$, and the formula (42) is therefore applicable so that

$$\delta J_3 = \left[\sum_{\mu=1}^{n} P_\mu \xi^{(\mu-1)} \right]_{t'}^{t_0-\alpha} + \int_{t'}^{t_0-\alpha} P\xi \, dt$$
$$+ \left[\sum_{\mu=1}^{n} P_\mu \xi^{(\mu-1)} \right]_{t_0+\beta}^{t''} + \int_{t_0+\beta}^{t''} P\xi \, dt ,$$

or, on account of $P = 0$, since otherwise, by II, there would certainly be no minimum, by means of (50)

$$\delta J_3 = \left[\sum_{\mu=1}^{n} P_\mu \xi^{(\mu-1)} \right]_{t_0+\beta}^{t_0-\alpha}$$

und

$$\delta J_3 = P_\lambda(t_0 - \alpha) - P_\lambda(t_0 + \beta) + \psi(\varrho, \alpha, \beta) \ , \tag{55}$$

wo

$$\psi(\varrho, \alpha, \beta) = \left[\sum_{\mu=1}^{n} P_\mu \left(\xi^{(\mu-1)} - e_{\mu,\lambda} \right) \right]_{t_0+\beta}^{t_0-\alpha} ,$$

wegen (50)

$$\lim_{\alpha=0} \xi^{(\mu-1)}(t_0 - \alpha) = \lim_{\beta=0} \xi^{(\mu-1)}(t_0 + \beta) = \xi^{(\mu-1)}(t_0) = e_{\mu,\lambda} \ ,$$

und weil auch die P_μ bei der Annäherung an t_0 nicht über alle Grenzen wachsen sollen, durch Verkleinerung von α und β dem Betrage nach beliebig klein gemacht werden kann, also auch

$$|\psi(\varrho, \alpha, \beta)| < \frac{\varepsilon}{3} \ . \tag{56}$$

44 | Nun soll aber: $\delta J = \delta J_1 + \delta J_2 + \delta J_3 = 0$ sein (51) auch für die hier betrachtete Variation ξ, also nach (55)

$$|P_\lambda(t_0 + \beta) - P_\lambda(t_0 - \alpha)| \leqq |\delta J_1| + |\delta J_2| + |\psi(\varrho, \alpha, \beta)|$$

und nach (53), (54), (56) $$< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \ ,$$

wie in (49) behauptet, wenn für ein beliebig klein vorgeschriebenes ε die obere Grenze α_0 von α und β hinreichend klein genommen wird.

Die Gültigkeit dieses jetzt vollständig bewiesenen Satzes III stützt sich jedoch auf die Voraussetzung, dass gemäss (34) (vergl. die dort gemachte Bemerkung) die $x^{(\mu)} = \varphi^{(\mu)}(t)$, $y^{(\mu)} = \psi^{(\mu)}(t)$ $(\mu = 0, 1, \ldots r)$ *selbst* im ganzen betrachteten Intervall von a überall *stetig* verlaufen, nicht nur ihre Osculations-Invarianten. Im anderen Falle würden die dem Beweise zu Grunde gelegten Variationen oft nicht mehr zu den „erlaubten“ gehören, und es wären nicht mehr die P_μ, Q_μ, sondern *andere Functionen*, deren ausnahmslose Stetigkeit behauptet werden könnte. Solche Ausdrücke aber erhält man z. B., wenn man für t die *Bogenlänge* s von a oder allgemeiner eine mit ihren r ersten Ableitungen stetige Function von s als unabhängige Variable einführt und demgemäss in den P_μ, Q_μ die Argumente $x^{(\mu)}$, $y^{(\mu)}$ durch die $\frac{d^\mu x}{ds^\mu}$, $\frac{d^\mu y}{ds^\mu}$ ersetzt, die sich ihrerseits wieder durch die α_μ und auch durch die $x^{(\mu)}$, $y^{(\mu)}$ ausdrücken lassen. Doch soll hierauf jetzt nicht näher eingegangen werden.

Bisher war die Anzahl r der ersten Ableitungen $x^{(\mu)}$, $y^{(\mu)}$, deren Stetigkeit in (34) für alle Curven A gefordert wurde, noch *unbestimmt* gelassen worden und die Sätze I, II, III für beliebiges r abgeleitet. Es wird jetzt aber gezeigt werden, dass

$$r \geqq n - 1$$

and

$$\delta J_3 = P_\lambda(t_0 - \alpha) - P_\lambda(t_0 + \beta) + \psi(\varrho, \alpha, \beta) \ , \tag{55}$$

where

$$\psi(\varrho, \alpha, \beta) = \left[\sum_{\mu=1}^{n} P_\mu \left(\xi^{(\mu-1)} - e_{\mu,\lambda} \right) \right]_{t_0+\beta}^{t_0-\alpha} ,$$

on account of (50),

$$\lim_{\alpha=0} \xi^{(\mu-1)}(t_0 - \alpha) = \lim_{\beta=0} \xi^{(\mu-1)}(t_0 + \beta) = \xi^{(\mu-1)}(t_0) = e_{\mu,\lambda} \ ,$$

and because the P_μ are not supposed to increase beyond all limits either as they approach t_0, can be made arbitrarily small with respect to its absolute value by decrease of α and β, and hence also

$$|\psi(\varrho, \alpha, \beta)| < \frac{\varepsilon}{3} \ . \tag{56}$$

But now we are supposed to have $\delta J = \delta J_1 + \delta J_2 + \delta J_3 = 0$ (51) also for the variation ξ under consideration here, and hence, by (55),

$$|P_\lambda(t_0 + \beta) - P_\lambda(t_0 - \alpha)| \leqq |\delta J_1| + |\delta J_2| + |\psi(\varrho, \alpha, \beta)|$$

and, by (53), (54), (56), $$< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \ ,$$

as asserted in (49), if, given an arbitrary small prescribed ε, the upper limit α_0 of α and β is taken sufficiently small.

But the validity of Theorem III, whose proof is now complete, is based on the assumption that, according to (34) (cf. the observation made there), not only their osculation invariants but the $x^{(\mu)} = \varphi^{(\mu)}(t)$, $y^{(\mu)} = \psi^{(\mu)}(t)$ $(\mu = 0, 1, \ldots r)$ *themselves* are everywhere *continuous* on the entire interval of a under consideration. For otherwise, the variations on which the proof is based oftentimes would no longer belong to the "admissible" ones, and it would no longer be the P_μ, Q_μ whose strict continuity could be asserted but that of *other functions*. But such expressions are obtained, e. g., when for t the *arc length* s of a, or, more generally, a function of s that, together with its first r derivatives, is continuous, is introduced as an independent variable and when, accordingly, the arguments $x^{(\mu)}$, $y^{(\mu)}$ are replaced by $\dfrac{d^\mu x}{ds^\mu}$, $\dfrac{d^\mu y}{ds^\mu}$ in the P_μ, Q_μ, which, in turn, can be expressed in terms of the α_μ as well as the $x^{(\mu)}$, $y^{(\mu)}$. We shall not, however, elaborate on this matter at this point.

Up to now, the number r of the first derivatives $x^{(\mu)}$, $y^{(\mu)}$ whose continuity was demanded in (34) for all curves A has remained *indeterminate*, and Theorems I, II, III have been derived for arbitrary r. Now, however, we shall show that

$$r \geqq n - 1$$

angenommen werden muss, wenn in allgemeineren Fällen *ein Minimum überhaupt möglich* sein soll.

Satz IV. *Für jedes Integral $J = \int F\,dt$ und jedes Curvenstück a lässt sich immer eine solche Variation ξ, η angeben, deren λ te Ableitungen $\xi^{(\lambda)}$, $\eta^{(\lambda)}$ $(\lambda \leqq n-1)$ an einer einzigen Stelle $t = t_0$ endliche Sprünge erleiden*
45 *dürfen, und welche der ersten Variation | einen von* 0 *verschiedenen Wert erteilt und damit ein Minimum unmöglich macht, vorausgesetzt, dass nicht in der ganzen Ausdehnung von a die Functionen P und Q unstetiq oder P_λ und Q_λ immer $= 0$ sind.*

Der Beweis kann auf den Fall beschränkt werden, wo die Forderungen der Sätze II und III erfüllt sind, da sonst die Unmöglichkeit eines Minimums schon feststände.

In einem Teil-Intervall $t' \ldots t''$, in welchem P und Q stetig sind, nach II also überall verschwinden, werde t_0 so gewählt, dass $P_\lambda \gtrless 0$ ist, und folgende Variation betrachtet:

$$\begin{array}{ll} \eta = 0 & (t' \leqq t \leqq t'') \\ \xi = \xi_1(t) = (t-t')^n \varphi_1(t) & (t' \leqq t < t_0) \\ \quad = \xi_2(t) = (t''-t)^n \varphi_2(t) & (t_0 < t \leqq t'')\,, \end{array}$$

wo $\varphi_1(t)$ und $\varphi_2(t)$ *ganze Functionen* sind und so beschaffen, dass ausser

$$\xi_1^{(\mu-1)}(t') = \xi_2^{(\mu-1)}(t'') = 0 \tag{57}$$

noch

$$\xi_1^{(\mu-1)}(t_0) - \xi_2^{(\mu-1)}(t_0) = \xi^{(\mu-1)}(t_0-0) - \xi^{(\mu-1)}(t_0+0) = e_{\mu,\lambda} \quad (\mu = 1, 2, \ldots n)\,, \tag{58}$$

was immer angenommen werden darf, da nach dem für (50) angewandten Verfahren den $\xi_1^{(\mu-1)}(t_0)$ und $\xi_2^{(\mu-1)}(t_0)$ beliebige Werte vorgeschrieben werden können.

Dann wird:

$$\begin{aligned} \delta J &= \int_{t'}^{t_0-0} \delta F\,dt + \int_{t_0+0}^{t''} \delta F\,dt \\ &= \left[\sum_{\mu=1}^{n} P_\mu \xi_1^{(\mu-1)}\right]_{t'}^{t_0-0} + \left[\sum_{\mu=1}^{n} P_\mu \xi_2^{(\mu-1)}\right]_{t_0+0}^{t''}\,, \end{aligned}$$

während die Integrale $\int P\xi\,dt$ wegen $P = 0$ verschwinden. Nach (57) und (58) mit Rücksicht auf die Stetigkeit der P_μ, die nach III auch für Unstetigkeits-

must be assumed for a *a minimum to be possible at all* in the more general cases.

Theorem IV. *For any integral* $J = \int F\,dt$ *and any curve segment a, it is always possible to specify a variation* ξ, η *whose* λ*th derivatives* $\xi^{(\lambda)}$, $\eta^{(\lambda)}$ $(\lambda \leqq n-1)$ *may suffer finite jump discontinuities at a unique point* $t = t_0$*, and which gives the first variation a value different from* 0*, thereby rendering a minimum impossible, provided that the functions* P *and* Q *are not discontinuous on the entire extension of a or* P_λ *and* Q_λ *always* $= 0$.

The proof is capable of restriction to the case in which the demands of the Theorems II and III are satisfied, since otherwise we would already have determined that a minimum is impossible.

Given a partial interval $t' \ldots t''$ on which P and Q are continuous, and hence, by II, vanish everywhere, let us choose t_0 so that $P_\lambda \gtrless 0$ and consider the following variation:

$$\begin{aligned} \eta &= 0 && (t' \leqq t \leqq t'') \\ \xi &= \xi_1(t) = (t-t')^n \varphi_1(t) && (t' \leqq t < t_0) \\ &= \xi_2(t) = (t''-t)^n \varphi_2(t) && (t_0 < t \leqq t'') \end{aligned}$$

where $\varphi_1(t)$ and $\varphi_2(t)$ are *entire functions* constituted so that, in addition to

$$\xi_1^{(\mu-1)}(t') = \xi_2^{(\mu-1)}(t'') = 0\,, \tag{57}$$

also

$$\xi_1^{(\mu-1)}(t_0) - \xi_2^{(\mu-1)}(t_0) = \xi^{(\mu-1)}(t_0-0) - \xi^{(\mu-1)}(t_0+0) = e_{\mu,\lambda} \quad (\mu = 1, 2, \ldots n)\,, \tag{58}$$

which may always be assumed, since, according to the procedure used for (50), it is possible to assign arbitrary values to the $\xi_1^{(\mu-1)}(t_0)$ and $\xi_2^{(\mu-1)}(t_0)$.

Then

$$\begin{aligned} \delta J &= \int_{t'}^{t_0-0} \delta F\,dt + \int_{t_0+0}^{t''} \delta F\,dt \\ &= \left[\sum_{\mu=1}^{n} P_\mu \xi_1^{(\mu-1)}\right]_{t'}^{t_0-0} + \left[\sum_{\mu=1}^{n} P_\mu \xi_2^{(\mu-1)}\right]_{t_0+0}^{t''}\,, \end{aligned}$$

while the integrals $\int P\xi\,dt$ vanish on account of $P = 0$. But, by (57) and (58), and considering the continuity of the P_μ, which, according to III, must be

stellen bestehen bleiben muss, wird dies aber

46 |
$$\delta J = \left[\sum_{\mu=1}^{n} P_\mu \xi^{(\mu-1)}\right]_{t_0+0}^{t_0-0}$$
$$= \sum_{\mu=1}^{n} P_\mu(t_0)\left(\xi^{(\mu-1)}(t_0 - 0) - \xi^{(\mu-1)}(t_0 + 0)\right)$$
$$= \sum_{\mu=1}^{n} P_\mu(t_0) e_{\mu,\lambda} = P_\lambda(t_0) \gtrless 0\ , \qquad \text{q. e. d.}$$

Der Beweis liesse sich noch verschärfen durch Anwendung des für III benutzten umständlicheren Verfahrens, was aber hier ohne Interesse wäre.

Auf Grund dieses Satzes IV wird im Folgenden von allen „erlaubten" Curven die *durchgängige Stetigkeit der Ableitungen* $x^{(\mu)}$, $y^{(\mu)}$ *bis zur* $\mu = r = n-1$ *ten Ordnung* immer vorausgesetzt werden. Berücksichtigt man aber auch hier wieder die Willkürlichkeit der „Darstellung", die von Teil zu Teil wechseln kann, so brauchen thatsächlich nur die Osculations-Invarianten $n-1$ ter Ordnung immer stetig zu sein; die einzelnen Teile müssen einander immer von mindestens $n-1$ ter Ordnung berühren.

Die Untersuchung der „ersten Variation", die zu den Sätzen I bis IV führte, wird in den meisten Fällen zur Auffindung der Curven genügen, die unter den vorgeschriebenen Bedingungen für ein Minimum *überhaupt in Betracht kommen*. Nach Satz II müssen sie im allgemeinen der „*Differentialgleichung des Problems*" $G = 0$ genügen, die von der Ordnung $2n$ ist und ein allgemeines Integral besitzt von der Form:

$$\begin{aligned} x &= \varphi(t; u_1, u_2, \ldots u_{2n}) = \varphi(t, u) \\ y &= \psi(t; u_1, u_2, \ldots u_{2n}) = \psi(t, u)\ , \end{aligned} \tag{59}$$

wo die u_ν $(\nu = 1, 2, \ldots 2n)$ willkürliche Constanten sind und in ihrer Gesamtheit durch u angedeutet werden mögen, während

$$\frac{\partial^\mu \varphi(t,u)}{\partial t^\mu} = \varphi^{(\mu)}(t,u)\ , \qquad \frac{\partial \varphi(t,u)}{\partial u_\nu} = \varphi_\nu(t,u)$$

u. s. w. gesetzt werde.

47 | Wird F als *analytische* Function der $x^{(\mu)}$, $y^{(\mu)}$ vorausgesetzt, so sind auch G und daher auch φ und ψ in Bezug auf ihre sämtlichen Argumente analytische Functionen.

Eine Curve a, die ein Minimum liefern soll, muss nach Satz II aus einer Anzahl *particulärer Lösungen der Differentialgleichung zusammengesetzt* sein, also von der Form (59), wo die u_ν in den einzelnen Teilen constante Werte besitzen. An den *Übergangsstellen* aber, an denen die u_ν endliche Sprünge erleiden können, müssen ausser den Bedingungen (34), wo nach IV $r = n-1$ zu setzen ist, noch gemäss III die P_μ, Q_μ $(\mu = 1, 2, \ldots n)$ stetig bleiben.

preserved also for discontinuities, this becomes

$$\begin{aligned}\delta J &= \left[\sum_{\mu=1}^{n} P_\mu \xi^{(\mu-1)}\right]_{t_0+0}^{t_0-0} \\ &= \sum_{\mu=1}^{n} P_\mu(t_0)\left(\xi^{(\mu-1)}(t_0 - 0) - \xi^{(\mu-1)}(t_0 + 0)\right) \\ &= \sum_{\mu=1}^{n} P_\mu(t_0) e_{\mu,\lambda} = P_\lambda(t_0) \gtrless 0\,, \qquad \text{q. e. d.}\end{aligned}$$

It is possible to further strengthen the proof by applying the more involved method used for III. But this is irrelevant to our present purposes.

Because of Theorem IV, we shall, in what follows, always assume that, given any "admissible" curve, *the derivatives* $x^{(\mu)}$, $y^{(\mu)}$ *up to the* $\mu = r = n-1$ *th order are always continuous.* In fact, taking into account again the arbitrariness of the "representation", which may vary from one part to another, only the osculation invariants of $n-1$ th order need to be always continuous; the individual parts must always have contact of at least $n-1$ th order with one another.

For most cases, the investigation of the "first variation", which led us to Theorems I– IV, suffices in order to find the curves that, given the prescribed conditions for a minimum, *are at all relevant.* Generally, by Theorem II, they must satisfy the "*differential equation of the problem*" $G = 0$, whose order is $2n$ and which possesses a general integral of the form

$$\begin{aligned} x &= \varphi(t; u_1, u_2, \ldots u_{2n}) = \varphi(t, u) \\ y &= \psi(t; u_1, u_2, \ldots u_{2n}) = \psi(t, u)\,, \end{aligned} \tag{59}$$

where the u_ν $(\nu = 1, 2, \ldots 2n)$ are arbitrary constants, which shall, summarily, be denoted by u, while we shall set

$$\frac{\partial^\mu \varphi(t,u)}{\partial t^\mu} = \varphi^{(\mu)}(t,u)\,, \qquad \frac{\partial \varphi(t,u)}{\partial u_\nu} = \varphi_\nu(t,u)$$

e. t. c.

If F is an *analytic* function of the $x^{(\mu)}$, $y^{(\mu)}$, then G, too, and hence also φ and ψ, are analytic functions with respect to all their arguments.

A curve a that is supposed to furnish a minimum must, by Theorem II, be *composed of particular solutions of the differential equation*, and hence must be of the form (59), where the u_ν have constant values in the individual parts. But at the *corners*, at which the u_ν may suffer finite jump discontinuities, the P_μ, Q_μ $(\mu = 1, 2, \ldots n)$ must also remain continuous according to III, in addition to the conditions (34), where, by IV, we have to set $r = n-1$.

Diesen Bedingungen zusammen mit den „Grenzbedingungen“ (32) wird, wie aus der Anzahl der verfügbaren Constanten hervorgeht, im allgemeinen nur auf eine *endliche Anzahl* von Weisen genügt werden können, die nunmehr einzeln untersucht werden müssen.

Unser Problem hat sich also auf die Frage reduciert, *ob ein vorgeschriebenes, den bisherigen Bedingungen genügendes Curvenstück a ein wirkliches Minimum des Integrales liefert.* Dann müssten es notwendig auch seine *einzelnen Teile*, die ja auf Grund von (34) ($r = n - 1$) unter festgehaltenen Osculations-Invarianten $n - 1$ ter Ordnung an den beiden Endpunkten, also unter den Grenzbedingungen (32), einzeln variiert werden können (der Fall $r > n - 1$ soll jetzt nicht mehr berücksichtigt werden), und erst wenn diese Eigenschaft von den einzelnen Teilen bewiesen ist, wird über die Zulässigkeit der Zusammensetzung entschieden werden können. Es kommt also alles an auf die Beantwortung der Frage:

Unter welchen Bedingungen entspricht einem von singulären Punkten freien Stück 1 2 *einer particulären Lösung* (59) *der Differentialgleichung des Problems gemäss der aufgestellten Definition ein wirkliches Minimum des zwischen seinen Grenzen erstreckten Integrales?*

Diese Aufgabe nach der Methode von Herrn Prof. *Weierstrass* zu lösen, soll in den folgenden Abschnitten versucht werden.

48 | Dritter Abschnitt.

Einführung der Function E.

Es sei die Curve a

$$x = \varphi(t) = \varphi(t, a) = \varphi(t; a_1, \ldots a_{2n})$$
$$y = \psi(t) = \psi(t, a) = \psi(t; a_1, \ldots a_{2n})$$

nach (59) ein durch $u_\nu = a_\nu$ ($\nu = 1, 2, \ldots 2n$) charakterisiertes *besonderes Integral der Differentialgleichung des Problems* und möge sich im ganzen Intervall 1 2 ($t_1 \leqq t \leqq t_2$) in der Weise *regulär* verhalten, dass $\varphi^{(\mu)}(t)$, $\psi^{(\mu)}(t)$, sowie auch $\varphi_\nu^{(\mu)}(t)$, $\psi_\nu^{(\mu)}(t)$ ($\mu, \nu = 0, 1, \ldots 2n$) überall eindeutig, endlich und stetig sind, und dass ausserdem noch für $n > 1$ bei constantem γ beständig

$$\varphi'^2(t) + \psi'^2(t) > \gamma^2 > 0 \,. \tag{35}$$

Innerhalb dieses Intervalls werden zwei Punkte 0 und 3 angenommen, für welche t die Werte t_0 und t_3 besitzen möge, also

$$t_1 \leqq t_0 < t_3 \leqq t_2$$

und, ebenso wie von Herrn Prof. *Weierstrass* für den Fall $n = 1$, folgende Variation 043 des Curvenstückes 03 betrachtet: Durch $x = \overline{\varphi}(\lambda)$, $y = \overline{\psi}(\lambda)$

As is evident from the number of available constants, these conditions, together with the "limit conditions" (32), can in general only be satisfied in a *finite number* of ways, each of which shall now be investigated.

Our problem has therefore been reduced to the question as to *whether a prescribed curve segment a satisfying the previous conditions furnishes a real minimum of the integral.* If so, then its *individual parts* would have do to do so as well, each of which can be varied individually because of (34) ($r = n - 1$) with fixed osculation invariants of $n - 1$ th order at the two endpoints, and hence under the limit conditions (32) (we shall no longer consider the case $r > n-1$). Only once this property has been proved for the individual parts will it be possible to decide on the question as to whether the composition is admissible. Hence, it all turns on the question,

What are the conditions under which to a piece 1 2 *of a particular solution* (59) *of the differential equation of the problem without any singular points there corresponds a real minimum of the integral taken between its limits in accordance with the present definition?*

In the following sections, we shall try to solve this problem using the method by Prof. *Weierstrass.*

Third section.

Introduction of the function E.

Suppose that the curve a

$$x = \varphi(t) = \varphi(t, a) = \varphi(t; a_1, \ldots a_{2n})$$
$$y = \psi(t) = \psi(t, a) = \psi(t; a_1, \ldots a_{2n})$$

is a *particular integral of the differential equation* characterized by $u_\nu = a_\nu$ ($\nu = 1, 2, \ldots 2n$) according to (59) and that it is *regular* on the entire interval 1 2 ($t_1 \leqq t \leqq t_2$) such that $\varphi^{(\mu)}(t)$, $\psi^{(\mu)}(t)$, as well as $\varphi_\nu^{(\mu)}(t)$, $\psi_\nu^{(\mu)}(t)$ ($\mu, \nu = 0, 1, \ldots 2n$), are everywhere single-valued, finite and continuous, and, for $n > 1$, with constant γ always

$$\varphi'^2(t) + \psi'^2(t) > \gamma^2 > 0\,. \tag{35}$$

Consider two points 0 and 3 in the interior of this interval for which t takes the values t_0 and t_3 respectively, and hence

$$t_1 \leqq t_0 < t_3 \leqq t_2\,.$$

Like Prof. *Weierstrass* did for the case $n = 1$, we now consider the following variation 043 of the curve segment 03: Suppose that $\overline{a}$ is a *second curve* given

werde eine *zweite Curve* $\overline{a}$ gegeben, welche die erste a im Punkte 3 *von* $n-1$ *ter Ordnung berühren* und sich in der Umgebung dieses Punktes ebenfalls bis zu den Ableitungen n ter Ordnung stetig verhalten soll. Dann lässt sich nach (30) eine mit beliebig vielen Ableitungen stetige Function $\lambda = \lambda(t)$, oder, was auf dasselbe hinauskommt, ein solches System von Werten $\lambda^{(\mu)} = \lambda^{(\mu)}(t_3)$ $(\mu = 0, 1, \ldots n-1; \lambda' > 0)$ angeben, dass für $t = t_3$:

49 |

$$\begin{gathered} \overline{x}_3^{(\mu)} = \overline{\varphi}_1^{(\mu)}(t_3) = D^\mu \overline{\varphi}(\lambda) = \varphi^{(\mu)}(t_3) = x_3^{(\mu)} \\ \overline{y}_3^{(\mu)} = \overline{\psi}_1^{(\mu)}(t_3) = D^\mu \overline{\psi}(\lambda) = \psi^{(\mu)}(t_3) = y_3^{(\mu)} \\ (\mu = 0, 1, \ldots n-1) \, , \end{gathered} \tag{60}$$

wenn

$$\overline{\varphi}(\lambda) = \varphi[\lambda(t)] = \overline{\varphi}_1(t) \, , \quad \overline{\psi}(\lambda) = \psi[\lambda(t)] = \overline{\psi}_1(t) \tag{61}$$

gesetzt wird, wo auch die Functionen $\overline{\varphi}_1$, $\overline{\psi}_1$ sich in der Umgebung von $t = t_3$ regulär verhalten. Diese Functionen $\overline{\varphi}_1$, $\overline{\psi}_1$ und damit die Curve $\overline{a}$ können immer so angenommen werden, dass nicht nur die Bedingungen (60) erfüllt werden, sondern auch die in Bezug auf das Argument t genommenen Ableitungen:

$$\overline{\varphi}_1^{(n)}(t_3) = \overline{x}_3^{(n)} \, , \qquad \overline{\psi}_1^{(n)}(t_3) = \overline{y}_3^{(n)} \tag{60a}$$

beliebig vorgeschriebene, von $\varphi^{(n)}(t_3) = x_3^{(n)}$, $\psi^{(n)}(t_3) = y_3^{(n)}$ verschiedene *Werte annehmen.*

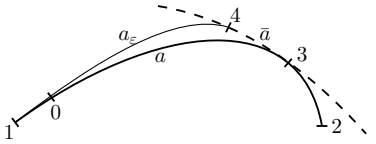


Fig. 2.

Nun werde in der Nähe von 3 auf der Curve $\overline{a}$ nach rückwärts ein Punkt 4 angenommen mit den Coordinaten:

$$\begin{aligned} x = x_4 &= \overline{\varphi}_1(t_3 - \varepsilon) = \overline{\varphi}\,[\lambda(t_3 - \varepsilon)] = \overline{\varphi}(\lambda - \iota) \\ y = y_4 &= \overline{\psi}_1(t_3 - \varepsilon) = \overline{\psi}\,[\lambda(t_3 - \varepsilon)] = \overline{\psi}(\lambda - \iota) \, , \end{aligned} \tag{62}$$

wo

$$\iota = \lambda(t_3) - \lambda(t_3 - \varepsilon) = \lambda'(t_3)\varepsilon + (\varepsilon)_2 = \lambda'\varepsilon + (\varepsilon)_2 \, ,$$

und mit 0 durch eine solche Curve α_ε:

$$\begin{aligned} x = \varphi_\varepsilon(t) &= \varphi(t) + \varepsilon\overline{\varphi}_\varepsilon(t) + (\varepsilon)_2 \, , \\ y = \psi_\varepsilon(t) &= \psi(t) + \varepsilon\overline{\psi}_\varepsilon(t) + (\varepsilon)_2 \end{aligned}$$

by $x = \overline{\varphi}(\lambda)$, $y = \overline{\psi}(\lambda)$ that *makes* $n-1$ *order contact with* the first one a at point 3 and is also continuous in the vicinity of that point up to the derivatives of n th order. Then, by (30), it is possible to specify a function $\lambda = \lambda(t)$ that, together with arbitrarily many derivatives, is continuous, or, what amounts to the same thing, a system of values $\lambda^{(\mu)} = \lambda^{(\mu)}(t_3)$ $(\mu = 0, 1, \ldots n-1; \lambda' > 0)$ such that for $t = t_3$

$$\begin{aligned} \overline{x}_3^{(\mu)} &= \overline{\varphi}_1^{(\mu)}(t_3) = D^\mu \overline{\varphi}(\lambda) = \varphi^{(\mu)}(t_3) = x_3^{(\mu)} \\ \overline{y}_3^{(\mu)} &= \overline{\psi}_1^{(\mu)}(t_3) = D^\mu \overline{\psi}(\lambda) = \psi^{(\mu)}(t_3) = y_3^{(\mu)} \\ &(\mu = 0, 1, \ldots n-1) \, , \end{aligned} \tag{60}$$

if we set

$$\overline{\varphi}(\lambda) = \varphi[\lambda(t)] = \overline{\varphi}_1(t) \, , \quad \overline{\psi}(\lambda) = \psi[\lambda(t)] = \overline{\psi}_1(t) \tag{61}$$

where the functions $\overline{\varphi}_1$, $\overline{\psi}_1$, too, are regular in the vicinity of $t = t_3$. It is always possible to take these functions $\overline{\varphi}_1$, $\overline{\psi}_1$, and hence the curve $\overline{a}$, so that not only the conditions (60) are satisfied, but also so that the derivatives with respect to the argument t,

$$\overline{\varphi}_1^{(n)}(t_3) = \overline{x}_3^{(n)} \, , \qquad \overline{\psi}_1^{(n)}(t_3) = \overline{y}_3^{(n)} \tag{60a}$$

assume arbitrarily prescribed values different from $\varphi^{(n)}(t_3) = x_3^{(n)}$, $\psi^{(n)}(t_3) = y_3^{(n)}$.

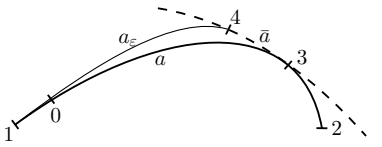


Fig. 2.

Let 4 be a point near 3 back on the curve $\overline{a}$ with the coordinates

$$\begin{aligned} x = x_4 &= \overline{\varphi}_1(t_3 - \varepsilon) = \overline{\varphi}\,[\lambda(t_3 - \varepsilon)] = \overline{\varphi}(\lambda - \iota) \\ y = y_4 &= \overline{\psi}_1(t_3 - \varepsilon) = \overline{\psi}\,[\lambda(t_3 - \varepsilon)] = \overline{\psi}(\lambda - \iota) \, , \end{aligned} \tag{62}$$

where

$$\iota = \lambda(t_3) - \lambda(t_3 - \varepsilon) = \lambda'(t_3)\varepsilon + (\varepsilon)_2 = \lambda'\varepsilon + (\varepsilon)_2 \, ,$$

that is connected with 0 by a curve α_ε given by

$$\begin{aligned} x = \varphi_\varepsilon(t) &= \varphi(t) + \varepsilon\overline{\varphi}_\varepsilon(t) + (\varepsilon)_2 \, , \\ y = \psi_\varepsilon(t) &= \psi(t) + \varepsilon\overline{\psi}_\varepsilon(t) + (\varepsilon)_2 \end{aligned}$$

verbunden, welche a in 0 und $\overline{a}$ in 4 von $n-1$ ter Ordnung berührt, während φ_ε und ψ_ε im ganzen Intervall 0 4 denselben Stetigkeitsbedingungen wie φ und ψ
50 genügen und sich für unendlich | kleine ε auf diese Functionen selbst reducieren. Dann müssen nach (30) für passende Functionen $\vartheta_0(t) = t + \tau_0(t)$, $\vartheta_3(t) = t + \tau_3(t)$ und für die Stellen $t = t_0$, $t = t_3$ die Gleichungen bestehen:

$$\begin{aligned} D^\mu\varphi_\varepsilon(\vartheta_0) &= \varphi^{(\mu)}(t_0) \qquad D^\mu\varphi_\varepsilon(\vartheta_3) = \overline{\varphi}_1^{(\mu)}(t_3 - \varepsilon) \\ D^\mu\psi_\varepsilon(\vartheta_0) &= \psi^{(\mu)}(t_0) \qquad D^\mu\psi_\varepsilon(\vartheta_3) = \overline{\psi}_1^{(\mu)}(t_3 - \varepsilon) \\ &(\mu = 0, 1, \ldots n-1)\,. \end{aligned} \tag{63}$$

Aus den ersten dieser Gleichungen (für $\mu = 0$) ergeben sich bei hinreichend kleinem ε die Grössen ϑ_0 und ϑ_3 eindeutig und beliebig wenig verschieden von t_0 und t_3:

$$\vartheta_0 = t_0 + \tau_0 = t_0 + \varepsilon\overline{\tau}_0 + (\varepsilon)_2\,, \quad \vartheta_3 = t_3 + \tau_3 = t_3 + \varepsilon\overline{\tau}_3 + (\varepsilon)_2\,,$$

wenn *nicht* etwa *die Curve a in* 0 *oder* 3 *sich selbst durchschneidet*, ein Fall, der von der Betrachtung *ausgeschlossen* werden möge. Wenn man nämlich zur Grenze $\varepsilon = 0$ übergeht, so wird nach (60) und (63)

$$\begin{aligned} \lim\varphi_\varepsilon(\vartheta_0) &= \lim\varphi(\vartheta_0) = \varphi(t_0) \\ \lim\psi_\varepsilon(\vartheta_0) &= \lim\psi(\vartheta_0) = \psi(t_0) \\ \lim\varphi_\varepsilon(\vartheta_3) &= \lim\varphi(\vartheta_3) = \overline{\varphi}_1(t_3) = \varphi(t_3) \\ \lim\psi_\varepsilon(\vartheta_3) &= \lim\psi(\vartheta_3) = \overline{\psi}_1(t_3) = \psi(t_3)\,, \end{aligned}$$

und daher (nach der eben gemachten Voraussetzung): $\lim\vartheta_0 = t_0$, $\lim\vartheta_3 = t_3$.

Da nun auch: $\lim\varphi_\varepsilon^{(\mu)}(\vartheta_0) = \lim\varphi^{(\mu)}(\vartheta_0) = \varphi^{(\mu)}(t_0)\,,$

$$\lim\psi_\varepsilon^{(\mu)}(\vartheta_0) = \psi^{(\mu)}(t_0)\,, \quad \lim\varphi_\varepsilon^{(\mu)}(\vartheta_3) = \psi^{(\mu)}(t_3) \quad \text{u. s. w.}$$

und mindestens eine der Grössen $\varphi'(t_0)$, $\psi'(t_0)$ und eine der $\varphi'(t_3)$, $\psi'(t_3)$, für kleine ε also auch eine der $\varphi'_\varepsilon(\vartheta_0)$, $\psi'_\varepsilon(\vartheta_0)$ und eine der $\varphi'_\varepsilon(\vartheta_0)$, $\psi'_\varepsilon(\vartheta_0)$ nicht verschwindet, so liefern die übrigen Gleichungen (63) ($\mu = 1, 2, \ldots n-1$) mit Hilfe von (2) durch successive, lineare Auflösung für die Unbekannten $\vartheta'_0, \vartheta''_0, \ldots \vartheta_0^{(n-1)}, \vartheta'_3, \ldots \vartheta_3^{(n-1)}$ ($\vartheta_0^{(\mu)} = \vartheta_0^{(\mu)}(t_0), \vartheta_3^{(\mu)} = \vartheta_3^{(\mu)}(t_3)$) eindeutig bestimmte Werte von der Beschaffenheit, dass auch die absoluten Beträge der Grössen:

51 | $$\tau_0^{(\mu)} = \vartheta_0^{(\mu)} - e_{\mu,1} = \tau_0^{(\mu)}(t_0)\,, \quad \tau_3^{(\mu)} = \vartheta_3^{(\mu)} - e_{\mu,1} = \tau_3^{(\mu)}(t_3)$$
$$(\mu = 1, 2, \ldots n-1)$$

mit ε gleichzeitig unendlich klein werden, also schliesslich:

$$\tau_0^{(\mu)} = \varepsilon\overline{\tau}_0^{(\mu)} + (\varepsilon)_2\,, \quad \tau_3^{(\mu)} = \varepsilon\overline{\tau}_3^{(\mu)} + (\varepsilon)_2$$
$$(\mu = 0, 1, \ldots n-1)\,.$$

which makes contact of the $n-1$ th order with a at 0 and with $\overline{a}$ at 4, while φ_ε and ψ_ε satisfy the same continuity conditions on the entire interval 0 4 as the functions φ and ψ to which they are reduced for infinitely small ε. Then, by (30), the following equations must hold for suitably chosen functions $\vartheta_0(t) = t + \tau_0(t)$, $\vartheta_3(t) = t + \tau_3(t)$ and for positions $t = t_0$, $t = t_3$:

$$\begin{aligned} D^\mu\varphi_\varepsilon(\vartheta_0) &= \varphi^{(\mu)}(t_0) \qquad D^\mu\varphi_\varepsilon(\vartheta_3) = \overline{\varphi}_1^{(\mu)}(t_3-\varepsilon) \\ D^\mu\psi_\varepsilon(\vartheta_0) &= \psi^{(\mu)}(t_0) \qquad D^\mu\psi_\varepsilon(\vartheta_3) = \overline{\psi}_1^{(\mu)}(t_3-\varepsilon) \\ &(\mu = 0, 1, \ldots n-1)\,. \end{aligned} \tag{63}$$

Given a sufficiently small ε, the first of these equations (for $\mu = 0$) yields the quantities ϑ_0 and ϑ_3 uniquely and of arbitrarily small difference from t_0 and t_3 respectively:

$$\vartheta_0 = t_0 + \tau_0 = t_0 + \varepsilon\overline{\tau}_0 + (\varepsilon)_2\,, \quad \vartheta_3 = t_3 + \tau_3 = t_3 + \varepsilon\overline{\tau}_3 + (\varepsilon)_2\,,$$

unless, say, *the curve a intersects itself at* 0 *or* 3, a case which shall be *excluded* from consideration. For if we take the limit $\varepsilon = 0$, we then have, by (60) and (63) ,

$$\begin{aligned} &\lim\varphi_\varepsilon(\vartheta_0) = \lim\varphi(\vartheta_0) = \varphi(t_0) \\ &\lim\psi_\varepsilon(\vartheta_0) = \lim\psi(\vartheta_0) = \psi(t_0) \\ &\lim\varphi_\varepsilon(\vartheta_3) = \lim\varphi(\vartheta_3) = \overline{\varphi}_1(t_3) = \varphi(t_3) \\ &\lim\psi_\varepsilon(\vartheta_3) = \lim\psi(\vartheta_3) = \overline{\psi}_1(t_3) = \psi(t_3)\,, \end{aligned}$$

and we therefore have (by the assumption just made) $\lim\vartheta_0 = t_0$, $\lim\vartheta_3 = t_3$.

Since now also: $\lim\varphi_\varepsilon^{(\mu)}(\vartheta_0) = \lim\varphi^{(\mu)}(\vartheta_0) = \varphi^{(\mu)}(t_0)$,

$$\lim\psi_\varepsilon^{(\mu)}(\vartheta_0) = \psi^{(\mu)}(t_0)\,, \quad \lim\varphi_\varepsilon^{(\mu)}(\vartheta_3) = \psi^{(\mu)}(t_3) \quad \text{e. t. c.}$$

and at least one of the quantities $\varphi'(t_0)$, $\psi'(t_0)$ and one of the $\varphi'(t_3)$, $\psi'(t_3)$, and hence, for small ε, also one of the $\varphi'_\varepsilon(\vartheta_0)$, $\psi'_\varepsilon(\vartheta_0)$ and one of the $\varphi'_\varepsilon(\vartheta_0)$, $\psi'_\varepsilon(\vartheta_0)$ do not vanish, the successive linear solution of the remaining equations (63) ($\mu = 1, 2, \ldots n-1$) in the unknowns $\vartheta'_0, \vartheta''_0, \ldots \vartheta_0^{(n-1)}, \vartheta'_3, \ldots \vartheta_3^{(n-1)}$ ($\vartheta_0^{(\mu)} = \vartheta_0^{(\mu)}(t_0), \vartheta_3^{(\mu)} = \vartheta_3^{(\mu)}(t_3)$) yields, by use of (2), uniquely determined values such that also the absolute values of the quantities

$$\tau_0^{(\mu)} = \vartheta_0^{(\mu)} - e_{\mu,1} = \tau_0^{(\mu)}(t_0)\,, \quad \tau_3^{(\mu)} = \vartheta_3^{(\mu)} - e_{\mu,1} = \tau_3^{(\mu)}(t_3)$$
$$(\mu = 1, 2, \ldots n-1)$$

become infinitely small together with ε simultaneously, and hence finally:

$$\tau_0^{(\mu)} = \varepsilon\overline{\tau}_0^{(\mu)} + (\varepsilon)_2\,, \quad \tau_3^{(\mu)} = \varepsilon\overline{\tau}_3^{(\mu)} + (\varepsilon)_2$$
$$(\mu = 0, 1, \ldots n-1)\,.$$

Nun kann man eine *einzige* Function:

$$\tau = \tau(t) = \tau(t;\varepsilon) = \varepsilon\overline{\tau}(t) + (\varepsilon)_2$$

bestimmen, welche im Intervall $t_0 \leqq t \leqq t_3$ mit beliebig vielen Ableitungen stetig ist, für $\varepsilon = 0$ sich auf 0 reduciert und den Bedingungen genügt:

$$\begin{gathered}\tau^{(\mu)}(t_0) = \tau_0^{(\mu)}\ , \quad \tau^{(\mu)}(t_3) = \tau_3^{(\mu)} \\ (\mu = 0, 1, \ldots n-1)\ ,\end{gathered} \tag{64}$$

z. B.

$$\tau = \sum_{i=0}^{n-1} \left\{ \tau_0^{(i)} k_i(t; t_0, t_3) + \tau_3^{(i)} k_i(t; t_3, t_0) \right\}\ ,$$

wo die ganze Function

$$k_i(t) = k_i(t; t_0, t_3) = (t - t_3)^n \chi_i(t)$$

nur den Bedingungen zu genügen braucht:

$$\begin{gathered}k_i^{(\mu)}(t_0) = e_{\mu,\,i}\ , \quad k_i^{(\mu)}(t_3) = 0 \\ (\mu = 0, 1, \ldots n-1)\ .\end{gathered}$$

Dies aber kann nach den zu (50) gemachten Bemerkungen durch entsprechende Bestimmung der $\chi_i^{(\mu)}(t_0)$ immer erreicht werden, nämlich durch

$$\chi_i^{(\mu)}(t_0) = D^\mu \left[\frac{(t-t_0)^i}{i!(t-t_3)^n} \right]_{t=t_0} \qquad (\mu = 0, 1, \ldots n-1)\ ,$$

da dann in der That:

$$k_i^{(\mu)}(t_0) = D^\mu \left[(t-t_3)^n \chi_i(t) \right]_{t_0} = D^\mu \left[\frac{(t-t_0)^i}{i!} \right]_{t_0} = e_{\mu,\,i}$$

52 | wird, wir erhalten also:

$$\begin{gathered}k_i(t; t_0, t_3) = (t-t_3)^n \sum_{\mu=0}^{n-1} \frac{(t-t_0)^\mu}{\mu!} D^\mu \left[\frac{(t-t_0)^i}{i!(t-t_3)^n} \right]_{t_0} \\ (i = 0, 1, \ldots n-1)\ ,\end{gathered}$$

ganze Functionen von t und unabhängig von ε, sodass der aus ihnen in der angegebenen Weise gebildete Ausdruck für τ in der That allen gestellten Forderungen genügt.

It is now possible to determine a *unique* function

$$\tau = \tau(t) = \tau(t;\varepsilon) = \varepsilon\overline{\tau}(t) + (\varepsilon)_2$$

that, together with arbitrarily many derivatives, is continuous on the interval $t_0 \leqq t \leqq t_3$ that is reduced to 0 when $\varepsilon = 0$ and that satisfies the conditions

$$\tau^{(\mu)}(t_0) = \tau_0^{(\mu)}\,, \quad \tau^{(\mu)}(t_3) = \tau_3^{(\mu)} \qquad (\mu = 0, 1, \ldots n-1)\,, \tag{64}$$

e. g.,

$$\tau = \sum_{i=0}^{n-1} \left\{ \tau_0^{(i)} k_i(t; t_0, t_3) + \tau_3^{(i)} k_i(t; t_3, t_0) \right\}\,,$$

where the entire function

$$k_i(t) = k_i(t; t_0, t_3) = (t - t_3)^n \chi_i(t)$$

only needs to satisfy the conditions

$$k_i^{(\mu)}(t_0) = e_{\mu,\,i}\,, \quad k_i^{(\mu)}(t_3) = 0 \qquad (\mu = 0, 1, \ldots n-1)\,.$$

But, according to the remarks made in connection with (50), this can always be attained by means of an appropriate determination of the $\chi_i^{(\mu)}(t_0)$, namely by means of

$$\chi_i^{(\mu)}(t_0) = D^\mu \left[\frac{(t-t_0)^i}{i!(t-t_3)^n} \right]_{t\,=\,t_0} \qquad (\mu = 0, 1, \ldots n-1)\,,$$

for then, in fact,

$$k_i^{(\mu)}(t_0) = D^\mu \left[(t-t_3)^n \chi_i(t) \right]_{t_0} = D^\mu \left[\frac{(t-t_0)^i}{i!} \right]_{t_0} = e_{\mu,\,i}\,,$$

and hence we obtain

$$k_i(t; t_0, t_3) = (t-t_3)^n \sum_{\mu\,=\,0}^{n-1} \frac{(t-t_0)^\mu}{\mu!} D^\mu \left[\frac{(t-t_0)^i}{i!(t-t_3)^n} \right]_{t_0} \qquad (i = 0, 1, \ldots n-1)\,,$$

entire functions of t independent of ε so that the expression for τ formed from them according to the manner specified really meets all requirements posed here.

Mit Hilfe von (64) lassen sich jetzt die Gleichungen (63) in der Form schreiben:

$$\left.\begin{aligned} D^\mu\varphi_\varepsilon(t+\tau) &= \varphi^{(\mu)}(t_0) \\ D^\mu\psi_\varepsilon(t+\tau) &= \psi^{(\mu)}(t_0) \end{aligned}\right\}\ (t=t_0)$$
$$\left.\begin{aligned} D^\mu\varphi_\varepsilon(t+\tau) &= \overline{\varphi}_1^{(\mu)}(t_3-\varepsilon) \\ D^\mu\psi_\varepsilon(t+\tau) &= \overline{\psi}_1^{(\mu)}(t_3-\varepsilon) \end{aligned}\right\}\ (t=t_3) \qquad (65)$$
$$(\mu = 0, 1, \ldots n-1)\ ,$$

und wenn man folgende gestattete Darstellung der Curve a_ε einführt:

$$\begin{aligned} x = \varphi_\varepsilon(t+\tau) &= \varphi(t) + \xi(t;\varepsilon) = \varphi(t) + \varepsilon\overline{\xi}(t) + (\varepsilon)_2 \\ y = \psi_\varepsilon(t+\tau) &= \psi(t) + \eta(t;\varepsilon) = \psi(t) + \varepsilon\overline{\eta}(t) + (\varepsilon)_2 \end{aligned} \qquad (66)$$
$$(t_0 \leqq t \leqq t_3)\ ,$$

mit Benutzung von (60) auch folgendermaassen:

$$\begin{aligned} &\xi^{(\mu)}(t_0;\varepsilon) = 0 \\ &\eta^{(\mu)}(t_0;\varepsilon) = 0 \\ &\xi^{(\mu)}(t_3;\varepsilon) = \overline{\varphi}_1^{(\mu)}(t_3-\varepsilon) - \varphi^{(\mu)}(t_3) = \overline{\varphi}_1^{(\mu)}(t_3-\varepsilon) - \overline{\varphi}_1^{(\mu)}(t_3) \\ &\qquad\qquad = -\varepsilon\overline{\varphi}_1^{(\mu+1)}(t_3) + (\varepsilon)_2 \qquad (67) \\ &\eta^{(\mu)}(t_3;\varepsilon) = \overline{\psi}_1^{(\mu)}(t_3-\varepsilon) - \psi^{(\mu)}(t_3) = \overline{\psi}_1^{(\mu)}(t_3-\varepsilon) - \overline{\psi}_1^{(\mu)}(t_3) \\ &\qquad\qquad = -\varepsilon\overline{\psi}_1^{(\mu+1)}(t_3) + (\varepsilon)_2 \\ &\qquad (\mu = 0, 1, \ldots n-1)\ , \end{aligned}$$

53 | schliesslich aber, durch Vergleichung der Coefficienten von ε:

$$\begin{aligned} &\overline{\xi}^{(\mu-1)}(t_0) = 0\ , \qquad \overline{\eta}^{(\mu-1)}(t_0) = 0\ , \\ &\overline{\xi}^{(\mu-1)}(t_3) = -\overline{x}_3^{(\mu)} = -x_3^{(\mu)} - e_{\mu,n}\left(\overline{x}_3^{(n)} - x_3^{(n)}\right)\ , \qquad (68) \\ &\overline{\eta}^{(\mu-1)}(t_3) = -\overline{y}_3^{(\mu)} = -y_3^{(\mu)} - e_{\mu,n}\left(\overline{y}_3^{(n)} - y_3^{(n)}\right) \\ &\qquad (\mu = 1, 2, \ldots n)\ . \end{aligned}$$

Den Gleichungen (67) kann durch passende Functionen $\xi(t;\varepsilon)$ und $\eta(t;\varepsilon)$ immer genügt werden, nach den Betrachtungen von (64) z. B. durch:

$$\xi(t,\varepsilon) = \sum_{\mu=0}^{n-1} \xi^{(\mu)}(t_3,\varepsilon) k_\mu(t;t_3,t_0)$$
$$\eta(t,\varepsilon) = \sum_{\mu=0}^{n-1} \eta^{(\mu)}(t_3,\varepsilon) k_\mu(t;t_3,t_0)$$

By means of (64), we may now write the equations (63) in the form

$$\left.\begin{aligned} D^\mu\varphi_\varepsilon(t+\tau) &= \varphi^{(\mu)}(t_0) \\ D^\mu\psi_\varepsilon(t+\tau) &= \psi^{(\mu)}(t_0) \end{aligned}\right\}\ (t=t_0)$$
$$\left.\begin{aligned} D^\mu\varphi_\varepsilon(t+\tau) &= \overline{\varphi}_1^{(\mu)}(t_3-\varepsilon) \\ D^\mu\psi_\varepsilon(t+\tau) &= \overline{\psi}_1^{(\mu)}(t_3-\varepsilon) \end{aligned}\right\}\ (t=t_3) \tag{65}$$
$$(\mu=0,1,\dots n-1)\ ,$$

and, by introducing the following admissible representation of the curve a_ε:

$$\begin{aligned} x=\varphi_\varepsilon(t+\tau)&=\varphi(t)+\xi(t;\varepsilon)=\varphi(t)+\varepsilon\overline{\xi}(t)+(\varepsilon)_2 \\ y=\psi_\varepsilon(t+\tau)&=\psi(t)+\eta(t;\varepsilon)=\psi(t)+\varepsilon\overline{\eta}(t)+(\varepsilon)_2 \end{aligned} \tag{66}$$
$$(t_0\leqq t\leqq t_3)\ ,$$

by use of (60), also as follows:

$$\begin{aligned} &\xi^{(\mu)}(t_0;\varepsilon)=0 \\ &\eta^{(\mu)}(t_0;\varepsilon)=0 \\ &\xi^{(\mu)}(t_3;\varepsilon)=\overline{\varphi}_1^{(\mu)}(t_3-\varepsilon)-\varphi^{(\mu)}(t_3)=\overline{\varphi}_1^{(\mu)}(t_3-\varepsilon)-\overline{\varphi}_1^{(\mu)}(t_3) \\ &\qquad\qquad =-\varepsilon\overline{\varphi}_1^{(\mu+1)}(t_3)+(\varepsilon)_2 \\ &\eta^{(\mu)}(t_3;\varepsilon)=\overline{\psi}_1^{(\mu)}(t_3-\varepsilon)-\psi^{(\mu)}(t_3)=\overline{\psi}_1^{(\mu)}(t_3-\varepsilon)-\overline{\psi}_1^{(\mu)}(t_3) \\ &\qquad\qquad =-\varepsilon\overline{\psi}_1^{(\mu+1)}(t_3)+(\varepsilon)_2 \end{aligned} \tag{67}$$
$$(\mu=0,1,\dots n-1)\ ,$$

and finally, by comparison of the coefficients of ε:

$$\begin{aligned} &\overline{\xi}^{(\mu-1)}(t_0)=0\ ,\qquad \overline{\eta}^{(\mu-1)}(t_0)=0\ , \\ &\overline{\xi}^{(\mu-1)}(t_3)=-\overline{x}_3^{(\mu)}=-x_3^{(\mu)}-e_{\mu,n}\left(\overline{x}_3^{(n)}-x_3^{(n)}\right)\ , \\ &\overline{\eta}^{(\mu-1)}(t_3)=-\overline{y}_3^{(\mu)}=-y_3^{(\mu)}-e_{\mu,n}\left(\overline{y}_3^{(n)}-y_3^{(n)}\right) \end{aligned} \tag{68}$$
$$(\mu=1,2,\dots n)\ .$$

It is always possible to satisfy the equations (67) by means of suitable functions $\xi(t;\varepsilon)$ and $\eta(t;\varepsilon)$; for example, according to the considerations of (64),

$$\xi(t,\varepsilon)=\sum_{\mu=0}^{n-1}\xi^{(\mu)}(t_3,\varepsilon)k_\mu(t;t_3,t_0)$$
$$\eta(t,\varepsilon)=\sum_{\mu=0}^{n-1}\eta^{(\mu)}(t_3,\varepsilon)k_\mu(t;t_3,t_0)$$

und dann liefert der Ansatz:

$$x = \varphi(t) + \xi(t,\varepsilon) = \varphi_\varepsilon(t)\ ,$$
$$y = \psi(t) + \eta(t,\varepsilon) = \psi_\varepsilon(t)$$

gemäss (66) (für $\tau = 0$) jedenfalls eine Curve a_ε von der verlangten Beschaffenheit, womit die *Ausführbarkeit* der vorausgesetzten Constructionen erwiesen ist.

Jetzt werden die über 0 3 und 0 4 erstreckten *Integrale*:

$$J_{03} = \int_{t_0}^{t_3} F\left(\varphi^{(\mu)}(t), \psi^{(\mu)}(t)\right)\, dt$$

und nach (66):

$$J_{04} = \int_{t_0}^{t_3} F\left(\varphi^{(\mu)}(t) + \varepsilon\overline{\xi}^{(\mu)}(t) + (\varepsilon)_2\ ,\ \ \psi^{(\mu)}(t) + \varepsilon\overline{\eta}^{(\mu)}(t) + (\varepsilon)_2\right) dt \qquad (69)$$
$$= J_{03} + \varepsilon\delta J_{03} + (\varepsilon)_2\ .$$

Hier ist nach (41):

54 |
$$\delta J_{03} = \int_{t_0}^{t_3} \sum_{\mu=0}^{n} \left(X_\mu \overline{\xi}^{(\mu)} + Y_\mu \overline{\eta}^{(\mu)}\right) dt$$

$$\left(x^{(\mu)} = \varphi^{(\mu)}(t), y^{(\mu)} = \psi^{(\mu)}(t), X_\mu = \frac{\partial F}{\partial x^{(\mu)}}\ \text{u. s. w.}\right)$$

oder, da unter den gemachten Voraussetzungen die Umformung (42) immer zulässig ist:

$$\delta J_{03} = \left[\sum_{\mu=1}^{n}\left(P_\mu \overline{\xi}^{(\mu-1)} + Q_\mu \overline{\eta}^{(\mu-1)}\right)\right]_{t_0}^{t_3} + \int_{t_0}^{t_3} (P\overline{\xi} + Q\overline{\eta})\, dt\ .$$

Da a der Differentialgleichung des Problems genügen soll, so verschwindet wegen $P = 0$ und $Q = 0$ das zweite Glied, und es wird nun wegen (68):

$$\delta J_{03} = -\sum_{\mu=1}^{n}\left(P_\mu(t_3)\overline{x}_3^{(\mu)} + Q_\mu(t_3)\overline{y}_3^{(\mu)}\right)$$
$$= -\sum_{\mu=1}^{n}\left[P_\mu x^{(\mu)} + Q_\mu y^{(\mu)}\right]_{t=t_3} - P_n(t_3)\left(\overline{x}_3^{(n)} - x_3^{(n)}\right)$$
$$- Q_n(t_3)\left(\overline{y}_3^{(n)} - y_3^{(n)}\right)$$

and the ansatz

$$x = \varphi(t) + \xi(t, \varepsilon) = \varphi_\varepsilon(t) \ ,$$
$$y = \psi(t) + \eta(t, \varepsilon) = \psi_\varepsilon(t) \ ,$$

in accordance with (66) (when $\tau = 0$), then certainly furnishes a curve a_ε constituted as required, whereby we have shown that it is *possible to carry out* the assumed constructions.

The *integrals* taken along 0 3 and 0 4 now become

$$J_{03} = \int_{t_0}^{t_3} F\left(\varphi^{(\mu)}(t), \psi^{(\mu)}(t)\right) dt$$

and, by (66),

$$J_{04} = \int_{t_0}^{t_3} F\left(\varphi^{(\mu)}(t) + \varepsilon\overline{\xi}^{(\mu)}(t) + (\varepsilon)_2 \ , \ \ \psi^{(\mu)}(t) + \varepsilon\overline{\eta}^{(\mu)}(t) + (\varepsilon)_2\right) dt \qquad (69)$$
$$= J_{03} + \varepsilon\delta J_{03} + (\varepsilon)_2 \ .$$

Here, by (41),

$$\delta J_{03} = \int_{t_0}^{t_3} \sum_{\mu=0}^{n} \left(X_\mu \overline{\xi}^{(\mu)} + Y_\mu \overline{\eta}^{(\mu)}\right) dt$$

$$\left(x^{(\mu)} = \varphi^{(\mu)}(t), y^{(\mu)} = \psi^{(\mu)}(t), X_\mu = \frac{\partial F}{\partial x^{(\mu)}} \text{ u. s. w.}\right)$$

or, since, by our assumptions, the transformation (42) is always admissible,

$$\delta J_{03} = \left[\sum_{\mu=1}^{n} \left(P_\mu \overline{\xi}^{(\mu-1)} + Q_\mu \overline{\eta}^{(\mu-1)}\right)\right]_{t_0}^{t_3} + \int_{t_0}^{t_3} (P\overline{\xi} + Q\overline{\eta})\, dt \ .$$

Since a is supposed to satisfy the differential equation of the problem, the second term vanishes on account of $P = 0$ and $Q = 0$, and now, on account of (68),

$$\delta J_{03} = -\sum_{\mu=1}^{n} \left(P_\mu(t_3)\overline{x}_3^{(\mu)} + Q_\mu(t_3)\overline{y}_3^{(\mu)}\right)$$
$$= -\sum_{\mu=1}^{n} \left[P_\mu x^{(\mu)} + Q_\mu y^{(\mu)}\right]_{t=t_3} - P_n(t_3)\left(\overline{x}_3^{(n)} - x_3^{(n)}\right)$$
$$- Q_n(t_3)\left(\overline{y}_3^{(n)} - y_3^{(n)}\right)$$

oder, mit Benutzung von $(18)_1$ und (8):

$$\delta J_{03} = -F(t_3) - X_n(t_3)\left(\overline{x}_3^{(n)} - x_3^{(n)}\right) - Y_n(t_3)\left(\overline{y}_3^{(n)} - y_3^{(n)}\right) . \qquad (70)$$

Das über 4 3 erstreckte Integral aber wird:

$$J_{43} = \int_{t_3-\varepsilon}^{t_3} F\left(\overline{\varphi}_1^{(\mu)}(t), \overline{\psi}_1^{(\mu)}(t)\right) dt \qquad (71)$$

gemäss der Form (61) für $\overline{a}$, oder

$$\begin{aligned} J_{43} &= \varepsilon F\left(\overline{\varphi}_1^{(\mu)}(t_3), \overline{\psi}_1^{(\mu)}(t_3)\right) + (\varepsilon)_2 \\ &= \varepsilon F\left(\overline{x}_3^{(\mu)}, \overline{y}_3^{(\mu)}\right) + (\varepsilon)_2 , \end{aligned}$$

55 | so dass schliesslich wegen (69), (70) und (71):

$$\begin{aligned} J_{04} - J_{03} + J_{43} &= J_{043} - J_{03} \qquad (72) \\ &= \varepsilon E\left(x_3^{(\mu)}, y_3^{(\mu)}; \overline{x}_3^{(n)}, \overline{y}_3^{(n)}\right) + (\varepsilon)_2 , \end{aligned}$$

oder nach (62)

$$= \frac{\iota}{\lambda'} E\left(x_3^{(\mu)}, y_3^{(\mu)}; \overline{x}_3^{(n)}, \overline{y}_3^{(n)}\right) + (\iota)_2 ,$$

wenn gesetzt wird:

$$\begin{aligned} E\big(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(n)}, \overline{y}^{(n)}\big) &= \overline{F} - F - X_n\big(\overline{x}^{(n)} - x^{(n)}\big) - Y_n\big(\overline{y}^{(n)} - y^{(n)}\big) \qquad (73) \\ &= F\big(\overline{x}^{(\mu)}, \overline{y}^{(\mu)}\big) - F\big(x^{(\mu)}, y^{(\mu)}\big) - X_n\big(x^{(\mu)}, y^{(\mu)}\big)\big(\overline{x}^{(n)} - x^{(n)}\big) \\ &\qquad - Y_n\big(x^{(\mu)}, y^{(\mu)}\big)\big(\overline{y}^{(n)} - y^{(n)}\big) , \end{aligned}$$

wo nach (62):

$$\overline{x}^{(\mu)} = x^{(\mu)},\ \overline{y}^{(\mu)} = y^{(\mu)} \qquad (\mu = 0, 1, \ldots n-1) ,$$

also:

$$\begin{aligned} \overline{F} &= F\left(\overline{x}^{(\mu)}, \overline{y}^{(\mu)}\right) \\ &= F\left(x, x', \ldots x^{(n-1)}, \overline{x}^{(n)};\ y, y', \ldots y^{(n-1)}, \overline{y}^{(n)}\right) . \end{aligned}$$

Lässt man hier die Argumente $x^{(\mu)}$, $y^{(\mu)}$ $(\mu = 0, 1, \ldots n-1)$ überall fort und ersetzt:

$$x^{(n)}, y^{(n)}; \quad \overline{x}^{(n)}, \overline{y}^{(n)}$$

or, by use of $(18)_1$ and (8),

$$\delta J_{03} = -F(t_3) - X_n(t_3)\left(\overline{x}_3^{(n)} - x_3^{(n)}\right) - Y_n(t_3)\left(\overline{y}_3^{(n)} - y_3^{(n)}\right) . \tag{70}$$

The integral taken along 4 3, however, becomes

$$J_{43} = \int_{t_3-\varepsilon}^{t_3} F\left(\overline{\varphi}_1^{(\mu)}(t), \overline{\psi}_1^{(\mu)}(t)\right) dt \tag{71}$$

in accordance with the form (61) for $\overline{a}$, or

$$\begin{aligned} J_{43} &= \varepsilon F\left(\overline{\varphi}_1^{(\mu)}(t_3), \overline{\psi}_1^{(\mu)}(t_3)\right) + (\varepsilon)_2 \\ &= \varepsilon F\left(\overline{x}_3^{(\mu)}, \overline{y}_3^{(\mu)}\right) + (\varepsilon)_2 , \end{aligned}$$

so that, finally, on account of (69), (70) and (71),

$$\begin{aligned} J_{04} - J_{03} + J_{43} &= J_{043} - J_{03} \\ &= \varepsilon E\left(x_3^{(\mu)}, y_3^{(\mu)}; \overline{x}_3^{(n)}, \overline{y}_3^{(n)}\right) + (\varepsilon)_2 , \end{aligned} \tag{72}$$

or, by (62),

$$= \frac{\iota}{\lambda'} E\left(x_3^{(\mu)}, y_3^{(\mu)}; \overline{x}_3^{(n)}, \overline{y}_3^{(n)}\right) + (\iota)_2 ,$$

if setting

$$\begin{aligned} E\big(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(n)}, \overline{y}^{(n)}\big) &= \overline{F} - F - X_n\big(\overline{x}^{(n)} - x^{(n)}\big) - Y_n\big(\overline{y}^{(n)} - y^{(n)}\big) \\ &= F\big(\overline{x}^{(\mu)}, \overline{y}^{(\mu)}\big) - F\big(x^{(\mu)}, y^{(\mu)}\big) - X_n\big(x^{(\mu)}, y^{(\mu)}\big)\big(\overline{x}^{(n)} - x^{(n)}\big) \\ &\qquad - Y_n\big(x^{(\mu)}, y^{(\mu)}\big)\big(\overline{y}^{(n)} - y^{(n)}\big) , \end{aligned} \tag{73}$$

where, by (62),

$$\overline{x}^{(\mu)} = x^{(\mu)},\ \overline{y}^{(\mu)} = y^{(\mu)} \qquad (\mu = 0, 1, \ldots n-1) ,$$

and hence

$$\begin{aligned} \overline{F} &= F\left(\overline{x}^{(\mu)}, \overline{y}^{(\mu)}\right) \\ &= F\left(x, x', \ldots x^{(n-1)}, \overline{x}^{(n)};\ y, y', \ldots y^{(n-1)}, \overline{y}^{(n)}\right) . \end{aligned}$$

If we now drop the arguments $x^{(\mu)}$, $y^{(\mu)}$ $(\mu = 0, 1, \ldots n-1)$ everywhere and replace

$$x^{(n)}, y^{(n)}; \quad \overline{x}^{(n)}, \overline{y}^{(n)}$$

der Reihe nach durch

$$p\,,\quad q\,;\quad \overline{p}\,,\quad \overline{q}\,,$$

so kann man abgekürzt schreiben:

$$E(p,q;\overline{p},\overline{q}) = \tag{73a}$$
$$F(\overline{p},\overline{q}) - F(p,q) - \frac{\partial F(p,q)}{\partial p}(\overline{p}-p) - \frac{\partial F(p,q)}{\partial q}(\overline{q}-q)\,.$$

Die Gleichung (72) entspricht genau der von Herrn Prof. *Weierstrass* für den Fall $n=1$ abgeleiteten, ebenso auch die Formel (73) oder (73a) einem der von ihm für die Function E aufgestellten Ausdrucke.

56 | Nach (60a) ist hier mit Hilfe von (2):

$$\overline{x}^{(n)} = \overline{\varphi}_1^{(n)}(t_3) = D^n\overline{\varphi}(\lambda) = R_n\left(\overline{\varphi}^{(\mu)}(\lambda),\lambda^{(\mu)}\right)$$
$$\overline{y}^{(n)} = \overline{\psi}_1^{(n)}(t_3) = D^n\overline{\psi}(\lambda) = R_n\left(\overline{\psi}^{(\mu)}(\lambda),\lambda^{(\mu)}\right)\,,$$

also abhängig von den Grössen $\lambda,\lambda',\ldots\lambda^{(n)}$, von denen durch (60) nur die ersten $\lambda,\lambda',\ldots\lambda^{(n-1)}$ bestimmt werden, während $\lambda^{(n)}$ noch gänzlich willkürlich bleibt. Da aber E seiner geometrischen Bedeutung wegen einen *bestimmten* Wert besitzen muss, so wird es *nur scheinbar* von $\lambda^{(n)}$ abhängen können und seine nach $\lambda^{(n)}$ genommene Ableitung identisch verschwinden müssen.

In der That wird nach (2) mit Hilfe von (60)

$$\frac{\partial\overline{x}^{(n)}}{\partial\lambda^{(n)}} = \overline{\varphi}'(\lambda) = \frac{\varphi'(t_3)}{\lambda'} = \frac{x'}{\lambda'}\,,$$
$$\frac{\partial\overline{y}^{(n)}}{\partial\lambda^{(n)}} = \overline{\psi}'(\lambda) = \frac{\psi'(t_3)}{\lambda'} = \frac{y'}{\lambda'}\,,$$

also nach (73):

$$\frac{\partial E}{\partial\lambda^{(n)}} = \frac{\partial\overline{F}}{\partial\lambda^{(n)}} - \frac{\partial\overline{x}^{(n)}}{\partial\lambda^{(n)}}\frac{\partial F}{\partial x^{(n)}} - \frac{\partial\overline{y}^{(n)}}{\partial\lambda^{(n)}}\frac{\partial F}{\partial y^{(n)}} = 0\,, \tag{74}$$

weil nach $(17)_n$:

$$\frac{\partial F}{\partial x^{(n)}}\frac{\partial\overline{x}^{(n)}}{\partial\lambda^{(n)}} + \frac{\partial F}{\partial y^{(n)}}\frac{\partial\overline{y}^{(n)}}{\partial\lambda^{(n)}} = \frac{1}{\lambda'}\left(\frac{\partial F}{\partial x^{(n)}}x' + \frac{\partial F}{\partial y^{(n)}}y'\right) = 0$$

und:

$$\frac{\partial\overline{F}}{\partial\lambda^{(n)}} = \frac{\partial\overline{F}}{\partial\overline{x}^{(n)}}\frac{\partial\overline{x}^{(n)}}{\partial\lambda^{(n)}} + \frac{\partial\overline{F}}{\partial\overline{y}^{(n)}}\frac{\partial\overline{y}^{(n)}}{\partial\lambda^{(n)}} = \frac{1}{\lambda'}\left(\frac{\partial\overline{F}}{\partial\overline{x}^{(n)}}x' + \frac{\partial\overline{F}}{\partial\overline{y}^{(n)}}y'\right) = 0\,.$$

successively by

$$p\,,\quad q\,;\quad \overline{p}\,,\quad \overline{q}\,,$$

we then may use the abbreviation

$$E(p,q;\overline{p},\overline{q}) = \tag{73a}$$
$$F(\overline{p},\overline{q}) - F(p,q) - \frac{\partial F(p,q)}{\partial p}(\overline{p}-p) - \frac{\partial F(p,q)}{\partial q}(\overline{q}-q)\,.$$

The equation (72) precisely corresponds to the one derived by Prof. *Weierstrass* for the case $n=1$, as does the formula (73) or (73a) to an expression used by him for the function E.

By (60a), we now have by means of (2)

$$\overline{x}^{(n)} = \overline{\varphi}_1^{(n)}(t_3) = D^n\overline{\varphi}(\lambda) = R_n\left(\overline{\varphi}^{(\mu)}(\lambda), \lambda^{(\mu)}\right)$$
$$\overline{y}^{(n)} = \overline{\psi}_1^{(n)}(t_3) = D^n\overline{\psi}(\lambda) = R_n\left(\overline{\psi}^{(\mu)}(\lambda), \lambda^{(\mu)}\right)\,,$$

and it thus depends on the quantities $\lambda, \lambda', \ldots \lambda^{(n)}$, of which, by (60), only the first $\lambda, \lambda', \ldots \lambda^{(n-1)}$, are determined, whereas $\lambda^{(n)}$ is still entirely arbitrary. But since E must have a *determinate* value on account of its geometric meaning, it can *only appear* to depend on $\lambda^{(n)}$ and its derivative with respect to $\lambda^{(n)}$ must vanish identically.

Indeed, by (2), we have by means of (60)

$$\frac{\partial\overline{x}^{(n)}}{\partial\lambda^{(n)}} = \overline{\varphi}'(\lambda) = \frac{\varphi'(t_3)}{\lambda'} = \frac{x'}{\lambda'}\,,$$
$$\frac{\partial\overline{y}^{(n)}}{\partial\lambda^{(n)}} = \overline{\psi}'(\lambda) = \frac{\psi'(t_3)}{\lambda'} = \frac{y'}{\lambda'}\,,$$

and hence, by (73),

$$\frac{\partial E}{\partial\lambda^{(n)}} = \frac{\partial\overline{F}}{\partial\lambda^{(n)}} - \frac{\partial\overline{x}^{(n)}}{\partial\lambda^{(n)}}\frac{\partial F}{\partial x^{(n)}} - \frac{\partial\overline{y}^{(n)}}{\partial\lambda^{(n)}}\frac{\partial F}{\partial y^{(n)}} = 0\,, \tag{74}$$

since, by $(17)_n$,

$$\frac{\partial F}{\partial x^{(n)}}\frac{\partial\overline{x}^{(n)}}{\partial\lambda^{(n)}} + \frac{\partial F}{\partial y^{(n)}}\frac{\partial\overline{y}^{(n)}}{\partial\lambda^{(n)}} = \frac{1}{\lambda'}\left(\frac{\partial F}{\partial x^{(n)}}x' + \frac{\partial F}{\partial y^{(n)}}y'\right) = 0$$

and

$$\frac{\partial\overline{F}}{\partial\lambda^{(n)}} = \frac{\partial\overline{F}}{\partial\overline{x}^{(n)}}\frac{\partial\overline{x}^{(n)}}{\partial\lambda^{(n)}} + \frac{\partial\overline{F}}{\partial\overline{y}^{(n)}}\frac{\partial\overline{y}^{(n)}}{\partial\lambda^{(n)}} = \frac{1}{\lambda'}\left(\frac{\partial\overline{F}}{\partial\overline{x}^{(n)}}x' + \frac{\partial\overline{F}}{\partial\overline{y}^{(n)}}y'\right) = 0\,.$$

Sei jetzt $\varrho = \varrho(t) = \varrho(t;\varepsilon) = \varepsilon\overline{\varrho}(t) + (\varepsilon)_2$ eine für alle Werte von t mit ε gleichzeitig unendlich klein werdende ganze Function von t, die den Bedingungen genügt:

$$\varrho^{(\mu)}(t_0) = \tau_0^{(\mu)} , \qquad \varrho^{(\mu)}(t_3 - \varepsilon) = \tau_3^{(\mu)} + e_{\mu,0}\,\varepsilon$$
$$(\mu = 0, 1, \ldots n-1) ,$$

57 wo $\tau_0^{(\mu)}$, $\tau_3^{(\mu)}$ dieselben Bedeutungen wie in (64) haben; eine solche | Function kann nach dem Verfahren von (64) immer bestimmt werden, nämlich

$$\varrho = \sum_{\mu=0}^{n-1} \left\{ \tau_0^{(\mu)} k_\mu(t; t_0, t_3 - \varepsilon) + \left(\tau_3^{(\mu)} + e_{\mu,0}\,\varepsilon\right) k_\mu(t; t_3 - \varepsilon,\, t_0) \right\} .$$

Dann werden die Gleichungen (65) übergehen in die folgenden:

$$\begin{aligned}
[D^\mu \varphi_\varepsilon(t+\varrho)]_{t=t_0} &= [D^\mu \varphi_\varepsilon(t+\tau)]_{t_0} = \varphi^{(\mu)}(t_0) \\
[D^\mu \psi_\varepsilon(t+\varrho)]_{t_0} &= [D^\mu \psi_\varepsilon(t+\tau)]_{t_0} = \psi^{(\mu)}(t_0) \\
[D^\mu \varphi_\varepsilon(t+\varrho)]_{t_3-\varepsilon} &= [D^\mu \varphi_\varepsilon(t+\tau)]_{t_3} = \overline{\varphi}_1^{(\mu)}(t_3 - \varepsilon) \\
[D^\mu \psi_\varepsilon(t+\varrho)]_{t_3-\varepsilon} &= [D^\mu \psi_\varepsilon(t+\tau)]_{t_3} = \overline{\psi}_1^{(\mu)}(t_3 - \varepsilon) \\
&(\mu = 0, 1, \ldots n-1) .
\end{aligned} \tag{75}$$

Ist daher, wie vorausgesetzt,

$$t_1 \leqq t_0 < t_3 \leqq t_2 ,$$

so kann der zusammengesetzte Linienzug 1 0 4 3 2 ausgedrückt werden durch die Gleichungen:

$$\begin{aligned}
&x = \varphi(t) , && y = \psi(t) && (t_1 \leqq t \leqq t_0) \\
&x = \varphi_\varepsilon(t+\varrho) , && y = \psi_\varepsilon(t+\varrho) && (t_0 \leqq t \leqq t_3 - \varepsilon) \\
&x = \overline{\varphi}_1(t) , && y = \overline{\psi}_1(t) && (t_3 - \varepsilon \leqq t \leqq t_3) \\
&x = \varphi(t) , && y = \psi(t) && (t_3 \leqq t \leqq t_2) ,
\end{aligned} \tag{76}$$

wo die Grössen

$$x^{(\mu)} = \frac{d^\mu x}{dt^\mu} , \quad y^{(\mu)} = \frac{d^\mu y}{dt^\mu} \qquad (\mu = 0, 1, \ldots n-1)$$

wegen (60) und (75) auch an den Übergangsstellen $t = t_0$, $t = t_3 - \varepsilon$, $t = t_3$ der Curven a, a_ε und $\overline{a}$ immer *stetig* bleiben und wo die entsprechenden absoluten Beträge

58 | $$\left| x^{(\mu)} - \varphi^{(\mu)}(t) \right| , \qquad \left| y^{(\mu)} - \psi^{(\mu)}(t) \right|$$

Now suppose that $\varrho = \varrho(t) = \varrho(t;\varepsilon) = \varepsilon\overline{\varrho}(t) + (\varepsilon)_2$ is an entire function of t that becomes infinitely small for all values of t together with ε and that satisfies the conditions

$$\varrho^{(\mu)}(t_0) = \tau_0^{(\mu)} , \qquad \varrho^{(\mu)}(t_3 - \varepsilon) = \tau_3^{(\mu)} + e_{\mu,0}\,\varepsilon$$
$$(\mu = 0, 1, \ldots n-1) ,$$

where $\tau_0^{(\mu)}$, $\tau_3^{(\mu)}$ have the same meaning as in (64); a function of this kind can always be determined by means of the procedure of (64), namely

$$\varrho = \sum_{\mu=0}^{n-1} \left\{ \tau_0^{(\mu)} k_\mu(t; t_0, t_3 - \varepsilon) + \left(\tau_3^{(\mu)} + e_{\mu,0}\,\varepsilon\right) k_\mu(t; t_3 - \varepsilon,\, t_0) \right\} .$$

The equations (65) are then transformed into the following ones:

$$\begin{aligned}
[D^\mu \varphi_\varepsilon(t+\varrho)]_{t=t_0} &= [D^\mu \varphi_\varepsilon(t+\tau)]_{t_0} = \varphi^{(\mu)}(t_0) \\
[D^\mu \psi_\varepsilon(t+\varrho)]_{t_0} &= [D^\mu \psi_\varepsilon(t+\tau)]_{t_0} = \psi^{(\mu)}(t_0) \\
[D^\mu \varphi_\varepsilon(t+\varrho)]_{t_3-\varepsilon} &= [D^\mu \varphi_\varepsilon(t+\tau)]_{t_3} = \overline{\varphi}_1^{(\mu)}(t_3 - \varepsilon) \\
[D^\mu \psi_\varepsilon(t+\varrho)]_{t_3-\varepsilon} &= [D^\mu \psi_\varepsilon(t+\tau)]_{t_3} = \overline{\psi}_1^{(\mu)}(t_3 - \varepsilon) \\
&(\mu = 0, 1, \ldots n-1) .
\end{aligned} \tag{75}$$

If, as assumed,

$$t_1 \leqq t_0 < t_3 \leqq t_2 ,$$

then the compound line 1 0 4 3 2 can be expressed by the equations

$$\begin{aligned}
&x = \varphi(t) , && y = \psi(t) && (t_1 \leqq t \leqq t_0) \\
&x = \varphi_\varepsilon(t+\varrho) , && y = \psi_\varepsilon(t+\varrho) && (t_0 \leqq t \leqq t_3 - \varepsilon) \\
&x = \overline{\varphi}_1(t) , && y = \overline{\psi}_1(t) && (t_3 - \varepsilon \leqq t \leqq t_3) \\
&x = \varphi(t) , && y = \psi(t) && (t_3 \leqq t \leqq t_2) ,
\end{aligned} \tag{76}$$

where the quantities

$$x^{(\mu)} = \frac{d^\mu x}{dt^\mu} , \quad y^{(\mu)} = \frac{d^\mu y}{dt^\mu} \qquad (\mu = 0, 1, \ldots n-1) ,$$

on account of (60) and (75), are always *continuous* at the corners $t = t_0$, $t = t_3 - \varepsilon$, $t = t_3$ of the curves a, a_ε and $\overline{a}$, and where the corresponding absolute values

$$\left| x^{(\mu)} - \varphi^{(\mu)}(t) \right| , \qquad \left| y^{(\mu)} - \psi^{(\mu)}(t) \right|$$

für alle Werte von t zwischen t_1 und t_2 mit ε gleichzeitig unendlich klein werden; denn da

$$\varrho = \varepsilon\overline{\varrho}(t) + (\varepsilon)_2$$

und

$$\varphi_\varepsilon(t) = \varphi(t) - \varepsilon\overline{\varphi}_\varepsilon(t) + (\varepsilon)_2 \ ,$$

so wird auch:

$$D^\mu\varphi_\varepsilon(t+\varrho) = \varphi^{(\mu)}(t) + (\varepsilon)$$

und für $t_3 - \varepsilon \leqq t \leqq t_3$ auch:

$$\overline{\varphi}_1^{(\mu)}(t) = \overline{\varphi}_1^{(\mu)}(t_3) + (\varepsilon) = \varphi^{(\mu)}(t_3) + (\varepsilon) = \varphi^{(\mu)}(t) + (\varepsilon)$$
$$(\mu = 0, 1, 2, \ldots n-1) \ .$$

Ausserdem berührt diese Curve 1 0 4 3 2 die erste a oder 1 0 3 2 in den Punkten 1 und 2 von der Ordnung $n-1$, da

$$\begin{array}{llll} x^{(\mu)} = \varphi^{(\mu)}(t_1) \ , & y^{(\mu)} = \psi^{(\mu)}(t_1) & (t = t_1) & \\ y^{(\mu)} = \varphi^{(\mu)}(t_2) \ , & y^{(\mu)} = \psi^{(\mu)}(t_2) & (t = t_2) & (\mu = 0, 1, \ldots n-1) \ , \end{array}$$

wegen (75) auch in den Grenzfällen, wo einer der Punkte 0 oder 3 mit 1 oder 2 zusammenfällt.

Es kann daher 1 0 4 3 2 als eine „erlaubte“ und für hinreichend kleine $\varepsilon > 0$ beliebig eng „benachbarte“ Variation von a oder 1 2 angesehen werden, wenn in den Definitions-Gleichungen (34) und (37) $r = m = n-1$ angenommen wird. Da nun alle Teil-Integrale unabhängig sind von der besonderen Darstellung der Curven, so kann man hier die Formel (72) benutzen und erhält:

$$\begin{aligned} \Delta J &= J_{10432} - J_{1032} = J_{043} - J_{03} \\ &= \varepsilon E\left(x_3^{(\mu)}, y_3^{(\mu)}; \overline{x}_3^{(n)}, \overline{y}_3^{(n)}\right) + (\varepsilon)_2 \ , \end{aligned}$$

ein Ausdruck, der nach (31) wenigstens für hinreichend kleine ε immer *positiv* sein muss, wenn dem Stück 1 2 von a ein wirkliches Minimum des betrachteten Integrales entsprechen soll, also wegen $\varepsilon > 0$ auch:

$$E\left(x_3^{(\mu)}, y_3^{(\mu)}; \overline{p}, \overline{q}\right) = E\left(t_3; \overline{p}, \overline{q}\right) \geqq 0 \tag{77}$$

unabhängig von ε, aber *für beliebige Werte der Grössen* $\overline{p}$, $\overline{q}$.

59 | Da nun die Punkte 0 und 3 zwischen 1 und 2 willkürlich angenommen werden durften, so muss auch allgemein:

$$E(t; \overline{p}, \overline{q}) \geqq 0$$

for all values of t between t_1 and t_2 become infinitely small together with ε; for, since

$$\varrho = \varepsilon\overline{\varrho}(t) + (\varepsilon)_2$$

and

$$\varphi_\varepsilon(t) = \varphi(t) - \varepsilon\overline{\varphi}_\varepsilon(t) + (\varepsilon)_2 \ ,$$

so also

$$D^\mu\varphi_\varepsilon(t+\varrho) = \varphi^{(\mu)}(t) + (\varepsilon)$$

and when $t_3 - \varepsilon \leqq t \leqq t_3$ also

$$\overline{\varphi}_1^{(\mu)}(t) = \overline{\varphi}_1^{(\mu)}(t_3) + (\varepsilon) = \varphi^{(\mu)}(t_3) + (\varepsilon) = \varphi^{(\mu)}(t) + (\varepsilon)$$
$$(\mu = 0, 1, 2, \ldots n-1) \ .$$

In addition, this curve 1 0 4 3 2 makes contact of order $n-1$ with the first one a or 1 0 3 2 at the points 1 and 2, since

$$\begin{array}{lll} x^{(\mu)} = \varphi^{(\mu)}(t_1) \ , & y^{(\mu)} = \psi^{(\mu)}(t_1) & (t = t_1) \\ y^{(\mu)} = \varphi^{(\mu)}(t_2) \ , & y^{(\mu)} = \psi^{(\mu)}(t_2) & (t = t_2) \end{array} \qquad (\mu = 0, 1, \ldots n-1) \ ,$$

on account of (75), also in the borderline cases where one of the points 0 and 3 coincides with 1 or 2.

We may therefore consider 1 0 4 3 2 an "admissible" and, for sufficiently small $\varepsilon > 0$, arbitrarily closely "neighboring" variation of a or 1 2, provided that $r = m = n-1$ is assumed in definition equations (34) and (37). Now since all partial integrals are independent of the particular representation of the curves, we may use formula (72) here, thereby obtaining

$$\begin{aligned} \Delta J &= J_{10432} - J_{1032} = J_{043} - J_{03} \\ &= \varepsilon E\left(x_3^{(\mu)}, y_3^{(\mu)}; \overline{x}_3^{(n)}, \overline{y}_3^{(n)}\right) + (\varepsilon)_2 \ , \end{aligned}$$

an expression that, by (31), must always be *positive*, at least for sufficiently small ε, if a real minimum of the integral under consideration is to correspond to the segment 1 2 of a, and hence also, by $\varepsilon > 0$,

$$E\left(x_3^{(\mu)}, y_3^{(\mu)}; \overline{p}, \overline{q}\right) = E\left(t_3; \overline{p}, \overline{q}\right) \geqq 0 \qquad (77)$$

independent of ε, but *for arbitrary values of the quantities* $\overline{p}$, $\overline{q}$.

Now since we could choose any points 0 and 3 between 1 and 2, it must, generally, be the case that

$$E(t; \overline{p}, \overline{q}) \geqq 0$$

sein für das ganze Intervall $t_1 < t < t_2$ als eine *notwendige Bedingung* für das Bestehen eines Minimums in dem angegebenen Sinne.

Satz V. *Soll unter der Annahme* $r = m = n - 1$ *in* (34) *und* (37) *ein unseren bisherigen Voraussetzunqen entsprechendes Stück* 1 2 *einer Lösung a der Differentialgleichung des Problems ein wirkliches Minimum des Integrales liefern, so darf die Function* $E(t; \overline{p}, \overline{q})$ *an keiner Stelle* t *des ganzen Intervalls* $t_1 \dots t_2$ *für irgend welche Werte der Grössen* $\overline{p}$, $\overline{q}$ *negativ werden können.*

Um nun das Vorzeichen der Function E zu untersuchen, wollen wir dieselbe einer gewissen für $n = 1$ ebenfalls schon von Herrn Prof. *Weierstrass* angegebenen *Umformung* unterziehen.

Es ist nach (73a):

$$E(p,q;\overline{p},\overline{q}) = F(\overline{p},\overline{q}) - F(p,q) - \frac{\partial F}{\partial p}(p,q)(\overline{p}-p) - \frac{\partial F}{\partial q}(p,q)(\overline{q}-q)\ .$$

Setzt man hier:

$$p_\varepsilon = p + \varepsilon(\overline{p}-p)\ , \quad q_\varepsilon = q + \varepsilon(\overline{q}-q)\ ,$$

so dass für $\varepsilon = 0$: $\quad p_\varepsilon = p\ , \quad q_\varepsilon = q\ ,$
und für $\varepsilon = 1$: $\quad p_\varepsilon = \overline{p}\ , \quad q_\varepsilon = \overline{q}\ ,$
so wird:

$$E(p,q;\overline{p},\overline{q}) = \Big[F(p_\varepsilon,q_\varepsilon) + (1-\varepsilon)(\overline{p}-p)\frac{\partial F}{\partial p_\varepsilon}(p_\varepsilon,q_\varepsilon) + (1-\varepsilon)(\overline{q}-q)\frac{\partial F}{\partial q_\varepsilon}(p_\varepsilon,q_\varepsilon)\Big]_{\varepsilon=0}^{\varepsilon=1}$$

$$= \int_0^1 (1-\varepsilon)\left\{\frac{\partial^2 F}{\partial p_\varepsilon^2}(p_\varepsilon,q_\varepsilon)(\overline{p}-p)^2 + 2\frac{\partial^2 F}{\partial p_\varepsilon \partial q_\varepsilon}(p_\varepsilon,q_\varepsilon)(\overline{p}-p)(\overline{q}-q) + \frac{\partial^2 F}{\partial q_\varepsilon^2}(p_\varepsilon,q_\varepsilon)(\overline{q}-q)^2\right\} d\varepsilon\ ,$$

60 | oder, da nach (19):

$$\frac{\partial^2 F}{\partial p_\varepsilon^2} = y'^2 F_1(p_\varepsilon,q_\varepsilon) \qquad \frac{\partial^2 F}{\partial q_\varepsilon^2} = x'^2 F_1(p_\varepsilon,q_\varepsilon)\ ,$$

$$\frac{\partial^2 F}{\partial p_\varepsilon \partial q_\varepsilon} = -x'y' F_1(p_\varepsilon,q_\varepsilon)\ ,$$

$$E(p,q;\overline{p},\overline{q}) = \big(y'(\overline{p}-p) - x'(\overline{q}-q)\big)^2 \int_0^1 F_1(p_\varepsilon,q_\varepsilon)(1-\varepsilon)\,d\varepsilon$$

$$= (\overline{k}-k)^2 E_1(p,q;\overline{p},\overline{q})\ , \tag{78}$$

for the entire interval $t_1 < t < t_2$, which is a *necessary condition* for the existence of a minimum in the specified sense.

Theorem V. *If, assuming that* $r = m = n - 1$, *in* (34) *and* (37), *a segment* 1 2 *satisfying our previous requirements of a solution a of the differential equation of the problem is supposed to furnish a real minimum of the integral, then the function* $E(t; \overline{p}, \overline{q})$ *must not be negative at any position* t *of the entire interval* $t_1 \dots t_2$ *for any values of the quantities* $\overline{p}$, $\overline{q}$.

In order to investigate the sign of the function E we shall subject the function to a certain transformation which, too, has already been specified by Prof. *Weierstrass* for $n = 1$.

By (73a),

$$E(p, q; \overline{p}, \overline{q}) = F(\overline{p}, \overline{q}) - F(p, q) - \frac{\partial F}{\partial p}(p, q)(\overline{p} - p) - \frac{\partial F}{\partial q}(p, q)(\overline{q} - q) \, .$$

If we now set

$$p_\varepsilon = p + \varepsilon(\overline{p} - p) \, , \quad q_\varepsilon = q + \varepsilon(\overline{q} - q) \, ,$$

so that for $\varepsilon = 0$: $\quad p_\varepsilon = p \, , \quad q_\varepsilon = q \, ,$
and for $\varepsilon = 1$: $\quad p_\varepsilon = \overline{p} \, , \quad q_\varepsilon = \overline{q} \, ,$
then

$$E(p, q; \overline{p}, \overline{q}) = \Big[F(p_\varepsilon, q_\varepsilon) + (1 - \varepsilon)(\overline{p} - p)\frac{\partial F}{\partial p_\varepsilon}(p_\varepsilon, q_\varepsilon) + (1 - \varepsilon)(\overline{q} - q)\frac{\partial F}{\partial q_\varepsilon}(p_\varepsilon, q_\varepsilon) \Big]_{\varepsilon = 0}^{\varepsilon = 1}$$

$$= \int_0^1 (1 - \varepsilon) \left\{ \frac{\partial^2 F}{\partial p_\varepsilon^2}(p_\varepsilon, q_\varepsilon)(\overline{p} - p)^2 + 2\frac{\partial^2 F}{\partial p_\varepsilon \partial q_\varepsilon}(p_\varepsilon, q_\varepsilon)(\overline{p} - p)(\overline{q} - q) + \frac{\partial^2 F}{\partial q_\varepsilon^2}(p_\varepsilon, q_\varepsilon)(\overline{q} - q)^2 \right\} d\varepsilon \, ,$$

or, since, by (19),

$$\frac{\partial^2 F}{\partial p_\varepsilon^2} = y'^2 F_1(p_\varepsilon, q_\varepsilon) \quad \frac{\partial^2 F}{\partial q_\varepsilon^2} = x'^2 F_1(p_\varepsilon, q_\varepsilon) \, ,$$

$$\frac{\partial^2 F}{\partial p_\varepsilon \partial q_\varepsilon} = -x' y' F_1(p_\varepsilon, q_\varepsilon) \, ,$$

$$E(p, q; \overline{p}, \overline{q}) = \big(y'(\overline{p} - p) - x'(\overline{q} - q)\big)^2 \int_0^1 F_1(p_\varepsilon, q_\varepsilon)(1 - \varepsilon) \, d\varepsilon$$

$$= (\overline{k} - k)^2 E_1(p, q; \overline{p}, \overline{q}) \, , \tag{78}$$

wenn:

$$y'p - x'q = k\ , \quad y'\overline{p} - x'\overline{q} = \overline{k}$$

gesetzt wird und das bestimmte Integral mit E_1 bezeichnet. Für $n > 1$ ist diese Formel immer anwendbar, so lange die nach p_ε und q_ε genommenen zweiten partiellen Ableitungen und daher auch F_1 endlich und stetig bleiben (vergl. die bei (19) gemachte Bemerkung).

Da in (78) der erste Factor immer positiv ist, so darf auch E_1 im Falle eines Minimums immer nur ein *positives* Vorzeichen besitzen.

Durch Anwendung des Mittelwertsatzes auf das Integral (78) erhält man, was sich noch einfacher direkt aus (73a) durch Berechnung des Restgliedes der Taylorschen Entwicklung ergeben hätte:

$$E_1(p, q; \overline{p}, \overline{q}) = F_1(p_\varkappa, q_\varkappa) \int_0^1 (1 - \varepsilon)\, d\varepsilon = \frac{1}{2} F_1(p_\varkappa, q_\varkappa)\ , \tag{79}$$

wo κ eine gewisse Grösse zwischen 0 und 1 bedeutet.

Die Formeln (78) und (79) gestatten nun folgende Schlüsse:

1. Für *hinreichend klein* gewählte $|\overline{p} - p|$ und $|\overline{q} - q|$ erhalten E_1 und E immer *dasselbe Vorzeichen* wie $F_1(p_\varepsilon, q_\varepsilon)$ und somit auch wie $F_1(p, q)$, falls dieser Ausdruck nicht verschwindet, während gleichzeitig $\overline{k} - k = y'(\overline{p} - p) - x'(\overline{q} - q)$ und daher auch E bei geeigneter Wahl von p und q von 0 verschieden angenommen werden kann.

2. Besitzt $F_1(p, q)$ für irgend eine Combination der fortgelassenen Argu-
61 mente $x, x', \ldots x^{(n-1)}$; $y, y', \ldots y^{(n-1)}$ ein *be-* | *stimmtes*, von p, q *unabhängiges* Vorzeichen, so besteht dasselbe Vorzeichen auch für alle zugehörigen $F_1(p_\varepsilon, q_\varepsilon)$ und damit auch für $E_1(p, q; \overline{p}, \overline{q})$ und $E(p, q; \overline{p}, \overline{q})$, unabhängig von $\overline{p}$ und $\overline{q}$.

Für das Stattfinden der in V verlangten Eigenschaft eines Curvenstückes a ist also:

1. *notwendig*, dass $F_1(x^{(\mu)}, y^{(\mu)}) \geqq 0$ ist auf dem ganzen Curvenstück, d. h. für (80)

$$x^{(\mu)} = \varphi^{(\mu)}(t)\ , \quad y^{(\mu)} = \psi^{(\mu)}(t) \qquad (\mu = 0, 1, \ldots n;\ t_1 \leqq t \leqq t_2)\ ,$$

2. *hinreichend*, dass $F_1\left(x^{(\mu)}, y^{(\mu)}\right) \geqq 0$ ebenfalls auf dem ganzen Curvenstück, für (81)

$$x^{(\mu)} = \varphi^{(\mu)}(t)\ , \quad y^{(\mu)} = \psi^{(\mu)}(t) \qquad (\mu = 0, 1, \ldots n - 1)\ ,$$

aber für *willkürliche* Werte der $x^{(n)} = p$, $y^{(n)} = q$.

In V ist also noch folgende *notwendige Bedingung* des Minimums enthalten, die, nur in anderer Form, gewöhnlich aus der Betrachtung der „zweiten Variation“ gewonnen wird:

if we set

$$y'p - x'q = k\ , \quad y'\overline{p} - x'\overline{q} = \overline{k}$$

and denote the definite integral by E_1. This formula may always be applied when $n > 1$ as long as the second partial derivatives with respect to p_ε and q_ε, and hence also F_1, remain finite and continuous (cf. the observation made in connection with (19)).

Since the first factor in (78) is always positive, E_1, too, may always only have a *positive* sign in case of a minimum.

Applying the mean value theorem to the integral (78), we obtain what we would have obtained directly and even more simply by the calculation from (73a) of the remainder term of the Taylor series:

$$E_1(p, q; \overline{p}, \overline{q}) = F_1(p_\varkappa, q_\varkappa) \int_0^1 (1 - \varepsilon)\, d\varepsilon = \frac{1}{2} F_1(p_\varkappa, q_\varkappa)\ , \tag{79}$$

where $\varkappa$ denotes a certain quantity between 0 and 1.

Formulas (78) and (79) now permit the following conclusions:

1. If $|\overline{p} - p|$ and $|\overline{q} - q|$ are chosen *sufficiently small*, then E_1 and E always get *the same sign* as $F_1(p_\varepsilon, q_\varepsilon)$, and hence also as $F_1(p, q)$, provided this expression does not vanish, whereas at the same time $\overline{k} - k = y'(\overline{p} - p) - x'(\overline{q} - q)$, and hence also E, may be taken to be different from 0 for a suitable choice of p and q.

2. If $F_1(p, q)$, for any combination of the omitted arguments $x, x', \ldots x^{(n-1)}; y, y', \ldots y^{(n-1)}$, possesses a *particular* sign *independent* of p, q, then this sign is also shared by all associated $F_1(p_\varepsilon, q_\varepsilon)$, and hence also by $E_1(p, q; \overline{p}, \overline{q})$ and $E(p, q; \overline{p}, \overline{q})$, independently of $\overline{p}$ and $\overline{q}$.

In order for a curve segment a to have the property required in V, it is therefore:

1. *necessary* that $F_1(x^{(\mu)}, y^{(\mu)}) \geqq 0$ on the entire curve segment, i. e., for (80)

$$x^{(\mu)} = \varphi^{(\mu)}(t)\ , \quad y^{(\mu)} = \psi^{(\mu)}(t) \qquad (\mu = 0, 1, \ldots n;\ t_1 \leqq t \leqq t_2)\ ,$$

2. *sufficient* that $F_1\left(x^{(\mu)}, y^{(\mu)}\right) \geqq 0$ also on the entire curve segment, for (81)

$$x^{(\mu)} = \varphi^{(\mu)}(t)\ , \quad y^{(\mu)} = \psi^{(\mu)}(t) \qquad (\mu = 0, 1, \ldots n - 1)\ ,$$

but for *arbitrary* values of the $x^{(n)} = p$, $y^{(n)} = q$.

Hence, V also contains the following *necessary condition* for the minimum, which, albeit in a different form, is usually obtained by considering the "second variation":

Satz VI. *Wenn a ein wirkliches Minimum liefern soll, so darf in der ganzen Ausdehnung des Curvenstückes die Function* F_1*, oder was nach* (19) *dasselbe bedeutet, die Function* $\frac{\partial^2 F}{\partial x^{(n)2}}$ *oder* $\frac{\partial^2 F}{\partial y^{(n)2}}$*, aus den Ableitungen* $x^{(\mu)}$*,* $y^{(\mu)}$ *der Coordinaten von a selbst gebildet, nirgendwo negative Werte annehmen.*

Hier ist der Beweis nur geführt für $r = m = n-1$ in (34) und (37); doch lässt er sich unmittelbar übertragen auf $m = n$, da, wie man sich leicht überzeugt, durch Verkleinerung von $|\overline{p} - p| = \left|\overline{\varphi}^{(n)}(t_3) - \varphi^{(n)}(t_3)\right|$ und $|\overline{q} - q| = \left|\overline{\psi}^{(n)}(t_3) - \psi^{(n)}(t_3)\right|$ in (76) auch überall $\left|x^{(n)} - \varphi^{(n)}(t)\right|$ und $\left|y^{(n)} - \psi^{(n)}(t)\right|$ beliebig klein gemacht werden können. Nur der Satz V, welcher die *unbeschränkt willkürliche* Wahl der $\overline{p}$, $\overline{q}$ voraussetzt, muss dann seine Gültigkeit verlieren, während VI gerade durch Betrachtung *kleiner* $|\overline{p} - p|$, $|\overline{q} - q|$ abgeleitet war.

Für den Fall $n = 1$ geht wegen

$$p\frac{\partial F}{\partial p}(p,q) + q\frac{\partial F}{\partial q}(p,q) = F(p,q) \qquad (p = x',\ q = y') \qquad (17)_n$$

62 | die Formel (73a) über in:

$$E(p,q;\overline{p},\overline{q}) = F(\overline{p},\overline{q}) - \overline{p}\frac{\partial F}{\partial p}(p,q) - \overline{q}\frac{\partial F}{\partial q}(p,q)\ ,$$

und es wird infolge der vorausgesetzten Homogeneität $(17)_n$ von $F(p,q)$, also auch von $F(\overline{p},\overline{q})$: $F(-\overline{p},-\overline{q}) = -F(\overline{p},\overline{q})$, ebenso auch $E(p,q;-\overline{p},-\overline{q}) = -E(p,q;\overline{p},\overline{q})$, wenn F als eine eindeutige analytische, z. B. als eine *rationale* Function von p und q angenommen wird. Es kann demnach $E(p,q;\overline{p},\overline{q})$, wenn es nicht etwa für beliebige $\overline{p}$, $\overline{q}$ immer verschwinden sollte, für passend gewählte Werte dieser Grössen beliebig auch *negativ* gemacht werden, und ein Minimum in dem angegebenen Sinne ($m = r = n-1 = 0$) ist *unmöglich.*

Dieser von Herrn Prof. *Weierstrass* in der dargestellten Weise allgemein bewiesene Satz lässt sich aber auf den Fall $n > 1$ *nicht* übertragen.

Auch für rationale Functionen $F\left(x^{(\mu)}, y^{(\mu)}\right)$ *kann die Function E ein von* $\overline{p},\overline{q}$ *unabhängiges Vorzeichen und das zugehörige Integral in einer Nachbarschaft von der Ordnung* $m \leqq n-1$ *ein wirkliches Minimum besitzen, wenn* $n > 1$ *ist.*

Dies soll zunächst durch ein *Beispiel* gezeigt werden.

Es sei nämlich:

$$F\left(x^{(\mu)}, y^{(\mu)}\right) = y_n^2 x' = \left(\frac{d^n y}{dx^n}\right)^2 x'\ ,$$

nach (20) eine gestattete Annahme. Hier kann, wie man sich durch vollständige Induction leicht überzeugt, geschrieben werden:

$$y_n = \frac{d^n y}{dx^n} = \frac{x' y^{(n)} - y' x^{(n)}}{x'^{n+1}} + \left(x^{(\mu)}, y^{(\mu)}\right)_{n-1}\ , \qquad (82)$$

Theorem VI. *For a to furnish a real minimum, the function* F_1*, or what, by* (19)*, amounts to the same thing, the function* $\dfrac{\partial^2 F}{\partial x^{(n)2}}$ *or* $\dfrac{\partial^2 F}{\partial y^{(n)2}}$*, formed from the derivatives* $x^{(\mu)}$*,* $y^{(\mu)}$ *of the coordinates of a itself, must never assume negative values in the entire extension of the curve segment.*

We only provide the proof for $r = m = n - 1$ in (34) and (37); but it can be immediately extended to $m = n$, since, as can be readily seen, we can make $\left|x^{(n)} - \varphi^{(n)}(t)\right|$ and $\left|y^{(n)} - \psi^{(n)}(t)\right|$ everywhere arbitrarily small by decreasing $|\overline{p} - p| = \left|\overline{\varphi}^{(n)}(t_3) - \varphi^{(n)}(t_3)\right|$ and $|\overline{q} - q| = \left|\overline{\psi}^{(n)}(t_3) - \psi^{(n)}(t_3)\right|$ in (76). Only Theorem V, which presupposes the *entirely arbitrary* choice of the $\overline{p}$, $\overline{q}$, must then lose its validity, whereas VI had been derived precisely by considering *small* $|\overline{p} - p|$, $|\overline{q} - q|$.

When $n = 1$, on account of

$$p\frac{\partial F}{\partial p}(p,q) + q\frac{\partial F}{\partial q}(p,q) = F(p,q) \qquad (p = x',\ q = y') \ , \qquad (17)_n$$

the formula (73a) is transformed into

$$E(p,q;\overline{p},\overline{q}) = F(\overline{p},\overline{q}) - \overline{p}\frac{\partial F}{\partial p}(p,q) - \overline{q}\frac{\partial F}{\partial q}(p,q) \ ,$$

and, by dint of the assumed homogeneity $(17)_n$ of $F(p,q)$, and hence also of $F(\overline{p},\overline{q})$, we have $F(-\overline{p},-\overline{q}) = -F(\overline{p},\overline{q})$, and also $E(p,q;-\overline{p},-\overline{q}) = -E(p,q;\overline{p},\overline{q})$, provided that F is a single-valued, analytic, e. g. a *rational*, function of p and q. Thus, unless $E(p,q;\overline{p},\overline{q})$ always vanishes for arbitrary $\overline{p}$, $\overline{q}$, we can also make it *negative* for suitably chosen values of these quantities, and a minimum in the sense specified ($m = r = n - 1 = 0$) is *impossible*.

It is, however, *not* possible to extend this theorem, which Prof. *Weierstrass* proved for the general case in the way presented here, to the case $n > 1$.

The function E may possess a sign independent of $\overline{p},\overline{q}$ *also for rational functions* $F\left(x^{(\mu)}, y^{(\mu)}\right)$*, and the associated integral a real minimum in a neighborhood of order* $m \leqq n - 1$ *when* $n > 1$.

At first, this will be demonstrated by an *example*.

For suppose that

$$F\left(x^{(\mu)}, y^{(\mu)}\right) = y_n^2 x' = \left(\frac{d^n y}{dx^n}\right)^2 x' \ ,$$

which, by (20), is a permissible assumption. In this case, as can be readily shown by mathematical induction, we can write

$$y_n = \frac{d^n y}{dx^n} = \frac{x' y^{(n)} - y' x^{(n)}}{x'^{n+1}} + \left(x^{(\mu)}, y^{(\mu)}\right)_{n-1} \ , \qquad (82)$$

wo der Klammerausdruck die Grössen $p = x^{(n)}$ und $q = y^{(n)}$ nicht mehr enthält. Daher ist:

$$\frac{\partial F}{\partial p} = 2y_n \frac{\partial y_n}{\partial p} x' = -\frac{2y'}{x'^{n+1}} y_n x' \,,$$

$$\frac{\partial^2 F}{\partial p^2} = y'^2 F_1 = \frac{2y'^2 x'}{x'^{2n+2}} \,,$$

$$F_1(p,q) = \frac{2}{x'^{2n+1}} = F_1(p_\varepsilon, q_\varepsilon)$$

63 | und nach (78):

$$E(p,q;\overline{p},\overline{q}) = (\overline{k}-k)^2 \int_0^1 \frac{2}{x'^{2n+1}}(1-\varepsilon)\,d\varepsilon = \frac{(\overline{k}-k)^2}{x'^{2n+1}}$$
$$= \left(\frac{x'(\overline{p}-p) - y'(\overline{q}-q)}{x'^{n+1}}\right)^2 x' \,,$$

ein Ausdruck, der immer ein *positives Vorzeichen* hat für $x' \geqq 0$, unabhängig von $\overline{p}$, $\overline{q}$.

In der That entspricht auch einer *Parabel* $2n$ *ter Ordnung*:

$$x = \varphi(t) = t \,, \quad y = \psi(t) = a_0 + a_1 t + \ldots a_{2n-1} t^{2n-1} \tag{83}$$

für $m < n$ ein *wirkliches Minimum* des zwischen beliebigen Grenzen t_1 und t_2 erstreckten Integrals: $J = \int y_n^2 x'\,dt$.

Für alle von einer Ordnung $m \geqq 1$ der Parabel benachbarten Curven $\overline{a}$

$$x = \overline{\varphi}(\lambda) \,, \qquad y = \overline{\psi}(\lambda)$$

kann nach (37) immer

$$|\overline{\varphi}'(\lambda)\lambda' - \varphi'(\varkappa)| = |\overline{\varphi}'(\lambda)\lambda' - 1| < 1 \qquad (\lambda' > 0) \,,$$

also

$$\overline{x}' = \overline{\varphi}'(\lambda)\lambda' > 0 \,, \qquad \overline{\varphi}'(\lambda) > 0$$

angenommen und daher auch immer eine Darstellung für $\overline{a}$ von der Form:

$$\overline{x} = t = x \,, \quad \overline{y} = \overline{\psi}(t) = \psi(t) + \eta(t) = y + \eta \,, \tag{84}$$

where the bracketed expression no longer contains the quantities $p = x^{(n)}$ and $q = y^{(n)}$. Hence, the expression:

$$\frac{\partial F}{\partial p} = 2y_n \frac{\partial y_n}{\partial p} x' = -\frac{2y'}{x'^{n+1}} y_n x' \ ,$$

$$\frac{\partial^2 F}{\partial p^2} = y'^2 F_1 = \frac{2y'^2 x'}{x'^{2n+2}} \ ,$$

$$F_1(p,q) = \frac{2}{x'^{2n+1}} = F_1(p_\varepsilon, q_\varepsilon)$$

and, by (78):

$$E(p,q;\overline{p},\overline{q}) = (\overline{k}-k)^2 \int_0^1 \frac{2}{x'^{2n+1}}(1-\varepsilon)\, d\varepsilon = \frac{(\overline{k}-k)^2}{x'^{2n+1}}$$

$$= \left(\frac{x'(\overline{p}-p) - y'(\overline{q}-q)}{x'^{n+1}} \right)^2 x' \ ,$$

always has a *positive sign* for $x' \geqq 0$ independent of $\overline{p}$, $\overline{q}$.

In fact, also to a *parabola of* $2n-1$ *th order*

$$x = \varphi(t) = t \ , \quad y = \psi(t) = a_0 + a_1 t + \ldots a_{2n-1} t^{2n-1} \tag{83}$$

for $m < n$ there corresponds a *real minimum* of the integral $J = \int y_n^2 x'\, dt$ taken between arbitrary boundaries t_1 and t_2.

For all curves $\overline{a}$ neighboring the parabola of order $m \geqq 1$ and given by

$$x = \overline{\varphi}(\lambda) \ , \qquad y = \overline{\psi}(\lambda) \ ,$$

we may always assume, by (37),

$$|\overline{\varphi}'(\lambda)\lambda' - \varphi'(\varkappa)| = |\overline{\varphi}'(\lambda)\lambda' - 1| < 1 \qquad (\lambda' > 0) \ ,$$

and hence

$$\overline{x}' = \overline{\varphi}'(\lambda)\lambda' > 0 \ , \qquad \overline{\varphi}'(\lambda) > 0 \ ,$$

and thus always use a representation for $\overline{a}$ of form

$$\overline{x} = t = x \ , \quad \overline{y} = \overline{\psi}(t) = \psi(t) + \eta(t) = y + \eta \ , \tag{84}$$

wo $\overline{\psi}(t)$ und ebenso $\eta(t)$ eine *eindeutige* Function bedeutet, zu Grunde gelegt werden, so dass:

$$\overline{x}' = 1\ , \quad \overline{y}_n = \frac{d^n\overline{y}}{dx^n} = \overline{y}^{(n)} = y^{(n)} + \eta^{(n)} \quad \text{und}$$

$$\overline{J} = J(\overline{a}) = \int_{t_1}^{t_2} \overline{y}_n^2 \overline{x}'\, dt = \int_{t_1}^{t_2} \left(y^{(n)} + \eta^{(n)}\right)^2 dt$$

$$= \int_{t_1}^{t_2} {y^{(n)}}^2\, dt + 2\int_{t_1}^{t_2} y^{(n)}\eta^{(n)}\, dt + \int_{t_1}^{t_2} {\eta^{(n)}}^2\, dt\ . \tag{85}$$

64 | Nun ergiebt sich, wenn in (34) gemäss IV $r \geqq n-1$ angenommen wird, also $\overline{y}$ und demnach η als eine mit ihren $n-1$ ersten Ableitungen stetige Function von t im ganzen Intervall, durch partielle Integration:

$$\int_{t_1}^{t_2} y^{(n)}\eta^{(n)}\, dt = \Big[y^{(n)}\eta^{(n-1)} - y^{(n+1)}\eta^{(n-2)} + \ldots + (-1)^{n-1}y^{(2n-1)}\eta\Big]_{t_1}^{t_2}$$

$$+ \int_{t_1}^{t_2} (-1)^n y^{(2n)}\eta\, dt = 0\ . \tag{86}$$

Hier verschwindet der *erste* Teil, weil nach (32) die beiden Curven in den Endpunkten t_1 und t_2 einander von $n-1$ ter Ordnung berühren, also

$$\eta^{(\mu)} = \overline{y}^{(\mu)} - y^{(\mu)} = \overline{y}_\mu - y_\mu = 0 \qquad (\mu = 0, 1, \ldots n-1)\ ,$$

und der *zweite* Teil wegen:

$$y^{(2n)} = \frac{d^{2n}}{dt^{2n}}\left(a_0 + a_1 t + \ldots + a_{2n-1}t^{2n-1}\right) = 0$$

für alle Werte von t. Es wird also nach (85) und (86):

$$\overline{J} - J = \int_{t_1}^{t_2} \overline{y}_n^2\, dt - \int_{t_1}^{t_2} y_n^2\, dt = \int_{t_1}^{t_2} \eta^{(n)2}\, dt > 0\ .$$

Verschwinden könnte die Differenz nur für $\eta^{(n)} = 0$ $(t_1 \leqq t \leqq t_2)$, dies aber ergäbe wegen $\eta^{(\mu)}(t_1) = 0$ $(\mu = 0, 1, \ldots n-1)$ durch n fache Integration oder nach dem Taylorschen Satze:

$$\eta = 0\ , \qquad \overline{y} = y \qquad (t_1 \leqq t \leqq t_2)\ ,$$

where $\overline{\psi}(t)$ as well as $\eta(t)$ denotes a *single-valued* function, such that

$$\overline{x}' = 1\,, \quad \overline{y}_n = \frac{d^n\overline{y}}{dx^n} = \overline{y}^{(n)} = y^{(n)} + \eta^{(n)} \quad \text{and}$$

$$\overline{J} = J(\overline{a}) = \int_{t_1}^{t_2} \overline{y}_n^2 \overline{x}'\, dt = \int_{t_1}^{t_2} \left(y^{(n)} + \eta^{(n)}\right)^2 dt$$

$$= \int_{t_1}^{t_2} {y^{(n)}}^2\, dt + 2\int_{t_1}^{t_2} y^{(n)}\eta^{(n)}\, dt + \int_{t_1}^{t_2} {\eta^{(n)}}^2\, dt\,. \tag{85}$$

Let us now assume that $r \geqq n-1$ in (34), in accordance with IV, and hence that $\overline{y}$, and thus η, is a function of t which, together with its first $n-1$ derivatives, is continuous on the entire interval. By partial integration, we then obtain

$$\int_{t_1}^{t_2} y^{(n)}\eta^{(n)}\, dt = \left[y^{(n)}\eta^{(n-1)} - y^{(n+1)}\eta^{(n-2)} + \ldots + (-1)^{n-1} y^{(2n-1)}\eta\right]_{t_1}^{t_2}$$

$$+ \int_{t_1}^{t_2} (-1)^n y^{(2n)}\eta\, dt = 0\,. \tag{86}$$

In this case, the *first* part vanishes since, by (32), the two curves make contact of $n-1$ th order with one another at the endpoints t_1 and t_2, and hence

$$\eta^{(\mu)} = \overline{y}^{(\mu)} - y^{(\mu)} = \overline{y}_\mu - y_\mu = 0 \qquad (\mu = 0, 1, \ldots n-1)\,,$$

and the *second* part, on account of

$$y^{(2n)} = \frac{d^{2n}}{dt^{2n}}\left(a_0 + a_1 t + \ldots + a_{2n-1} t^{2n-1}\right) = 0\,,$$

for all values of t. By (85) and (86), we therefore have

$$\overline{J} - J = \int_{t_1}^{t_2} \overline{y}_n^2\, dt - \int_{t_1}^{t_2} y_n^2\, dt = \int_{t_1}^{t_2} \eta^{(n)2}\, dt > 0\,.$$

The difference could only vanish when $\eta^{(n)} = 0$ $(t_1 \leqq t \leqq t_2)$. But then we would obtain, on account of $\eta^{(\mu)}(t_1) = 0$ $(\mu = 0, 1, \ldots n-1)$, by n-fold integration or by Taylor's theorem

$$\eta = 0\,, \qquad \overline{y} = y \qquad (t_1 \leqq t \leqq t_2)\,,$$

d. h. beide Curven fielen zusammen. Also:

Das über ein Stück einer Parabel $2n-1$ ter Ordnung erstreckte Integral $\int_{t_1}^{t_2} y_n^2\, dt$ ist immer *kleiner* als dasselbe Integral, erstreckt über irgend eine andere, die Parabel in den Endpunkten 1 und 2 von $n-1$ ter Ordnung be-
65 rührende Curve mit beständig | wachsender x-Coordinate und stetig sich ändernden Ableitungen bis zur $n-1$ ten Ordnung.

Ausser der Möglichkeit eines Minimums (für $m \leqq n-1$) auch für rationale Functionen F zeigt dieses Beispiel, dass in der Definitionsgleichung (37) unter Umständen auch $m < n-1$, hier nämlich $m = 1$ bei $n > 2$, angenommen werden darf, dass also *für die Ordnungszahl* m *der Nachbarschaft n i c h t etwa, dem Satze IV* $(r \geqq n-1)$ *analog, eine von* 0 *oder* 1 *verschiedene untere Grenze allqemein festqestellt werden kann.*

Der Unterschied der Fälle $n = 1$ und $n > 1$ tritt noch klarer hervor, wenn man gemäss

$$F\left(x^{(\mu)}, y^{(\mu)}\right) = f\,(x, y_\mu)\, x'\,, \quad y_\mu = \frac{d^\mu y}{dx^\mu} \quad (\mu = 0, 1, \ldots n) \tag{20}$$

auch die Function E einer entsprechenden *Umformung* unterzieht.

Für $n = 1$ nämlich ist:

$$\begin{aligned}
F\,(x, y; x', y') &= f\,(x, y, y_1)\, x' = f\left(x, y, \frac{y'}{x'}\right) x'\,, \\
\frac{\partial F}{\partial x'} &= f - \frac{\partial f}{\partial y_1}\frac{y'}{x'}\,, \quad \frac{\partial F}{\partial y'} = \frac{\partial f}{\partial y_1} = f'(y_1)\,, \\
&\frac{\partial F}{\partial x'}\,(\overline{x}' - x') + \frac{\partial F}{\partial y'}\,(\overline{y}' - y') \\
&= f \cdot (\overline{x}' - x') + f'\,(y_1)\left(-\frac{y'}{x'}\,(\overline{x}' - x') + \overline{y}' - y'\right) \\
&= f \cdot (\overline{x}' - x') + f'\,(y_1)\,\frac{\overline{y}' x' - \overline{x}' y'}{x' \overline{x}'}\,\overline{x}' \\
&= f \cdot (\overline{x}' - x') + f'\,(y_1)\,(\overline{y}_1 - y_1)\,\overline{x}'\,, \\
F\,(x, y; \overline{x}', \overline{y}') - F\,(x, y; x', y') &= f\,(x, y, \overline{y}_1)\,\overline{x}' - f\,(x, y, y_1)\, x'\,.
\end{aligned}$$

Also ist nach (73):

$$\begin{aligned}
&E\,(x, y; x', y'; \overline{x}', \overline{y}') \\
&= \left\{ f\,(x, y, \overline{y}_1) - f\,(x, y, y_1) - (\overline{y}_1 - y_1)\,\frac{\partial f}{\partial y_1} \right\} \overline{x}'
\end{aligned} \tag{87}$$

66 und wechselt für eindeutige Functionen f sein Vorzeichen, wenn | man $\overline{x}'$, $\overline{y}'$ durch $-\overline{x}'$, $-\overline{y}'$ ersetzt, wobei $\overline{y}_1 = \dfrac{\overline{y}'}{\overline{x}'}$ unverändert bleibt.

i. e., the two curves would coincide. Hence:

The integral $\int_{t_1}^{t_2} y_n^2\, dt$ taken along a segment of a parabola of $2n-1$ th order is always *smaller* than the same integral taken along any other curve that has contact of $n-1$ th order with the parabola at the endpoints 1 and 2 and has continuously increasing x coordinate and continuously varying derivatives up of the $n-1$ th order.

What this example illustrates, besides the possibility that rational functions F, too, have a minimum (when $m \leqq n-1$), is the fact that, in the definition equation (37), we may also assume $m < n-1$, namely, in this case, $m = 1$ when $n > 2$, and hence that *it is n o t possible, along the lines of Theorem IV* ($r \geqq n-1$)*, to generally find a lower limit different from* 0 *or* 1 *for the order number m of the neighborhood.*

The difference between the cases $n = 1$ and $n > 1$ becomes even more evident once we, in accordance with

$$F\left(x^{(\mu)}, y^{(\mu)}\right) = f\,(x, y_\mu)\, x'\,, \quad y_\mu = \frac{d^\mu y}{dx^\mu} \quad (\mu = 0, 1, \ldots n) \tag{20}$$

subject also the function E to a corresponding *transformation.*

For when $n = 1$,

$$\begin{aligned}
F\,(x, y; x', y') &= f\,(x, y, y_1)\, x' = f\left(x, y, \frac{y'}{x'}\right) x'\,, \\
&\frac{\partial F}{\partial x'} = f - \frac{\partial f}{\partial y_1}\frac{y'}{x'}\,, \quad \frac{\partial F}{\partial y'} = \frac{\partial f}{\partial y_1} = f'(y_1)\,, \\
&\frac{\partial F}{\partial x'}\,(\overline{x}' - x') + \frac{\partial F}{\partial y'}\,(\overline{y}' - y') \\
&= f \cdot (\overline{x}' - x') + f'\,(y_1)\left(-\frac{y'}{x'}\,(\overline{x}' - x') + \overline{y}' - y'\right) \\
&= f \cdot (\overline{x}' - x') + f'\,(y_1)\,\frac{\overline{y}'x' - \overline{x}'y'}{x'\overline{x}'}\overline{x}' \\
&= f \cdot (\overline{x}' - x') + f'\,(y_1)\,(\overline{y}_1 - y_1)\,\overline{x}'\,, \\
F\,(x, y; \overline{x}', \overline{y}') - F\,(x, y; x', y') &= f\,(x, y, \overline{y}_1)\,\overline{x}' - f\,(x, y, y_1)\,x'\,.
\end{aligned}$$

Hence, by (73),

$$\begin{aligned}
&E\,(x, y; x', y'; \overline{x}', \overline{y}') \\
&= \left\{ f\,(x, y, \overline{y}_1) - f\,(x, y, y_1) - (\overline{y}_1 - y_1)\,\frac{\partial f}{\partial y_1} \right\} \overline{x}'
\end{aligned} \tag{87}$$

and it switches its sign for single-valued functions f when $\overline{x}'$, $\overline{y}'$ are replaced by $-\overline{x}'$, $-\overline{y}'$, where $\overline{y}_1 = \dfrac{\overline{y}'}{\overline{x}'}$ remains unaltered.

Für $n > 1$ dagegen ist nach (82):

$$y_n = \frac{x'y^{(n)} - y'x^{(n)}}{x'^{n+1}} + \left(x^{(\mu)}, y^{(\mu)}\right)_{n-1} ,$$

$$\overline{y}_n = \frac{x'\overline{y}^{(n)} - y'\overline{x}^{(n)}}{x'^{n+1}} + \left(x^{(\mu)}, y^{(\mu)}\right)_{n-1} ,$$

$$\frac{\partial F}{\partial x^{(n)}} = -\frac{y'}{x'^{n+1}} \frac{\partial f}{\partial y_n} x' , \quad \frac{\partial F}{\partial y^{(n)}} = \frac{x'}{x'^{n+1}} \frac{\partial f}{\partial y_n} x' ,$$

$$\frac{\partial F}{\partial x^{(n)}} \left(\overline{x}^{(n)} - x^{(n)}\right) + \frac{\partial F}{\partial y^{(n)}} \left(\overline{y}^{(n)} - y^{(n)}\right)$$

$$= \frac{\partial f}{\partial y_n} \left\{ \frac{x'\overline{y}^{(n)} - y'\overline{x}^{(n)}}{x'^{n+1}} - \frac{x'y^{(n)} - y'x^{(n)}}{x'^{n+1}} \right\} x'$$

$$= \frac{\partial f}{\partial y_n} \left(\overline{y}_n - y_n\right) x' .$$

Also wird:

$$E\left(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(n)}, \overline{y}^{(n)}\right) = E\left(y_n, \overline{y}_n\right) \tag{88}$$
$$= \left\{ f\left(\overline{y}_n\right) - f\left(y_n\right) - f'\left(y_n\right)\left(\overline{y}_n - y_n\right) \right\} x'$$

bei Fortlassung der Argumente $x^{(\mu)}$, $y^{(\mu)}$ $(\mu = 0, 1, \ldots n-1)$.

Hier enthält der zweite Factor x' keine der willkürlichen Grössen $\overline{x}^{(n)}$, $\overline{y}^{(n)}$, so dass der frühere Schluss nicht mehr anwendbar ist, während der erste Factor ebenso wie der in (87) sehr wohl ein von $\overline{x}^{(n)}$, $\overline{y}^{(n)}$ unabhängiges Vorzeichen besitzen kann.

Setzt man nämlich $y_n = z$, $\overline{y}_n = \overline{z}$, so wird die Function

$$E_2(z, \overline{z}) = f(\overline{z}) - f(z) - (\overline{z} - z) f'(z)$$

an einer Stelle z dann und nur dann für *alle* Werte von $\overline{z}$ positiv sein, *wenn die durch* $\overline{w} = f(\overline{z})$ *dargestellte Curve ganz oberhalb der zum Punkte* $\overline{z} = z$, $\overline{w} = w = f(z)$ *gehörigen Tangente verläuft.*

67 | Dazu aber ist

1. *notwendig*, dass sie an der Stelle z, w *nach oben gekrümmt* ist: $f''(z) \geqq 0$,

2. *hinreichend,* dass sie *allenthalben* nur nach oben gekrümmt ist, $f''(z) \geqq 0$, und ihre Tangentenrichtung $f'(z)$ nur *stetig* verändert.

Fig. 3 Offenbar ist weder die erste Bedingung zugleich hinreichend noch auch die zweite zugleich notwendig. Doch genügt schon die erste Bedingung, wenn $E_2(z, \overline{z})$ nur für alle $\overline{z}$ einer gewissen *Umgebung* von z positiv zu sein braucht.

When $n > 1$ one has, on the other hand, by (82),

$$
\begin{aligned}
y_n &= \frac{x'y^{(n)} - y'x^{(n)}}{x'^{n+1}} + \left(x^{(\mu)}, y^{(\mu)}\right)_{n-1} , \\
\overline{y}_n &= \frac{x'\overline{y}^{(n)} - y'\overline{x}^{(n)}}{x'^{n+1}} + \left(x^{(\mu)}, y^{(\mu)}\right)_{n-1} , \\
\frac{\partial F}{\partial x^{(n)}} &= -\frac{y'}{x'^{n+1}} \frac{\partial f}{\partial y_n} x' , \quad \frac{\partial F}{\partial y^{(n)}} = \frac{x'}{x'^{n+1}} \frac{\partial f}{\partial y_n} x' , \\
&\frac{\partial F}{\partial x^{(n)}} \left(\overline{x}^{(n)} - x^{(n)}\right) + \frac{\partial F}{\partial y^{(n)}} \left(\overline{y}^{(n)} - y^{(n)}\right) \\
&= \frac{\partial f}{\partial y_n} \left\{ \frac{x'\overline{y}^{(n)} - y'\overline{x}^{(n)}}{x'^{n+1}} - \frac{x'y^{(n)} - y'x^{(n)}}{x'^{n+1}} \right\} x' \\
&= \frac{\partial f}{\partial y_n} \left(\overline{y}_n - y_n\right) x' .
\end{aligned}
$$

And hence

$$
\begin{aligned}
&E\left(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(n)}, \overline{y}^{(n)}\right) = E\left(y_n, \overline{y}_n\right) \\
&= \left\{ f\left(\overline{y}_n\right) - f\left(y_n\right) - f'\left(y_n\right)\left(\overline{y}_n - y_n\right) \right\} x'
\end{aligned} \tag{88}
$$

when the arguments $x^{(\mu)}$, $y^{(\mu)}$ $(\mu = 0, 1, \ldots n-1)$ are omitted.

In this case, the second factor x' contains no arbitrary quantities $\overline{x}^{(n)}$, $\overline{y}^{(n)}$ so that the previous conclusion is no longer applicable, whereas the sign of the first factor, as well as that of the one in (87), may certainly be independent of $\overline{x}^{(n)}$, $\overline{y}^{(n)}$.

For if we set $y_n = z$, $\overline{y}_n = \overline{z}$, the function

$$E_2(z, \overline{z}) = f(\overline{z}) - f(z) - (\overline{z} - z) f'(z)$$

is positive at a position z for *all* values of $\overline{z}$ if and only *if the curve represented by* $\overline{w} = f(\overline{z})$ *runs entirely above the tangent belonging to the point* $\overline{z} = z$, $\overline{w} = w = f(z)$.

But for this to be the case, it is

1. *necessary* that it *bends upwards* at the position z, w: $f''(z) \geqq 0$,

2. *sufficient* that it only bends upwards *everywhere*, $f''(z) \geqq 0$, and that the direction of its tangent $f'(z)$ changes only *continuously.*

Fig. 3 Evidently, neither is the first condition also sufficient nor the second also necessary. But the first condition already suffices if $E_2(z, \overline{z})$ only needs to be positive for any $\overline{z}$ in a certain *vicinity* of z.

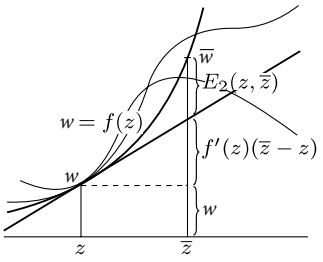


Fig. 3.

Diese Regeln ergeben sich analytisch aus der zu (78) analogen Umformung:

$$\begin{aligned} E_2(z,\overline{z}) &= \left[f(z_\varepsilon) + (1-\varepsilon)f'(z_\varepsilon)\,(\overline{z}-z)\right]_{\varepsilon=0}^{\varepsilon=1} \\ &= (\overline{z}-z)^2 \int_0^1 (1-\varepsilon) f''(z_\varepsilon)\, d\varepsilon \\ &= \frac{1}{2}(\overline{z}-z)^2 f''(z_\varkappa) \qquad (0<\varkappa<1)\,, \end{aligned}$$

wenn $z + \varepsilon(\overline{z}-z) = z_\varepsilon$ gesetzt wird, also:

$$E\left(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(\mu)}, \overline{y}^{(\mu)}\right) = E_2\,(y_n, \overline{y}_n)\, x' \tag{89}$$

$$= x'\,(\overline{y}_n - y_n)^2 \int_0^1 (1-\varepsilon) f_2\,(x, y_\mu, z_\varepsilon)\, d\varepsilon\,,$$

wo:
$$f_2\,(x, y_\mu, y_n) = \frac{\partial^2 f\,(x, y_\mu, y_n)}{\partial y_n^2}$$

68 | und

$$z_\varepsilon = y_n + \varepsilon\,(\overline{y}_n - y_n)\,.$$

Der *wahre Grund* des abweichenden Verhaltens von E in den Fällen $n = 1$ und $n > 1$ liegt, (87) und (88) zufolge, darin, dass im ersten Fall der durch

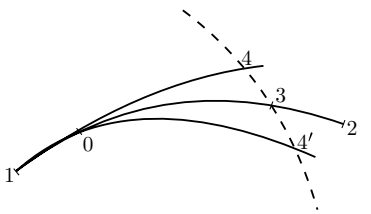


Fig. 4.

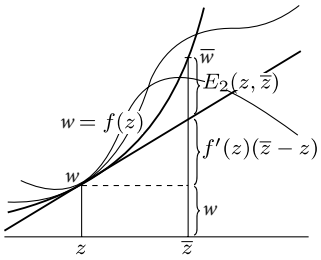


Fig. 3.

These rules analytically arise from the transformation along the lines of (78):

$$E_2(z,\overline{z}) = \left[f(z_\varepsilon) + (1-\varepsilon)f'(z_\varepsilon)(\overline{z}-z)\right]_{\varepsilon=0}^{\varepsilon=1}$$
$$= (\overline{z}-z)^2 \int_0^1 (1-\varepsilon)f''(z_\varepsilon)\, d\varepsilon$$
$$= \frac{1}{2}(\overline{z}-z)^2 f''(z_\varkappa) \qquad (0<\varkappa<1)\,,$$

when we set $z+\varepsilon(\overline{z}-z)=z_\varepsilon$, and hence

$$E\left(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(\mu)}, \overline{y}^{(\mu)}\right) = E_2(y_n, \overline{y}_n)\, x' \tag{89}$$
$$= x'(\overline{y}_n - y_n)^2 \int_0^1 (1-\varepsilon) f_2(x, y_\mu, z_\varepsilon)\, d\varepsilon\,,$$

where
$$f_2(x, y_\mu, y_n) = \frac{\partial^2 f(x, y_\mu, y_n)}{\partial y_n^2}$$

and

$$z_\varepsilon = y_n + \varepsilon(\overline{y}_n - y_n)\,.$$

According to (87) and (88), the *true reason* for the deviant behavior of E in the cases $n=1$ and $n>1$ lies in the fact that, in the first case, the

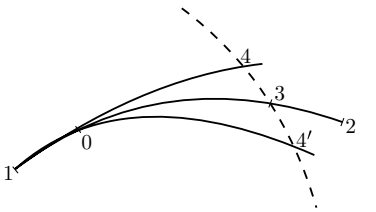


Fig. 4.

(62) definierte Punkt 4 auf derselben Curve $\overline{a}$ von 3 aus *nach entgegengesetzten Richtungen* fortschreiten kann, wobei die $\overline{x}'$, $\overline{y}'$ und mit ihnen auch E entgegengesetzte Vorzeichen annehmen, während für $n > 1$ dies *nicht* mehr gestattet ist, da nach Satz IV die Ableitungen bis zur Ordnung $n - 1 \geqq 1$ als *stetig* vorausgesetzt werden müssen.

Die Formeln (87) und (88) kann man auch direct aus der Definition von E ableiten, wenn man sich des bekannten Ausdrucks für die „erste Variation" mit x als unabhängiger Variable bedient, dessen Entwicklung hier noch kurz wiedergegeben werden möge.

Es ist nach (20):

$$F\left(x^{(\mu)}, y^{(\mu)}\right) = f\,(x, y_\mu)\,x' \ ,$$

wenn $y_\mu \dfrac{d^\mu y}{dx^\mu}$ als Function der $x^{(\nu)}$, $y^{(\nu)}$ $(\nu \leqq \mu)$ aufgefasst wird. Betrachtet man hier $\xi = \delta x$ und $\eta = \delta y$ als unabhängige Variationen, so wird:

$$\begin{aligned} \delta y_\mu &= \delta \frac{y'_{\mu-1}}{x'} = \frac{\delta y'_{\mu-1}}{x'} - \frac{y'_{\mu-1}\delta x'}{x'^2} = \frac{d\delta y_{\mu-1}}{dx} - y_\mu \frac{d\delta x}{dx} \ , \\ \delta y_\mu - y_{\mu+1}\delta x &= \frac{d}{dx}\left(\delta y_{\mu-1} - y_\mu \delta x\right) = \frac{d^\mu}{dx^\mu}\left(\delta y - y_1 \delta x\right) \ . \end{aligned} \tag{90}$$

Nun ist:

69 |
$$\begin{aligned} \delta F = \delta\left\{f\,(x, y_\mu)\,x'\right\} &= f\delta x' + x' \frac{\partial f}{\partial x}\delta x + x' \sum_{\mu=0}^{n} \frac{\partial f}{\partial y_\mu}\delta y_\mu \\ &= f\delta x' + \frac{df}{dt}\delta x + x' \sum_{\mu=0}^{n} \frac{\partial f}{\partial y_\mu}\left(\delta y_\mu - y_{\mu+1}\delta x\right) \\ &= \frac{d}{dt}\left(f\delta x\right) + x' \sum_{\mu=0}^{n} \frac{\partial f}{\partial y_\mu}\frac{d^\mu}{dx^\mu}\left(\delta y - y_1 \delta x\right) \ . \end{aligned}$$

Es ist aber, wenn man für den Augenblick x als unabhängige Variable ansieht, nach einer zu (42) analogen Umformung:

$$\begin{aligned} \sum_{\mu=0}^{n} \frac{\partial f}{\partial y_\mu}\frac{d^\mu}{dx^\mu}\left(\delta y - y_1 \delta x\right) = \frac{d}{dx}\sum_{\mu=1}^{n} L_\mu \frac{d^{\mu-1}}{dx^{\mu-1}}\left(\delta y - y_1 \delta x\right) \\ + L\left(\delta y - y_1 \delta x\right) \ , \end{aligned}$$

wo:
$$L_\mu = \sum_{i=0}^{n-\mu} (-1)^i \frac{d^i}{dx^i}\frac{\partial f}{\partial y_{\mu+i}} \quad (\mu = 0, 1, \ldots n)$$

gesetzt wird. Also:

$$\delta F = \frac{d}{dt}\Big[f\delta x + \sum_{\mu=1}^{n} L_\mu\left(\delta y_{\mu-1} - y_\mu \delta x\right)\Big] + L\left(\delta y - y_1 \delta x\right) x' \ .$$

point 4 defined by (62) may move along the same curve $\overline{a}$ from 3 in *opposite directions*, $\overline{x}'$, $\overline{y}'$, and E with them, assume opposite signs, whereas this is *no* longer permissible when $n > 1$ since, by Theorem IV, it is to be assumed that the derivatives up to order $n-1 \geqq 1$ are *continuous*.

We may also derive the formulas (87) and (88) directly from the definition of E by using the known expression for the "first variation" with x as independent variable, whose expansion will be outlined here briefly.

By (20),

$$F\left(x^{(\mu)}, y^{(\mu)}\right) = f\left(x, y_\mu\right) x' \ ,$$

when $y_\mu \dfrac{d^\mu y}{dx^\mu}$ is considered a function of $x^{(\nu)}$, $y^{(\nu)}$ $(\nu \leqq \mu)$. Considering $\xi = \delta x$ and $\eta = \delta y$ as independent variations, we obtain

$$\begin{aligned} \delta y_\mu &= \delta \frac{y'_{\mu-1}}{x'} = \frac{\delta y'_{\mu-1}}{x'} - \frac{y'_{\mu-1}\delta x'}{x'^2} = \frac{d\delta y_{\mu-1}}{dx} - y_\mu \frac{d\delta x}{dx} \ , \\ \delta y_\mu - y_{\mu+1}\delta x &= \frac{d}{dx}\left(\delta y_{\mu-1} - y_\mu \delta x\right) = \frac{d^\mu}{dx^\mu}\left(\delta y - y_1 \delta x\right) \ . \end{aligned} \tag{90}$$

Now we have

$$\begin{aligned} \delta F = \delta\left\{f\left(x, y_\mu\right) x'\right\} &= f\delta x' + x' \frac{\partial f}{\partial x}\delta x + x' \sum_{\mu=0}^{n} \frac{\partial f}{\partial y_\mu}\delta y_\mu \\ &= f\delta x' + \frac{df}{dt}\delta x + x' \sum_{\mu=0}^{n} \frac{\partial f}{\partial y_\mu}\left(\delta y_\mu - y_{\mu+1}\delta x\right) \\ &= \frac{d}{dt}\left(f\delta x\right) + x' \sum_{\mu=0}^{n} \frac{\partial f}{\partial y_\mu}\frac{d^\mu}{dx^\mu}\left(\delta y - y_1 \delta x\right) \ . \end{aligned}$$

Considering x as an independent variable for the present, we have, by a transformation along the lines of (42),

$$\begin{aligned} \sum_{\mu=0}^{n} \frac{\partial f}{\partial y_\mu}\frac{d^\mu}{dx^\mu}\left(\delta y - y_1 \delta x\right) = \frac{d}{dx}\sum_{\mu=1}^{n} L_\mu \frac{d^{\mu-1}}{dx^{\mu-1}}\left(\delta y - y_1 \delta x\right) \\ + L\left(\delta y - y_1 \delta x\right) \ , \end{aligned}$$

where:

$$L_\mu = \sum_{i=0}^{n-\mu} (-1)^i \frac{d^i}{dx^i}\frac{\partial f}{\partial y_{\mu+i}} \quad (\mu = 0, 1, \ldots n)$$

is set. Hence

$$\delta F = \frac{d}{dt}\Big[f\delta x + \sum_{\mu=1}^{n} L_\mu\left(\delta y_{\mu-1} - y_\mu \delta x\right)\Big] + L\left(\delta y - y_1 \delta x\right) x' \ .$$

Ist nun $x = \varphi(t)$, $y = \psi(t)$ eine Lösung der Differentialgleichung des Problems und im Intervall $t_1 \leqq t \leqq t_3$ von singulären Punkten frei, so verschwindet wegen $L = 0$ das zweite Glied und es wird:

$$\delta J_{03} = \Big[f\delta x + \sum_{\mu=1}^{n} L_\mu \left(\delta y_{\mu-1} - y_\mu \delta x\right) \Big]_{t_0}^{t_3} .$$

Wenn aber $x + \varepsilon\delta x + (\varepsilon)_2$, $y + \varepsilon\delta y + (\varepsilon)_2$ die Coordinaten der Curve a_ε oder 0 4 und $\overline{x}$, $\overline{y}$ die von $\overline{a}$ oder 4 3 darstellen, in der im Anfang des Abschnittes definierten Bedeutung, so wird:

$$\text{für } t = t_0 \colon \quad \delta x = 0 \,, \quad \delta y_\mu = 0 \qquad\qquad (\mu = 0, 1, \ldots n-1) \,,$$

weil a und a_ε einander in 0 von $n-1$ ter Ordnung berühren;

70 | $$\text{für } t = t_3 \colon \qquad \delta x = -\overline{x}' \,, \quad \delta y_\mu = -\overline{y}_{\mu+1}\overline{x}' \quad (\mu = 0, 1, \ldots n-1) \,,$$

wo $\overline{x}'$, $\overline{y}_\mu$ die Werte dieser Functionen im Punkte 3 von $\overline{a}$ bezeichnen. Es wird daher:

$$\begin{aligned} J_{04} - J_{03} &= \varepsilon\delta J_{03} + (\varepsilon)_2 \\ &= \varepsilon\Big[- f\left(x, y_\mu\right)\overline{x}' - \sum_{\mu=1}^{n} L_\mu \left(\overline{y}_\mu - y_\mu\right)\overline{x}' \Big]_{t=t_3} + (\varepsilon)_2 \\ &= -\varepsilon\overline{x}'\left[f\left(x, y_\mu\right) + L_n\left(x, y_\mu\right)\left(\overline{y}_n - y_n\right)\right] + (\varepsilon)_2 \,, \end{aligned}$$

wegen $$\overline{y}_\mu = y_\mu \qquad\qquad (\mu = 0, 1, \ldots n-1) \,.$$

Da ferner: $$J_{43} = f\left(x, \overline{y}_\mu\right)\overline{x}'\varepsilon + (\varepsilon)_2$$

$$\text{und} \quad L_n\left(x, y_\mu\right) = \frac{\partial f}{\partial y_n} \quad \text{ist,}$$

so wird schliesslich:

$$\begin{aligned} J_{043} - J_{03} &= \varepsilon E + (\varepsilon)_2 \\ &= \Big[f\left(x, \overline{y}_\mu\right) - f\left(x, y_\mu\right) - \frac{\partial f}{\partial y_n}\left(\overline{y}_n - y_n\right)\Big]\overline{x}'\varepsilon + (\varepsilon)_2 \,, \end{aligned}$$

also für $n = 1$

$$E = \Big[f\left(x, y, \overline{y}_1\right) - f\left(x, y, y_1\right) - \frac{\partial f\left(x, y, y_1\right)}{\partial y_1}\left(\overline{y}_1 - y_1\right)\Big]\overline{x}' \,. \qquad (87)$$

Für $n > 1$ aber wegen $\overline{x}' = x'$:

$$E = \Big[f\left(x, \overline{y}_\mu\right) - f\left(x, y_\mu\right) - \frac{\partial f}{\partial y_n}\left(\overline{y}_n - y_n\right)\Big]x' \,. \qquad (88)$$

Now if $x = \varphi(t)$, $y = \psi(t)$ is a solution of the differential equation of the problem and without singular points in the interval $t_1 \leqq t \leqq t_3$, then, on account of $L = 0$, the second term vanishes, and

$$\delta J_{03} = \Big[f\delta x + \sum_{\mu=1}^{n} L_\mu \left(\delta y_{\mu-1} - y_\mu \delta x \right) \Big]_{t_0}^{t_3} .$$

But if $x+\varepsilon\delta x+(\varepsilon)_2$, $y+\varepsilon\delta y+(\varepsilon)_2$ represent the coordinates of the curve a_ε or 0 4, and $\overline{x}$, $\overline{y}$ those of $\overline{a}$ or 4 3, in the sense defined at the beginning of the section, then

$$\text{when } t = t_0 : \quad \delta x = 0 , \quad \delta y_\mu = 0 \qquad (\mu = 0, 1, \ldots n-1) ,$$

since a and a_ε have contact of $n-1$ th order at 0 with one another;

$$\text{when } t = t_3 : \qquad \delta x = -\overline{x}' , \quad \delta y_\mu = -\overline{y}_{\mu+1}\overline{x}' \quad (\mu = 0, 1, \ldots n-1) ,$$

where $\overline{x}'$, $\overline{y}_\mu$ denote the values of these functions at the point 3 of $\overline{a}$. Thus

$$\begin{aligned} J_{04} - J_{03} &= \varepsilon\delta J_{03} + (\varepsilon)_2 \\ &= \varepsilon\Big[- f\left(x, y_\mu\right)\overline{x}' - \sum_{\mu=1}^{n} L_\mu \left(\overline{y}_\mu - y_\mu \right) \overline{x}' \Big]_{t=t_3} + (\varepsilon)_2 \\ &= -\varepsilon\overline{x}' \left[f\left(x, y_\mu\right) + L_n \left(x, y_\mu\right) \left(\overline{y}_n - y_n \right) \right] + (\varepsilon)_2 , \end{aligned}$$

on account of $\qquad \overline{y}_\mu = y_\mu \qquad (\mu = 0, 1, \ldots n-1) .$

Since furthermore $\qquad J_{43} = f\left(x, \overline{y}_\mu\right)\overline{x}'\varepsilon + (\varepsilon)_2$

$$\text{and} \quad L_n \left(x, y_\mu\right) = \frac{\partial f}{\partial y_n} ,$$

we finally have

$$\begin{aligned} J_{043} - J_{03} &= \varepsilon E + (\varepsilon)_2 \\ = \Big[f\left(x, \overline{y}_\mu\right) - f\left(x, y_\mu\right) - \frac{\partial f}{\partial y_n} \left(\overline{y}_n - y_n \right) \Big] \overline{x}'\varepsilon + (\varepsilon)_2 , \end{aligned}$$

and hence when $n = 1$

$$E = \Big[f\left(x, y, \overline{y}_1\right) - f\left(x, y, y_1\right) - \frac{\partial f\left(x, y, y_1\right)}{\partial y_1} \left(\overline{y}_1 - y_1 \right) \Big] \overline{x}' . \qquad (87)$$

But when $n > 1$, on account of $\overline{x}' = x'$,

$$E = \Big[f\left(x, \overline{y}_\mu\right) - f\left(x, y_\mu\right) - \frac{\partial f}{\partial y_n} \left(\overline{y}_n - y_n \right) \Big] x' . \qquad (88)$$

In derselben Weise kann man auch den zur Darstellung:

$$F\left(x^{(\mu)}, y^{(\mu)}\right) = \Phi\left(x, y, \alpha_\mu\right) \sqrt{x'^2 + y'^2}\,, \tag{26}$$
$$J = \int_{s_1}^{s_2} \Phi\left(x, y, \alpha_\mu\right)\, ds\,,$$

gehörigen Ausdruck für E bilden.

Hier wird für $n > 1$ ähnlich wie (82):

71 |
$$\alpha_n = \frac{d^{n-2}\alpha_2}{ds^{n-2}} = \frac{d^{n-2}}{ds^{n-2}} \frac{x'y'' - y'x''}{\left(x'^2 + y'^2\right)^{\frac{3}{2}}}$$
$$= \frac{x'y^{(n)} - y'x^{(n)}}{\left(x'^2 + y'^2\right)^{\frac{n+1}{2}}} + \left(x^{(\mu)}, y^{(\mu)}\right)_{n-1}\,,$$

$$\frac{\partial F}{\partial x^{(n)}} = -\frac{y'}{\left(x'^2 + y'^2\right)^{\frac{n+1}{2}}} \frac{\partial \Phi}{\partial \alpha_n} \sqrt{x'^2 + y'^2}\,,$$
$$\frac{\partial F}{\partial y^{(n)}} = \frac{x'}{\left(x'^2 + y'^2\right)^{\frac{n+1}{2}}} \frac{\partial \Phi}{\partial \alpha_n} \sqrt{x'^2 + y'^2}\,,$$
$$\frac{\partial F}{\partial x^{(n)}} \left(\overline{x}^{(n)} - x^{(n)}\right) + \frac{\partial F}{\partial y^{(n)}} \left(\overline{y}^{(n)} - y^{(n)}\right)$$
$$= \left\{ \frac{x'\overline{y}^{(n)} - y'\overline{x}^{(n)}}{\left(x'^2 + y'^2\right)^{\frac{n+1}{2}}} - \frac{x'y^{(n)} - y'x^{(n)}}{\left(x'^2 + y'^2\right)^{\frac{n+1}{2}}} \right\} \frac{\partial \Phi}{\partial \alpha_n} \sqrt{x'^2 + y'^2}$$
$$= \left(\overline{\alpha}_n - \alpha_n\right) \frac{\partial \Phi}{\partial \alpha_n} \sqrt{x'^2 + y'^2}\,.$$

Also, wenn

$$\overline{F} = F\left(\overline{x}^{(\mu)}, \overline{y}^{(\mu)}\right) = \overline{\Phi} \sqrt{x'^2 + y'^2} = \Phi\left(x, y, \overline{\alpha}_\mu\right) \sqrt{x'^2 + y'^2}$$
$$\left(\overline{x}^{(\mu)} = x^{(\mu)}, \quad \overline{y}^{(\mu)} = y^{(\mu)}, \quad \overline{\alpha}_\mu = \alpha_\mu \quad \text{für} \quad \mu = 0, 1, \ldots n-1\right)$$

gesetzt wird, so ist nach (73):

$$E = E\left(x, y, \alpha_\mu; \overline{\alpha}_n\right) = \left\{ \overline{\Phi} - \Phi - \frac{\partial \Phi}{\partial \alpha_n} \left(\overline{\alpha}_n - \alpha_n\right) \right\} \frac{ds}{dt}\,. \tag{91}$$

Hier ist der erste Factor von derselben Form wie der in (88) und (89), es gelten also für ihn dieselben Betrachtungen bezüglich des Vorzeichens wie

In the same way, it is possible to form the expression E belonging to the representation

$$F\left(x^{(\mu)}, y^{(\mu)}\right) = \Phi\left(x, y, \alpha_\mu\right) \sqrt{x'^2 + y'^2}\,, \tag{26}$$

$$J = \int_{s_1}^{s_2} \Phi\left(x, y, \alpha_\mu\right)\, ds\,.$$

In this case, for $n > 1$, similar to (82),

$$\alpha_n = \frac{d^{n-2}\alpha_2}{ds^{n-2}} = \frac{d^{n-2}}{ds^{n-2}} \frac{x'y'' - y'x''}{\left(x'^2 + y'^2\right)^{\frac{3}{2}}}$$

$$= \frac{x'y^{(n)} - y'x^{(n)}}{\left(x'^2 + y'^2\right)^{\frac{n+1}{2}}} + \left(x^{(\mu)}, y^{(\mu)}\right)_{n-1}\,,$$

$$\frac{\partial F}{\partial x^{(n)}} = -\frac{y'}{\left(x'^2 + y'^2\right)^{\frac{n+1}{2}}} \frac{\partial \Phi}{\partial \alpha_n} \sqrt{x'^2 + y'^2}\,,$$

$$\frac{\partial F}{\partial y^{(n)}} = \frac{x'}{\left(x'^2 + y'^2\right)^{\frac{n+1}{2}}} \frac{\partial \Phi}{\partial \alpha_n} \sqrt{x'^2 + y'^2}\,,$$

$$\frac{\partial F}{\partial x^{(n)}} \left(\overline{x}^{(n)} - x^{(n)}\right) + \frac{\partial F}{\partial y^{(n)}} \left(\overline{y}^{(n)} - y^{(n)}\right)$$

$$= \left\{ \frac{x'\overline{y}^{(n)} - y'\overline{x}^{(n)}}{\left(x'^2 + y'^2\right)^{\frac{n+1}{2}}} - \frac{x'y^{(n)} - y'x^{(n)}}{\left(x'^2 + y'^2\right)^{\frac{n+1}{2}}} \right\} \frac{\partial \Phi}{\partial \alpha_n} \sqrt{x'^2 + y'^2}$$

$$= \left(\overline{\alpha}_n - \alpha_n\right) \frac{\partial \Phi}{\partial \alpha_n} \sqrt{x'^2 + y'^2}\,.$$

Hence, if we set

$$\overline{F} = F\left(\overline{x}^{(\mu)}, \overline{y}^{(\mu)}\right) = \overline{\Phi} \sqrt{x'^2 + y'^2} = \Phi\left(x, y, \overline{\alpha}_\mu\right) \sqrt{x'^2 + y'^2}$$

$$\left(\overline{x}^{(\mu)} = x^{(\mu)}, \quad \overline{y}^{(\mu)} = y^{(\mu)}, \quad \overline{\alpha}_\mu = \alpha_\mu \quad \text{for} \quad \mu = 0, 1, \ldots n-1\right)$$

then, by (73),

$$E = E\left(x, y, \alpha_\mu; \overline{\alpha}_n\right) = \left\{ \overline{\Phi} - \Phi - \frac{\partial \Phi}{\partial \alpha_n} \left(\overline{\alpha}_n - \alpha_n\right) \right\} \frac{ds}{dt}\,. \tag{91}$$

Here the first factor has the same form as that in (88) and (89), and hence it is subject to the same considerations with respect to its sign as E_2 and

für E_2 und auch die analoge Umformung:

$$E = \frac{ds}{dt}(\overline{\alpha}_n - \alpha_n)^2 \int_0^1 \frac{\partial^2 \Phi_\varepsilon}{\partial \alpha_\varepsilon^2}(1-\varepsilon)\, d\varepsilon \tag{92}$$

$$= \frac{1}{2}\frac{ds}{dt}(\overline{\alpha}_n - \alpha_n)^2 \frac{\partial^2 \Phi_\varkappa}{\partial \alpha_\varkappa^2} \qquad (0 < \varkappa < 1)\,,$$

72 | wo $\frac{\partial^2 \Phi_\varepsilon}{\partial \alpha_\varepsilon^2}$ den Wert von $\frac{\partial^2 \Phi}{\partial \alpha_n^2}$ bezeichnet, wenn $\alpha_n + \varepsilon(\overline{\alpha}_n - \alpha_n)$ an Stelle von α_n gesetzt wird.

Der andere Factor $\frac{ds}{dt} = \sqrt{x'^2 + y'^2}$ aber ist unter allen Umständen positiv.

Auch die Formel (91) lässt sich direct aus einer neuen Form der ersten Variation ableiten, deren Entwicklung hier vollständig Platz finden möge, da die Darstellung (26) bisher in der Variationsrechnung, soweit mir bekannt ist, keine Anwendung gefunden hat, für viele, namentlich geometrische Aufgaben jedoch manche Vorzüge zu besitzen scheint.

Nach (26) ist:

$$\delta F = \delta\left(\Phi \frac{ds}{dt}\right) \tag{93}$$

$$= \Phi\delta\frac{ds}{dt} + \left[\frac{\partial \Phi}{\partial x}\delta x + \frac{\partial \Phi}{\partial y}\delta y + \sum_{\mu=1}^{n} \Phi_\mu \delta\alpha_\mu\right]\frac{ds}{dt}\,,$$

wenn $\frac{\partial \Phi}{\partial \alpha_\mu} = \Phi_\mu$ gesetzt wird. Statt der unabhängigen Variationen δx und δy führen wir die neuen δv und δw ein:

$$\left.\begin{aligned} \delta v &= \cos\alpha_1 \delta x + \sin\alpha_1 \delta y & \delta x &= \cos\alpha_1 \delta v + \sin\alpha_1 \delta w \\ \delta w &= \sin\alpha_1 \delta x - \cos\alpha_1 \delta y\,, & \delta y &= \sin\alpha_1 \delta v - \cos\alpha_1 \delta w \end{aligned}\right\} \tag{94}$$

$$\frac{d\delta x}{ds} = \cos\alpha_1 \frac{d\delta v}{ds} + \sin\alpha_1 \frac{d\delta w}{ds} + \alpha_2(-\sin\alpha_1 \delta v + \cos\alpha_1 \delta w)\,,$$

$$\frac{d\delta y}{ds} = \sin\alpha_1 \frac{d\delta v}{ds} - \cos\alpha_1 \frac{d\delta w}{ds} + \alpha_2(\cos\alpha_1 \delta v + \sin\alpha_1 \delta w)\,,$$

$$\frac{d\delta s}{ds} = \frac{\delta\sqrt{x'^2 + y'^2}}{\sqrt{x'^2 + y'^2}} = \frac{x'\delta x' + y'\delta y'}{x'^2 + y'^2} = \cos\alpha_1 \frac{d\delta x}{ds} + \sin\alpha_1 \frac{d\delta y}{ds}\,,$$

$$\frac{d\delta s}{ds} = \frac{d\delta v}{ds} + \alpha_2 \delta w\,, \tag{95}$$

$$\delta\alpha_1 = \delta\,\mathrm{arctg}\,\frac{y'}{x'} = \frac{x'\delta y' - y'\delta x'}{x'^2 + y'^2} = \cos\alpha_1 \frac{d\delta y}{ds} - \sin\alpha_1 \frac{d\delta x}{ds}\,,$$

also the analogous transformation:

$$E = \frac{ds}{dt}(\overline{\alpha}_n - \alpha_n)^2 \int_0^1 \frac{\partial^2 \Phi_\varepsilon}{\partial \alpha_\varepsilon^2}(1-\varepsilon)\, d\varepsilon \tag{92}$$

$$= \frac{1}{2}\frac{ds}{dt}(\overline{\alpha}_n - \alpha_n)^2 \frac{\partial^2 \Phi_\varkappa}{\partial \alpha_\varkappa^2} \qquad (0 < \varkappa < 1)\,,$$

where $\frac{\partial^2 \Phi_\varepsilon}{\partial \alpha_\varepsilon^2}$ denotes the value of $\frac{\partial^2 \Phi}{\partial \alpha_n^2}$ when we substitute $\alpha_n + \varepsilon(\overline{\alpha}_n - \alpha_n)$ for α_n.

The other factor $\frac{ds}{dt} = \sqrt{x'^2 + y'^2}$, however, is always positive.

The formula (91), too, is capable of direct derivation from a new form of the first variation whose development shall be given here completely, since the representation (26) has, to the best of my knowledge, never been applied before in the calculus of variations, but appears to have many an advantage for a great number of problems, and in particular geometric ones.

By (26),

$$\delta F = \delta\left(\Phi \frac{ds}{dt}\right) \tag{93}$$

$$= \Phi\delta\frac{ds}{dt} + \left[\frac{\partial \Phi}{\partial x}\delta x + \frac{\partial \Phi}{\partial y}\delta y + \sum_{\mu=1}^{n} \Phi_\mu \delta\alpha_\mu\right]\frac{ds}{dt}\,,$$

assuming we set $\frac{\partial \Phi}{\partial \alpha_\mu} = \Phi_\mu$. Instead of the independent variations δx and δy we introduce the new ones δv and δw:

$$\left.\begin{aligned} \delta v &= \cos\alpha_1 \delta x + \sin\alpha_1 \delta y & \delta x &= \cos\alpha_1 \delta v + \sin\alpha_1 \delta w \\ \delta w &= \sin\alpha_1 \delta x - \cos\alpha_1 \delta y\,, & \delta y &= \sin\alpha_1 \delta v - \cos\alpha_1 \delta w \end{aligned}\right\} \tag{94}$$

$$\frac{d\delta x}{ds} = \cos\alpha_1 \frac{d\delta v}{ds} + \sin\alpha_1 \frac{d\delta w}{ds} + \alpha_2(-\sin\alpha_1 \delta v + \cos\alpha_1 \delta w)\,,$$

$$\frac{d\delta y}{ds} = \sin\alpha_1 \frac{d\delta v}{ds} - \cos\alpha_1 \frac{d\delta w}{ds} + \alpha_2(\cos\alpha_1 \delta v + \sin\alpha_1 \delta w)\,,$$

$$\frac{d\delta s}{ds} = \frac{\delta\sqrt{x'^2 + y'^2}}{\sqrt{x'^2 + y'^2}} = \frac{x'\delta x' + y'\delta y'}{x'^2 + y'^2} = \cos\alpha_1 \frac{d\delta x}{ds} + \sin\alpha_1 \frac{d\delta y}{ds}\,,$$

$$\frac{d\delta s}{ds} = \frac{d\delta v}{ds} + \alpha_2 \delta w\,, \tag{95}$$

$$\delta\alpha_1 = \delta\,\mathrm{arctg}\,\frac{y'}{x'} = \frac{x'\delta y' - y'\delta x'}{x'^2 + y'^2} = \cos\alpha_1 \frac{d\delta y}{ds} - \sin\alpha_1 \frac{d\delta x}{ds}\,,$$

73 |
$$\delta\alpha_1 = -\frac{d\delta w}{ds} + \alpha_2\, dv\ , \tag{96}$$

$$\frac{d\Phi}{ds} = \frac{\partial\Phi}{\partial x}\cos\alpha_1 + \frac{\partial\Phi}{\partial y}\sin\alpha_1 + \sum_{\mu=1}^{n}\Phi_\mu\alpha_{\mu+1}\ .$$

Nun geht (93) über in:

$$\begin{aligned}\frac{\delta(\Phi\, ds)}{ds} &= \Phi\left(\frac{d\delta v}{ds} + \alpha_2\delta w\right) + \frac{\partial\Phi}{\partial x}\left(\cos\alpha_1\delta v + \sin\alpha_1\delta w\right)\\ &\quad + \frac{\partial\Phi}{\partial y}\left(\sin\alpha_1\delta v - \cos\alpha_1\delta w\right) + \sum_{\mu=1}^{n}\Phi_\mu\delta\alpha_\mu\ ,\\ \frac{\delta(\Phi\, ds)}{ds} &= \frac{d}{ds}(\Phi\delta v) + \left(\Phi\alpha_2 + \frac{\partial\Phi}{\partial x}\sin\alpha_1 - \frac{\partial\Phi}{\partial y}\cos\alpha_1\right)\delta w\\ &\quad + \sum_{\mu=1}^{n}\Phi_\mu\left(\delta\alpha_\mu - \alpha_{\mu+1}\delta v\right)\ .\end{aligned} \tag{97}$$

Nun ist aber für $\mu \geqq 2$:

$$\begin{aligned}\delta\alpha_\mu = \delta\frac{d\alpha_{\mu-1}}{ds} &= \frac{\delta d\alpha_{\mu-1}}{ds} - \frac{d\alpha_{\mu-1}}{ds}\frac{\delta ds}{ds}\\ &= \frac{d\alpha_{\mu-1}}{ds} - \alpha_\mu\left(\frac{d\delta v}{ds} + \alpha_2\delta w\right)\ ,\\ \delta\alpha_\mu - \alpha_{\mu+1}\delta v &= \frac{d}{ds}\left(\delta\alpha_{\mu-1} - \alpha_\mu\delta v\right) - \alpha_\mu\alpha_2\delta w\ .\end{aligned}$$

Also ist für eine beliebige Function A_μ von t oder s:

$$\begin{aligned}A_\mu\left(\delta\alpha_\mu - \alpha_{\mu+1}\delta v\right) &= \frac{d}{ds}\left[A_\mu\left(\delta\alpha_{\mu-1} - \alpha_\mu\delta v\right)\right]\\ &\quad - \frac{dA_\mu}{ds}\left(\delta\alpha_{\mu-1} - \alpha_\mu\delta v\right) - A_\mu\alpha_\mu\alpha_2\delta w\ .\end{aligned}$$

Werden aber die A_μ der Reihe nach definiert durch die Gleichungen:

$$A_\mu + \frac{dA_{\mu+1}}{ds} = \Phi_\mu\ ,\quad (\mu = 1, 2, \ldots n,\quad A_{n+1} = 0)\ , \tag{98}$$

74 |
$$A_\mu = \sum_{i=0}^{n-\mu}(-1)^i\frac{d^i}{ds^i}\Phi_{\mu+i}\ ,$$

so geht die vorhergehende Gleichung durch Hinzufügung von $\frac{dA_{\mu+1}}{ds}\left(\delta\alpha_\mu - \alpha_{\mu+1}\delta v\right)$ auf beiden Seiten über in:

$$\begin{aligned}\Phi_\mu\left(\delta\alpha_\mu - \alpha_{\mu+1}\delta v\right) &= \frac{d}{ds}\left[A_\mu\left(\delta\alpha_{\mu-1} - \alpha_\mu\delta v\right)\right]\\ +\frac{dA_{\mu+1}}{ds}\left(\delta\alpha_\mu - \alpha_{\mu+1}\delta v\right) &- \frac{dA_\mu}{ds}\left(\delta\alpha_{\mu-1} - \alpha_\mu\delta v\right) - A_\mu\alpha_\mu\alpha_2\delta w\ ,\end{aligned}$$

$$\delta\alpha_1 = -\frac{d\delta w}{ds} + \alpha_2\, dv\ , \tag{96}$$

$$\frac{d\Phi}{ds} = \frac{\partial\Phi}{\partial x}\cos\alpha_1 + \frac{\partial\Phi}{\partial y}\sin\alpha_1 + \sum_{\mu=1}^{n}\Phi_\mu\alpha_{\mu+1}\ .$$

Now (93) is transformed into

$$\begin{aligned}\frac{\delta(\Phi\, ds)}{ds} &= \Phi\left(\frac{d\delta v}{ds} + \alpha_2\delta w\right) + \frac{\partial\Phi}{\partial x}\left(\cos\alpha_1\delta v + \sin\alpha_1\delta w\right)\\ &\quad + \frac{\partial\Phi}{\partial y}\left(\sin\alpha_1\delta v - \cos\alpha_1\delta w\right) + \sum_{\mu=1}^{n}\Phi_\mu\delta\alpha_\mu\ ,\\ \frac{\delta(\Phi\, ds)}{ds} &= \frac{d}{ds}(\Phi\delta v) + \left(\Phi\alpha_2 + \frac{\partial\Phi}{\partial x}\sin\alpha_1 - \frac{\partial\Phi}{\partial y}\cos\alpha_1\right)\delta w\\ &\quad + \sum_{\mu=1}^{n}\Phi_\mu\left(\delta\alpha_\mu - \alpha_{\mu+1}\delta v\right)\ .\end{aligned} \tag{97}$$

But now for $\mu \geqq 2$

$$\begin{aligned}\delta\alpha_\mu &= \delta\frac{d\alpha_{\mu-1}}{ds} = \frac{\delta d\alpha_{\mu-1}}{ds} - \frac{d\alpha_{\mu-1}}{ds}\frac{\delta ds}{ds}\\ &= \frac{d\alpha_{\mu-1}}{ds} - \alpha_\mu\left(\frac{d\delta v}{ds} + \alpha_2\delta w\right)\ ,\\ \delta\alpha_\mu - \alpha_{\mu+1}\delta v &= \frac{d}{ds}\left(\delta\alpha_{\mu-1} - \alpha_\mu\delta v\right) - \alpha_\mu\alpha_2\delta w\ .\end{aligned}$$

Thus, for any function A_μ of t or s,

$$\begin{aligned}A_\mu\left(\delta\alpha_\mu - \alpha_{\mu+1}\delta v\right) &= \frac{d}{ds}\left[A_\mu\left(\delta\alpha_{\mu-1} - \alpha_\mu\delta v\right)\right]\\ &\quad - \frac{dA_\mu}{ds}\left(\delta\alpha_{\mu-1} - \alpha_\mu\delta v\right) - A_\mu\alpha_\mu\alpha_2\delta w\ .\end{aligned}$$

But if the A_μ are defined in succession by the equations

$$A_\mu + \frac{dA_{\mu+1}}{ds} = \Phi_\mu\ , \quad (\mu = 1, 2, \ldots n, \quad A_{n+1} = 0)\ , \tag{98}$$

$$A_\mu = \sum_{i=0}^{n-\mu}(-1)^i\frac{d^i}{ds^i}\Phi_{\mu+i}\ ,$$

then, by adding $\frac{dA_{\mu+1}}{ds}\left(\delta\alpha_\mu - \alpha_{\mu+1}\delta v\right)$ on both sides of the previous equation, we obtain

$$\begin{aligned}\Phi_\mu\left(\delta\alpha_\mu - \alpha_{\mu+1}\delta v\right) &= \frac{d}{ds}\left[A_\mu\left(\delta\alpha_{\mu-1} - \alpha_\mu\delta v\right)\right]\\ + \frac{dA_{\mu+1}}{ds}\left(\delta\alpha_\mu - \alpha_{\mu+1}\delta v\right) &- \frac{dA_\mu}{ds}\left(\delta\alpha_{\mu-1} - \alpha_\mu\delta v\right) - A_\mu\alpha_\mu\alpha_2\delta w\ ,\end{aligned}$$

und wenn man auf beiden Seiten über $\mu = 2, 3, \ldots n$ summiert und das Glied hinzufügt:

$$\begin{aligned} \Phi_1 (\delta\alpha_1 - \alpha_2\delta v) &= \left(A_1 + \frac{dA_2}{ds} \right) (\delta\alpha_1 - \alpha_2\delta v) \ , \\ \sum_{\mu=1}^{n} \Phi_\mu (\delta\alpha_\mu - \alpha_{\mu+1}\delta v) &= \frac{d}{ds} \sum_{\mu=2}^{n} A_\mu (\delta\alpha_{\mu-1} - \alpha_\mu \delta v) \\ &\quad + A_1 (\delta\alpha_1 - \alpha_2\delta v) - \sum_{\mu=2}^{n} A_\mu \alpha_\mu \alpha_2 \delta w \ . \end{aligned} \tag{99}$$

Hier ist aber nach (96):

$$A_1 (\delta\alpha_1 - \alpha_2\delta v) = -A_1 \frac{d\delta w}{ds} = -\frac{d}{ds} (A_1 \delta w) + \frac{dA_1}{ds} \delta w \ , \tag{100}$$

so dass nun aus (97), (99) und (100) durch Addition folgt:

$$\begin{aligned} \frac{\delta(\Phi\, ds)}{ds} &= \frac{d}{ds} \left\{ \Phi\delta v - A_1 \delta w + \sum_{\mu=2}^{n} A_\mu (\delta\alpha_{\mu-1} - \alpha_\mu \delta v) \right\} \\ &\quad + \left\{ \frac{\partial\Phi}{\partial x} \sin\alpha_1 - \frac{\partial\Phi}{\partial y} \cos\alpha_1 + \alpha_2 \Big(\Phi - \sum_{\mu=2}^{n} A_\mu \alpha_\mu \Big) + \frac{dA_1}{ds} \right\} \delta w \\ &= \frac{d\delta K}{ds} + G' \delta w \ , \end{aligned} \tag{101}$$

75 | $$\delta J = \delta \int_{(1)}^{(2)} \Phi\, ds = [\delta K]_{(1)}^{(2)} + \int_{(1)}^{(2)} G' \delta w\, ds \ , \quad \text{wo:}$$

$$G' = \frac{\partial\Phi}{\partial x} \sin\alpha_1 - \frac{\partial\Phi}{\partial y} \cos\alpha_1 + \alpha_2 A_0 + \frac{dA_1}{ds} = 0 \tag{102}$$

$$\Big[A_0 = \Phi - \sum_{\mu=2}^{n} A_\mu \alpha_\mu \Big]$$

die *Differentialgleichung des Problems* ist und $A_0, A_1, \ldots A_n$, die Coefficienten in:

$$\delta K = A_0 \delta v - A_1 \delta w + \sum_{\mu=2}^{n} A_\mu \delta\alpha_{\mu-1} \ ,$$

die Ausdrücke sind, welche nach einem zu III analogen Satze immer *stetig* bleiben müssen für alle Curven a, denen ein Minimum entsprechen soll. Die Beweise werden analog geführt wie im zweiten Abschnitt für die dort zu Grunde gelegte Darstellung.

and, by summing over $\mu = 2, 3, \ldots n$ on both sides and adding the term:

$$\begin{aligned}
\Phi_1 (\delta\alpha_1 - \alpha_2\delta v) &= \left(A_1 + \frac{dA_2}{ds} \right) (\delta\alpha_1 - \alpha_2\delta v) \ , \\
\sum_{\mu=1}^{n} \Phi_\mu (\delta\alpha_\mu - \alpha_{\mu+1}\delta v) &= \frac{d}{ds} \sum_{\mu=2}^{n} A_\mu (\delta\alpha_{\mu-1} - \alpha_\mu \delta v) \\
&\quad + A_1 (\delta\alpha_1 - \alpha_2\delta v) - \sum_{\mu=2}^{n} A_\mu \alpha_\mu \alpha_2 \delta w \ .
\end{aligned} \tag{99}$$

But here, by (96),

$$A_1 (\delta\alpha_1 - \alpha_2\delta v) = -A_1 \frac{d\delta w}{ds} = -\frac{d}{ds} (A_1 \delta w) + \frac{dA_1}{ds} \delta w \ , \tag{100}$$

so that it now follows from (97), (99) and (100) by addition that

$$\begin{aligned}
\frac{\delta(\Phi\, ds)}{ds} &= \frac{d}{ds} \left\{ \Phi\delta v - A_1 \delta w + \sum_{\mu=2}^{n} A_\mu (\delta\alpha_{\mu-1} - \alpha_\mu \delta v) \right\} \\
&+ \left\{ \frac{\partial\Phi}{\partial x} \sin\alpha_1 - \frac{\partial\Phi}{\partial y} \cos\alpha_1 + \alpha_2 \Big(\Phi - \sum_{\mu=2}^{n} A_\mu \alpha_\mu \Big) + \frac{dA_1}{ds} \right\} \delta w \\
&= \frac{d\delta K}{ds} + G' \delta w \ ,
\end{aligned} \tag{101}$$

$$\begin{aligned}
\delta J &= \delta \int_{(1)}^{(2)} \Phi\, ds = [\delta K]_{(1)}^{(2)} + \int_{(1)}^{(2)} G' \delta w\, ds \ , \quad \text{where} \\
G' &= \frac{\partial\Phi}{\partial x} \sin\alpha_1 - \frac{\partial\Phi}{\partial y} \cos\alpha_1 + \alpha_2 A_0 + \frac{dA_1}{ds} = 0 \\
&\Big[A_0 = \Phi - \sum_{\mu=2}^{n} A_\mu \alpha_\mu \Big]
\end{aligned} \tag{102}$$

is the *differential equation of the problem*, and $A_0, A_1, \ldots A_n$, the coefficients in

$$\delta K = A_0 \delta v - A_1 \delta w + \sum_{\mu=2}^{n} A_\mu \delta\alpha_{\mu-1} \ ,$$

are those expressions that, according to a theorem along the lines of III, must always remain *continuous* for all curves a to which a minimum is to correspond. The proofs are carried out in the same way as those for the representation underlying the second section.

Für den Fall $n = 2$ z. B., wo $\alpha_1 = \alpha$ den *Winkel zwischen Tangente* und *x-Achse*, $\alpha_2 = k$ die Krümmung bedeutet, wird:

$$F = \Phi(x, y, \alpha, k)\frac{ds}{dt} ,$$

$$A_2 = \Phi_2 = \frac{\partial \Phi}{\partial k} , \quad A_0 = \Phi - \Phi_2 k , \quad A_1 = \frac{\partial \Phi}{\partial \alpha} - \frac{d}{ds}\frac{\partial \Phi}{\partial k} , \quad \text{also}$$

$$G' = \frac{\partial \Phi}{\partial x}\sin\alpha - \frac{\partial \Phi}{\partial y}\cos\alpha + k\Phi - k^2\frac{\partial \Phi}{\partial k} + \frac{d}{ds}\frac{\partial \Phi}{\partial \alpha} - \frac{d^2}{ds^2}\frac{\partial \Phi}{\partial k} ,$$

$$\delta K = \left(\Phi - \frac{\partial \Phi}{\partial k}k\right)\delta v - \left(\frac{\partial \Phi}{\partial \alpha} - \frac{d}{ds}\frac{\partial \Phi}{\partial k}\right)\delta w + \frac{\partial \Phi}{\partial k}\delta\alpha .$$

Zur Berechnung von E nehmen wir an, dass sich der Punkt 4 unserer Fig. 2 auf der Curve $\overline{a}$ um das Bogenelement σ von 3 aus rückwärts verschoben habe, und es seien $x + \sigma\delta x + (\sigma)_2$, $y + \sigma\delta y + (\sigma)_2$ die Coordinaten von 0 4 oder α_ε, wenn die entsprechenden von a oder 0 3 mit x, y bezeichnet werden. Dann ist für den Punkt 0:

$$\delta x = 0 , \quad \delta y = 0 ; \quad \delta v = 0 , \quad \delta w = 0 ,$$
$$\delta\alpha_1 = 0, \ldots \delta\alpha_{n-1} = 0 , \quad \text{also} \quad \delta K = 0 ,$$

76 | für den Punkt 4 aber, den wir durch Variation von 3 entstanden denken:

$$\delta x = -\frac{\overline{dx}}{ds} = -\cos\overline{\alpha}_1 , \qquad \delta y = -\frac{\overline{dy}}{ds} = -\sin\overline{\alpha}_1 ,$$
$$\delta v = -\cos(\alpha_1 - \overline{\alpha}_1) , \qquad \delta w = \sin(\overline{\alpha}_1 - \alpha_1) ,$$

und für $n > 1$ wegen $\overline{\alpha}_\mu = \alpha_\mu \qquad (\mu < n)$:

$$\delta x = -\cos\alpha_1 , \quad \delta y = -\sin\alpha_1 , \quad \delta v = -1 , \quad \delta w = 0 ,$$

$$\delta\alpha_{\mu-1} = -\overline{\alpha}_\mu = -\alpha_\mu - e_{\mu,n}(\overline{\alpha}_n - \alpha_n) \qquad (\mu = 2, 3, \ldots n) ,$$

$$\delta K = -\Phi - \sum_{\mu=2}^{n} A_\mu e_{\mu,n}(\overline{\alpha}_n - \alpha_n) = \Phi - \Phi_n(\overline{\alpha}_n - \alpha_n) ,$$

wenn die gestrichenen Buchstaben die auf $\overline{a}$, die ungestrichenen die auf a genommenen Werte der Ausdrücke im Punkte 3 bezeichnen.

Es wird nun nach (101), da a der Gleichung $G' = 0$ genügt:

$$\delta J_{03} = [\delta K]_{(0)}^{(3)} = -\Phi - \Phi_n(\overline{\alpha}_n - \alpha_n)$$

und
$$\begin{aligned} J_{043} - J_{03} &= J_{04} - J_{03} + J_{43} = \sigma\delta J_{03} + \overline{\Phi}\sigma + (\sigma)_2 \\ &= \sigma\left\{\overline{\Phi} - \Phi - \Phi_n(\overline{\alpha}_n - \alpha_n)\right\} + (\sigma)_2 \\ &= \varepsilon\frac{ds}{dt}\left\{\overline{\Phi} - \Phi - \frac{\partial \Phi}{\Phi\alpha_n}(\overline{\alpha}_n - \alpha_n)\right\} + (\varepsilon)_2 , \end{aligned}$$

For instance, for $n = 2$,

$$F = \Phi(x, y, \alpha, k)\frac{ds}{dt} ,$$

$$A_2 = \Phi_2 = \frac{\partial \Phi}{\partial k} , \quad A_0 = \Phi - \Phi_2 k , \quad A_1 = \frac{\partial \Phi}{\partial \alpha} - \frac{d}{ds}\frac{\partial \Phi}{\partial k} , \quad \text{hence}$$

$$G' = \frac{\partial \Phi}{\partial x}\sin\alpha - \frac{\partial \Phi}{\partial y}\cos\alpha + k\Phi - k^2\frac{\partial \Phi}{\partial k} + \frac{d}{ds}\frac{\partial \Phi}{\partial \alpha} - \frac{d^2}{ds^2}\frac{\partial \Phi}{\partial k} ,$$

$$\delta K = \left(\Phi - \frac{\partial \Phi}{\partial k}k\right)\delta v - \left(\frac{\partial \Phi}{\partial \alpha} - \frac{d}{ds}\frac{\partial \Phi}{\partial k}\right)\delta w + \frac{\partial \Phi}{\partial k}\delta\alpha ,$$

where $\alpha_1 = \alpha$ denotes the *angle between tangent* and x *axis*, and $\alpha_2 = k$ the curvature.

To evaluate E we assume that point 4 in our Fig. 2 has slid backwards on the curve $\overline{a}$ from 3 by the element of arc σ. Let $x + \sigma\delta x + (\sigma)_2$, $y + \sigma\delta y + (\sigma)_2$ be the coordinates of 0 4 or α_ε, where the corresponding ones of a or 0 3 are denoted by x, y. Then, for the point 0,

$$\delta x = 0 , \quad \delta y = 0 ; \quad \delta v = 0 , \quad \delta w = 0 ,$$
$$\delta\alpha_1 = 0, \ldots \delta\alpha_{n-1} = 0 , \quad \text{also} \quad \delta K = 0 ,$$

but for the point 4, which we take to be generated by variation of 3,

$$\delta x = -\frac{\overline{dx}}{ds} = -\cos\overline{\alpha}_1 , \qquad \delta y = -\frac{\overline{dy}}{ds} = -\sin\overline{\alpha}_1 ,$$
$$\delta v = -\cos(\alpha_1 - \overline{\alpha}_1) , \qquad \delta w = \sin(\overline{\alpha}_1 - \alpha_1) ,$$

and for $n > 1$, on account of $\overline{\alpha}_\mu = \alpha_\mu \qquad (\mu < n)$,

$$\delta x = -\cos\alpha_1 , \quad \delta y = -\sin\alpha_1 , \quad \delta v = -1 , \quad \delta w = 0 ,$$

$$\delta\alpha_{\mu-1} = -\overline{\alpha}_\mu = -\alpha_\mu - e_{\mu,n}(\overline{\alpha}_n - \alpha_n) \qquad (\mu = 2, 3, \ldots n) ,$$

$$\delta K = -\Phi - \sum_{\mu=2}^{n} A_\mu e_{\mu,n}(\overline{\alpha}_n - \alpha_n) = \Phi - \Phi_n(\overline{\alpha}_n - \alpha_n) ,$$

where the barred letters denote those values of the expressions taken on $\overline{a}$ at the point 3 , the non-barred letters those taken on a.

Since a satisfies the equation $G' = 0$, we now have, by (101),

$$\delta J_{03} = [\delta K]_{(0)}^{(3)} = -\Phi - \Phi_n(\overline{\alpha}_n - \alpha_n)$$

and

$$\begin{aligned} J_{043} - J_{03} &= J_{04} - J_{03} + J_{43} = \sigma\delta J_{03} + \overline{\Phi}\sigma + (\sigma)_2 \\ &= \sigma\left\{\overline{\Phi} - \Phi - \Phi_n(\overline{\alpha}_n - \alpha_n)\right\} + (\sigma)_2 \\ &= \varepsilon\frac{ds}{dt}\left\{\overline{\Phi} - \Phi - \frac{\partial \Phi}{\Phi\alpha_n}(\overline{\alpha}_n - \alpha_n)\right\} + (\varepsilon)_2 , \end{aligned}$$

wenn ε die Differenz der Werte bezeichnet, welche die unabhängige Variable t von $\overline{a}$ in 4 und in 3 annimmt, so dass:

$$\sigma = \frac{ds}{dt}\varepsilon + (\varepsilon)_2 \, .$$

Es ergiebt sich also wieder:

$$E = \left\{ \overline{\Phi} - \Phi - \frac{\partial \Phi}{\partial \alpha_n} (\overline{\alpha}_n - \alpha_n) \right\} \frac{ds}{dt} \, . \tag{91}$$

77 |

Vierter Abschnitt.

Aufstellung hinreichender Bedingungen.

Es sei 1 2 ein solches Stück einer Curve $\overline{a}$: $x = \overline{\varphi}(\lambda)$, $y = \overline{\psi}(\lambda)$, dass in seiner ganzen Ausdehnung $\lambda_1 \leqq \lambda \leqq \lambda_2$ die Functionen $\overline{\varphi}$, $\overline{\psi}$ mit ihren Ableitungen bis zur mindestens n ten Ordnung stetig sind und nirgends $\overline{\varphi'}(\lambda)$ und $\overline{\psi'}(\lambda)$ gleichzeitig verschwinden.

Nun werde weiter vorausgesetzt, es gebe eine *einfach unendliche stetige Schar U von particulären Lösungen u* der Differentialgleichung des Problems:

$$\begin{aligned} x &= \varphi(t,u) = \varphi\,(t; u_1, \ldots u_{2n}) \\ y &= \psi(t,u) = \psi\,(t; u_1, \ldots u_{2n}) \; , \\ u_\nu &= u_\nu(\lambda) \qquad (\nu = 1, 2, \ldots 2n) \, , \end{aligned} \tag{59}$$

welche in einem festen Punkte 0 alle mit einander eine Berührung $n-1$ ter Ordnung eingehen und ausserdem die Curve $\overline{a}$ der Reihe nach in allen ihren Punkten 3 zwischen 1 und 2 von $n-1$ ter Ordnung berühren, zwischen 0 und 3 aber sich ebenfalls bis auf Ableitungen n ter Ordnung stetig verhalten.

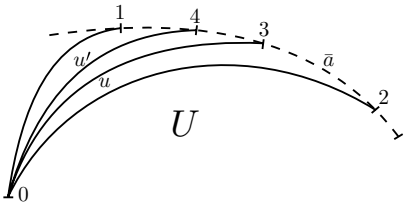


Fig. 5.

Es giebt dann für jeden Wert von λ des Intervalls eine solche Function $\lambda = \lambda(t) = \lambda(t;\lambda)$ und eine solche Function: $t'_3 = t'_3(\lambda)$, dass wie in (60) und

where ε denotes the difference of the values assumed by the independent variable t of $\overline{a}$ in 4 and in 3 so that

$$\sigma = \frac{ds}{dt}\varepsilon + (\varepsilon)_2 \ .$$

Hence, we again obtain

$$E = \left\{ \overline{\Phi} - \Phi - \frac{\partial \Phi}{\partial \alpha_n} (\overline{\alpha}_n - \alpha_n) \right\} \frac{ds}{dt} \ . \tag{91}$$

Fourth section.

Specification of sufficient conditions.

Let 1 2 be a segment of a curve given by $\overline{a}$: $x = \overline{\varphi}(\lambda)$, $y = \overline{\psi}(\lambda)$ such that the functions $\overline{\varphi}$, $\overline{\psi}$ and their derivatives up to at least the n th order are continuous on its entire extension $\lambda_1 \leqq \lambda \leqq \lambda_2$, and $\overline{\varphi}'(\lambda)$ and $\overline{\psi}'(\lambda)$ nowhere vanish simultaneously.

Furthermore, let there be a *simply infinite continuous family U of particular solutions u* of the differential equation of the problem

$$\begin{aligned} x &= \varphi(t,u) = \varphi\,(t; u_1, \ldots u_{2n}) \\ y &= \psi(t,u) = \psi\,(t; u_1, \ldots u_{2n}) \ , \\ u_\nu &= u_\nu(\lambda) \qquad (\nu = 1, 2, \ldots 2n) \ , \end{aligned} \tag{59}$$

all of which have contact of $n-1$ th order with one another at a fixed point 0 and, moreover, have contact of $n-1$ th order with the curve $\overline{a}$ successively at all its points 3 between 1 and 2, but are also continuous between 0 and 3 except for derivatives of n th order.

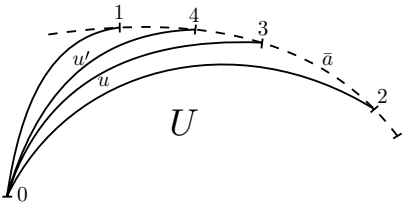


Fig. 5.

Then, for any value of λ of the interval, there is a function $\lambda = \lambda(t) = \lambda(t;\lambda)$ and a function $t_3' = t_3'(\lambda)$ such that, as in (60) and according to the

in der dort angegebenen Bedeutung für $t = t_3'$:

78 |
$$\begin{gathered} D^\mu \overline{\varphi}(\lambda) = \overline{\varphi}_1^{(\mu)}(t_3') = \varphi^{(\mu)}(t_3', u) \\ D^\mu \overline{\psi}(\lambda) = \overline{\psi}_1^{(\mu)}(t_3') = \psi^{(\mu)}(t_3', u) \\ (\mu = 0, 1, \ldots n-1)\,, \end{gathered} \tag{103}$$

während das zu $\lambda - \iota$ gehörige particuläre Integral u' die Curve u in 0 und $\overline{a}$ in $4[x = \overline{\varphi}(\lambda - \iota),\ y = \overline{\psi}(\lambda - \iota)]$ von $n-1$ ter Ordnung berührt, für $\iota = 0$ aber sich auf u reduciert, also allen im vorigen Abschnitte an a_ε gestellten Forderungen genügt. Daher ist die Formel (72) anwendbar und ergiebt, wenn die zusammengesetzten Integrale

$$J_{032} = J_{03} + J_{32} = S(\lambda)\,, \qquad J_{042} = J_{04} + J_{42} = S(\lambda - \iota)$$

gesetzt werden,

$$\begin{gathered} S(\lambda - \iota) - S(\lambda) = J_{04} + J_{43} - J_{03} \\ = \frac{\iota}{\lambda'} E\left(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(n)}, \overline{y}^{(n)}\right) + (\iota)_2 = \frac{\iota}{\lambda'} E(\lambda) + (\iota)_2\,, \end{gathered}$$

wo
$$x^{(\mu)} = \varphi^{(\mu)}(t_3', u)\ ,\ y^{(\mu)} = \psi^{(\mu)}(t_3', u)\ ,$$
$$\overline{x}^{(n)} = D^n \overline{\varphi}(\lambda) = \overline{\varphi}_1^{(n)}(t_3')\ ,\ \overline{y}^{(n)} = D^\mu \overline{\psi}(\lambda) = \overline{\psi}_1^{(n)}(t_3')$$

anzunehmen ist. Es wird daher:

$$S'(\lambda) = \lim_{\iota = 0} \frac{S(\lambda - \iota) - S(\lambda)}{-\iota} = -\frac{E(\lambda)}{\lambda'}\,, \tag{104}$$

also durch Integration:

$$S(\lambda) - S(\lambda_2) = J_{032} - J_{02} = \int_\lambda^{\lambda_2} E(\lambda) \frac{d\lambda}{\lambda'}\,,$$

$$J_{012} - J_{02} = J_{01} - J_{02} + J_{12} = S(\lambda_1) - S(\lambda_2) = \int_{\lambda_1}^{\lambda_2} E(\lambda) \frac{d\lambda}{\lambda'}\,. \tag{105}$$

Sind im besonderen Falle die Curvenstücke 0 1 und 0 2 Teile einer einzigen particulären Lösung a oder 0 1 2, welche von $\overline{a}$ in den Punkten 1 und 2 von $n-1$ ter Ordnung berührt wird, so ist wegen $J_{02} = J_{01} + J_{12}$ und $\lambda' > 0$:

79 |
$$\Delta J = \overline{J} - J = J(\overline{a}) - J(a) = \int_{\lambda_1}^{\lambda_2} E(\lambda) \frac{d\lambda}{\lambda'} > 0\,, \tag{105a}$$

wenn im ganzen Intervall $\lambda_1 \leqq \lambda \leqq \lambda_2$ *beständig* $E(\lambda) \geqq 0$ und nicht überall $E(\lambda) = 0$ ist.

meaning for $t = t_3'$ specified there,

$$\begin{gathered} D^\mu \overline{\varphi}(\lambda) = \overline{\varphi}_1^{(\mu)}(t_3') = \varphi^{(\mu)}(t_3', u) \\ D^\mu \overline{\psi}(\lambda) = \overline{\psi}_1^{(\mu)}(t_3') = \psi^{(\mu)}(t_3', u) \\ (\mu = 0, 1, \ldots n-1) , \end{gathered} \tag{103}$$

while the particular integral u' belonging to $\lambda - \iota$ has contact of $n-1$ th order with the curve u at 0 and with $\overline{a}$ at $4[x = \overline{\varphi}(\lambda - \iota),\ y = \overline{\psi}(\lambda - \iota)]$ but is reduced to u when $\iota = 0$, and hence meets all requirements specified for a_ε in the previous section. The formula (72) is therefore applicable and yields, provided we set the composite integrals

$$J_{032} = J_{03} + J_{32} = S(\lambda) , \qquad J_{042} = J_{04} + J_{42} = S(\lambda - \iota) ,$$

$$\begin{gathered} S(\lambda - \iota) - S(\lambda) = J_{04} + J_{43} - J_{03} \\ = \frac{\iota}{\lambda'} E\left(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(n)}, \overline{y}^{(n)}\right) + (\iota)_2 = \frac{\iota}{\lambda'} E(\lambda) + (\iota)_2 , \end{gathered}$$

where

$$\begin{gathered} x^{(\mu)} = \varphi^{(\mu)}(t_3', u) , \ y^{(\mu)} = \psi^{(\mu)}(t_3', u) , \\ \overline{x}^{(n)} = D^n \overline{\varphi}(\lambda) = \overline{\varphi}_1^{(n)}(t_3') , \ \overline{y}^{(n)} = D^\mu \overline{\psi}(\lambda) = \overline{\psi}_1^{(n)}(t_3') \end{gathered}$$

is to be assumed. Therefore

$$S'(\lambda) = \lim_{\iota = 0} \frac{S(\lambda - \iota) - S(\lambda)}{-\iota} = -\frac{E(\lambda)}{\lambda'} , \tag{104}$$

and hence by integration

$$S(\lambda) - S(\lambda_2) = J_{032} - J_{02} = \int_\lambda^{\lambda_2} E(\lambda) \frac{d\lambda}{\lambda'} ,$$

$$J_{012} - J_{02} = J_{01} - J_{02} + J_{12} = S(\lambda_1) - S(\lambda_2) = \int_{\lambda_1}^{\lambda_2} E(\lambda) \frac{d\lambda}{\lambda'} . \tag{105}$$

If, in particular, the curve segments 0 1 and 0 2 are parts of a single particular solution a or 0 1 2 with which $\overline{a}$ makes contact of the $n-1$ th order at the points 1 and 2, then, on account of $J_{02} = J_{01} + J_{12}$ and $\lambda' > 0$,

$$\Delta J = \overline{J} - J = J(\overline{a}) - J(a) = \int_{\lambda_1}^{\lambda_2} E(\lambda) \frac{d\lambda}{\lambda'} > 0 , \tag{105a}$$

provided that in the entire interval $\lambda_1 \leqq \lambda \leqq \lambda_2$ we *always* have $E(\lambda) \geqq 0$ and not everywhere $E(\lambda) = 0$.

Die Formel (105) mit der daran geknüpften Folgerung gilt auch dann noch, wenn $\overline{\varphi}^{(n)}(\lambda)$, $\overline{\psi}^{(n)}(\lambda)$ in einzelnen Punkten des Intervalls endliche Sprünge erleiden, wie man sich durch Zerlegung von 1 2 in Teil-Intervalle leicht überzeugt, vorausgesetzt nur, dass die Ableitungen bis zur $n-1$ten Ordnung und mit ihnen die u_ν überall stetig bleiben, d. h. die Lösungen der Differentialgleichung stetig in einander übergehen. Sei nämlich z. B. 3 ein solcher Unstetigkeitspunkt, so wird, wenn er der einzige ist:

$$J_{01} + J_{13} - J_{03} = \int_{\lambda_1}^{\lambda_3} E(\lambda)\frac{d\lambda}{\lambda'} ,$$

$$J_{03} + J_{32} - J_{02} = \int_{\lambda_3}^{\lambda_2} E(\lambda)\frac{d\lambda}{\lambda'} ,$$

also durch Addition:

$$J_{01} + J_{12} - J_{02} = \int_{\lambda_1}^{\lambda_2} E(\lambda)\frac{d\lambda}{\lambda'} , \quad \text{q. e. d.}$$

Ganz analog verfährt man für mehrere Unstetigkeitspunkte. $\overline{a}$ braucht daher nur den Stetigkeitsbedingungen (33) und (34) zu genügen, und wir erhalten den Satz:

Satz VII. *Das über ein reguläres Stück 1 2 einer Curve a, die der Differentialgleichung des Problemes genügt, erstreckte Integral J besitzt einen grösseren Wert, als alle über solche erlaubten Variationen $\overline{a}$ erstreckten Integrale $\overline{J}$, für welche eine Schar U von Lösungen der Differentialgleichung der betrachteten Art existiert und für welche immer $E(\lambda) \geqq 0$ und nicht beständig $= 0$ ist.*

Dieser Satz soll nun benutzt werden, um *hinreichende Bedingungen* für das Bestehen eines Minimums herzuleiten.

80 | Zunächst handelt es sich um die *Existenz* einer solchen Schar U von particulären Lösungen der Differentialgleichung, wie sie oben vorausgesetzt wurde. Da die Curvenstücke 0 3, 3 4 und 0 4 als Teile der Curven u, $\overline{a}$ und u' in denselben Beziehungen zu einander stehen wie im vorigen Abschnitte als Teile von a, $\overline{a}$ und a_ε mit dem einzigen Unterschiede, dass jetzt auch $a_\varepsilon = u'$ eine Lösung der Differentialgleichung sein soll, so möge zunächst untersucht werden, *unter welchen Bedingungen eine Curve a_ε mit den in* (60) *bis* (63) *angegebenen Eigenschaften existiert, welche ausserdem noch gleichfalls der Differentialgleichung des Problemes genügt, also eine zu a „benachbarte Lösung" derselben darstellt.*

Nach (59) muss eine solche Curve, wenn sie existiert, von der Form sein:

$$\begin{aligned} x &= \varphi(t, a+\omega) = \varphi\,(t; a_1+\omega_1, \ldots a_{2n}+\omega_{2n}) = \varphi_\varepsilon(t) \\ y &= \psi(t, a+\omega) = \psi\,(t; a_1+\omega_1, \ldots a_{2n}+\omega_{2n}) = \psi_\varepsilon(t) \end{aligned} \qquad (106)$$

The formula (105) together with its associated conclusion even holds when $\overline{\varphi}^{(n)}(\lambda)$, $\overline{\psi}^{(n)}(\lambda)$ suffer finite jump discontinuities at some points of the interval, as can readily be seen from the decomposition of 1 2 into partial intervals, provided only that the derivatives up to the $n-1$ th order, together with the u_ν, are continuous everywhere, i. e., the solutions of the differential equation are continuously transformed into one another. For if we suppose that, e. g., 3 is such a point of discontinuity, then, assuming it is the only one,

$$J_{01} + J_{13} - J_{03} = \int_{\lambda_1}^{\lambda_3} E(\lambda) \frac{d\lambda}{\lambda'} ,$$

$$J_{03} + J_{32} - J_{02} = \int_{\lambda_3}^{\lambda_2} E(\lambda) \frac{d\lambda}{\lambda'} ,$$

and hence, by addition,

$$J_{01} + J_{12} - J_{02} = \int_{\lambda_1}^{\lambda_2} E(\lambda) \frac{d\lambda}{\lambda'} , \quad \text{q. e. d.}$$

The analogous procedure is used for several points of discontinuity. Thus, $\overline{a}$ only needs to satisfy the continuity conditions (33) and (34), and we obtain the theorem

Theorem VII. *The interval J taken along a regular segment 1 2 of a curve a satisfying the differential equation of the problem has a greater value than all integrals $\overline{J}$ taken along admissible variations $\overline{a}$ for which there exists a family U of solutions of the kind of differential equation under consideration and for which always $E(\lambda) \geqq 0$ and not always $= 0$.*

We will now use this theorem in order to deduce *sufficient conditions* for the existence of a minimum.

At first, we are concerned with the *existence* of such a family U of particular solutions of the differential equation as it was assumed above. Since the same relations obtain among the curve segments 0 3, 3 4 and 0 4, whether they are parts of the curves u, $\overline{a}$ and u' or, as in the previous section, parts of the curves a, $\overline{a}$ and a_ε, with the only exception that now $a_\varepsilon = u'$, too, is supposed to be a solution of the differential equation, we shall, at first, investigate *the conditions under which there exists a curve a_ε with the properties specified in* (60) – (63) *that, in addition, satisfies the differential equation of the problem, and hence constitutes a "solution" thereof "neighboring" a.*

By (59), a curve of this kind, assuming it exists, must have the form

$$\begin{aligned} x &= \varphi(t, a + \omega) = \varphi\,(t; a_1 + \omega_1, \ldots a_{2n} + \omega_{2n}) = \varphi_\varepsilon(t) \\ y &= \psi(t, a + \omega) = \psi\,(t; a_1 + \omega_1, \ldots a_{2n} + \omega_{2n}) = \psi_\varepsilon(t) \end{aligned} \tag{106}$$

und für passend bestimmte Werte der Grössen

$$\tau_0^{(\mu)}(t_0) = \tau_0^{(\mu)}\ , \qquad \tau_3^{(\mu)}(t_3) = \tau_3^{(\mu)}\ , \quad \omega_\nu$$
$$(\mu = 0, 1, \ldots n-1\,; \quad \nu = 1, 2, \ldots 2n)\ ,$$

die mit ε gleichzeitig unendlich klein werden müssen, dem Gleichungs-Systeme (65) genügen, das sich jetzt in der Form schreiben lässt:

$$\left.\begin{aligned} D^\mu\left[\varphi\left(t+\tau_0, a+\omega\right) - \varphi(t)\right] &= 0 \\ D^\mu\left[\psi\left(t+\tau_0, a+\omega\right) - \psi(t)\right] &= 0 \end{aligned}\right\} \quad (t = t_0)$$
$$\left.\begin{aligned} D^\mu\left[\varphi\left(t+\tau_3, a+\omega\right) - \varphi(t)\right] &= \xi_3^{(\mu)} \\ D^\mu\left[\psi\left(t+\tau_3, a+\omega\right) - \psi(t)\right] &= \eta_3^{(\mu)} \end{aligned}\right\} \quad (t = t_3)$$
$$(\mu = 0, 1, \ldots n-1)\ , \tag{107}$$

wo wegen (60) auch die Grössen

$$\begin{aligned} \xi_3^{(\mu)} &= \overline{\varphi}_1^{(\mu)}\left(t_3 - \varepsilon\right) - \varphi^{(\mu)}\left(t_3\right) = \overline{\varphi}_1^{(\mu)}\left(t_3 - \varepsilon\right) - \overline{\varphi}_1^{(\mu)}\left(t_3\right)\ , \\ \eta_3^{(\mu)} &= \overline{\psi}_1^{(\mu)}\left(t_3 - \varepsilon\right) - \psi^{(\mu)}\left(t_3\right) = \overline{\psi}_1^{(\mu)}\left(t_3 - \varepsilon\right) - \overline{\psi}_1^{(\mu)}\left(t_3\right) \end{aligned} \tag{108}$$

mit ε gleichzeitig unendlich klein werden.

81 | Für hinreichend kleine $|\tau_0|$, $|\omega_\nu|$ aber gilt die Entwicklung:

$$\begin{aligned} \varphi\left(t+\tau_0, a+\omega\right) &= \varphi'(t,a)\tau_0 + \sum_\nu \varphi_\nu(t,a)\omega_\nu + (\tau_0, \omega_\nu)_2 \\ &= \varphi'(t)\tau_0 + \sum_\nu \varphi_\nu(t)\omega_\nu + (\tau_0, \omega_\nu)_2\ . \end{aligned}$$

Diese Potenzreihe kann man unter der Voraussetzung, dass $\tau_0 = \tau_0(t)$ in der Umgebung von $t = t_0$ den Charakter einer ganzen Function besitzt, *gliedweise differentiieren*, wobei die Dimension jedes Gliedes in Bezug auf die Gesamtheit der Grössen $\tau_0^{(\mu)}$, ω_ν ungeändert bleibt, und die so entstehenden Reihen werden für hinreichend kleine Beträge dieser Grössen in einem Bereiche $|t - t_0| < \delta$ unbedingt und gleichmässig convergieren und die Ableitungen der ursprünglichen Reihe darstellen. Behandelt man ebenso auch die anderen, durch Vertauschung von φ mit ψ oder von t_0 mit t_3 entstehenden Reihen, so lassen die Gleichungen (107) sich schreiben:

$$\begin{aligned} D^\mu\left[\varphi'(t)\tau_0\right]_{t_0} + \sum_\nu \varphi_\nu^{(\mu)}\left(t_0\right)\omega_\nu + \left(\tau_0^{(\mu)}, \omega_\nu\right)_2 &= 0 \\ D^\mu\left[\psi'(t)\tau_0\right]_{t_0} + \sum_\nu \psi_\nu^{(\mu)}\left(t_0\right)\omega_\nu + \left(\tau_0^{(\mu)}, \omega_\nu\right)_2 &= 0 \\ D^\mu\left[\varphi'(t)\tau_3\right]_{t_3} + \sum_\nu \varphi_\nu^{(\mu)}\left(t_3\right)\omega_\nu + \left(\tau_3^{(\mu)}, \omega_\nu\right)_2 &= \xi_3^{(\mu)} \\ D^\mu\left[\psi'(t)\tau_3\right]_{t_3} + \sum_\nu \psi_\nu^{(\mu)}\left(t_3\right)\omega_\nu + \left(\tau_3^{(\mu)}, \omega_\nu\right)_2 &= \eta_3^{(\mu)} \end{aligned} \tag{107a}$$
$$(\mu = 0, 1, \ldots n-1)\ ,$$

and, for suitably determined values of the quantities

$$\begin{aligned} &\tau_0^{(\mu)}(t_0) = \tau_0^{(\mu)}\,, \qquad \tau_3^{(\mu)}(t_3) = \tau_3^{(\mu)}\,, \quad \omega_\nu \\ &(\mu = 0, 1, \ldots n-1\,; \quad \nu = 1, 2, \ldots 2n)\,, \end{aligned}$$

which have to become infinitely small together with ε, must satisfy the system of equations (65), which can now be written in the form

$$\left.\begin{aligned} &D^\mu\left[\varphi\left(t+\tau_0, a+\omega\right) - \varphi(t)\right] = 0 \\ &D^\mu\left[\psi\left(t+\tau_0, a+\omega\right) - \psi(t)\right] = 0 \\ &D^\mu\left[\varphi\left(t+\tau_3, a+\omega\right) - \varphi(t)\right] = \xi_3^{(\mu)} \\ &D^\mu\left[\psi\left(t+\tau_3, a+\omega\right) - \psi(t)\right] = \eta_3^{(\mu)} \end{aligned}\right\} \begin{aligned} &(t = t_0) \\ &\\ &(t = t_3) \\ &\\ &(\mu = 0, 1, \ldots n-1)\,, \end{aligned} \tag{107}$$

where, on account of (60), the quantities

$$\begin{aligned} \xi_3^{(\mu)} &= \overline{\varphi}_1^{(\mu)}\left(t_3 - \varepsilon\right) - \varphi^{(\mu)}\left(t_3\right) = \overline{\varphi}_1^{(\mu)}\left(t_3 - \varepsilon\right) - \overline{\varphi}_1^{(\mu)}\left(t_3\right)\,, \\ \eta_3^{(\mu)} &= \overline{\psi}_1^{(\mu)}\left(t_3 - \varepsilon\right) - \psi^{(\mu)}\left(t_3\right) = \overline{\psi}_1^{(\mu)}\left(t_3 - \varepsilon\right) - \overline{\psi}_1^{(\mu)}\left(t_3\right) \end{aligned} \tag{108}$$

also become infinitely small together with ε.

For sufficiently small $|\tau_0|$, $|\omega_\nu|$, however, the following expansion holds:

$$\begin{aligned} \varphi\left(t+\tau_0, a+\omega\right) &= \varphi'(t,a)\tau_0 + \sum_\nu \varphi_\nu(t,a)\omega_\nu + (\tau_0, \omega_\nu)_2 \\ &= \varphi'(t)\tau_0 + \sum_\nu \varphi_\nu(t)\omega_\nu + (\tau_0, \omega_\nu)_2\,. \end{aligned}$$

Assuming that $\tau_0 = \tau_0(t)$ has the character of an entire function in the vicinity of $t = t_0$, we can *differentiate* this power series *term by term*, where the dimension of each term with respect to the totality of the quantities $\tau_0^{(\mu)}$, ω_ν remains unaltered. The series thus obtained are unconditionally and uniformly convergent for sufficiently small values of these quantities in a domain $|t - t_0| < \delta$ and represent the derivatives of the original series. If we treat the other series obtained by replacing φ by ψ or t_0 by t_3, then we can write the equations (107) as follows:

$$\left.\begin{aligned} &D^\mu\left[\varphi'(t)\tau_0\right]_{t_0} + \sum_\nu \varphi_\nu^{(\mu)}\left(t_0\right)\omega_\nu + \left(\tau_0^{(\mu)}, \omega_\nu\right)_2 = 0 \\ &D^\mu\left[\psi'(t)\tau_0\right]_{t_0} + \sum_\nu \psi_\nu^{(\mu)}\left(t_0\right)\omega_\nu + \left(\tau_0^{(\mu)}, \omega_\nu\right)_2 = 0 \\ &D^\mu\left[\varphi'(t)\tau_3\right]_{t_3} + \sum_\nu \varphi_\nu^{(\mu)}\left(t_3\right)\omega_\nu + \left(\tau_3^{(\mu)}, \omega_\nu\right)_2 = \xi_3^{(\mu)} \\ &D^\mu\left[\psi'(t)\tau_3\right]_{t_3} + \sum_\nu \psi_\nu^{(\mu)}\left(t_3\right)\omega_\nu + \left(\tau_3^{(\mu)}, \omega_\nu\right)_2 = \eta_3^{(\mu)} \end{aligned}\right\} \begin{aligned} &\\ &\\ &\\ &\\ &(\mu = 0, 1, \ldots n-1)\,, \end{aligned} \tag{107a}$$

wo nach der Leibnitzschen Formel

$$D^{\mu}\left[\varphi'(t)\tau_0\right]_{t_0} = \sum_{i=0}^{\mu}\binom{\mu}{i}\varphi^{(\mu-i+1)}(t_0)\,\tau_0^{(i)} \quad \text{u. s. w.}$$

gesetzt werden kann. Nun lässt sich nach einem bekannten Satze der Functionen-Theorie die *Auflösung* dieser $4n$ Gleichungen nach den $4n$ Unbekannten $\tau_0^{(\mu)}$, $\tau_3^{(\mu)}$, ω_ν $(\mu = 0, 1, \ldots n-1;\ \nu = 1, 2, \ldots 2n)$ für hinreichend kleine Beträge der $\xi_3^{(\mu)}$, $\eta_3^{(\mu)}$ durch Potenzreihen, die nach ganzen positiven Potenzen dieser Grössen fortschreiten und mit ihnen zugleich unendlich klein
82 werden, auf eine einzige Weise immer ausführen, vorausgesetzt, dass | die *Determinante*, gebildet aus den Coefficienten der Glieder erster Dimension, *nicht verschwindet*, d. h. dass

$$\Theta = \Theta(t_0, t_3) = \Theta(t_0, t_3; a) \tag{109}$$

$$= \begin{vmatrix} \binom{\mu}{i}\varphi^{(\mu-i+1)}(t_0) & 0 & \varphi_\nu^{(\mu)}(t_0) & \varphi_{n+\nu}^{(\mu)}(t_0) \\ \binom{\mu}{i}\psi^{(\mu-i+1)}(t_0) & 0 & \psi_\nu^{(\mu)}(t_0) & \psi_{n+\nu}^{(\mu)}(t_0) \\ 0 & \binom{\mu}{i}\varphi^{(\mu-i+1)}(t_3) & \varphi_\nu^{(\mu)}(t_3) & \varphi_{n+\nu}^{(\mu)}(t_3) \\ 0 & \binom{\mu}{i}\psi^{(\mu-i+1)}(t_3) & \psi_\nu^{(\mu)}(t_3) & \psi_{n+\nu}^{(\mu)}(t_3) \end{vmatrix} \gtrless 0\ ,$$

wo die einzelnen Glieder immer ganze Systeme von je n^2 Elementen vertreten:

$$\mu = 0, 1, 2, \ldots n-1\ ; \quad i, \nu = 1, 2, \ldots n\ .$$

Durch eine kleine Umformung des nach Fortlassung der Glieder höherer Dimension von (107a) übrig bleibenden Systemes lässt sich nun die Determinante auf eine übersichtlichere Form und auf den Grad $2n$ statt $4n$ reducieren. Dieses System nämlich lässt sich schreiben, indem man zugleich den Buchstaben μ durch i ersetzt:

$$\begin{vmatrix} A_i) \quad 0 = D^i\left[\varphi'\tau_0 + \sum_\nu \varphi_\nu\omega_\nu\right]_0 & \binom{\mu}{i}\psi^{(\mu-i+1)}(t_0) \\ B_i) \quad 0 = D^i\left[\psi'\tau_0 + \sum_\nu \psi_\nu\omega_\nu\right]_0 & -\binom{\mu}{i}\varphi^{(\mu-i+1)}(t_0) \\ C_i) \ \xi_3^{(i)} = D^i\left[\varphi'\tau_3 + \sum_\nu \varphi_\nu\omega_\nu\right]_3 & \binom{\mu}{i}\psi^{(\mu-i+1)}(t_3) \\ D_i) \ \eta_3^{(i)} = D^i\left[\psi'\tau_3 + \sum_\nu \psi_\nu\omega_\nu\right]_3 & -\binom{\mu}{i}\varphi^{(\mu-i+1)}\ . \end{vmatrix} \tag{110}$$

Multipliciert man hier beide Seiten von $A_i)$ und ebenso von $B_i)$ mit den dahinter stehenden Factoren und summiert alle diese Producte über $i = 0, 1, 2, \ldots \mu$ und verfährt dann ebenso mit $C_i)$ und $D_i)$ und den zugehörigen

where, by Leibniz's formula, we can set

$$D^{\mu}\left[\varphi'(t)\tau_0\right]_{t_0} = \sum_{i=0}^{\mu} \binom{\mu}{i} \varphi^{(\mu-i+1)}(t_0)\, \tau_0^{(i)} \quad \text{e. t. c.}$$

Now, by a well-known theorem of the theory of functions, it is always possible to uniquely *solve* these $4n$ equations for the $4n$ unknowns $\tau_0^{(\mu)}$, $\tau_3^{(\mu)}$, ω_ν $(\mu = 0, 1, \ldots n-1;\ \nu = 1, 2, \ldots 2n)$ for sufficiently small values of the $\xi_3^{(\mu)}$, $\eta_3^{(\mu)}$ by means of power series that progress according to positive integral powers of these magnitudes and become infinitely small together with them, provided that the *determinant*, formed from the coefficients of the terms of first dimension, *does not vanish*, i. e., that

$$\Theta = \Theta(t_0, t_3) = \Theta(t_0, t_3; a) \tag{109}$$

$$= \begin{vmatrix} \binom{\mu}{i}\varphi^{(\mu-i+1)}(t_0) & 0 & \varphi_\nu^{(\mu)}(t_0) & \varphi_{n+\nu}^{(\mu)}(t_0) \\ \binom{\mu}{i}\psi^{(\mu-i+1)}(t_0) & 0 & \psi_\nu^{(\mu)}(t_0) & \psi_{n+\nu}^{(\mu)}(t_0) \\ 0 & \binom{\mu}{i}\varphi^{(\mu-i+1)}(t_3) & \varphi_\nu^{(\mu)}(t_3) & \varphi_{n+\nu}^{(\mu)}(t_3) \\ 0 & \binom{\mu}{i}\psi^{(\mu-i+1)}(t_3) & \psi_\nu^{(\mu)}(t_3) & \psi_{n+\nu}^{(\mu)}(t_3) \end{vmatrix} \gtrless 0\ ,$$

where the individual terms always represent entire systems of n^2 elements each:

$$\mu = 0, 1, 2, \ldots n-1\ ; \quad i, \nu = 1, 2, \ldots n\ .$$

By a minor transformation of the system that is obtained from (107a) by omitting the terms of higher dimension it is now possible to bring the determinant into a clearer form and reduce its order from $4n$ to $2n$. For we can write this system down by also replacing the letter μ by i:

$$\begin{vmatrix} A_i) & 0 = D^i\left[\varphi'\tau_0 + \sum_\nu \varphi_\nu\omega_\nu\right]_0 & \binom{\mu}{i}\psi^{(\mu-i+1)}(t_0) \\ B_i) & 0 = D^i\left[\psi'\tau_0 + \sum_\nu \psi_\nu\omega_\nu\right]_0 & -\binom{\mu}{i}\varphi^{(\mu-i+1)}(t_0) \\ C_i) & \xi_3^{(i)} = D^i\left[\varphi'\tau_3 + \sum_\nu \varphi_\nu\omega_\nu\right]_3 & \binom{\mu}{i}\psi^{(\mu-i+1)}(t_3) \\ D_i) & \eta_3^{(i)} = D^i\left[\psi'\tau_3 + \sum_\nu \psi_\nu\omega_\nu\right]_3 & -\binom{\mu}{i}\varphi^{(\mu-i+1)}\ . \end{vmatrix} \tag{110}$$

Let us now multiply both sides of $A_i)$ and also of $B_i)$ by the factors after them, and sum all these products over $i = 0, 1, 2, \ldots \mu$. And let us proceed likewise with $C_i)$ and $D_i)$ and the corresponding factors by simultaneously considering the $\xi_3^{(i)}$, $\eta_3^{(i)}$ as the successive derivatives of functions $\xi_3(t)$

Factoren, indem man zugleich die $\xi_3^{(i)}$, $\eta_3^{(i)}$, was man immer kann, als die successiven Ableitungen von Functionen $\xi_3(t)$ und $\eta_3(t)$ für $t = t_3$ auffasst und die Beziehungen

83 |
$$\sum_{i=0}^{\mu} \binom{\mu}{i} \psi^{(\mu - i + 1)} D^i \mu = D^\mu (\psi' u) \quad \text{u. s. w.}$$

benutzt, so erhält man die folgenden Gleichungen:

$$\left.\begin{aligned} &E_\mu)\ 0 = D^\mu \left[\psi'\varphi'\tau_0 - \varphi'\psi'\tau_0 + \sum_\nu (\psi'\varphi_\nu - \varphi'\psi_\nu)\,\omega_\nu \right]_0 , \\ &F_\mu)\ D^\mu \left[\psi'\xi_3 - \varphi'\eta_3\right]_3 \\ &\qquad = D^\mu \left[\psi'\varphi'\tau_3 - \varphi'\psi'\tau_3 + \sum_\nu (\psi'\varphi_\nu - \varphi'\psi_\nu)\,\omega_\nu \right]_3 \end{aligned}\right\} \tag{111}$$

oder, durch Einführung der Bezeichnungen:

$$\begin{aligned} w^{(\mu)} = D^\mu \left[\psi'\xi_3 - \varphi'\eta_3\right]_{t_3} , \quad w_\nu(t) = \psi'(t)\varphi_\nu(t) - \varphi'(t)\psi_\nu(t) , \\ E_\mu) \qquad 0 = \sum_\nu w_\nu^{(\mu)}(t_0)\omega_\nu \\ F_\mu)\ w^{(\mu)} = \sum_\nu w_\nu^{(\mu)}(t_3)\omega_\nu . \end{aligned} \tag{111a}$$

Für den Fall $\psi'(t_0) \gtrless 0$ wird durch dieses Verfahren die Gleichung $A_\mu)$, die nur den Factor $\psi'(t_0)$ erhält, durch Hinzufügung der vorhergehenden $A_i)$ $(i < \mu)$ sowie der $B_i)$ mit ihren Factoren, ersetzt werden durch $E_\mu)$, wobei die Determinante sich nur mit $\psi'(t_0)$ multipliciert; ist aber $\varphi'(t_0) \gtrless 0$, was nach unseren Voraussetzungen notwendig eintreten müsste für $\psi'(t_0) = 0$, so wird statt dessen ganz analog $B_\mu)$ ersetzt werden durch $E_\mu)$, wobei die Determinante sich mit $-\varphi'(t_0)$ multipliciert. Ebenso kann $C_\mu)$ oder $D_\mu)$ durch $F_\mu)$ ersetzt werden, je nachdem $\psi'(t_3)$ oder $\varphi'(t_3)$ von 0 verschieden ist, und die Determinante ändert sich dabei um den Factor $\psi'(t_3)$ oder $-\varphi(t_3)$. Wird dieses Verfahren successive angewandt für $\mu = n-1, n-2, \ldots 2, 1, 0$, wobei immer wegen $i \leqq \mu$ die einmal umgeformten Gleichungen $A_\mu)$, $B_\mu)$, $C_\mu)$, $D_\mu)$ niemals zum zweiten Male verwendet werden, so wird im Falle $\psi'(t_0) \gtrless 0$, $\psi'(t_3) \gtrless 0$ das ganze System

$$A_\mu)\,, \quad B_\mu)\,, \quad C_\mu)\,, \quad D_\mu)$$

ersetzt werden durch: $\quad E_\mu)\,, \quad B_\mu)\,, \quad F_\mu)\,, \quad D_\mu)$

mit der neuen Determinante:

$$= \pm\, \psi'(t_0)^n \psi'(t_3)^n \Theta(t_0, t_3)$$

84 | und analog in den übrigen Fällen. Da aber $E_\mu)$ und $F_\mu)$ die Unbekannten $\tau_0^{(i)}$, $\tau_3^{(i)}$ gar nicht mehr enthalten und $\tau_0^{(i)}$ nur noch in $B_\mu)$, $\tau_3^{(i)}$ nur

and $\eta_3(t)$ for $t = t_3$, which is always possible. Then, by use of the relations

$$\sum_{i=0}^{\mu} \binom{\mu}{i} \psi^{(\mu-i+1)} D^i \mu = D^\mu (\psi' u) \quad \text{e. t. c.},$$

we obtain the following equations:

$$\left.\begin{aligned} E_\mu)\ 0 &= D^\mu \left[\psi'\varphi'\tau_0 - \varphi'\psi'\tau_0 + \sum_\nu (\psi'\varphi_\nu - \varphi'\psi_\nu)\,\omega_\nu \right]_0 , \\ F_\mu)\ D^\mu &[\psi'\xi_3 - \varphi'\eta_3]_3 \\ &= D^\mu \left[\psi'\varphi'\tau_3 - \varphi'\psi'\tau_3 + \sum_\nu (\psi'\varphi_\nu - \varphi'\psi_\nu)\,\omega_\nu \right]_3 \end{aligned}\right\} \tag{111}$$

or, by introduction of the denotations

$$\left.\begin{aligned} w^{(\mu)} = D^\mu [\psi'\xi_3 - \varphi'\eta_3]_{t_3} , \quad w_\nu(t) &= \psi'(t)\varphi_\nu(t) - \varphi'(t)\psi_\nu(t) , \\ E_\mu) \qquad 0 &= \sum_\nu w_\nu^{(\mu)}(t_0)\omega_\nu \\ F_\mu)\ w^{(\mu)} &= \sum_\nu w_\nu^{(\mu)}(t_3)\omega_\nu . \end{aligned}\right\} \tag{111a}$$

When $\psi'(t_0) \gtrless 0$, the equation $A_\mu)$, which only contains the factor $\psi'(t_0)$, will be replaced by $E_\mu)$ according to this procedure, through addition of the preceding $A_i)$ $(i < \mu)$ as well as of the $B_i)$ together with their factors, where the determinant is only multiplied by $\psi'(t_0)$; but if $\varphi'(t_0) \gtrless 0$, which, by our assumptions, should be the case when $\psi'(t_0) = 0$, then, instead, $B_\mu)$ is replaced by $E_\mu)$ along the very same lines, where the determinant is multiplied by $-\varphi'(t_0)$. Likewise, it is possible to replace $C_\mu)$ or $D_\mu)$ by $F_\mu)$, depending on whether $\psi'(t_3)$ or $\varphi'(t_3)$ differs from 0. In those cases, the determinant is altered by the factor $\psi'(t_3)$ or $-\varphi(t_3)$. Applying this procedure successively for $\mu = n-1, n-2, \ldots 2, 1, 0$, where, on account of $i \leqq \mu$, the equations $A_\mu)$, $B_\mu)$, $C_\mu)$, $D_\mu)$ can never be used again once they have been transformed, then, for $\psi'(t_0) \gtrless 0$, $\psi'(t_3) \gtrless 0$, the entire system

$$A_\mu), \quad B_\mu), \quad C_\mu), \quad D_\mu)$$

is replaced by $\quad E_\mu), \quad B_\mu), \quad F_\mu), \quad D_\mu)$

with the new determinant

$$= \pm\, \psi'(t_0)^n \psi'(t_3)^n \Theta(t_0, t_3)$$

and the same holds for the remaining cases. But since $E_\mu)$ and $F_\mu)$ no longer contain the unknowns $\tau_0^{(i)}$, $\tau_3^{(i)}$ and since $\tau_0^{(i)}$ and $\tau_3^{(i)}$ only still occur in $B_\mu)$

noch in D_μ) erscheint, so zerfällt die neue Determinante in ein Product:

$$\pm\Theta\psi'(t_0)^n\psi'(t_3)^n$$

$$= \begin{vmatrix} 0 & 0 & w_\nu^{(\mu)}(t_0)\ w_{n+\nu}^{(\mu)}(t_0) \\ \binom{\mu}{i}\psi^{(\mu-i+1)}(t_0) & 0 & \psi_\nu^{(\mu)}(t_0)\ \psi_{n+\nu}^{(\mu)}(t_0) \\ 0 & 0 & w_\nu^{(\mu)}(t_3)\ w_{n+\nu}^{(\mu)}(t_3) \\ 0 & \binom{\mu}{i}\psi^{(\mu-i+1)}(t_3) & \psi_\nu^{(\mu)}(t_3)\ \psi_{n+\nu}^{(\mu)}(t_3) \end{vmatrix}$$

$$= \pm\left|\binom{\mu}{i}\psi^{(\mu-i+1)}(t_0)\right| \cdot \left|\binom{\mu}{i}\psi^{(\mu-i+1)}(t_3)\right| \cdot \begin{vmatrix} w_\nu^{(\mu)}(t_0)\ w_{n+\nu}^{(\mu)}(t_0) \\ w_\nu^{(\mu)}(t_3)\ w_{n+\nu}^{(\mu)}(t_3) \end{vmatrix},$$

wo die beiden Determinanten n ten Grades sich wegen $\binom{\mu}{i} = 0$ $(\mu < i)$ auf ihre Diagonalglieder $\psi'(t_0)^n$ und $\psi'(t_3)^n$ reducieren und sich gegen die Factoren links aufheben, so dass schliesslich:

$$\Theta(t_0, t_3; a) = \pm\left|w_\nu^{(\mu)}(t_0)\ w_\nu^{(\mu)}(t_3)\right| \tag{112}$$
$$(\mu = 0, 1, \ldots n-1;\ \nu = 1, 2, \ldots 2n)\ ,$$

wo zuletzt ν den Colonnenindex und jedes hingeschriebene Glied das System von n $2n$ gliedrigen Colonnen bedeutet, und

$$w_\nu(t) = \psi'(t)\varphi'(t)\ \psi_\nu(t)\ .$$

Dieselbe Formel (112) ergiebt sich auf ganz analogem Wege, wenn eine der Grössen $\psi'(t_0)$ und $\psi'(t_3)$ verschwindet und dafür sicher $\varphi'(t_0)$ oder $\varphi'(t_3)$ von Null verschieden ist.

Wir können immer annehmen, *dass* $\Theta(t_0, t)$ *nicht für alle Werte* t *eines endlichen Intervalles* und daher, wenn F und somit auch φ, ψ, Θ als „analytische" Functionen vorausgesetzt werden, für das *ganze* particuläre Integral a der Differentialgleichung *beständig verschwindet*. Sonst müsste es nämlich $2n$ Constanten $\omega_1, \ldots \omega_{2n}$ geben, für welche überall

85 |
$$\sum_{\nu=1}^{2n} \omega_\nu w_\nu(t) = 0$$

wird, falls nicht alle Determinanten n ten Grades des Systemes $w_\nu^{(\mu)}(t_0)$ $(\mu = 0, 1, \ldots n-1; \nu = 1, 2, \ldots 2n)$ verschwinden. Diesen Ausnahmefall kann man aber vermeiden, indem man t_0 durch einen beliebig nahen Punkt t_0' ersetzt, oder es müssten alle diese Determinanten $\left|w_\nu^{(\mu)}(t)\right|$ in einem ganzen Intervall verschwinden, und daraus folgte dann wieder eine Relation der behaupteten Form. Es wäre also immer:

$$\sum_\nu \omega_\nu \left(\psi'(t)\varphi_\nu(t) - \varphi'(t)\psi_\nu(t)\right)$$
$$= \psi'(t)\sum_\nu \varphi_\nu(t)\omega_\nu - \varphi'(t)\sum_\nu \psi_\nu(t)\omega_\nu = 0\ .$$

and D_μ) respectively, the new determinant decomposes into a product:

$$\pm\Theta\psi'(t_0)^n\psi'(t_3)^n$$

$$= \begin{vmatrix} 0 & 0 & w_\nu^{(\mu)}(t_0)\ w_{n+\nu}^{(\mu)}(t_0) \\ \binom{\mu}{i}\psi^{(\mu-i+1)}(t_0) & 0 & \psi_\nu^{(\mu)}(t_0)\ \psi_{n+\nu}^{(\mu)}(t_0) \\ 0 & 0 & w_\nu^{(\mu)}(t_3)\ w_{n+\nu}^{(\mu)}(t_3) \\ 0 & \binom{\mu}{i}\psi^{(\mu-i+1)}(t_3) & \psi_\nu^{(\mu)}(t_3)\ \psi_{n+\nu}^{(\mu)}(t_3) \end{vmatrix}$$

$$= \pm\left|\binom{\mu}{i}\psi^{(\mu-i+1)}(t_0)\right| \cdot \left|\binom{\mu}{i}\psi^{(\mu-i+1)}(t_3)\right| \cdot \begin{vmatrix} w_\nu^{(\mu)}(t_0)\ w_{n+\nu}^{(\mu)}(t_0) \\ w_\nu^{(\mu)}(t_3)\ w_{n+\nu}^{(\mu)}(t_3) \end{vmatrix},$$

where the two determinants of n th order are reduced to their diagonal terms $\psi'(t_0)^n$ and $\psi'(t_3)^n$ on account of $\binom{\mu}{i} = 0$ $(\mu < i)$ and canceled out by the factors on the left side so that finally

$$\Theta(t_0, t_3; a) = \pm\left|w_\nu^{(\mu)}(t_0)\ w_\nu^{(\mu)}(t_3)\right| \tag{112}$$
$$(\mu = 0, 1, \ldots n-1;\ \nu = 1, 2, \ldots 2n)\ ,$$

where ν denotes the column index and every entered term the system of n columns of $2n$ terms, and

$$w_\nu(t) = \psi'(t)\varphi'(t)\ \psi_\nu(t)\ .$$

The same formula (112) is obtained along the same lines when, on the one hand, one of the quantities $\psi'(t_0)$ and $\psi'(t_3)$ vanishes and, on the other hand, $\varphi'(t_0)$ or $\varphi'(t_3)$ is different from zero.

We may always assume *that $\Theta(t_0, t)$ does not always vanish for all values t of a finite interval* and hence, provided F and thus also φ, ψ, Θ are taken to be "analytic" functions, for the *entire* particular integral a of the differential equation. For otherwise there would have to be $2n$ constants $\omega_1, \ldots \omega_{2n}$ such that everywhere

$$\sum_{\nu=1}^{2n} \omega_\nu w_\nu(t) = 0\ ,$$

provided that not all determinants of n th order of the system $w_\nu^{(\mu)}(t_0)$ $(\mu = 0, 1, \ldots n-1; \nu = 1, 2, \ldots 2n)$ vanish. But we can avoid this exceptional case by replacing t_0 by an arbitrarily close point t_0', or all these determinants $\left|w_\nu^{(\mu)}(t)\right|$ would have to vanish in an entire interval, from which, in turn, a relation of the asserted form would follow. Hence we would always have

$$\sum_\nu \omega_\nu\left(\psi'(t)\varphi_\nu(t) - \varphi'(t)\psi_\nu(t)\right)$$
$$= \psi'(t)\sum_\nu \varphi_\nu(t)\omega_\nu - \varphi'(t)\sum_\nu \psi_\nu(t)\omega_\nu = 0\ .$$

Da aber hier nirgends die $\varphi_\nu(t)$, $\psi_\nu(t)$ unendlich werden oder die $\varphi'(t)$, $\psi'(t)$ gleichzeitig verschwinden sollen, so wird die Function:

$$\frac{\sum_\nu \varphi_\nu(t)\omega_\nu}{\varphi'(t)} = \frac{\sum_\nu \psi_\nu(t)\omega_\nu}{\psi'(t)} = \tau(t)$$

im ganzen Intervall endlich bleiben, so dass man schreiben kann:

$$\begin{aligned}\varphi(t, a+\varepsilon\omega) &= \varphi(t) + \varepsilon\sum_\nu \varphi_\nu(t)\omega_\nu + (\varepsilon)_2 \\ &= \varphi(t) + \varepsilon\tau\varphi'(t) + (\varepsilon)_2 = \varphi(t+\varepsilon\tau) + (\varepsilon)_2 \,, \\ \psi(t, a+\varepsilon\omega) &= \psi(t) + \varepsilon\tau\psi'(t) + (\varepsilon)_2 = \psi(t+\varepsilon\tau) + (\varepsilon)_2 \,.\end{aligned}$$

Nun ist aber der Ansatz

$$x = \varphi(t+\varepsilon\tau)\,, \qquad y = \psi(t+\varepsilon\tau)$$

nur eine andere Darstellung der Curve a:

$$x = \varphi(t)\,, \qquad y = \psi(t)\,;$$

es gäbe also eine durch $u_\nu = a_\nu + \varepsilon\omega_\nu$ charakterisierte Lösung der Differentialgleichung, die sich für kleine ε von a selbst nur um Glieder höherer Dimension unterschiede, oder, anders ausgedrückt: *in a fielen zwei unendlich benachbarte Lösungen der Differentialgleichung zusammen*, a wäre also eine *singuläre Curve* aus der Schar der Integrale und müsste besonders untersucht werden.

86 Schliessen wir diesen Fall aus, so muss die Gleichung $\mid \Theta(t_0, t; a) = 0$, wenn sie überhaupt bestehen kann, eine *kleinste Wurzel* $t = t_0' > t_0$ besitzen. Dann mögen nach der Bezeichnung des Herrn Prof. *Weierstrass* die beiden durch $t = t_0$ und $t = t_0'$ definierten Stellen der Curve zwei „auf a *conjugierte Punkte*“ genannt werden.

Für ein solches Paar conjugierter Punkte t_0, $t_3 = t_0'$ müssen sich wegen $\Theta = 0$ die linearen Gleichungen (110) befriedigen lassen für $\xi_3^{(\mu)} = 0$, $\eta_3^{(\mu)} = 0$, so dass für gewisse Werte der Grössen $\tau_0^{(\mu)}, \tau_3^{(\mu)}, \omega_\nu$, die nicht sämtlich verschwinden, die Gleichungen (107) die Form annehmen:

$$\begin{aligned} D^\mu\left[\varphi\left(t+\varepsilon\tau_0, a+\varepsilon\omega\right)\right]_{t_0} &= \varphi^{(\mu)}(t_0) + (\varepsilon)_2 \\ D^\mu\left[\psi\left(t+\varepsilon\tau_0, a+\varepsilon\omega\right)\right]_{t_0} &= \psi^{(\mu)}(t_0) + (\varepsilon)_2 \\ D^\mu\left[\varphi\left(t+\varepsilon\tau_3, a+\varepsilon\omega\right)\right]_{t_3} &= \varphi^{(\mu)}(t_3) + (\varepsilon)_2 \\ D^\mu\left[\psi\left(t+\varepsilon\tau_3, a+\varepsilon\omega\right)\right]_{t_3} &= \psi^{(\mu)}(t_3) + (\varepsilon)_2 \\ (\mu = 0, 1, \ldots n-1)& \end{aligned} \tag{113}$$

But since, in this case, neither the $\varphi_\nu(t)$, $\psi_\nu(t)$ are ever supposed to become infinite nor the $\varphi'(t)$, $\psi'(t)$ to vanish simultaneously, the function

$$\frac{\sum_\nu \varphi_\nu(t)\omega_\nu}{\varphi'(t)} = \frac{\sum_\nu \psi_\nu(t)\omega_\nu}{\psi'(t)} = \tau(t)$$

remains finite in the entire interval, so that we can write

$$\begin{aligned}\varphi(t, a+\varepsilon\omega) &= \varphi(t) + \varepsilon\sum_\nu \varphi_\nu(t)\omega_\nu + (\varepsilon)_2 \\ &= \varphi(t) + \varepsilon\tau\varphi'(t) + (\varepsilon)_2 = \varphi(t+\varepsilon\tau) + (\varepsilon)_2\ , \\ \psi(t, a+\varepsilon\omega) &= \psi(t) + \varepsilon\tau\psi'(t) + (\varepsilon)_2 = \psi(t+\varepsilon\tau) + (\varepsilon)_2\ .\end{aligned}$$

Now the ansatz

$$x = \varphi(t+\varepsilon\tau)\ , \qquad y = \psi(t+\varepsilon\tau)$$

is but a different representation of the curve a:

$$x = \varphi(t)\ , \qquad y = \psi(t)\ ;$$

and hence there would be a solution of the differential equation characterized by $u_\nu = a_\nu + \varepsilon\omega_\nu$ which, for small ε, differs from a itself only by terms of higher dimension, or, put differently: *two infinitely neighboring solutions of the differential equations would coincide on a*. Hence a would be a *singular curve* from the family of the integrals and would require a separate investigation.

If we exclude this case, then the equation $\Theta(t_0, t; a) = 0$, provided that it can obtain at all, must have a *smallest root* $t = t_0' > t_0$. Then, following the terminology used by Prof. *Weierstrass*, the two positions of the curve defined by $t = t_0$ and $t = t_0'$ shall be called two "*points conjugate* on a".

Given such a pair of conjugate points $t_0, t_3 = t_0'$, the linear equations (110) must be capable of satisfaction on account of $\Theta = 0$ for $\xi_3^{(\mu)} = 0$, $\eta_3^{(\mu)} = 0$ so that, for certain values of the quantities $\tau_0^{(\mu)}, \tau_3^{(\mu)}, \omega_\nu$ not all of which vanish, the equations (107) take the form

$$\begin{aligned} D^\mu\left[\varphi\left(t+\varepsilon\tau_0, a+\varepsilon\omega\right)\right]_{t_0} &= \varphi^{(\mu)}(t_0) + (\varepsilon)_2 \\ D^\mu\left[\psi\left(t+\varepsilon\tau_0, a+\varepsilon\omega\right)\right]_{t_0} &= \psi^{(\mu)}(t_0) + (\varepsilon)_2 \\ D^\mu\left[\varphi\left(t+\varepsilon\tau_3, a+\varepsilon\omega\right)\right]_{t_3} &= \varphi^{(\mu)}(t_3) + (\varepsilon)_2 \\ D^\mu\left[\psi\left(t+\varepsilon\tau_3, a+\varepsilon\omega\right)\right]_{t_3} &= \psi^{(\mu)}(t_3) + (\varepsilon)_2 \\ (\mu = 0, 1, \ldots n-1)& \end{aligned} \tag{113}$$

für beliebig kleine ε, die unter einer gewissen Grenze liegen. Dann ist nach (21) auch für alle „Osculations-Invarianten“ $\Phi\left(x^{(\mu)}, y^{(\mu)}\right)$ bis zur $n-1$ ten Ordnung:

$$\begin{aligned}\Phi\left(\varphi^{(\mu)}(t_0+\varepsilon\tau_0, a+\varepsilon\omega), \psi^{(\mu)}(t_0+\varepsilon\tau_0,\, a+\varepsilon\omega)\right)\\ = \Phi\left(\varphi^{(\mu)}(t_0), \psi^{(\mu)}(t_0)\right) + (\varepsilon)_2\\ \Phi\left(\varphi^{(\mu)}(t_3+\varepsilon\tau_3, a+\varepsilon\omega), \psi^{(\mu)}(t_3+\varepsilon\tau_3,\, a+\varepsilon\omega)\right)\\ = \Phi\left(\varphi^{(\mu)}(t_3), \psi^{(\mu)}(t_3)\right) + (\varepsilon)_2\,,\end{aligned} \tag{113a}$$

was man so ausdrücken kann: *es giebt eine zu a unendlich benachbarte Lösung der Differentialgleichung des Problems* ($u_\nu = a_\nu + \varepsilon\,\omega_\nu$ *für ein unendlich kleines* ε), *welche a in den beiden Punkten* $t = t_0$ *und* $t = t_3$ *von* $n-1$ *ter Ordnung berührt*, eine Eigenschaft, die als Definition der „conjugierten Punkte“ angesehen werden kann, wenn sie in dem hier entwickelten Sinne verstanden wird. Oft giebt es wirklich eine Schar von reellen Lösungen u, die sich an a beliebig eng anschliessen und diese Curve ausser in 0 immer noch in einem zweiten Punkte 3 von $n-1$ ter Ordnung berühren, dann wird die *Grenzlage*
87 dieser | Punkte 3, wenn man u mit a zusammenfallen lässt, durch den zu 0 „conjugierten“ Punkt $0'$ dargestellt. Doch braucht dieses Verhalten nicht für jedes Paar conjugierter Punkte stattzufinden, worauf aber hier nicht weiter eingegangen werden soll.

Auch die ursprünglichste und gebräuchlichste Definition der „conjugierten Punkte“ durch die Existenz gewisser Integrale einer linearen Differentialgleichung ($\delta G = 0$) soll hier ausser Betracht bleiben, da sie für die vorliegende Untersuchung ohne wesentliche Bedeutung ist.

Sind nun 0 und 3 *nicht* zwei auf a conjugierte Punkte, also

$$\pm\,\Theta(t_0, t_3; a) > \zeta > 0\,, \tag{114}$$

so kann man positive Grössen g'_μ so klein angeben, dass für

$$\left|\xi_3^{(\mu)}\right| < g'_\mu\,, \quad \left|\eta_3^{(\mu)}\right| < g'_\mu \qquad (\mu = 0, 1, \ldots n-1) \tag{115}$$

die Gleichungen (107) in der dort angegebenen Weise durch Potenzreihen-Entwicklung nach den Unbekannten $\tau_0^{(\mu)}$, $\tau_3^{(\mu)}$, ω_ν aufgelöst werden können, während die Beträge dieser Grössen mit den g'_μ gleichzeitig unendlich klein werden. Das so gefundene Werte-System ist aber auch *die einzige Lösung*, für welche die Beträge der Unbekannten unter eine gewisse Grenze α' fallen, ja die einzige, für welche nur $|\omega_\nu| < \alpha \leqq \alpha'$ ($\nu = 1, 2, \ldots 2n$) sein soll, ohne dass die $\tau_0^{(\mu)}$, $\tau_3^{(\mu)}$ von vorn herein irgendwelchen Beschränkungen unterworfen würden. Denn den Bemerkungen von (63) zufolge können allein durch hinreichende Verkleinerung der g'_μ und $|\omega_\nu|$ oder der g'_μ und α auch die $\left|\tau_0^{(\mu)}\right|$, $\left|\tau_3^{(\mu)}\right|$ immer beliebig klein, also auch $< \alpha'$ gemacht werden, vorausgesetzt, dass a sich im betrachteten Intervall nicht selbst durchschneidet. Lässt man jetzt t_3 das Intervall $t_1 \leqq t_3 \leqq t_2$ durchlaufen, in welchem alle unsere Vorausset-

for arbitrarily small ε below a certain limit. Then also for all "osculation invariants" $\Phi\left(x^{(\mu)}, y^{(\mu)}\right)$ up to the $n-1$ th order, by (21),

$$\begin{aligned}
&\Phi\left(\varphi^{(\mu)}(t_0+\varepsilon\tau_0, a+\varepsilon\omega), \psi^{(\mu)}(t_0+\varepsilon\tau_0,\, a+\varepsilon\omega)\right) \\
&\qquad = \Phi\left(\varphi^{(\mu)}(t_0), \psi^{(\mu)}(t_0)\right) + (\varepsilon)_2 \\
&\Phi\left(\varphi^{(\mu)}(t_3+\varepsilon\tau_3, a+\varepsilon\omega), \psi^{(\mu)}(t_3+\varepsilon\tau_3,\, a+\varepsilon\omega)\right) \\
&\qquad = \Phi\left(\varphi^{(\mu)}(t_3), \psi^{(\mu)}(t_3)\right) + (\varepsilon)_2\ ,
\end{aligned} \tag{113a}$$

which may be expressed as follows: *There is a solution of the differential equation of the problem ($u_\nu = a_\nu + \varepsilon\,\omega_\nu$ for an infinitely small ε) infinitely neighboring a that has contact of $n-1$ th order with a at the two points $t = t_0$ and $t = t_3$.* This property can be regarded as a definition of the "conjugate points", provided it is understood in the sense developed here. There often really is a family of real solutions u that follow a arbitrarily closely and that still have contact of $n-1$ th order with this curve at a second point 3 besides 0. In this case, the *limit position* of these points 3 is represented by the point $0'$ "conjugate" to 0, if we let u coincide with a. That this behavior, however, need not arise for every pair of conjugate points will not be discussed any further here.

Nor shall we discuss the original and most common definition of "conjugate points" in terms of the existence of certain integrals of a linear differential equation ($\delta G = 0$), since it has no real bearing on our present investigation.

Now if 0 and 3 are *not* two points conjugate on a, hence

$$\pm\,\Theta(t_0, t_3; a) > \zeta > 0\ , \tag{114}$$

then we can take positive quantities g'_μ sufficiently small so that, for

$$\left|\xi_3^{(\mu)}\right| < g'_\mu\ , \quad \left|\eta_3^{(\mu)}\right| < g'_\mu \qquad (\mu = 0, 1, \ldots n-1)\ , \tag{115}$$

the equations (107) can be solved in the way specified there for the unknowns $\tau_0^{(\mu)}$, $\tau_3^{(\mu)}$, ω_ν by power series expansion, where the values of these quantities become infinitely small together with the g'_μ. The value system thus found is also *the unique solution* for which the values of the unknowns are below under a certain limit α', and even the unique solution for which we should only have $|\omega_\nu| < \alpha \leqq \alpha'$ $(\nu = 1, 2, \ldots 2n)$ without imposing some restrictions on the $\tau_0^{(\mu)}$, $\tau_3^{(\mu)}$ from the outset. For, according to the observations concerning (63), we can make the $\left|\tau_0^{(\mu)}\right|$, $\left|\tau_3^{(\mu)}\right|$ arbitrarily small, and hence also $< \alpha'$, only by making the g'_μ and $|\omega_\nu|$ or the g'_μ and α sufficiently small, provided that a does not intersect itself in the interval under consideration. If we now let t_3 pass through the interval $t_1 \leqq t_3 \leqq t_2$, in which our assumptions are all valid without exception, then the ζ belonging to the individual

zungen ausnahmslos gültig sind, so können die zu den einzelnen Punkten t_3 gehörigen ζ durch ihre *constante untere Grenze* und ebenso die g'_μ und α durch ihre *oberen Grenzen* im Intervall ersetzt werden.

Die hier entwickelte Auflösung der Gleichungen (107) ist immer anwendbar, wenn die $\xi_3^{(\mu)}$, $\eta_3^{(\mu)}$ *beliebige* Grössen sind, die den Bedingungen (115) genügen. Haben sie aber die Bedeutung (108)

$$\xi_3^{(\mu)} = \overline{\varphi}_1^{(\mu)}(t_3 - \varepsilon) - \overline{\varphi}_1^{(\mu)}(t_3) \ , \quad \eta_3^{(\mu)} = \overline{\psi}_1^{(\mu)}(t_3 - \varepsilon) - \overline{\psi}_1^{(\mu)}(t_3) \ ,$$

88 | so kann (115) und damit auch (107) für ein hinreichend kleines positives ε_0 durch $\varepsilon < \varepsilon_0$ immer befriedigt werden.

Dann giebt es eine eindeutig bestimmte und an a sich stetig anschliessende *Schar U von Lösungen der Differentialgleichungen $G = 0$*, welche für das Stück der Curve $\overline{a}$ vom Punkte 3 an bis 4 die Coordinaten:

$$x = \overline{\varphi}_1(t_3 - \varepsilon) = \overline{\varphi}(\lambda - \iota) \ , \quad y = \overline{\psi}_1(t_3 = \varepsilon) = \overline{\psi}(\lambda - \iota)$$

und die im Anfang des Abschnittes geforderten Eigenschaften besitzt.

Dieselben Betrachtungen wie für a gelten auch für jede andere Lösung u der Differentialgleichung, $x = \varphi(t, u)$, $y = \psi(t, u)$, welche denselben Bedingungen wie a genügt. So lange also $\Theta(t'_0, t'_3, u) \gtrless 0$ bleibt, wenn mit t'_0 und t'_3 die zu den Endpunkten 0 und 3 gehörigen Werte von t bezeichnet werden, und die Curven u fortfahren, sich zwischen 0 und 3 in der angegebenen Weise regulär zu verhalten, kann die Schar U, und zwar nur auf eine einzige Weise, längs der Curve $\overline{a}$ *stetig fortgesetzt werden*. Die zu 0 „conjugierten" Punkte 3 der Curven u, in denen die eindeutige stetige Fortsetzung der Schar im allgemeinen ein Ende findet, und nach dem Früheren zwei unendlich benachbarte u einander von $n-1$ ter Ordnung berühren, können als „Verzweigungsstellen" der Schar U aufgefasst werden.

Von der Curve a werde nun vorausgesetzt, dass sie sich im Intervall $t_1 \leqq t \leqq t_2$ (einschliesslich der Grenzen) regulär verhalte und keinen zum Anfangspunkt $t = t_1$ conjugierten Punkt $t = t'_1 \leqq t_2$ besitze, so dass die Determinante $\Theta(t_1, t)$ immer von Null verschieden ist für $t_1 < t \leqq t_2$ und nur für $t = t_1$ verschwindet. Dann kann aber wegen der Stetigkeit der Function $\Theta(t', t'')$ immer eine Grösse $t_0 < t_1$ so nahe an t_1 und damit auf a ein Punkt 0 so nahe vor 1 angenommen werden, dass das reguläre Verhalten von a sich auch auf das grössere Intervall $t_0 \leqq t \leqq t_2$ erstreckt, gleichzeitig aber $\Theta(t_0, t)$ im ganzen Intervall $t_1 \ldots t_2$ nirgends, auch an keiner der Grenzen mehr, verschwindet, der absolute Betrag also eine positive *untere Grenze* ζ besitzt:

$$|\Theta(t_0, t)| > \zeta > 0 \qquad (t_1 \leqq t \leqq t_2) \ . \tag{116}$$

89 Mithin können nach den früheren Betrachtungen positive Con-|stanten α, g'_μ so klein angenommen werden, dass die Gleichungen (107) unter den Bedingungen (115) für $t_1 \leqq t_3 \leqq t_2$ immer und nur auf eine einzige Weise durch

points t_3 can be replaced by their *constant lower limit*, and likewise the g'_μ and α by their *upper limits* in the interval.

The solution of the equations (107) developed here is applicable whenever the $\xi_3^{(\mu)}$, $\eta_3^{(\mu)}$ are *arbitrary* quantities satisfying the conditions (115). But when they have the meaning (108)

$$\xi_3^{(\mu)} = \overline{\varphi}_1^{(\mu)}(t_3 - \varepsilon) - \overline{\varphi}_1^{(\mu)}(t_3) \,, \quad \eta_3^{(\mu)} = \overline{\psi}_1^{(\mu)}(t_3 - \varepsilon) - \overline{\psi}_1^{(\mu)}(t_3) \,,$$

then (115), and hence also (107), can always be satisfied by $\varepsilon < \varepsilon_0$ for a sufficiently small positive ε_0.

In this case, there is a uniquely determined *family* U continuously following *a of solutions of the differential equations* $G = 0$ which for the segment of the curve $\overline{a}$ from point 3 up to point 4 has the coordinates

$$x = \overline{\varphi}_1(t_3 - \varepsilon) = \overline{\varphi}(\lambda - \iota) \,, \quad y = \overline{\psi}_1(t_3 = \varepsilon) = \overline{\psi}(\lambda - \iota)$$

and which possesses the properties called for at the beginning of this section.

The considerations valid for a are also valid for any other solution u of the differential equation, $x = \varphi(t, u)$, $y = \psi(t, u)$, that meets the same requirements as a. Hence, as long as still $\Theta(t'_0, t'_3, u) \gtrless 0$, where t'_0 and t'_3 denote the values of t belonging to the endpoints 0 and 3 and as long as the curves u continue to be regular between 0 and 3 in the specified way, the family U can be *continuously continued* along the curve $\overline{a}$ in one way only. The points 3 of the curves u which are "conjugate" to 0 and in which the unique continuous continuation of the family usually ends, and at which, according to what was said above, two infinitely neighboring u have contact of the $n-1$ th order with one another may be considered "branching positions" of the family U.

Now suppose that the curve a is regular in the interval $t_1 \leqq t \leqq t_2$ (including the boundaries) and possesses no point $t = t'_1 \leqq t_2$ conjugate to the starting point $t = t_1$ so that the determinant $\Theta(t_1, t)$ always differs from zero for $t_1 < t \leqq t_2$ and only vanishes when $t = t_1$. But then, on account of the continuity of the function $\Theta(t', t'')$, we can always take a quantity $t_0 < t_1$ close enough to t_1, and hence a point 0 close enough in front of 1 on a so that the regularity of a also extends to the greater interval $t_0 \leqq t \leqq t_2$, while, at the same time, $\Theta(t_0, t)$ does not vanish anywhere in the entire interval $t_1 \ldots t_2$, including even its boundaries, and the absolute value thus possesses a positive *lower limit* ζ:

$$|\Theta(t_0, t)| > \zeta > 0 \qquad (t_1 \leqq t \leqq t_2) \,. \tag{116}$$

Therefore, according to the previous considerations, we can take positive constants α, g'_μ sufficiently small so that, under the conditions (115), the equations (107) are always and only in one single way satisfied by values of

Werte der Unbekannten $\tau_0^{(\mu)}$, $\tau_3^{(\mu)}$, ω_ν mit der Nebenbedingung $|\omega_\nu| < \alpha$ befriedigt werden, dass ferner die Curven $x = \varphi(t, a+\omega)$, $y = \psi(t, a+\omega)$ für $t_0 + \tau_0 \leqq t \leqq t_3 + \tau_3$ sich ebenso wie a regulär verhalten und dass endlich immer

$$\pm\, \Theta(t_0 + \tau_0, t_3 + \tau_3; a + \omega) > 0 \tag{117}$$

wird. Diese Gleichungen (107) können für:

$$\xi_3^{(\mu)} = D^\mu \overline{\varphi}(\lambda) - \varphi^{(\mu)}(t_3)\,, \quad \eta_3^{(\mu)} = D^\mu \overline{\psi}(\lambda) - \psi^{(\mu)}(t_3)\,,$$

also nach (115) für:

$$\begin{gathered} \left|D^\mu \overline{\varphi}(\lambda) - \varphi^{(\mu)}(t_3)\right| < g_\mu \leqq g'_\mu \\ \left|D^\mu \overline{\psi}(\lambda) - \psi^{(\mu)}(t_3)\right| < g_\mu \leqq g'_\mu \\ (\mu = 0, 1, \ldots n-1) \qquad\qquad [\text{vergl. (37)}] \end{gathered} \tag{118}$$

in der Form geschrieben werden:

$$\begin{gathered} \left[D^\mu \varphi(t + \tau_0, a + \omega)\right]_{t_0} = \varphi^{(\mu)}(t_0) \\ \left[D^\mu \psi(t + \tau_0, a + \omega)\right]_{t_0} = \psi^{(\mu)}(t_0) \\ \left[D^\mu \varphi(t + \tau_3, a + \omega)\right]_{t_3} = D^\mu \overline{\varphi}(\lambda) \\ \left[D^\mu \psi(t + \tau_3, a + \omega)\right]_{t_3} = D^\mu \overline{\psi}(\lambda) \\ (\mu = 0, 1, \ldots n-1)\,, \end{gathered} \tag{119}$$

wo alle Bezeichnungen ihre früheren Bedeutungen haben. Dann drücken sie aus, dass das particuläre Integral u:

$$x = \varphi(t, a+\omega)\,, \quad y = \psi(t, a+\omega)\,,$$

einerseits mit a im Punkte 0 ($x = \varphi(t_0), y = \psi(t_0)$), andrerseits aber im Punkte 3:

$$x = \varphi(t_3 + \tau_3, a+\omega)\,, \quad y = \psi(t_3 + \tau_3, a+\omega)\,,$$

mit einer zweiten Curve $\overline{a}$ $\big(x = \overline{\varphi}(\lambda), y = \overline{\psi}(\lambda)\big)$ eine Berührung von $n-1$ ter Ordnung eingeht. Wenn also $\overline{a}$ im ganzen Intervall $\lambda_1 \leqq \lambda \leqq \lambda_2$ überall
90 für passende Werte der t_3 zwischen t_1 und t_2 | den Bedingungen (118) in der Bedeutung von (37) genügt, so bilden die so bestimmten Curven u *eine Schar U mit den sämtlichen im Anfang des Abschnittes geforderten Eigenschaften.* Denn sie müssen sich auch *stetig* an einanderschliessen, weil die Auflösung der Gleichungen (119) unter der Bedingung $|\omega_\nu| < \alpha$ eine *eindeutige* ist und daher die „stetige Fortsetzung" der Schar, die wegen (117) immer
Fig. 6 möglich ist, niemals zu anderen Ergebnissen führen kann.

the unknown $\tau_0^{(\mu)}$, $\tau_3^{(\mu)}$, ω_ν under the ancillary condition $|\omega_\nu| < \alpha$, for $t_1 \leqq t_3 \leqq t_2$, and so that, furthermore, the curves $x = \varphi(t, a+\omega)$, $y = \psi(t, a+\omega)$ are regular for $t_0 + \tau_0 \leqq t \leqq t_3 + \tau_3$ just like a, and so that, finally, always

$$\pm\, \Theta(t_0 + \tau_0, t_3 + \tau_3; a + \omega) > 0\,. \tag{117}$$

Thus, when

$$\xi_3^{(\mu)} = D^\mu \overline{\varphi}(\lambda) - \varphi^{(\mu)}(t_3)\,, \quad \eta_3^{(\mu)} = D^\mu \overline{\psi}(\lambda) - \psi^{(\mu)}(t_3)\,,$$

and hence, by (115), when

$$\begin{gathered} \left|D^\mu \overline{\varphi}(\lambda) - \varphi^{(\mu)}(t_3)\right| < g_\mu \leqq g'_\mu \\ \left|D^\mu \overline{\psi}(\lambda) - \psi^{(\mu)}(t_3)\right| < g_\mu \leqq g'_\mu \\ (\mu = 0, 1, \ldots n-1) \qquad\qquad [\text{comp. (37)}] \end{gathered} \tag{118}$$

these equations (107) can be written in the form

$$\begin{aligned} \left[D^\mu \varphi(t + \tau_0, a + \omega)\right]_{t_0} &= \varphi^{(\mu)}(t_0) \\ \left[D^\mu \psi(t + \tau_0, a + \omega)\right]_{t_0} &= \psi^{(\mu)}(t_0) \\ \left[D^\mu \varphi(t + \tau_3, a + \omega)\right]_{t_3} &= D^\mu \overline{\varphi}(\lambda) \\ \left[D^\mu \psi(t + \tau_3, a + \omega)\right]_{t_3} &= D^\mu \overline{\psi}(\lambda) \\ (\mu = 0, 1, \ldots n-1)\,, & \end{aligned} \tag{119}$$

where all denotations have their previous meaning. They then express the fact that the particular integral u:

$$x = \varphi(t, a+\omega)\,, \quad y = \psi(t, a+\omega)\,,$$

on the one hand, has contact of $n-1$ th order with a at the point 0 ($x = \varphi(t_0)$, $y = \psi(t_0)$) but, on the other hand, with a second curve $\overline{a}$ ($x = \overline{\varphi}(\lambda)$, $y = \overline{\psi}(\lambda)$) at the point 3:

$$x = \varphi(t_3 + \tau_3, a+\omega)\,, \quad y = \psi(t_3 + \tau_3, a+\omega)\,.$$

Hence, if $\overline{a}$ satisfies the conditions (118) in the sense of (37) everywhere in the entire interval $\lambda_1 \leqq \lambda \leqq \lambda_2$ for suitable values of the t_3 between t_1 and t_2, then the curves u so determined form *a family U possessing all properties called for at the beginning of the section*. For they also must *continuously* follow one another, since the solution of the equations (119) is *unique* under the condition $|\omega_\nu| < \alpha$, and hence the "continuous continuation" of the family, which, on account of (117), is always possible, can never lead to different results.

Fig. 6

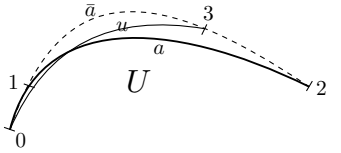


Fig. 6.

Wenn ferner die Curven $\overline{a}$ und a in den Endpunkten 1 und 2 einander von $n-1$ ter Ordnung berühren, so müssen auf Grund derselben Eindeutigkeit die zu diesen Endpunkten gehörigen Curven u mit a selbst zusammenfallen, sodass nunmehr der Satz VII angewandt werden kann.

Ist nämlich gleichzeitig für ein constantes positives ϱ beständig

$$E_1\left(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(n)}, \overline{y}^{(n)}\right) > \varrho > 0 \tag{120}$$
$$\left(x^{(\mu)} = \varphi^{(\mu)}(t), y^{(\mu)} = \psi^{(\mu)}(t), t_1 \leqq t \leqq t_2\right)$$

und für willkürliche Werte der Variablen $\overline{x}^{(n)}, \overline{y}^{(n)}$, wo E_1 die Bedeutung (78) besitzt, so können in (118) oder in (37) die Constanten g_μ so klein gewählt werden, dass auch immer

$$E_1(\lambda) = E_1\left(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(n)}, \overline{y}^{(n)}\right) > 0$$

und somit auch $$E(\lambda) = E\left(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(n)}, \overline{y}^{(n)}\right) \geqq 0$$

ist für:

$$x^{(\mu)} = D^\mu \varphi(t+\tau_3, a+\omega)_{t_3}\ , \quad y^{(\mu)} = D^\mu \psi(t+\tau_3, a+\omega)_{t_3}$$

in der Bedeutung (119) und für beliebige $\overline{x}^{(n)}$, $\overline{y}^{(n)}$, so dass nach VII dem Stück 1 2 von *a ein wirkliches Minimum des Integrales in einer „Nachbarschaft $n-1$ ter Ordnung“ entspricht.*

91 | Ist aber statt dessen nur beständig

$$F_1\left(x^{(\mu)}, y^{(\mu)}\right) > \varrho > 0 \tag{121}$$
$$\left(x^{(\mu)} = \varphi^{(\mu)}(t),\ y^{(\mu)} = \psi^{(\mu)}(t),\ t_1 \leqq t \leqq t_2\right)\ ,$$

F_1 in der Bedeutung (19) genommen, so können durch Hinzufügung der neuen Bedingungen

$$\begin{aligned} \left|D^n\overline{\varphi}(\lambda) - \varphi^{(n)}(t_3)\right| &< g_n \\ \left|D^n\overline{\psi}(\lambda) - \psi^{(n)}(t_3)\right| &< g_n \end{aligned} \tag{118)$_n$}$$

auch die Grössen

$$\left|\overline{x}^{(n)} - x^{(n)}\right| = \left|D^n\overline{\varphi}(\lambda) - D^n\varphi(t+\tau_3, a+\omega)_{t_3}\right| < g'_n\ ,$$
$$\left|\overline{y}^{(n)} - y^{(n)}\right| = \left|D^n\overline{\psi}(\lambda) - D^n\psi(t+\tau_3, a+\omega)_{t_3}\right| < g'_n$$

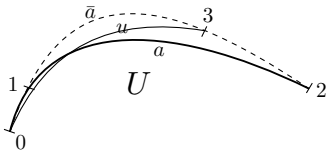


Fig. 6.

If, furthermore, the curves $\overline{a}$ and a have contact of $n-1$ th order at the endpoints 1 and 2, then, on account of that very uniqueness, the curves u belonging to these endpoints must coincide with a itself so that it is now possible to apply Theorem VII.

For if, at the same time, we always have, for some constant positive ϱ,

$$E_1\left(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(n)}, \overline{y}^{(n)}\right) > \varrho > 0 \tag{120}$$
$$\left(x^{(\mu)} = \varphi^{(\mu)}(t), y^{(\mu)} = \psi^{(\mu)}(t), t_1 \leqq t \leqq t_2\right)$$

and for arbitrary values of the variables $\overline{x}^{(n)}$, $\overline{y}^{(n)}$, where E_1 has the sense of (78), then, in (118) or in (37), we may choose the constants g_μ so small that always also

$$E_1(\lambda) = E_1\left(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(n)}, \overline{y}^{(n)}\right) > 0$$

and hence also $$E(\lambda) = E\left(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(n)}, \overline{y}^{(n)}\right) \geqq 0$$

for

$$x^{(\mu)} = D^\mu \varphi(t+\tau_3, a+\omega)_{t_3}\ , \quad y^{(\mu)} = D^\mu \psi(t+\tau_3, a+\omega)_{t_3}$$

in the sense of (119) and for arbitrary $\overline{x}^{(n)}$, $\overline{y}^{(n)}$, so that, by VII, there corresponds to the segment 1 2 of a *a real minimum of the integral in a "vicinity of* $n-1$ *th order*".

But if, instead, always only

$$F_1\left(x^{(\mu)}, y^{(\mu)}\right) > \varrho > 0 \tag{121}$$
$$\left(x^{(\mu)} = \varphi^{(\mu)}(t),\ y^{(\mu)} = \psi^{(\mu)}(t),\ t_1 \leqq t \leqq t_2\right)\ ,$$

where F_1 is taken in the sense of (19), then, by adding the new conditions

$$\begin{aligned} \left|D^n\overline{\varphi}(\lambda) - \varphi^{(n)}(t_3)\right| &< g_n \\ \left|D^n\overline{\psi}(\lambda) - \psi^{(n)}(t_3)\right| &< g_n\ , \end{aligned} \tag{118}_n$$

we can also make the quantities

$$\left|\overline{x}^{(n)} - x^{(n)}\right| = \left|D^n\overline{\varphi}(\lambda) - D^n\varphi(t+\tau_3, a+\omega)_{t_3}\right| < g_n'\ ,$$
$$\left|\overline{y}^{(n)} - y^{(n)}\right| = \left|D^n\overline{\psi}(\lambda) - D^n\psi(t+\tau_3, a+\omega)_{t_3}\right| < g_n'$$

so klein gemacht werden, dass nach (78) wieder

$$E_1\left(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(n)}, \overline{y}^{(n)}\right) > 0 \qquad \text{wird für}$$
$$x^{(\mu)} = D^\mu \varphi(t+\tau_3, a+\omega)_{t_3}\,, \quad y^{(\mu)} = D^\mu \psi(t+\tau_3, a+\omega)_{t_3}$$
$$(\mu = 0, 1, \ldots n) \quad \text{und} \quad \overline{x}^{(n)} = D^n \overline{\varphi}(\lambda)\,,\ \overline{y}^{(n)} = D^n \overline{\psi}(\lambda)\,,$$

dass also auch $E(\lambda) \geqq 0$ wird und somit *a wieder ein Minimum* des Integrales liefert, jetzt aber *in einer „Nachbarschaft n ter Ordnung“*.

Die aus (120) und (121) gezogenen Schlussfolgerungen bleiben gültig, wenn hier die positive Grösse ϱ durch 0 ersetzt wird, d. h. auch wenn E_1 oder F_1 in einzelnen Punkten von *a verschwindet*, ohne jedoch, als Function von $x^{(\mu)}$, $y^{(\mu)}$ betrachtet, an einer dieser Stellen sein Vorzeichen zu *wechseln*, also auch für alle Werte-Combinationen der Umgebung immer positiv bleibt.

Dagegen ist noch zu untersuchen, ob nicht der andere Factor $(\overline{k}-k)^2$ von E in (78) und damit auch E selbst auf $\overline{a}$ *beständig verschwinden* könne; dann würde nämlich $J(\overline{a}) = J(a)$ werden, weshalb auch dieser Fall im Satze VII ausdrücklich ausgeschlossen wurde. Es wird sich aber zeigen, dass er unter den bereits gemachten Voraussetzungen *nur für a selbst* eintritt.

Ist nämlich in (78)

$$\overline{k} - k = y'\left(\overline{x}^{(n)} - x^{(n)}\right) - x'\left(\overline{y}^{(n)} - y^{(n)}\right) = 0\,, \tag{122}$$

92 | so ist wegen $x'^2 + y'^2 > 0$

$$\frac{\overline{x}^{(n)} - x^{(n)}}{x'} = \frac{\overline{y}^{(n)} - y^{(n)}}{y'} = h$$

eine endliche Grösse, und daher ausser (119) noch:

$$x^{(n)} = D^n \varphi(t+\tau_3, a+\omega)_{t_3} = \overline{x}^{(n)} - hx' = D^n \overline{\varphi}(\lambda)\,,$$
$$y^{(n)} = D^n \psi(t+\tau_3, a+\omega)_{t_3} = \overline{y}^{(n)} - hy' = D^n \overline{\psi}(\lambda)\,,$$

wenn gemäss den Bemerkungen von S. 56 die in (119) gar nicht vorkommende und in Bezug auf den Wert von E willkürliche Grösse $\lambda^{(n)}$ durch $\lambda^{(n)} - h\lambda'$ ersetzt wird. Das heisst aber: Die beiden Curven $\overline{a}$ und u gehen im Punkte 3 eine Berührung *von n ter statt nur von n − 1ter Ordnung* mit einander ein.

Dann kann die Curve $\overline{a}$ in der Umgebung des Berührungspunktes 3 $(t = t'_3)$ in der Form ausgedrückt werden:

$$x = \overline{\varphi}_1(t)\,, \quad y = \overline{\psi}_1(t)\,, \qquad \text{so dass}$$
$$\overline{\varphi}_1^{(\mu)}(t'_3) = \varphi^{(\mu)}(t'_3, u)\,, \quad \overline{\psi}_1^{(\mu)}(t'_3) = \psi^{(\mu)}(t'_3, u) \tag{123}$$
$$(\mu = 0, 1, \ldots n)\,.$$

Da aber nach (117) auf $\overline{a}$ auch immer $|\Theta(t'_0, t'_3; u)| > 0$ angenommen werden kann, wo der Wert t'_0 zum Punkte 0 gehört, so sind auch die Formeln (107)

so small that, by (78), again

$$E_1\left(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(n)}, \overline{y}^{(n)}\right) > 0 \qquad \text{for}$$
$$x^{(\mu)} = D^\mu \varphi(t+\tau_3, a+\omega)_{t_3}\,, \quad y^{(\mu)} = D^\mu \psi(t+\tau_3, a+\omega)_{t_3}$$
$$(\mu = 0, 1, \ldots n) \quad \text{and} \quad \overline{x}^{(n)} = D^n \overline{\varphi}(\lambda)\,,\ \overline{y}^{(n)} = D^n \overline{\psi}(\lambda)\,,$$

and thus also $E(\lambda) \geqq 0$, and hence that a furnishes *again a minimum* of the integral, but now *in a "vicinity of n th order"*.

The consequences deduced from (120) and (121) retain their validity if the positive quantity ϱ is replaced here by 0, i. e., even if E_1, or F_1, *vanishes* in individual points of a, without, however, when considered as a function of $x^{(\mu)}$, $y^{(\mu)}$, *switching* its sign at one of these positions, and hence always remains positive for all combinations of values in the vicinity.

By contrast, we still need to investigate whether it is not possible that the other factor $(\overline{k}-k)^2$ of E in (78), and hence also E itself on $\overline{a}$, *always vanishes*; for, in this case, we would have $J(\overline{a}) = J(a)$, which is the very reason why this case was expressly excluded in Theorem VII. But, as we shall see, this case *only* arises *for a itself* under the assumptions already made.

For if in (78)

$$\overline{k} - k = y'\left(\overline{x}^{(n)} - x^{(n)}\right) - x'\left(\overline{y}^{(n)} - y^{(n)}\right) = 0\,, \tag{122}$$

then, on account of $x'^2 + y'^2 > 0$,

$$\frac{\overline{x}^{(n)} - x^{(n)}}{x'} = \frac{\overline{y}^{(n)} - y^{(n)}}{y'} = h$$

is a finite quantity, and hence, besides (119), also

$$x^{(n)} = D^n \varphi(t+\tau_3, a+\omega)_{t_3} = \overline{x}^{(n)} - hx' = D^n \overline{\varphi}(\lambda)\,,$$
$$y^{(n)} = D^n \psi(t+\tau_3, a+\omega)_{t_3} = \overline{y}^{(n)} - hy' = D^n \overline{\psi}(\lambda)\,,$$

if, according to remarks on p. 56, the quantity $\lambda^{(n)}$, which occurs nowhere in (119) and is arbitrary with respect to the value of E, is replaced by $\lambda^{(n)} - h\lambda'$. But this means that the contact that the two curve $\overline{a}$ and u have with one another at point 3 is *of n th, instead of only* $n-1$ *th, order*.

The curve $\overline{a}$ can then be expressed as follows in the vicinity of the point of contact 3 $(t = t'_3)$:

$$x = \overline{\varphi}_1(t)\,, \quad y = \overline{\psi}_1(t)\,, \qquad \text{so that}$$
$$\overline{\varphi}_1^{(\mu)}(t'_3) = \varphi^{(\mu)}(t'_3, u)\,, \quad \overline{\psi}_1^{(\mu)}(t'_3) = \psi^{(\mu)}(t'_3, u) \tag{123}$$
$$(\mu = 0, 1, \ldots n)\,.$$

But since, by (117), we may always also assume $|\Theta(t'_0, t'_3; u)| > 0$ on $\overline{a}$, where the value t'_0 belongs to point 0, the formulas (107) of the "continuous

der „stetigen Fortsetzung“ der Schar U hier anwendbar, wenn jetzt die constanten Grössen a_ν, t_0, t_3 durch die variablen u_ν, t'_0, t'_3 ersetzt werden, so dass mit Rücksicht auf (123) für geeignete Werte der Grössen $\tau_0^{(\mu)}$, $\tau_3^{(\mu)}$, ω_ν:

$$\begin{aligned} D^\mu\big[\varphi(t+\tau_0,u+\omega)-\varphi(t)\big]_{t'_0} &= 0 \\ D^\mu\big[\psi(t+\tau_0,u+\omega)-\psi(t)\big]_{t'_0} &= 0 \\ D^\mu\big[\varphi(t+\tau_3,u+\omega)-\varphi(t)\big]_{t'_3} &= \overline{\varphi}_1^{(\mu)}(t'_3+\varepsilon)-\overline{\varphi}_1^{(\mu)}(t'_3) \\ D^\mu\big[\psi(t+\tau_3,u+\omega)-\psi(t)\big]_{t'_3} &= \overline{\psi}_1^{(\mu)}(t'_3+\varepsilon)-\overline{\psi}_1^{(\mu)}(t'_3) \\ (\mu=0,1,\dots n-1)\ , & \end{aligned} \tag{124}$$

oder, wenn man für kleine $|\varepsilon|$ setzt:

93 | $$\omega_\nu=\varepsilon\overline{\omega}_\nu+(\varepsilon)_2\ ,\quad \tau_0^{(\mu)}=\varepsilon\overline{\tau}_0^{(\mu)}+(\varepsilon)_2\ ,\quad \tau_3^{(\mu)}=\varepsilon\overline{\tau}_3^{(\mu)}+(\varepsilon)_2\ ,$$

gemäss (107a) beiderseits nach Potenzen von ε entwickelt und die Coefficienten von ε^1 einander gleich setzt:

$$\begin{aligned} D^\mu\Big[\varphi'(t,u)\overline{\tau}_0+\sum_\nu\varphi_\nu(t,u)\overline{\omega}_\nu\Big]_{t'_0} &= 0 \\ D^\mu\Big[\psi'(t,u)\overline{\tau}_0+\sum_\nu\psi_\nu(t,u)\overline{\omega}_\nu\Big]_{t'_0} &= 0 \\ D^\mu\Big[\varphi'(t,u)\overline{\tau}_3+\sum_\nu\varphi_\nu(t,u)\overline{\omega}_\nu\Big]_{t'_3} &= \overline{\varphi}_1^{(\mu+1)}(t'_3) \\ D^\mu\Big[\psi'(t,u)\overline{\tau}_3+\sum_\nu\psi_\nu(t,u)\overline{\omega}_\nu\Big]_{t'_3} &= \overline{\psi}_1^{(\mu+1)}(t'_3) \\ (\mu=0,1,\dots n-1)\ , & \end{aligned} \tag{124a}$$

wo wegen (123):

$$\begin{aligned} \overline{\varphi}_1^{(\mu+1)}(t'_3)=\varphi^{(\mu+1)}(t'_3,u)=D^\mu\varphi'(t,u)_{t'_3}\ , \\ \overline{\psi}_1^{(\mu+1)}(t'_3)=\psi^{(\mu+1)}(t'_3,u)=D^\mu\psi'(t,u)_{t'_3}\ , \end{aligned}$$

die beiden letzten Zeilen in der Form geschrieben werden können:

$$\begin{aligned} D^\mu\Big[\varphi'(t,u)(\overline{\tau}_3-1)+\sum_\nu\varphi_\nu(t,u)\overline{\omega}_\nu\Big]_{t'_3} &= 0 \\ D^\mu\Big[\psi'(t,u)(\overline{\tau}_3-1)+\sum_\nu\psi_\nu(t,u)\overline{\omega}_\nu\Big]_{t'_3} &= 0 \end{aligned}$$

Dieses System (124a) von $4n$ linearen und homogenen Gleichungen mit den $4n$ Unbekannten: $\overline{\omega}_\nu$, $\overline{\tau}_0^{(\mu)}$, $\overline{\tau}_3^{(\mu)}$ $-\,e_{\mu,0}$ kann aber nur bestehen, wenn

continuation" of the family U apply here, if the constant quantities a_ν, t_0, t_3 are replaced by the variable ones u_ν, t'_0, t'_3, so that, considering (123), for suitable values of the quantities $\tau_0^{(\mu)}$, $\tau_3^{(\mu)}$, ω_ν,

$$\begin{aligned}
D^\mu\big[\varphi(t+\tau_0,u+\omega)-\varphi(t)\big]_{t'_0} &= 0\\
D^\mu\big[\psi(t+\tau_0,u+\omega)-\psi(t)\big]_{t'_0} &= 0\\
D^\mu\big[\varphi(t+\tau_3,u+\omega)-\varphi(t)\big]_{t'_3} &= \overline{\varphi}_1^{(\mu)}(t'_3+\varepsilon)-\overline{\varphi}_1^{(\mu)}(t'_3)\\
D^\mu\big[\psi(t+\tau_3,u+\omega)-\psi(t)\big]_{t'_3} &= \overline{\psi}_1^{(\mu)}(t'_3+\varepsilon)-\overline{\psi}_1^{(\mu)}(t'_3)\\
(\mu=0,1,\dots n-1)\ ,&
\end{aligned} \tag{124}$$

or, if, for small $|\varepsilon|$, we put

$$\omega_\nu=\varepsilon\overline{\omega}_\nu+(\varepsilon)_2\ ,\quad \tau_0^{(\mu)}=\varepsilon\overline{\tau}_0^{(\mu)}+(\varepsilon)_2\ ,\quad \tau_3^{(\mu)}=\varepsilon\overline{\tau}_3^{(\mu)}+(\varepsilon)_2\ ,$$

expand, according to (107a), in powers of ε on both sides and equate the coefficients of ε^1 with one another:

$$\begin{aligned}
D^\mu\Big[\varphi'(t,u)\overline{\tau}_0+\sum_\nu\varphi_\nu(t,u)\overline{\omega}_\nu\Big]_{t'_0} &= 0\\
D^\mu\Big[\psi'(t,u)\overline{\tau}_0+\sum_\nu\psi_\nu(t,u)\overline{\omega}_\nu\Big]_{t'_0} &= 0\\
D^\mu\Big[\varphi'(t,u)\overline{\tau}_3+\sum_\nu\varphi_\nu(t,u)\overline{\omega}_\nu\Big]_{t'_3} &= \overline{\varphi}_1^{(\mu+1)}(t'_3)\\
D^\mu\Big[\psi'(t,u)\overline{\tau}_3+\sum_\nu\psi_\nu(t,u)\overline{\omega}_\nu\Big]_{t'_3} &= \overline{\psi}_1^{(\mu+1)}(t'_3)\\
(\mu=0,1,\dots n-1)\ ,&
\end{aligned} \tag{124a}$$

where, on account of (123):

$$\begin{aligned}
\overline{\varphi}_1^{(\mu+1)}(t'_3)&=\varphi^{(\mu+1)}(t'_3,u)=D^\mu\varphi'(t,u)_{t'_3}\ ,\\
\overline{\psi}_1^{(\mu+1)}(t'_3)&=\psi^{(\mu+1)}(t'_3,u)=D^\mu\psi'(t,u)_{t'_3}\ ,
\end{aligned}$$

the last two lines can be written in the following form:

$$\begin{aligned}
D^\mu\Big[\varphi'(t,u)(\overline{\tau}_3-1)+\sum_\nu\varphi_\nu(t,u)\overline{\omega}_\nu\Big]_{t'_3} &= 0\\
D^\mu\Big[\psi'(t,u)(\overline{\tau}_3-1)+\sum_\nu\psi_\nu(t,u)\overline{\omega}_\nu\Big]_{t'_3} &= 0\ .
\end{aligned}$$

This system (124a) of $4n$ linear and homogeneous equations in the $4n$ unknowns: $\overline{\omega}_\nu$, $\overline{\tau}_0^{(\mu)}$, $\overline{\tau}_3^{(\mu)}-e_{\mu,0}$ can only obtain, if either the determinant

entweder die Determinante $\Theta(t_0', t_3'; u) = 0$ ist, oder die Unbekannten sämtlich verschwinden. Da hier das erste ausgeschlossen ist, so muss

$$\overline{\omega}_\nu = \lim_{\varepsilon=0} \frac{\omega_\nu}{\varepsilon} = \left(\frac{du_\nu}{dt}\right)_{t_3'} = \lambda' \left(\frac{du_\nu}{d\lambda}\right)_3 = 0$$

sein für $\nu = 1, 2, \ldots 2n$.

Sollte dies aber für ein endliches Stück von $\overline{a}$ überall stattfinden, so müsste $u_\nu = \text{const.}$ sein, d. h. dieser Teil von $\overline{a}$ fiele wegen (123) mit einer particulären Lösung u der Differentialgleichung, *mit einem Individuum der Schar U, vollständig zusammen*, also mit a selbst, wenn $\overline{a}$ diese Eigenschaft (122) in seiner ganzen Ausdehnung zwischen 1 und 2 besitzen soll. Für *andere* Curven $\overline{a}$ unserer „Nachbarschaft" ausser a ist also diese Annahme mit unseren Voraussetzungen *unverträglich*, q. e. d.

94 | Wir sind daher zur Aufstellung des folgenden Satzes berechtigt:

Satz VIII. *Eine particuläre Lösung a der Differentialgleichung des Problems:* $x = \varphi(t) = \varphi(t, a)$, $y = \psi(t) = \psi(t, a)$, *für welche zwischen den Punkten* 1 *und* 2, *d. h. im Intervall* $t_1 \leqq t \leqq t_2$ *die Functionen* $\varphi^{(\mu)}(t)$, $\psi^{(\mu)}(t)$, *sowie auch* $\varphi_\nu^{(\mu)}(t)$, $\psi_\nu^{(\mu)}(t)$ $(\mu = 0, 1, \ldots n;\ \nu = 1, 2, \ldots 2n)$ *eindeutig, endlich und stetig sind,* $\varphi'(t)$ *und* $\psi'(t)$ *niemals gleichzeitig verschwinden und endlich die Function* $\Theta(t_1, t; a)$ [*cf.* (109) *und* (112)] *von* 0 *verschieden ist, liefert nach der Definition des zweiten Abschnittes und für* $r = n - 1$ *in* (34) ein wirkliches Minimum des Integrales

$$J = \int_{t_1}^{t_2} F\left(x^{(\mu)}, y^{(\mu)}\right) dt$$

1. in einer Nachbarschaft $m = n - 1$ ter Ordnung [*cf.* (37)], *wenn gleichzeitig die Function* $E(t) = E\left(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(n)}, \overline{y}^{(n)}\right)$ *in der Bedeutung* (73) *und* (78) *für* $x^{(\mu)} = \varphi^{(\mu)}(t)$, $y^{(\mu)} = \psi^{(\mu)}(t)$ *und für beliebige Werte der* $\overline{x}^{(n)}$, $\overline{y}^{(n)}$ *im ganzen Intervall* nur positive Werte besitzt *und auch an den Stellen, wo sie etwa verschwindet, ihr Vorzeichen nicht wechseln kann;*

2. in einer Nachbarschaft $m = n$ ter Ordnung, *wenn statt dessen die Function* $F_1\left(x^{(\mu)}, y^{(\mu)}\right)$ *für* $x^{(\mu)} = \varphi^{(\mu)}(t)$, $y^{(\mu)} = \psi^{(\mu)}(t)$ *im ganzen Intervall* nur positiv ist *und, auch wo sie verschwindet, ihr Zeichen nicht wechseln kann.*

In den Sätzen V und VI wurde die Erfüllung der Bedingungen (77) $E\left(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(\mu)}, \overline{y}^{(\mu)}\right) \geqq 0$ und (80) $F_1\left(x^{(\mu)}, y^{(\mu)}\right) \geqq 0$ in der ganzen Ausdehnung von a als *notwendig* für das Bestehen eines Minimums bei $r = m = n - 1$, von (80) auch bei $m = n$, erwiesen. Wir sind nunmehr im stande, die Notwendigkeit der letzten Bedingung (80) auch unter der allgemeineren Annahme $r \geqq n - 1$, $m \geqq n$ darzuthun, während $r < n - 1$ durch den Satz IV ausgeschlossen ist. Wäre nämlich auf einer Curve a, welche nicht zu den singulären im Sinne von S. 85 gehört, an irgend einer Stelle t des

$\Theta(t_0', t_3'; u) = 0$ or the unknowns all vanish. But since the first case is excluded here, we must have

$$\overline{\omega}_\nu = \lim_{\varepsilon=0} \frac{\omega_\nu}{\varepsilon} = \left(\frac{du_\nu}{dt}\right)_{t_3'} = \lambda' \left(\frac{du_\nu}{d\lambda}\right)_3 = 0$$

for $\nu = 1, 2, \ldots 2n$.

But if this should be the case everywhere for a finite segment of $\overline{a}$, then we would have to have $u_\nu = \text{const.}$, i. e., this part of $\overline{a}$ would *completely coincide* with a particular solution u of the differential equation, *with an individual of the family* U, on account of (123), and hence with a itself, provided that $\overline{a}$ is to possess this property (122) in its entire extension between 1 and 2. Hence, as far as *other* curves $\overline{a}$ of your "vicinity" except a are concerned, this assumption is *incompatible* with our presuppositions, q. e. d.

We are thus entitled to present

Theorem VIII. *A particular solution* a *of the differential equation of the problem:* $x = \varphi(t) = \varphi(t, a)$, $y = \psi(t) = \psi(t, a)$, *for which, between the points* 1 *and* 2, *i. e., on the interval* $t_1 \leqq t \leqq t_2$, *the functions* $\varphi^{(\mu)}(t)$, $\psi^{(\mu)}(t)$, *and also* $\varphi_\nu^{(\mu)}(t)$, $\psi_\nu^{(\mu)}(t)$ $(\mu = 0, 1, \ldots n;\ \nu = 1, 2, \ldots 2n)$, *are single-valued, finite and continuous,* $\varphi'(t)$ *and* $\psi'(t)$ *never vanish simultaneously and, finally, the function* $\Theta(t_1, t; a)$ [*cf.* (109) *and* (112)] *is everywhere different from* 0, *furnishes, according to the definition in the second section and for* $r = n - 1$ *in* (34), a real minimum of the integral

$$J = \int_{t_1}^{t_2} F\left(x^{(\mu)}, y^{(\mu)}\right) dt$$

1. in a vicinity of $m = n - 1$ th order [*cf.* (37)], *if, at the same time, the function* $E(t) = E\left(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(n)}, \overline{y}^{(n)}\right)$ *in the sense of* (73) *and* (78) *for* $x^{(\mu)} = \varphi^{(\mu)}(t)$, $y^{(\mu)} = \psi^{(\mu)}(t)$ *and for arbitrary values of the* $\overline{x}^{(n)}$, $\overline{y}^{(n)}$ takes only positive values *in the entire interval and never changes its sign, not even at positions where it vanishes;*

2. in a vicinity of $m = n$ th order, *if, instead, the function* $F_1\left(x^{(\mu)}, y^{(\mu)}\right)$ *is* only positive *for* $x^{(\mu)} = \varphi^{(\mu)}(t)$, $y^{(\mu)} = \psi^{(\mu)}(t)$ *in the entire interval and does not change its sign, not even where it vanishes.*

In Theorems V and VI, we proved that the satisfaction of the conditions (77) $E\left(x^{(\mu)}, y^{(\mu)}; \overline{x}^{(\mu)}, \overline{y}^{(\mu)}\right) \geqq 0$ and (80) $F_1\left(x^{(\mu)}, y^{(\mu)}\right) \geqq 0$ in the entire extension of a is *necessary* for the existence of a minimum when $r = m = n - 1$, and, with respect to (80) also when $m = n$. We are now able to show that, even under the more general assumption $r \geqq n - 1$, $m \geqq n$, the last condition (80) is necessary, while $r < n - 1$ is excluded by Theorem IV. For if we had $F_1\left(x^{(\mu)}, y^{(\mu)}\right) < 0$ at any position t of the interval $t_1 \leqq t \leqq t_2$ on a curve a that does not belong among the singular curves in the sense

Intervalls $t_1 \leqq t \leqq t_2$ $F_1\left(x^{(\mu)}, y^{(\mu)}\right) < 0$, so bestände diese Ungleichheit we-
95 gen der Stetigkeit von F_1 auch für ein ganzes | Intervall $t' \leqq t \leqq t''$. Hier könnte aber $t'' - t'$ so klein und eine neue Stelle $t_0 < t'$ so nahe an t' angenommen werden, dass nicht nur beständig $\pm\Theta(t', t) > 0$, sondern auch $\pm\Theta(t_0, t) > \zeta > 0$ wäre für das ganze Teil-Intervall $t' \leqq t \leqq t''$. Ferner könn-

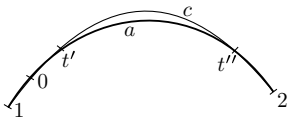


Fig. 7.

te man eine Curve c bestimmen, welche a in den Punkten t' und t'' von der Ordnung $r \geqq n-1$ berührt, bis zu den Ableitungen r ter Ordnung stetig ist und einer beliebig engen Nachbarschaft m ter Ordnung (37) von a angehört. Nun könnten in (37) die Grössen g_μ $(\mu = 0, 1, \ldots n)$ so klein gewählt werden, dass auch (vergl. die Bemerkungen bei (121)) $E(\lambda) < 0$ wäre in der ganzen Ausdehnung von c und daher nach (105a) $\overline{J} - J < 0$, wenn unter J das über a erstreckte Integral verstanden wird, unter $\overline{J}$ aber dasselbe Integral, in welchem nur der zwischen t' und t'' liegende Teil von a durch c ersetzt ist. Somit würde a die Eigenschaft des Minimums nicht besitzen, da die betrachtete Variation jedenfalls zu den „erlaubten" zu zählen ist.

In Bezug auf die „*conjugierten Punkte*", deren Vorhandensein für VIII ausgeschlossen wurde, ist es mir allerdings noch nicht gelungen, mittelst der hier überall angewandten „Methode der benachbarten Lösungen" zu sicheren Kriterien zu gelangen, wann sie mit dem Bestehen eines Minimums vereinbar sind, und ich muss hierüber auf die in der Einleitung erwähnte Arbeit von *Scheeffer* (Math. Ann. XXV) verweisen; doch mögen hier noch die folgenden Bemerkungen Platz finden.

In der Nähe eines zu 0 auf a conjugierten Punktes 3, für welchen $\Theta(t_0, t_3; a) = 0$ ist, existiert unter gewissen, sehr allgemeinen Bedingungen eine Curve $\overline{a} = g\left(x = \overline{\varphi}(\lambda),\, y = \overline{\psi}(\lambda)\right)$, welche *alle Curven* u, darunter a, einer Schar U von der definierten Eigenschaft der Reihe nach von *n ter*, *statt* nur von $n-1$ ter *Ordnung berührt*, ohne selbst mit einer Curve u zusammenzufallen, und die als „*Enveloppe*" der Schar U bezeichnet werden möge
96 in einem von dem gewöhnlichen etwas abweichenden Sinne | des Wortes. Auf dieser „Enveloppe" muss einerseits, unseren Betrachtungen von (124a) zufolge, beständig $\Theta(t_0', t_3'; u) = 0$ sein, d. h. sie enthält die zu 0 conjugierten Punkte $3'$ aller u, andererseits aber ist nach (122) hier beständig $\overline{k} - k = 0$ und daher $E(\lambda) = 0$.

Nimmt man also $t_1 \leqq t_0 < t_3 \leqq t_2$ an und betrachtet die erlaubte Variation 0 4 3 von 0 3 oder 1 0 4 3 2 von 1 2, wo 4 einen Punkt von g und 0 4

specified on p. 85, then, on account of the continuity of F_1, this inequality would also obtain for an entire interval $t' \leqq t \leqq t''$. But, in this case, it would be possible to choose $t'' - t'$ so small and a new position $t_0 < t'$ so close to t' that we would not only always have $\pm\Theta(t', t) > 0$ but also $\pm\Theta(t_0, t) > \zeta > 0$ for the entire partial interval $t' \leqq t \leqq t''$. Furthermore, it would be possible

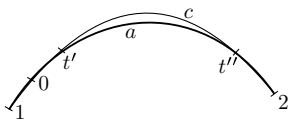


Fig. 7.

to determine a curve c that has contact of the order $r \geqq n-1$ with a at the points t' and t'', is continuous up to the derivatives of r th order and belongs to an arbitrarily narrow neighborhood of m th order (37) of a. It would now be possible to choose the quantities g_μ $(\mu = 0, 1, \ldots n)$ in (37) so small that we would also have $E(\lambda) < 0$ (cf. the remarks at (121)) in the entire extension of c, and hence, by (105a), $\overline{J} - J < 0$, assuming that J is the integral taken along a, while $\overline{J}$ is the same integral in which only the part a lying between t' and t'' is replaced by c. Thus, a would lack the property of the minimum, since there is no doubt that the variation under consideration is to be considered "admissible".

But in regard to the "*conjugate points*" whose existence was excluded for VIII, I have failed thus far to arrive at reliable criteria by means of the "method of neighboring solutions", which is used everywhere here, to determine when they are compatible with the existence of a minimum, and, on this matter, I have to refer to the paper by *Scheeffer (1885)*, which I mentioned in the introduction; nevertheless, let me add the following observations here.

Nearby a point 3 conjugate to 0 on a for which $\Theta(t_0, t_3; a) = 0$ there exists, under certain, very general conditions, a curve $\overline{a} = g\left(x = \overline{\varphi}(\lambda),\ y = \overline{\psi}(\lambda)\right)$ that *has contact of the n th*, *instead* of only $n-1$ th, *order* with *all curves* u, including a, of a family U of the defined property successively, without itself coinciding with a curve u, which shall be called the "*envelope*" of the family U, in a slight departure from customary usage of that word. According to our considerations concerning (124a), we must, on the one hand, always have $\Theta(t_0', t_3'; u) = 0$ on this "envelope", i. e., it contains the points $3'$ conjugate to 0 of all u, but, on the other hand, by (122), always $\overline{k} - k = 0$ here, and hence $E(\lambda) = 0$.

Thus, if we suppose that $t_1 \leqq t_0 < t_3 \leqq t_2$ and consider the admissible variation 0 4 3 of 0 3, or 1 0 4 3 2 of 1 2, where 4 denotes a point of g and 0 4

die zu 4 gehörige Curve u bedeutet, so ist nach (105)

$$J_{043} - J_{03} = \int_{\lambda_4}^{\lambda_3} E(\lambda)\, d\lambda = 0$$

und daher $$J_{10432} = J_{12}\ ,$$

während 4 immer so nahe an 3 angenommen werden kann, dass 0 4 3 einer beliebig engen Nachbarschaft $n-1$ ter, ja, weil die Berührung in 3 von n ter Ordnung ist, auch von n ter Ordnung angehört. Da aber g der Differential-

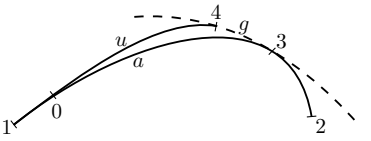


Fig. 8.

gleichung des Problems (von der Ordnung $2n$) im allgemeinen *nicht* genügt, so kann immer J_{43} und damit auch $J_{10432} = J_{12}$ nach II durch beliebig kleine Variationen noch *verkleinert* werden, so dass für a ein Minimum sicher *nicht* besteht.

Dieses Ergebnis ist im wesentlichen die Verallgemeinerung einer zuerst von Herrn *Lindelöf* (Moigno et Lindelöf, Leçons de Calcul Différentiel et Intégral; IV Calcul des Variations) entdeckten Eigenschaft der *Kettenlinie* in Bezug auf den Inhalt der Rotationsfläche $\int y\, ds$, bei welcher zwei auf der Rotationsachse sich schneidende Tangenten die Rolle der „Enveloppe“ spielen. Dagegen fehlt mir noch ein einfaches *Kriterium* für die Existenz einer solchen allgemeineren „Enveloppe“ von den vorausgesetzten Eigenschaften.

97 | Die in der vorliegenden Arbeit abgeleiteten Sätze, namentlich II, V, VI und VIII, gestatten jetzt eine *vollständige Entscheidung* über das Vorhandensein eines Minimums in dem betrachteten Sinne und damit die Lösung der gestellten Aufgabe mit alleiniger Ausnahme der folgenden Fälle:

1. wenn $m < n-1$ sein soll, d. h. wenn ein Minimum in einer „Nachbarschaft“ von niedrigerer als der $n-1$ ten Ordnung verlangt wird,

2. wenn $m = n-1$ und gleichzeitig $r > n-1$ sein soll (cf. V),

3. wenn auf a die Functionen P, Q in ganzen Intervallen unstetig sind (cf. II),

4. wenn a gemäss IV aus einer Anzahl verschiedener particulärer Lösungen der Differentialgleichung zusammengesetzt ist,

5. wenn a nicht sämmtlichen in VIII vorausgesetzten Stetigkeitsbedingungen, sondern nur den durch (32) bis (35) ausgedrückten genügt, d. h. wenn die $\varphi_\nu^{(\mu)}(t)$ nicht alle stetig sind oder wenn a sich zwischen 1 und 2 irgendwo selbst durchschneidet,

the curve u belonging to 4, then, by (105),

$$J_{043} - J_{03} = \int_{\lambda_4}^{\lambda_3} E(\lambda)\, d\lambda = 0$$

and hence $$J_{10432} = J_{12}\ ,$$

while we can always take 4 so close to 3 that 0 4 3 belongs to a arbitrarily narrow neighborhood of the $n-1$ th, and even n th, order since the contact

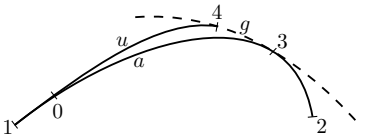


Fig. 8.

at 3 is of n th order. But since g does *not*, in general, satisfy the differential equation of the problem (of order $2n$), J_{43}, and hence also $J_{10432} = J_{12}$ can, by II, always be *scaled down* further by means of arbitrarily small variations so that there is certainly *no* minimum for a.

This result is essentially a generalization of the property, which was first discovered by *Lindelöf* (*Lindelöf and Moigno 1861*, IV: Calcul des Variations), of the *catenary* with respect to the volume of the surface of revolution $\int y\, ds$, where two tangents intersecting one another on the axis of rotation play the part of the "envelope". By contrast, I still have no simple criterion for the *existence* of such a more general "envelope" of the properties assumed.

The theorems derived in the present work, and in particular II, V, VI and VIII, now provide for a *complete decision* on the existence of a minimum in the sense under consideration, and hence the solution of the task before us, except only in the following cases:

1. when we are supposed to have $m < n-1$, i. e., when a minimum is required in a "neighborhood" of an order lower than $n-1$,

2. when we are supposed to simultaneously have $m = n-1$ and $r > n-1$ (cf. V),

3. when the functions P, Q are discontinuous on a on the entire interval (cf. II),

4. when a is composed of several different particular solutions of the differential equation in accordance with IV,

5. when a does not satisfy all continuity conditions assumed in VIII, but only those expressed by (32) to (35), i. e., when the $\varphi_\nu^{(\mu)}(t)$ are not all continuous or when a intersects itself somewhere between 1 and 2,

6. wenn a schon im betrachteten Intervall den zum Anfangspunkt „conjugierten“ Punkt t_1' enthält: $\Theta(t_1, t_1'; a) = 0$, ohne dass dort eine stetige „Enveloppe“ g existiert,

7. wenn F_1 oder E_1 an einzelnen Stellen von a verschwindet, in deren Umgebung die Function negativ werden kann (cf. V, VI und VIII), sonst aber auf a immer positiv ist.

Nur in einem dieser Ausnahmefälle würde die Entscheidung unter Umständen erst von einer Special-Untersuchung abhängig zu machen sein.

Zum Schlusse ist es mir Bedürfnis, Herrn Prof. H. A. *Schwarz*, dem ich die erste Anregung zu meiner Arbeit und vielfache Unterstützung durch warmes Interesse und gütige Ratschläge verdanke, meinen ergebensten Dank an dieser Stelle auszusprechen.

98 |

Thesen.

I.

In der Variations-Rechnung ist auf eine genaue Definition des Maximums oder Minimums grösserer Wert als bisher zu legen.

II.

Mit Unrecht wird der Physik die Aufgabe gestellt, alle Naturerscheinungen auf Mechanik der Atome zurückzuführen.

III.

Die Messung ist aufzufassen als das überall anwendbare Hilfsmittel, stetig veränderliche Qualitäten zu unterscheiden und zu vergleichen.

99 |

Vita.

Natus sum *Ernestus Zermelo* Berolini a. d. VI Kal. Aug. anni MDCCCLXXI patre Theodoro, matre Augusta e gente Zieger, quos morte praematura mihi ereptos valde lugeo. Fidei addictus sum evangelicae. Primis litterarum elementis imbutus postquam per novem annos gymnasium, quod dicitur „Luisenstädtisches“, Berolini frequentavi, maturitatis ibi testimonium vere anni h. s. LXXXIX adeptus sum. Tum in philosophorum ordinem receptus per quinque annos et sex menses studiis praecipue mathematicis, physicis, philosophicis operam dedi, primum Berolini, deinde sex menses Halis Saxonum,

6. when a contains the point t_1': $\Theta(t_1, t_1'; a) = 0$, that is "conjugate" to the starting point, already in the interval under consideration without there being a continuous "envelope" g,

7. when F_1 or E_1 vanishes at particular positions of a in whose vicinity the function may become negative (cf. V, VI and VIII), while otherwise it is always positive on a.

It is only in these exceptional cases that the decision may depend on the results of a separate investigation.

In closing, I would like to express my deepest gratitude to Prof. *H. A. Schwarz*, who got me started in this project in the first place and supported my work in many ways, for his warm regard and gracious counsel.

Theses.

I.

It is necessary to place greater emphasis than before on a precise definition of the maximum or minimum in the calculus of variations.

II.

It is wrong to charge physics with the task of reducing all phenomena of nature to the mechanics of atoms.

III.

Measurement is to be conceived of as a universally applicable expedient for the distinction and comparison of continuously varying qualities.

Vita.

I was born *Ernst Zermelo* in Berlin on the 27th of July of the year 1871, of father Theodor and mother Auguste née Zieger, taken from me by premature deaths that caused me great sadness. I am of Protestant faith. After I had attended the Luisenstädtisches Gymnasium in Berlin for nine years and had become well-acquainted with the first elements of literature and science, I attained my certificate of graduation in the spring of the year 1889. I was then admitted to the philosophical faculty. For five years and six months I studied mathematics, physics, and philosophy in particular, first in Berlin, then for six months in Halle in Saxony, then for six months in Freiburg im

tum sex menses Friburgi Brisgaviae, postremo iterum Berolini. Audivi viros doctissimos:

Aron, Dilthey, Ebbinghaus, Förster, Frobenius, Fuchs, Glan, Hettner, Knoblauch, E. Kötter, Kundt (†), Lehmann-Filhés, Paulsen, Planck, de Richthofen, Schlesinger, E. Schmidt, Schwarz, Wien (Berolini), Cantor, Erdmann, Haym, Husserl, Ule, Wangerin (Halis), Lüroth, Münsterberg, Riehl, Stickelberger, Warburg, Weissenfels (Friburgi). Berolini seminarii mathematici, cui praesunt Fuchs, Schwarz, Frobenius, per duos annos et sex menses sodalis fui, exercitationibus, quas Planck de rebus mathematicis physicis instituere solet, per quinque semestria interfui.

Omnibus viris, qui me docuerunt, imprimis Lazaro Fuchs, Hermanno Amando Schwarz, Maximiliano Planck, summas gratias ago semperque habebo, nec non viro illustrissimo Aemilio Lampe, cuius consiliis et benevolentia in studiis meis magnopere adiutus sum. Societati quoque mathematicae, cuius, quamdiu Berolini aderam, sodalis fui, multum me debere confiteor.

Breisgau, and finally again in Berlin. I attended lectures given by the very learned men:

Aron, Dilthey, Ebbinghaus, Förster, Frobenius, Fuchs, Glan, Hettner, Knoblauch, E. Kötter, Kundt (†), Lehmann-Filhés, Paulsen, Planck, von Richthofen, Schlesinger, E. Schmidt, Schwarz, Wien (Berlin), Cantor, Erdmann, Haym, Husserl, Ule, Wangerin (Halle), Lüroth, Münsterberg, Riehl, Stickelberger, Warburg, Weissenfels (Freiburg). For two years and six months I was a member of the Berlin Mathematical Seminar run by Fuchs, Schwarz, and Frobenius. For five semesters I took part in the exercises of mathematical physics offered by Planck.

I am, and always will be, very grateful to all the men who taught me, in particular to Lazarus Fuchs, Hermann Amandus Schwarz, Max Planck, and, last not least, to the very excellent Emil Lampe whose advice and goodwill were very helpful and encouraging during my studies. I affirm that I also acknowledge a great debt to the Mathematical Society of which I was a member during my time in Berlin.

Introductory note to 1896a, 1896b, and Boltzmann 1896, 1897

Jos Uffink

In 1896, Zermelo published a paper containing what is commonly called the recurrence objection (*Umkehreinwand*) to Ludwig Boltzmann's approach to statistical physics. This paper immediately gave rise to a heated dispute with Boltzmann, a controversy comprising the four papers discussed here, two from each author.[1] Zermelo returned to the disputed topic in 1899, when he chose the foundations of statistical physics as the subject of his *Habilitation* lecture (*1900*) at the University of Göttingen.

Apart from this encounter with Boltzmann, Zermelo was also strongly interested in Josiah Willard Gibbs' approach to statistical physics. Indeed, he translated Gibbs' book, *Elementary Principles in Statistical Mechanics* (*1902*), into German (*1905*) and wrote a critical review (*1906*) of the book, which concludes his papers devoted to statistical physics.

This commentary provides some background, both historical and with respect to the current understanding, of Zermelo's involvement with the foundations of statistical physics. Section 1 provides some historical background to the Zermelo-Boltzmann controversy; Section 2 deals with the conceptual preliminaries to the controversy; and Section 3 is devoted to a discussion of the controversy itself.

1. The background to the Zermelo-Boltzmann controversy

1.1. Theoretical physics in the late 19th century: Boltzmann and Planck

When Zermelo wrote his *1896a,b*, he was working in Berlin as an assistant to Max Planck, at that time already a prominent theoretical physicist. The relations between Planck and Boltzmann, and their different positions in the problems facing theoretical physics in that era, seem crucial to the understanding of the ensuing dispute.

In the late nineteenth century, the foundations of physics was in a state of flux. Many new developments from thermodynamics, chemistry, electrodynamics, and various other areas, both theoretical and experimental, were

[1] Arguably a third paper by Boltzmann (*1897b*), entitled "On a mechanical theorem by Poincaré" and not included in this volume, also contains a response by Boltzmann to Zermelo.

H.-D. Ebbinghaus, A. Kanamori (Eds.), *Ernst Zermelo - Collected Works/ Gesammelte Werke II*, DOI 10.1007/978-3-540-70856-8_3, Schriften der Mathematisch-naturwissenschaftlichen Klasse der Heidelberger Akademie der Wissenschaften 23,

challenging traditional physical views, and one might well say that a lively competition was going on between different world views. Many physicists discussed the question of which world view would provide the best prospect for the unity and progress of physics.

Some authors, in particular Boltzmann, argued that the most promising way to move forward was to adopt the view that matter consists of molecules or atoms moving in accordance with the laws of classical mechanics. This view is sometimes called "mechanicism" (or "the mechanistic view of nature", etc.) or "atomism", depending on which of these two independent ingredients one wishes to emphasize, and had been successfully employed in the theory of gases by James Clerk Maxwell and Boltzmann in the 1860s and 1870s.

Other authors, like Planck and Pierre Duhem, believed that thermodynamics provided the most trustworthy framework for the future of theoretical physics. This theory refrained from asking whether matter is composed of atoms, moving in accordance with classical mechanics. Rather, it concentrated on describing empirical relations between various observable quantities, like temperature, pressure, volume, energy, and so forth. And whenever the issue arose about how such quantities would vary from one part of a material system to another, the usual approach was to model the system as a continuum rather than as being composed of atoms. Thermodynamics had emerged in the 1850s through the work of Rudolf Clausius and Lord Kelvin, but developed rapidly in the late 19th century to become a most versatile theory with a wide range of applications in chemistry, magnetism, optics and other areas far removed from the original study of heat engines.

Still others placed their trust in electrodynamics and proposed a world view in which the electromagnetic fields were considered as the most fundamental entities, and particles were taken as mere singularities in these fields. A further view, the ill-fated but briefly popular view of energeticism advocated by Wilhelm Ostwald and Georg Helm, argued that energy, and its various ways of transformation, was be the sole universal concept needed in physics.

This struggle between the various world views in theoretical physics and the positions taken by the various authors inevitably came to depend not only on physical arguments, but also on a variety of philosophical themes, in particular the question of what the goal of a physical theory ought to be: to aim for a literally true description of the physical world (realism) or rather to describe and predict the results of experiments (empiricism).

1.2. Boltzmann and Planck's views

In this confusing situation, Boltzmann and Planck came to hold opposing views. Boltzmann championed the mechanistic-atomistic approach in his work on the theory of gases and physics in general. This view clearly embraced the atomic hypothesis which in those days could not (yet) be checked by empirical evidence. However, this should not lead one to jump to the

conclusion that Boltzmann was a naïve realist about the existence of unobservable atoms. He was well aware of the point, and indeed one of the first physicists to emphasize, that concepts like "atoms" are mere images (*Bilder*) that we impose on natural phenomena, and that we can never be sure about whether our images actually correspond to physical reality (cf. *de Regt 2005*). His arguments for defending mechanicism were epistemological, or perhaps strategic, rather than ontological: Boltzmann argued (*1897a,c*) that the alternative view of matter as a continuum, as was common in thermodynamics, involved an equally untestable hypothesis, and claimed that the mechanistic-atomistic view provided the best prospect for moving forward in physics, even if we are not sure that atoms really exist.[2]

However, in the course of his lifelong work on the subject, Boltzmann came to recognize that the hypothesis that matter is constituted of atoms of finite size and mass, moving in accordance with classical mechanics, would not suffice to explain the thermal behavior of macroscopic bodies, even in the case of gases. The result by Boltzmann which had most impressed his contemporaries was his famous H-theorem of 1872 (to be discussed in more detail below). This theorem seemed to explain the irreversible approach towards thermal equilibrium for gases and was first presented by Boltzmann as being derived from purely mechanical suppositions. However, the derivation of this result turned out to be more subtle than Boltzmann had originally presented it. Additional suppositions in the form of arguments from probability theory or statistical considerations were needed, and Boltzmann gradually came to appreciate and admit these suppositions. Indeed, it seems fair to say that Boltzmann never produced a clear account of how the mechanical and statistical ingredients would interrelate to provide the desired derivation of the approach to equilibrium and prove his H-theorem. How Boltzmann's H-theorem would explain the irreversible behavior of a gas towards equilibrium and what assumptions were needed for such an explanation turns out to be an extraordinarily subtle matter, even today. However, this issue was not the focus of his disagreements with Planck.

Planck, on the other hand, took a position that is less easy to describe. He certainly took thermodynamics, which had been the mainstay of his own work since the 1880s, as a more trustworthy theory than the mechanistic-atomistic view. However, he did not take the continuum view of matter as an essential part of his world view. For example, in *1887* he had advanced, simultaneously with (but independent of) Svante Arrhenius, the proposal that salts, upon dissolution in water, split into electrically charged ions, in order to explain the empirical fact that the change in the boiling or freezing temperature and electrical conductivity of a solvent is strikingly different when a salt rather than a neutral substance (like sugar) is dissolved. Clearly, Planck had no problem in entertaining the hypothesis that matter is made

[2] The Boltzmann biography by Carlo Cercignani (*1998*) is therefore aptly called *The man who trusted atoms*, rather than *The man who believed in atoms*.

of unobservable particles. But he was certainly not dogmatic about atomism, and always stressed that one should be most careful in assuming too many hypotheses about their properties. In particular, the assumption that they move in accordance with the laws of classical mechanics and that this would suffice to explain the thermal behavior of macroscopic matter seems to have been a bridge too far for Planck in the 1890s. Further, Planck agreed with Boltzmann in rejecting the position of the instrumentalists like Ernst Mach who argued that atomism was mistaken as a matter of methodological principle. Moreover, Planck and Boltzmann basically agreed in their rejection of energeticism in the debate on this view in 1894-5.

Instead, Planck's main disagreement with Boltzmann in the 1890s centered on two issues. First of all there is the issue of irreversibility. Throughout his writings, at least until 1909, Planck regarded the irreversibility of thermal phenomena, as encapsulated by the second law of thermodynamics, as having absolute validity, and he criticized mechanistic views of nature for making the irreversible behavior of matter a merely statistical, rather than an absolute certainty. Indeed, Planck's "personal reminiscences" describe his differences with Boltzmann (*1946*):

> Boltzmann knew quite well that my point of view was actually rather different from his. In fact, he became angry that I was not only indifferent towards the atomistic theory, but even a little bit negative. The reason was that at this time I attributed the same exceptionless validity to the principle of the increase of entropy as to the principle of the conservation of energy, whereas with Boltzmann the former principle appears only as a probabilistic law which as such also admits exceptions.

Furthermore, Planck argued that the mechanistic-atomistic view adopted in the kinetic theory of gases had simply failed to yield enough recent significant results in the 1890s, and contrasted this to the progress in thermodynamics. Thus he wrote (*1891*):

> Anyone who has studied the works of Maxwell and Boltzmann—the two scientists who have penetrated most deeply into the analysis of molecular motion—will scarcely be able to escape the impression that the remarkable physical insight and mathematical skill exhibited in conquering these problems is inadequately rewarded by the fruitfulness of the results gained.

Similarly, his election address to the Prussian Academy stated (*1894b*):

> At present, the theoretical physicist is faced with problems of a higher difficulty than a generation ago. In those days, there was for everyone who searched for a big encompassing idea in the exact natural sciences, or an all-embracing world view, only a single [. . .] goal: the reduction of all natural processes to mechanics. This view has contributed many rich results to science, even when the audacious hope

> of following every single molecule or even every atom by measurement could not be realized. Still, in the irregular to-and-fro which reigns even in the smallest observable spaces in a gas containing billions of gas molecules, the statistical method has delivered many corresponding results.
> Today, however, this effort, directed at this ultimate goal, has come to a stand-still, and has given rise to a certain disillusionment. Indeed, the mere mathematical analysis needed to penetrate further into these complicated kinds of motion already meets with unsurmountable difficulties [...].

In contrast, Planck continues:

> Recent physical research has witnessed a breakthrough in the tendency to forego the attempt to search for the connection of the phenomena in mechanics. [...] The whole recent development of thermodynamics has been achieved by relying just on the two fundamental principles of the theory of heat. In particular, the fundamental relations between electrodynamics and optics, among electric phenomena, chemical affinity, and thermodynamics were obtained without taking recourse to the mechanical view of the nature of such processes. Similarly, one expects that in the dependence of electrodynamic processes on temperature, as appears in the theory of radiation, one will come closer to an explanation without taking the tedious detour through the mechanical conception of electricity.

Planck emphasized that his endorsement of the thermodynamical approach should not be seen as a rejection of the idea that the ultimate goal of physics is the reduction of all phenomena to mechanics. He did not reject this goal as a matter of principle, but simply claimed that recent attempts to attain this ultimate goal had been fruitless.

Planck's judgment of the recent developments in theoretical physics in the 1890s is actually not hard to understand. The mechanistic view, which underlies the kinetic theory of gases, had produced quite remarkable and unexpected results in the 1860s and 1870s, e.g. the prediction by Maxwell (*1860*) that the viscosity of a gas should be independent of its density, or the explanation of phase transitions in terms of intermolecular forces by Johannes Diderik van der Waals (*1873*). But in the 1880s and 1890s there was little to follow up these early successes. Boltzmann himself, for example, devoted great effort in the 1880s to calculating the viscosity of gases from the perspective of the mechanistic-atomistic view, taking his Boltzmann equation as a starting point. But these calculations got him nowhere near to the observed values for gas viscosity (*1880, 1881a,b*). By contrast, the thermodynamical approach that Planck championed had been quite fertile in the 1890s, establishing fruitful connections between the theory of heat and *prima facie* unrelated areas like magnetism and chemistry.

1.3. The succession of Kirchhoff

Apart from their different positions on the question of what world view would be the most promising for the development of physics, the lives of Boltzmann and Planck also crossed at a professional level.[3] In October 1887 Gustav Kirchhoff died and left vacant the prestigious chair in theoretical physics at the University of Berlin. The university decided in December to offer this chair to Boltzmann, who, at that time, was employed as an ordinary professor in Graz, but also served as Rector of the University. This offer started what must have been one of the most peculiar job negotiations in academia.

Boltzmann went to Berlin in the same month to negotiate the offer and signed an agreement on January 3, 1888 to take up his new chair in October 1888.[4] News about this new contract soon leaked out to the local press, and the university administrators in Graz, who were eager to keep Boltzmann, started to question him. A report of such a conversation on January 8 states that Boltzmann denied having reached an agreement with Berlin, claiming instead that the negotiations were still only at an explorative stage. Of course, the University of Graz took this message as an indication that they could still be able to keep Boltzmann.

Meanwhile, Berlin went ahead with Boltzmann's intended appointment. On March 19 his appointment was made official by the Prussian king. On April 23 followed his election as full member of the Prussian Academy of Sciences. However, in contacts with the authorities in Graz who were inquiring under which conditions Boltzmann was willing to stay there, he suggested that he would seek to forego the position in Berlin on the basis of his bad eyes.

On June 6 Boltzmann wrote to Berlin to inquire when his employment would start and expressed his willingness to resign from Graz. But after a request from Berlin that he send proof of his resignation, he sent a letter on June 24 requesting that he be relieved from his commitment to Berlin, explaining that he had gone through terrible agony and that he was suffering from poor eyesight and nervous disorder and felt unable to take up the prestigious chair. On June 27 he sent a telegram requesting that his previous letter be left unopened. On June 28 he sent another telegram requesting that his last telegram be ignored and that his letter be opened.

Thereupon, Berlin discretely approached Boltzmann's wife, Henriette, about how to proceed. She replied on July 2 that the many agonizing and sad affairs Boltzmann had been dealing with as Rector in Graz had wrecked his nerves, and that the effort to decide on his resignation had excited his nerves further so much so that the physicians feared the worse for his health. The

[3] All the historical claims in this subsection are based on *Höflechner 1994*.

[4] The prestige involved in this offer may perhaps be judged from the salaries mentioned by Höflechner in *1994*: In Graz, Boltzmann earned a yearly income equivalent to around 30,000 Mark. In his new position in Berlin, Boltzmann would receive a yearly income of 137,000 Mark.

upshot was that Boltzmann had not resigned in Graz, although she added that she had witnessed how terribly difficult it had been for him to forego the offer from Berlin and that he was already in deepest grief about rejecting the offer for a position that would suit him so much. In a letter of July 9 the Berlin authorities confirmed Boltzmann's release from his contract. On July 16 Boltzmann sent another message to Berlin, stating that he was day and night in the most bitter resentment over a step he had taken in a moment of excitation and asked whether it was still possible to change his mind (i.e. to accept the offer after all) and offered to travel to Berlin and explain and apologize for his behavior. However, the authorities replied that his release was final. Yet, in August, Boltzmann inquired again, expressing his eagerness to come to Berlin. And even six years later, in 1895, Henriette Boltzmann wrote to the University of Berlin to inquire whether the position was still open.

Berlin, however, had already made their decision. On September 27, 1888 the Ministry of Education requested that the University of Berlin propose a successor to Kirchhoff, and suggested the names of Heinrich Hertz and Max Planck. The appointment went to Planck.

It is, of course, a delicate matter to speculate on how historical actors felt in a given situation. It seems likely, however, that Boltzmann felt torn between the attractive offer from Berlin, where he would have been in a circle of other physicists he could talk to; fears that he would not be able to live up to expectations, due to his ailing health; and feelings of responsibility towards Graz. How Boltzmann felt about the failure of these negotiations we do not know.[5] But it is not hard to imagine how he might have felt when Planck, invested with the authority of his new position, which (at least part of) Boltzmann had sincerely desired, proceeded to state (*1891,1894b*) that Boltzmann's approach to physics seemed fruitless.

Amongst the first tasks that Planck undertook as Kirchhoff's successor was to edit Kirchhoff's collected works, among which were Kirchhoff's lectures on the theory of heat, which touched on the kinetic theory of gases (*Kirchhoff 1894*). Kirchhoff's aims in these lectures had been much more modest than Boltzmann's: in particular, Kirchhoff only considered equilibrium states and left the whole issue of the approach to equilibrium or the increase of entropy during irreversible processes untouched. However, Kirchhoff's lectures did contain an argument to show that a system of gas molecules in equilibrium would be described by the Maxwell distribution law.

[5] There is, however, one passage in his later writing where Boltzmann returned to the episode. In *1905b*, he eagerly quotes an anonymous American colleague stating that Berlin had recently declined in reputation, and that "a lot of things would be in much better shape if I had accepted the offer from Berlin". Boltzmann added: "Many seemingly unavailable persons could have been made to come, if one had really wanted to have them".

Boltzmann (*1894*) reviewed *Kirchhoff 1894* and pointed out an objection against Kirchhoff's argument, and along the way, implicitly commented on Planck's verdict (*1891*) on his own approach:

> Even those who, like the editor of the volume under consideration on Kirchhoff's lectures on the theory of gases [[i.e., Planck]], maintain that this theory is unworthy of the acumen that has been applied to it, will not wish that those who write on this topic will do so with less acumen.

Planck (*1894a*) responded to Boltzmann's review in a manner which showed that he felt hurt by these remarks by Boltzmann. He extensively argued that his job as editor of Kirchhoff's papers had nothing to do with his own opinions on the theory of gases, and that he took no responsibility for their content. However, Planck did take the occasion to mention an idea of his own on how Kirchhoff's proof could be improved and saved against Boltzmann's objection.

Unfortunately, Planck's idea was untenable, and Boltzmann (*1895a*) proceeded, after an admission that he had only intended to criticize the contents of Kirchhoff's ideas and never the person of the editor—and calling Planck's idea "promising"— to make short shrift of Planck's proposal.

My main intention so far has been to argue that the relation between Boltzmann and Planck must have been tense in the period preceding the controversy with Zermelo. Of course, relations between Planck and Boltzmann did not end here. Planck went on to do epoch-making work in black-body radiation in 1899, and, along the way, came closer to Boltzmann's point of view than he could have imagined in the mid-1890s. Indeed, Planck's address at the University of Leiden (*1909*) has often been seen as a complete turn of mind in favour of Boltzmann's viewpoint. But it would take us too far afield to discuss this here, since, by then, Zermelo had long left Berlin and Planck's guidance. Boltzmann, on the other hand, received (and accepted) offers from the universities of Munich (1889), Vienna (1894), Leipzig (1900) and moved back to Vienna in 1902.

Given these strained relations between Boltzmann and Planck at the time when Zermelo published his critique in 1896 and Zermelo's close connection to Planck, Boltzmann saw Zermelo as Planck's mouthpiece. In fact, Zermelo was not only employed that time as Planck's assistant, but Planck was also editor of the *Annalen der Physik und Chemie*, the journal in which Zermelo's *1896a* was published. Boltzmann wrote a letter to the editor-in-chief of the *Annalen*, Eilhard Wiedemann, on March 20, 1896 accompanying his reply (*1896*):

> Most esteemed Colleague!
> I enclose a reply to a paper that appeared in the previous issue of your highly esteemed Annals by a certain Mr. Zermelo. I wrote it within 2 days after I received this issue, but it is not long, and I

> would like to ask you kindly that it will appear as soon as possible. Because it seems that, today, since Maxwell, Clausius, Helmholtz etc. are all dead, I am the only advocate who opposes the view that the mechanical explanation of nature has to be given up, it appears to me that I am, I might say in the interest of science, obliged to take care that at least my voice does not die away, and therefore an answer, as quick as possible, is essential.
> Now I come to a delicate point. Prof. Max Planck is explicitly mentioned as a collaborator of the Annals, and Zermelo is his student. I believe to be justified in demanding: 1. that Mr. Planck will not delay the appearance of my reply, 2. that no word in this reply is altered, 3. that no reply appears in the same issue; later he can answer what he wants and can.

It is clear that Boltzmann felt Planck's blessing behind Zermelo's *1896a*. Planck himself later reminisced feeling the same way (*1946*):

> In any case, he ⟦i.e., Boltzmann⟧ answered young Zermelo with scathing remarks that also hit me, because actually Zermelo's paper had appeared with my permission.

Also, in a letter to his friend Leo Graetz of May 23, 1897, Planck expressed his views on the Zermelo-Boltzmann controversy (cf. *Kuhn 1978*, 27):

> On the main point I side with Zermelo, in that I think it altogether hopeless to [attempt to] derive the speed of irreversible processes—e.g. viscosity or heat conduction in gases—in a really rigorous way from contemporary gas theory. Since Boltzmann himself admits that even the *direction* in which viscosity and heat conduction act can be derived only from considerations of probability, how can it happen that under all conditions the *magnitude* of these effects has entirely determinate values? Probability can serve, if nothing is known in advance, to determine the most probable state. But it cannot serve, if an improbable [initial] state is given, to compute the following [state]. That is determined not by probability but by mechanics. To maintain that change in nature always proceeds from [states of] lower to higher probability would be totally without foundation. [...] Zermelo, however, goes further [than I], and I think incorrect⟦ly⟧. He believes that the second law, considered as a law of nature, is incompatible with any mechanical view of nature. The problem becomes essentially different, however, if one considers continuous matter instead of discrete mass-points like the molecules of gas theory. I believe and hope that a strict mechanical interpretation can be found for the second law along this path, but the problem is obviously difficult and requires time.

It is clear from this letter that Planck and Zermelo did not agree completely, and that Zermelo was not in fact merely Planck's mouthpiece.[6]

However, although Boltzmann was right in recognizing Planck's patronage behind Zermelo's *1896a*, and given that Planck had made it explicit how he regarded Boltzmann's mechanistic approach as lacking in fruifulness, Zermelo's paper is remarkably even-handed in treating the mechanistic and the thermodynamic viewpoints: it simply presents a logical dilemma between these two viewpoints without taking sides.

2. Boltzmann's work in the kinetic theory of gases

Before going into the subject of the Zermelo-Boltzmann controversy, I will now try to review how Boltzmann's work in the kinetic theory of gases evolved in the preceding years. From 1866 onwards, Boltzmann wrote an impressive number of papers and books dealing with statistical physics, in particular on the kinetic theory of gases.

The basic assumption of the kinetic theory of gases, as it developed in the 19th century, is that a gas may be modeled as a large but finite number N of tiny particles moving in accordance with the laws of mechanics, occasionally colliding with each other as well as with the walls of the container in which it is enclosed. The basic tool of the theory, introduced by Maxwell (*1860*), is to represent the state of such a gas by a smooth distribution function $f(\mathbf{v})$ such that $f(\mathbf{v})d\mathbf{v}$ provides the relative number of particles moving with a velocity between $\mathbf{v}$ and $\mathbf{v}+d\mathbf{v}$. Of course, by modern standards, one would say that such a mode of representation involves some approximations or idealizations, since the actual number of particles with their velocity in some given range will always be a natural number, and hence $f(\mathbf{v})$ cannot be smooth. Nevertheless, the idea seems reasonable enough, under the approximative assumptions that N is really huge, and the diameter (or range of interaction) of the molecules is very tiny. In any case, none of the 19th century authors in kinetic theory worried too much about how or under what limit this mode of representation could be made exact.

Maxwell had famously argued in *1860* and *1867* that for the special case when the gas was in thermal equilibrium, the corresponding distribution function takes the form of a normal or Gaussian distribution (in this context often referred to as the Maxwell distribution law):

$$f_{\mathrm{eq}}(\mathbf{v}) = Ae^{-\mathbf{v}^2/B}, \tag{1}$$

[6] However, I can find no written statement by Zermelo endorsing the view attributed to him by Planck that the strict validity of the second law is incompatible with mechanicism *tout court*, i.e. even for continuum mechanics. It is quite clear that the very basis of Zermelo's objection, i.e. the recurrence theorem, does not hold in that context.

where A is just a normalization constant and B is proportional to the absolute temperature of the gas. Although Maxwell investigated many inferences drawn from this distribution law, some of which were quite unexpected and yet turned out to be confirmed by subsequent experiment, he never investigated systematically what would happen if the gas was not in thermal equilibrium.

Ever since 1868, Boltzmann had been closely studying Maxwell's work on kinetic theory, and extended it in various directions (e.g. to the case where the molecules are subject to an external force field). In *1872* he proposed an argument to deal with the case when the gas was not in thermal equilibrium. Boltzmann assumed that in such a case one could still represent the state of the gas by a distribution function $f_t(\mathbf{v})$, where the index t indicates that this function may change in the course of time. He argued that the evolution of f_t in time, as a consequence of the collisions between the particles, should obey a particular evolution equation, now known as the Boltzmann equation. And although this equation was far too difficult to solve, Boltzmann was nevertheless able to show that his evolution equation implied a most important result: For the functional H defined on the distribution functions by $H[f_t] := \int f_t(\mathbf{v}) \ln f_t(\mathbf{v}) d\mathbf{v}$, the equation implies that $H[f_t]$ can only change monotonically in time, and becomes stationary in time only if f is the Maxwell distribution (1). In short, Boltzmann claimed to have shown that if a gas was not yet in a state of thermal equilibrium, it would necessarily evolve towards a final state of equilibrium in the course of time.

It is clear that this particular result of Boltzmann, which is now known as the H-theorem, was considered the most impressive of his achievements by his contemporaries. For example, the proposal to elect Boltzmann to the Prussian Academy of Sciences (1888), written by the academicians Hermann von Helmholtz, Leopold Kronecker, Wilhelm von Bezold and Werner von Siemens, states what these authors considered to be Boltzmann's main achievement (*Kirsten and Körber 1975*, 109):

> The main work in his life is the kinetic theory of gases. In particular, he has proved that the law of distribution of the various values of the velocities, which Maxwell had only verified as a correctly guessed hypothesis, must actually be the necessary form of the final state, as a consequence of the collisions between the molecules. In this work he has shown a high degree of capability of abstraction, and the ability of conquering extremely difficult and involved problems.

Nevertheless, not all of Boltzmann's contemporaries accepted the H-theorem at face value. Loschmidt (*1876a,b, 1877a,b*) was the first to question the general validity of this theorem, and later in 1894-5 a flurry of papers appeared in *Nature* after Edmund P. Culverwell (*1894*) asked the seemingly innocent question "Will someone say exactly what the H-theorem proves?".

I cannot do justice here to the debates about the H-theorem that preceded the Zermelo-Boltzmann controversy. I limit myself to only a few remarks.

First, taking the ahistorical perspective of hindsight, in the derivation of the Boltzmann equation, on which the H-theorem depends, Boltzmann had to rely on a crucial assumption about the collisions between pairs of particles in the gas, which is now known as the *Stoßzahlansatz*. Roughly speaking, the *Stoßzahlansatz* states that the velocities of any two particles entering into a collision are statistically independent *before* the collision—although a similar independence is not demanded for their velocities *after* the collision. This assumption had already been explicitly made by Maxwell in *1867* (which might explain why Boltzmann regarded it as uncontroversial or self-evident). But then, Maxwell did not enter into the question of the evolution of the distribution function but focused on the problem of characterizing the distribution corresponding to the thermal equilibrium state, so that in Maxwell's case the assumption was not used for explaining time-asymmetric phenomena.

However, Boltzmann (*1872*) had extended Maxwell's investigations to the question of describing the evolution of the distribution function, considered at any time t, and the claim of his H-theorem is that this function must approach a Maxwellian distribution at later times, but not at earlier times, i.e., the H-theorem is not time-reversal invariant.

And so, when the issue is raised of how the asymmetry under time reversal embodied in the H-theorem could be reconciled with a perfectly time reversal invariant theory as the mechanics of particles, modern commentators will point out that Boltzmann had relied on the *Stoßzahlansatz*, which treats pre-collision coordinates differently than post-collision coordinates, and that this assumption is responsible for breaking the underlying symmetry between the past and future directions of time in mechanics and is crucial to understanding the H-theorem. But Boltzmann had not stated his assumption in any detail. Indeed, here is the only passage in his 1872 paper devoted to the *Stoßzahlansatz* (*1872*, 323):

> The determination [of the number of collisions] can only be obtained in a truly tedious manner. [...] But since this determination has, apart from its tediousness, not the slightest difficulty, nor any special interest, and because the result is so simple that one might almost say it is self-evident I will only state the result.

Obviously, by avoiding details and calling the issue "self-evident" and claiming that nothing of interest was at stake, Boltzmann did not recognize, or at least failed to alert his readers to, the crucial significance of the *Stoßzahlansatz* to his H-theorem.

Indeed, when Boltzmann was subsequently challenged by Loschmidt in 1876 on the question how the time-asymmetry of the H-theorem could be reconciled with the time-symmetry of the mechanical laws of motion, he did not point to the *Stoßzahlansatz*, but only claimed, without proof, that although there were conceivable states of a gas that would violate the conclusion of his H-theorem, such exceptions were extremely improbable. Thus, this exchange

did not bring along any clearer recognition of the assumptions involved in the derivation of the H-theorem.

It is only in *Boltzmann 1894*, in the very paper that reviewed Planck's edition of Kirchhoff's work in the theory of gases and just prior to the 1895 debate in *Nature*, that I can find Boltzmann stating the claim clearly that, in gas theory, for any pair of molecules entering into a collision, their pre-collision velocities should be regarded as independent, but their post-collision velocities should not. As mentioned earlier, Kirchhoff's *Vorlesungen über Wärmetheorie* (*1894*) discusses the kinetic theory of gases (but only in equilibrium) and assumes that the probability that a pair of molecules simultaneously have positions and velocities in the regions $\delta\mathbf{x}_1\delta\mathbf{v}_1$ and $\delta\mathbf{x}_2\delta\mathbf{v}_2$ is proportional to

$$f(\mathbf{x}_1, \mathbf{v}_1) f(\mathbf{x}_2, \mathbf{v}_2)\,. \tag{2}$$

Boltzmann's critique in *1894* distinguishes three cases:

1. These molecules are about to collide with each other.
2. These molecules have just collided with each other.
3. All other cases.

He then writes:

> In the second case [. . .] one cannot consider the presence of a molecule in its region as independent of the presence of the other particle in its region.

Thus, according to *Boltzmann 1894*, the assumption of statistical independence between pairs of particles is all right immediately *before* collisions, but not immediately *after*. However, the paper does not discuss the H-theorem, so it would still not have been evident to his readers that the assumption stated in this claim is in fact crucial to the proof of that theorem. And apart from this, Boltzmann did not discuss the motivation for his claim, and it might be that he still regarded its validity as self-evident.

The next occasion at which the problem was discussed of how the H-theorem could be reconciled with, or even derived from, the time-reversal invariant laws of motion is the exchange that took place in the columns of *Nature* in 1895. This exchange occurred after Boltzmann's visit to Oxford, where he received a honorary doctorate and gave a lecture at a meeting of the British Association for the Advancement of Science. In the wake of this meeting, Culverwell published a short paper (*1894*) in *Nature* in which he pointed out a "palpable absurdity" in the statement of the H-theorem, as he had understood it from a presentation in Watson's textbook *1893*. The "absurdity" was that any time-asymmetrical result could be obtained from a derivation that only contained time-symmetrical assumptions. Culverwell says that he has not seen Boltzmann's own proof, but assumed it to be all right. Yet, he remained worried about the idea that such a proof could exist at

all, and ended his letter with the innocently-sounding question: "Will someone say exactly what the H-theorem proves?"

This question gave rise to about a dozen letters to *Nature*, each attempting to explain the nature of the proof. However, only the contributions of Samuel H. Burbury, George Hartley Bryan, and Boltzmann himself are worth further consideration. Burbury (*1894a,b*) is the first author to state clearly the logic behind Boltzmann's H-theorem: this theorem depended on a special assumption, which he called "Condition A", that he regarded as breaking the underlying invariance of time-reversal invariant mechanics, and therefore could not itself be grounded in classical mechanics. However, Burbury's statement of his Condition A was obscure, and actually does not provide the required break in time-reversal invariance. Moreover, in later contributions, Burbury showed a remarkable flexibility in what he actually meant by it (cf. *Dias 1994*).

While Burbury thus succeeded in clarifying the logic behind Boltzmann's H-theorem but failed in pinpointing the relevant physical content of the assumption needed, the contribution by Bryan may be said to have the opposite qualities. Bryan (*1894*) was the first to identify what we now call the *Stoßzahlansatz* as the missing ingredient in the proof of the theorem. He argued that a violation of this condition would require that the particles colliding be endowed with "the power of forethought" to regulate their motions so as to move away from equilibrium. As he saw it, the condition was the only natural and reasonable one to be imposed, if the particles "are allowed to take their own natural course and nothing special is known about them".

When Boltzmann entered this exchange in 1895, he stated for the first time explicitly that his H-theorem did not rely on mechanics alone (*1895b*, 414):

> Though interesting and striking at the first moment, Mr. Culverwell's arguments rest, as I think, only upon a mistake of my assumptions. It can never be proved from the equations of motion alone, that the minimum function H must always decrease. It can only be deduced from the laws of probability that if the initial state is not selected for some special purpose, but haphazard⟦ness⟧ governs freely, the probability that H decreases is always greater than that it increases.

Indeed, in what I shall call the statistical reading of the H-theorem, he now makes the following claims: If the number of molecules is very large but finite, the H-function will almost always have the following properties:

1. Most of the time it remains close to its minimum value.
2. Only in the rarest cases does the curve rise to a peak (or "hump") above this minimum value.
3. The probability of a peak decreases rapidly with height.
4. Whenever the value of H is very close to its minimum, the velocity distribution is almost Maxwellian.

In his last contribution to the exchange (*1895c*), Boltzmann praised Burbury for pointing out that a non-mechanical assumption (Condition A) was needed and that this constituted "the weakest point in the derivation of the H-theorem". However, his formulation of the additional assumption was even less clear than Burbury's:

> Condition A is simply this: that the laws of probability are applicable for finding the number of collisions.

The upshot of this exchange was thus that Boltzmann explicitly stated that his H-theorem was not a result of pure mechanics (although Boltzmann's claims that he had already said this in his earlier papers may remain disputable). But the exchange failed, at least in Boltzmann's writing, to bring about a clear recognition of exactly what additional ingredient was needed. Burbury's statement of Condition A had been obscure and flexible. Boltzmann's own formulation of what ingredient was needed (viz. that "haphazard⟦ness⟧ governs freely" or that "the laws of probability are applicable") also lacked sufficient clarity.

Boltzmann's most definitive statement about what assumption is needed in the derivation of the H-theorem is his discussion presented in his *Vorlesungen über Gastheorie*, which were published in two installments: *1896a* (written before his controversy with Zermelo) and *1898* (written afterwards). In *1896a*, Boltzmann again admits that the derivation of the H-theorem requires the assumption of a special condition and credits Burbury for pointing this out. Boltzmann now calls this condition the hypothesis of molecular disorder. Unfortunately, Boltzmann is still not very clear about what this assumption amounts to.

On the one hand, *1896a* first introduces the notion of molecular disorder without giving a formal definition, but as stating that a molecularly ordered distribution is one in which "groups of two or a small number of molecules exhibit definite regularities." On the other hand, Boltzmann also discusses an equation which (apart from notational differences) expresses the idea that pairs of particles that are about to collide should be regarded as independent *before* they collide, and argues that the validity of that equation could be taken as the definition of the statement "the distribution is molecularly disordered". In this latter case, he was formulating the assumption in a manner very close to what we today call the *Stoßzahlansatz*.

Summing up, we can say that Boltzmann had been emphasizing more and more strongly in various places that his H-theorem should be given a statistical reading and did not rely on mechanical assumptions alone. However, apart from Bryan (*1894*), none of the contributors to this exchange succeeded in stating the required additional assumption clearly, nor pointed out how the additional assumption broke time-reversal symmetry. Indeed, it is questionable whether Boltzmann himself fully recognized that his hypothesis of molecular disorder involved a time-asymmetrical element. The litmus test for this issue is, of course, the question of whether a molecularly (dis)ordered

distribution will transform into another (dis)ordered distibution if we reverse all the velocities of all the particles. But even here Boltzmann seemed to be of two minds. In his *1896a* he wrote (*Boltzmann 1964*, 60):

> A molecular disordered distribution after reversal of all velocities can transform into a molecular-ordered one.

which suggests that molecular disorder in not invariant under the transformation. But in the second part of his *Vorlesungen über Gastheorie* (*1898*) he expressed the opposite viewpoint (*Boltzmann 1964*, 442):

> The [molecular] ordered states are not related to the disordered ones in the way that a definite state is to the opposite state (arising from the mere reversal of the directions of all motions), but rather the opposite of each ordered state is again an ordered state.

In any case, Zermelo ignored much of these developments. It is not possible for me to judge whether that was because he was not fully aware of this recent gloss Boltzmann had been putting on his H-theorem, or whether he regarded those formulations of the H-theorem too vague to analyze and for that reason chose to focus exclusively on that what could be stated clearly.

3. The Zermelo-Boltzmann controversy

3.1. *Zermelo 1896a*

In his *1896a*, Zermelo aims to present a discussion of Poincaré's recurrence theorem (*Poincaré 1890*) and to point out its consequences for statistical physics, in particular for the kinetic theory of gases. In modern terms, this theorem is commonly presented as follows:

Theorem 3.1 (Recurrence theorem). *Let* $\langle \Gamma, \aleph, \mu, \{T_t\}\rangle$ *be any dynamical system such that* $\mu(\Gamma) < \infty$*, and let* $A \in \aleph$ *be any measurable set in* Γ*. Consider any* $\tau > 0$ *and let*

$$B = \{x \in \Gamma : x \in A \;\&\; \forall t > \tau \; (T_t x \notin A)\}\,, \tag{3}$$

the set of states in A *that after time* τ *have left* A *and will never return to* A*. Then*

$$\mu(B) = 0\,. \tag{4}$$

Here, a dynamical system $\langle \Gamma, \aleph, \mu, \{T_t\}\rangle$ consists of a measure space $\langle \Gamma, \aleph, \mu\rangle$ in the sense of measure theory.[7] To turn a measure space into a dynamical

[7] That is, $\aleph$ is a σ-algebra consisting of all measurable subsets of Γ, and μ is a non-negative valued-function on $\aleph$ satisfying $\mu(\emptyset) = 0$ and $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ for all countable sequences of mutually disjoint measurable subsets A_i of Γ.

system, one further assumes the existence of a one-parameter group of measurable evolution operators $T_t : \Gamma \to \Gamma$ for all $t \in \mathbb{R}$ which has the group property $T_t \circ T_{t'} = T_{t+t'}$ and is measure preserving:

$$\forall t \in \mathbb{R}\,,\ \forall A \in \aleph\ :\ \mu(T_t A) = \mu(A)\,. \tag{5}$$

Of course, neither Poincaré nor Zermelo had recourse to concepts of measure theory nor to the theory of dynamical systems, but their statements that only "exceptional states" x would have the property that they never return to a set where they were initially located, and their elucidation that the qualification "exceptional" means that those points only make up an "extension" (where we would say "measure") zero, indicate that their grasp of the concept was already firm.[8] Poincaré and Zermelo both state their theorem only for the special case of the Lebesgue measure, i.e. the special choice of measure that makes open sets in a Euclidean phase space measurable and assigns a measure value to such sets which equals their Euclidean volume. However, the validity of the recurrence theorem does not depend on that choice.

Zermelo notes immediately that this theorem implies an objection to the kinetic theory of gases and argues that the latter needs to be "fundamentally revised". He further states that Poincaré had failed to draw attention to the consequences of this theorem for the kinetic theory of gases. He is wrong about this: Poincaré had done so in his *1893c*, in the *Revue de Metaphysique et Morale*, a journal that may not have been part of the staple literature that Zermelo or other physicists regularly read. Actually, Poincaré's *1893c* version of the objection is in an important sense different from the way Zermelo framed it: Poincaré phrases the objection as a conflict between mechanicism and experience; whereas Zermelo presents it as a contradiction between two physical theories: mechanics and thermodynamics. Indeed, Zermelo's paper never brings empirical considerations into the discussion.

Zermelo starts in *1896a* by giving a proof of the theorem. He considers a system of N material points and assumes that their motion is governed by some first-order differential equations [(1)].[9] He argues that these equations of motion will be integrable, i.e., that they guarantee that for each time t a later state P_t will correspond to every initial P_0 at time $t = 0$, as given by the solutions [(3)]. In the formulation given above (where states are denoted by x rather than P), this corresponds to the idea that the evolution operators $T_t : \Gamma \to \Gamma$ are defined with $T_t : x \mapsto T_t x = x_t$.

Actually, Zermelo's treatment here is a bit sloppy: the conditions he mentions on the equations of motion do not actually guarantee the correspondence

[8] Actually, not all the ingredients of the theory of dynamical systems are needed to guarantee the validity of the recurrence theorem. Essential in this theorem are the assumptions that $\mu(\Gamma)$ is finite and that T_t preserves measure (i.e. equation (5)); but otherwise, one could, e.g., also employ the weaker assumption that $\{T_t\}$ is just a one-parameter semigroup instead of a group (cf. *Brown et al. 2009*).

[9] The notation "[(n)]" is used to refer to the equations in *Zermelo 1896a*.

of a unique later state to each initial state; for that purpose the functions X_μ in [(1)] (or in the Hamiltonian equations introduced later) must not only be continuous, but also Lipschitz-continuous. (Poincaré (*1890*) had also failed to mention this condition.) This omission was pointed out, however, in *Boltzmann 1897b*.

Next, Zermelo introduces a set g_0 of initial states obeying inequalities of the form $F(x_0) < 0$ for some unspecified, but presumably continuous, functions F. If continuity is assumed, this makes g_0 (i.e. A in the formulation above) the "continuously extended area", as Zermelo calls it, i.e., an arbitrary open set in Γ. Actually, nothing in the proof of the recurrence theorem hinges on this assumption; the theorem as formulated above holds for any measurable set A and does not depend on topological ingredients. However, taking the set A to be open (in the Borel topology on Γ) implies that it has positive (Lebesgue) measure $\mu(A) > 0$. This consequence is important in the subsequent argument that the subset $B \subseteq A$ of initial states that never return to A after time τ must be "exceptional" or "singular" because $\mu(B) = 0$. After all, if $\mu(A) = 0$ is allowed to begin with, the result that $\mu(B) = 0$ would not sustain the claim that the states in B would have to be exceptional in A.

Zermelo considers the "extension" γ_0 of g_0, i.e. the (Lebesgue) measure of the set g_0, $\gamma_0 = \mu(g_0)$, and notes that by Liouville's theorem[10] this measure is preserved through time, i.e., if $g_t = T_t g_0$:

$$\gamma_t := \mu(g_t) = \mu(g_0) = \gamma_0 \,. \tag{6}$$

Next, Zermelo considers the union of the sets g_t,[11]

$$G_0 = \bigcup_{t \geq 0} g_t \,. \tag{7}$$

Now consider the time evolution of this set G_0, i.e. $G_\tau = T_\tau G_0$, for $\tau > 0$. On the one hand, it is easy to show that $G_\tau = \bigcup_{t \geq \tau} g_t$, which implies that $G_\tau \subseteq G_0$. As Zermelo puts it: "this change [of G_τ during its evolution in time] is due always to the disappearance of early states, but never to the appearance of new states." On the other hand, by Liouville's theorem, the measure of G_0 is also conserved under the evolution, so that $\mu(G_\tau) = \mu(G_0)$.

[10] Actually, Zermelo's paper may well have been one of the first times that the result is called "Liouville's theorem". Boltzmann used the theorem on many previous occasions, but referred to it as "Jacobi's theorem on the last multiplier".

[11] By modern standards, one might prefer instead of (7) to consider a discrete version of this union, i.e. something like $G_0 = \bigcup_{n=0}^{\infty} g_{t+n\tau}$ for some $\tau > 0$, since otherwise the set G_0 might fail to be measurable, i.e. not belong to the σ-algebra $\aleph$. However, if we specialize to Lebesgue measure, and assume γ_t open for all $t > 0$, as seems to be Zermelo's intention, the problem evaporates, since even in an uncountable union such as (7), G_0 is still an open set and hence Lebesgue measurable. Indeed, these observations do not affect the validity of Zermelo's proof.

It follows that the difference, the set of "disappearing states",

$$B := G_0 \setminus G_\tau = \bigcup_{0 \leq t < \tau} g_t \,, \tag{8}$$

is a measure zero set. But this set contains exactly those states that were in g_0 but after a time lapse τ never return to g_0.

After providing this proof of the recurrence theorem, which is essentially the same as Poincaré's proof (even to the point of both failing to mention the need for Lipschitz-continuity), Zermelo provides a corollary:

Theorem 3.2 *There exists no single-valued and continuous function S on Γ that has the following property: One can find an open set g, however small, such that $\forall x \in g$, $S(T_t x)$ is monotonically increasing as a function of t.*

Interestingly, even before he had found the recurrence theorem, Poincaré (*1889*) also had considered the question whether it was possible to define a function of the state that would have the property that it increases along the dynamicalal trajectories and claimed this was not possible. This paper, which was noticed by Zermelo, has been translated and analyzed by Elwood Olsen (*1993*). Olsen concluded that this attempt by Poincaré was unconvincing. Poincaré himself, apparently did not realize that his own recurrence theorem allowed a more secure basis for such a claim.

Harvey Brown et al. (*2009*) gave a slightly extended formulation of Zermelo's corollary in which the assumption of continuity of the function S is weakened to the assumption that S is an integrable non-negative function, and the open set g is replaced by an arbitrary measurable set. Their formulation is as follows:

Theorem 3.3 *Let $\langle \Gamma, \aleph, \mu, \{T_t\}\rangle$ be any dynamical system such that $\mu(\Gamma) < \infty$, and let g_0 be any measurable set in Γ and $G = \bigcup_{n=0}^{\infty} T_{n\tau} g_0$ for some time $\tau > 0$. Then there exists no integrable non-negative function S on Γ such that*

$$\int_{T_t(T_t G)} S d\mu$$

is monotonically increasing as a function of t.

Remarkably, this extended version is not only implied by, but is fully equivalent to, the recurrence theorem. Indeed, as long as we assume that the dynamical flow is smooth (which is guaranteed by assuming a Lipshitz-continuous Hamiltonian), for g_0 an arbitrary measurable set in phase space, we can define a function S on Γ by

$$S(x) := \inf\{t \geq \tau \ : \ T_t x \in g_0\} \,. \tag{9}$$

In words, this function simply indicates, for any choice of x, how much time has elapsed (after a threshold time τ) since the trajectory through x last passed through region g_0. If the trajectory through x never passes through

g_0 (and the infimum is over the empty set), we simply stipulate that $S(x) = 0$. Clearly, $S(T_t x)$ will be monotonically increasing in t iff the trajectory through x never recurs to g_0.

After proving the two theorems mentioned above, Zermelo replaces the abstract differential equations of motion [(1)] by the familiar Hamiltonian equations of motion. He argues that all attempts to give an atomistic, mechanical account of the behavior of matter, e.g. the kinetic theory of gases, fall under this Hamiltonian framework.

Zermelo formulates the conclusion of his theorems as an obstacle for "irreversible" processes. Unfortunately, he does not define that term. He may have meant by "irreversible" processes those that display a permanent approach towards equilibrium, i.e. processes in which any initial non-equilibrium state eventually move towards, and permanently remain within, a region of phase space associated with equilibrium. Another reading might be that he means by "irreversible" processes evolutions such that some continuous entropy-like function on phase space increases monotonically with time. Either way, it is clear that the recurrence theorem or its corollary imply that no Hamiltonian dynamical system can be irreversible.

Zermelo carefully discusses various options for avoiding this conclusion:

1. We can assume that the system has no bounded phase space. This could be achieved by allowing the particles (a) to reach unbounded positions in space, or (b) to attain unbounded velocities. Option (a) is however excluded by the assumption that a gas is contained in a finite volume. Option (b) could be achieved when the gas consists of point particles which attract each other at small distances (e.g. an $F \propto r^{-2}$ inter-particle attractive force can accelerate them toward arbitrarily high velocities). However, Zermelo argues on physical grounds that one ought to assume that there is always repulsion between particles at very small distances.

2. We can assume that the particles act upon each other by velocity-dependent forces. This, however would lead either to a violation of the conservation of energy or the law of action and reaction, both of which Zermelo regards as essential to atomic theory.

3. The H-theorem holds only for those special initial states which are the exception to the recurrence theorem, and we assume that only those states are realized in nature. This option would be unrefutable, says Zermelo. Indeed, the reversibility objection has already shown that not all initial states can correspond to the second law. However, here we would have to exclude the overwhelming majority of all imaginable initial states, since the exceptions to the recurrence theorem only make up a set of total extension (i.e. measure) zero. Moreover, the smallest change in the state variables would transform a singular state into a recurring state, and thus suffice to destroy the assumption. Therefore, this assumption would be quite unique in physics and I do not believe that anyone would be satisfied with it for very long.

This leaves only two major options.

4. The Carnot-Clausius principle is to be altered.

5. The kinetic theory is to be formulated in an essentially different way, or even be given up altogether.

Zermelo's *1896a* does not express any preference between these last two options. He concludes that his aim has been to explain as clearly as possible what can be proved rigorously and hopes that this will contribute to a renewed discussion and final solution of the problem. (Of course, his next paper shows that his own preferences lie along option 5.)

I would like to emphasize that, in my opinion, Zermelo's argument is fair and entirely correct. If he can be faulted for anything, it is that he had not noticed that Boltzmann, in his very recent papers, had already been putting a different gloss on the H-theorem.

3.2. *Boltzmann 1896*

Boltzmann first states that he has repeatedly pointed out that the theorems of the kinetic theory of gases have the character of "statistical truths", and refers to several of his previous works to substantiate the claim.

The claim is fair enough: Boltzmann had pointed out the statistical aspect of his understanding of the theorems of the theory of gases (in particular the H-theorem) already in his reply (*1877a*) to Loschmidt, and with much more emphasis in his contribution to the debate in *Nature* (*1895b*) just a year before Zermelo's article was written. However, I think it is also fair to note that Boltzmann never explained clearly what the exact nature or status of such "statistical truths" were, or in what sense they were independent from mechanical considerations. Moreover, it also seems fair to note that in spite of Boltzmann's repeated warnings that such theorems had to be taken in a statistical sense, there are also occasions, not just in 1872 but also fairly recently before the debate with Zermelo, where he did not refer to statistical considerations, but claimed that the theorems of the theory of gases were rigorous analytical theorems from mechanics alone. For example, in *1892*, after giving yet another proof of the Maxwell distribution in thermal equilibrium, he concluded (p. 432): "I believe therefore that its correctness [i.e. that of the Maxwell distribution] as a theorem of analytical mechanics can hardly be doubted."

As I said earlier, Boltzmann's writings are not easy to interpret, even today. Zermelo may well have chosen to avoid the labor of going into an analysis of what Boltzmann might have meant and to focus instead on stating what could be deduced from Poincaré's theorem. Yet it is equally understandable that Boltzmann might have felt misunderstood by Zermelo's paper, since he had occasionally emphasized that the H-theorem and the permanence of Maxwell distribution should not be conceived of as rigorous theorems in mechanics.

Boltzmann's rejoinder, which he apparently wrote within two days, starts with a sarcastic remark, claiming that he "cannot but take delight" in Zermelo's essay as "the first proof that these works of mine receive any attention at all in Germany".[12]

Boltzmann correctly points out that if the number of molecules is infinite, the recurrence theorem does not apply (because the phase space is not bounded in this case).

Boltzmann then repeats the claims 1. to 4. (cf. p. 201) he made in his contribution *1895b* to the *Nature* debate, and which I called the statistical reading of the H-theorem. The only remark I can add is that although Boltzmann presented these claims in his *1895b* as having been "proved in my papers" and in his present reply introduced these claims by the phrase, "As I have already shown in the contribution to *Nature*", he actually never gave a demonstration of the validity of these claims. Indeed, most of these claims are too vague to admit of a rigorous demonstration. In particular, an ambiguity which persists in much of Boltzmann's writings is the question whether he intends "probability" to be measured by duration in time or by measure in phase space. Thus, when he speaks about a hump in the H-curve occurring "only in the rarest cases" above, the intended meaning could be that such humps are rare in time, but also that they only occur for very rare choices of the initial states—or both. And then there is the question of whether and how the *Stoßzahlansatz* or similar conditions might be needed to demonstrate their validity, which Boltzmann leaves completely untouched. Thus, up to this point, Boltzmann describes how his own views on the interpretation of his 1872 H-theorem had developed.

Next, Boltzmann states a first disagreement with Zermelo: Zermelo believes that it is only for exceptional initial states that the gas comes ever closer to satisfying Maxwell's law of distribution, and that this does not seem right to him. Instead, Boltzmann claims, it is only for exceptional initial states that the Maxwell distribution never holds while for the vast majority of initial states the H-curve has the properties just stated. This disagreement, however, is only an optical illusion. When Zermelo stated (*1896a*) that "it is [...] impossible to show [...] the well-known law of velocity distribution among gas molecules to be the stationary final state regularly reached after some time, as its discoverers, Maxwell and Boltzmann intended to", he clearly meant with a "stationary final state" a condition that was permanent, and would never be changed at any later time. When Boltzmann claimed that it is only for exceptional states that the Maxwell distribution will never be reached for some period of time, he is of course also right, at least for typical gas models. But this claim does not contradict Zermelo's at all, because the quantifiers are in a different order.

[12] This remark seems somewhat odd, since Boltzmann's previous call to the University of Berlin, his election to the Prussian Academy, and his previous debate with Planck made clear that Boltzmann's work did not escape notice in Germany.

In fact, Boltzmann subsequently admits the validity of the recurrence theorem, but sees this as fully consistent with his own approach. However, he disputes the conclusion that the mechanical approach should be modified or even abandoned. He argues that this conclusion would only be justified if the mechanical approach violates our experience. To show that this is not the case, Boltzmann argues by means of a thought experiment, elaborated in the appendix of his paper, that the recurrence time for even a cubic cm of gas could be truly enormous ($10^{10^{19}}$ seconds) and hence utterly escape observation.

From a historical point of view it is interesting that Boltzmann points out other examples of improbable yet not impossible states of gas that would not require a quasi-recurrence to an original state, like fluctuations in pressure or chemical transformations at temperatures below the reaction threshold. Boltzmann argues that such improbable transitions have actually been observed. The most pregnant of his remarks is that "observations were made of movements of very small corpuscles, which may be due to the fact that in such cases a pressure which is sometimes a little greater, sometimes a little smaller, really acts on a part of their surface that no longer vanishes compared to their entire surface." Here, Boltzmann seems to be referring to what we now know as Brownian motion, the subject that Einstein would deal with in much more detail in 1905 and that would eventually contribute to a much wider acceptance of the reality of molecules. If Boltzmann had paid more attention to how the predictions of the theory of gases relate to such observations and had played down this empirical card, which he held in his sleeve, with more emphasis, he might have anticipated Einstein in this regard.

Boltzmann ends with a biting remark:

> All objections against the mechanical approach to nature are therefore unfounded and based on mistakes. Anyone unable to overcome the difficulties attendant on a clear understanding of the principles of the theory of gases really ought to heed Mr. Zermelo's advice and resolve to abandon the theory altogether.

Given the fact that Zermelo had set out to ascertain what can be proved rigorously and what can not, Boltzmann's hasty response may be rather disappointing, since he provided no clear statement of how he understood the statistical reading of the H-theorem or of the conditions on which it relied. Still, Boltzmann is obviously correct when he says that Zermelo's objection did not lead to a conflict between the theory of gases and experience. Thus, his argument would have been more successful as a counter-argument to Poincaré than to Zermelo.

3.3. Zermelo's response *1896b*

In his *1896b*, Zermelo notes that Boltzmann's response confirms his views by admitting that Poincaré's theorem is correct and applicable to a closed system

of a finite number of gas molecules. Hence, in such a system, all [sic!] motions "are *periodic* and hence strictly *non-irreversible*". Thus, the kinetic theory of gases cannot assert that there is a strict monotonic increase of entropy as the second law would require. He adds that this general clarification was not at all superfluous.

Therefore, Zermelo argues, his main point had been conceded: there is indeed a conflict between thermodynamics and kinetic theory, and it remains a matter of taste which of the two is to be abandoned. Zermelo admits that observation of the Poincaré recurrences may well fall beyond the bounds of human experience. He points out (correctly) that Boltzmann's estimate of the recurrence time presupposes that the system visits *all* other cells in phase space before recurring to an initial state. This estimate is inconclusive, since the latter assumption is somewhat ad hoc. In general, these recurrence times need not "come out so 'comfortingly' large". But, as I stressed before, the relation with experience simply was not an issue in Zermelo's objection.

The main part of Zermelo's reply analyzes the justification of and consequences drawn from Boltzmann's assumption that the initial state is very improbable, i.e., that H_0 is very high. Zermelo argues that even in order to obtain an approximate or empirical analogue of the second law as Boltzmann envisaged, i.e. an approach to a long-lasting, but not permanent equilibrium state, it would not suffice to establish this result for one particular initial state. Rather, one would have to show that evolutions *always* take place in the same sense, at least during observable time spans.

As Zermelo understands it, Boltzmann does not merely assume that the initial state has a very high value for H, but also that, as a rule, the initial state lies on a maximum or has just passed a maximum. If this assumption is granted, then it is obvious that one can only observe a decreasing flank of the H-curve. However, Zermelo protests, one could have chosen any time as the initial time. In order to obtain a satisfactorily general result, the additional assumption would thus have to apply at all times. But then the H-curve would have to consist entirely of maxima. This leads to nonsense, Zermelo argues, since the curve cannot be constant. Zermelo concludes that Boltzmann's assumptions about the initial state are thus in need of further *physical* explanation.

Further, Zermelo points out that probability theory, by itself, is neutral with respect to the direction of time, so that no preference for evolutions in a particular sense can be derived from it. He also points out that Boltzmann apparently equates the duration of a state and its extension (i.e. the relative time spent in a region and the relative volume of that region in phase space). "That he has actually *demonstrated* this property for his function H [...] I fail to see, since, in my view, probability and duration of a state are not identical" (*1896b*, 796).

3.4. Boltzmann's second reply *1897*

In his second reply Boltzmann rebuts Zermelo's demand for a physical explanation of his assumptions about the initial state of the system with the claim that the question is not what will happen to an arbitrarily chosen initial state, but rather what will happen to a system in the present state of the universe.

He argues that one may depart from the (admittedly unprovable) assumption that the universe (or at least a very large part of the universe) that surrounds us started in a very improbable state and still is in an improbable state. If one then considers a small system (e.g. a gas) that is suddenly isolated from the rest of the universe, there are the following possibilities: (i) The system may already be in equilibrium, i.e. H is close to its minimum value. This, Boltzmann says, is by far the most probable case. But among the few cases in which the system is not in equilibrium, the most probable case is (ii) that H will be on a maximum of the H-curve, so that it will decrease in both directions of time. Even more rare is the case in which (iii) the initial value of H will fall on a decreasing flank of the H curve. But such cases are just as frequent as those in which (iv) H falls on an increasing flank.[13]

Thus, Boltzmann's explanation for the claim that H is initially on a maximum is that this is the most likely case for a system in a non-equilibrium state which becomes isolated from the rest of the universe in its present, improbable state.

Boltzmann does not respond to Zermelo's requests for more definite proofs of his claims, in particular the equality of averages over phase space volume and time averages. He bluntly states that he has thirty years of priority in measuring probabilities by means of phase space volume (which is true) and adds that he always had done so (which is false). Indeed, Boltzmann continues to equivocate between these two senses of averaging: A few lines below, he claims that the most probable states will also occur most frequently, except for a vanishingly small number of initial states, but does not attempt to justify such a claim, in spite of Zermelo's complaint that he could not find a proof of such claims.

3.5. Concluding remarks

Boltzmann's replies to Zermelo have been recommended as "superbly clear and right on the money" (*Lebowitz 1999*, 347). However, as will be clear from the above, I do not share this view. See also *M. Klein 1973, Curd 1982, Batterman 1990, Cercignani 1998, Brush 1999, Earman 2006*, and *Frigg 2008* for other commentaries on the Zermelo-Boltzmann controversy.

[13] The Ehrenfests later (*Ehrenfest and Ehrenfest 1912*) added a final possible case (v): H may initially be on a local minimum of the H-curve, so that it increases in both directions of time. But by a similar reasoning, that case is even less probable than the cases mentioned by Boltzmann.

It is clear that, in at least one main point of the dispute, Boltzmann and Zermelo had been talking past each other. When Zermelo argued that in the kinetic theory of gases there can be no continual approach towards a final stationary state, he obviously meant this in the sense of a limit $t \to \infty$. But Boltzmann's reply indicates that he took the "approach" as something that is not certain but only probable, and as lasting for a very long, but finite time. His graph of the H-curve (*1897*, 398) makes abundantly clear that he does not intend to claim that $\lim_{t\to\infty} H(t)$ even exists.

It is true that Boltzmann's statistical reading of the H-theorem, which he stressed in his second reply, had already been explicit in *1895b*, and thus he could claim with some justification that this development in his thinking had been overlooked by Zermelo. But in fairness, one must note that, even only just before the debate, Boltzmann had also expressed views which seemed to contradict this statistical reading of the H-theorem. Indeed, the first volume of Boltzmann's *Vorlesungen über Gastheorie* (*1896a*) stressed, much like his original *1872* on the H-theorem, the necessity and exceptionless generality of the H-theorem, adding only that the theorem depended on the assumption of molecular disorder (*1896a*, § 5, 38):

> [T]he quantity designated as H can only decrease; at most it can remain constant. [...] The only assumption we have made here is that the distribution of velocities was initially 'molecularly disordered' and remains disordered. Under this condition we have therefore proved that the quantity called H can only decrease and that the distribution of velocities must necessarily approach the Maxwell distribution ever more closely.[14]

The point here is not whether the claim that H can only decrease in the course of time could be derived from pure analytical mechanics: Boltzmann states clearly in this quote that an additional assumption of molecular disorder is involved, which, although its precise meaning might be unclear, seems to involve a motivation that goes beyond mechanics. Rather, the point is that in this quote Boltzmann still expressed the claim that, with this assumption in place, one could prove a necessary and steady approach towards equilibrium, without exceptions. This claim, of course, is quite different from Boltzmann's statistical reading of the H-theorem in *1895b*. Zermelo was thus at least equally justified in claiming that Boltzmann's clarification "was not at all superfluous" (*1896b*, 793).

[14] In his reply to Zermelo, Boltzmann claimed that his discussion of the H-theorem in the *Vorlesungen über Gastheorie* was intended under the explicitly emphasized assumption that the number of molecules was infinite, so that the recurrence theorem did not apply. However, I can find no mention of such an assumption in this context. On the contrary, the first occasion on which this latter assumption appears is in §6 on page 46 where it is introduced as "an assumption we shall make later", suggesting that the previous discussion did *not* depend on it.

To repeat, Boltzmann's various presentations of claims about the H-theorem, and the assumptions involved, are not easy to digest, even today. To date, the most elaborate attempt to provide a version of a proof of a statistical H-theorem which remains close to Boltzmann's original intentions is *Lanford 1975* and the subsequent work by authors elaborating on this approach. However, it would take us to far afield to delve into that.

To conclude, Zermelo's dispute with Boltzmann in 1896–7 is rather well-known and has often been commented on before, both from a historical and from a philosophical foundations of physics perspective. In the popular literature dealing with this episode, a picture has often been painted in which Boltzmann had the upper hand and Zermelo's objections are described as

Ueber einen Satz der Dynamik und die mechanische Wärmetheorie

1896a

Im zweiten Kapitel der *Poincaré*'schen Preisschrift über das Dreikörperproblem[1] findet sich ein Satz bewiesen, aus welchem hervorgeht, dass die verbreiteten Vorstellungen von der Wärmebewegung der Molecüle, wie sie z. B. der kinetischen Gastheorie zu Grunde liegen, einer wesentlichen Abänderung bedürften, um mit dem thermodynamischen Hauptsatze von der Vermehrung der Entropie vereinbar zu werden. Dieses *Poincaré*'sche Theorem sagt aus, *dass in einem System von materiellen Punkten unter Einwirkung von Kräften, die allein von der Lage im Raume abhängen, im allgemeinen ein einmal angenommener durch Configuration und Geschwindigkeiten charakterisirter Bewegungszustand im Laufe der Zeit, wenn auch nicht genau, so doch mit beliebiger Annäherung noch einmal, ja beliebig oft wiederkehren muss, vorausgesetzt, dass die Coordinaten, sowie die Geschwindigkeiten nicht ins Unendliche wachsen.* In einem solchen System sind daher, von singulären Anfangszuständen abgesehen, *irreversible Vorgänge unmöglich*, es kann keine eindeutige und stetige Function der Zustandsgrössen wie die Entropie fortwährend zunehmen, da jeder endlichen Zunahme bei der Rückkehr in den Anfangszustand wieder eine Abnahme entsprechen müsste. Hr. *Poincaré* bedient sich in der genannten Abhandlung seines Satzes zu astronomischen Erörterungen über die Stabilität des Sonnensystemes, er scheint aber seine Anwendbarkeit auf Systeme von Molecülen oder Atomen und damit auf die mechanische Wärmetheorie nicht bemerkt zu haben, wiewohl er gerade

[1] *Poincaré*, „Sur les équations de la dynamique et le problème des trois corps", Acta Mathematica **13.** p. 1–270. 1890; der betreffende Satz p. 67–72.

hostile, misguided, or wrongheaded. However, Zermelo clearly had the better arguments in this debate. Although it seems clear that he did not sympathize with Boltzmann's approach, his objections were stated fairly and precisely. By contrast, Boltzmann's own responses added a sense of hostility to the controversy, and failed to answer or fully elucidate his position on the questions Zermelo was asking. Of course, this view on the Zermelo-Boltzmann controversy does not deny the fact that in the early decades of the 20th century, the mechanistic-atomistic approach championed by Boltzmann gained a clear victory over its alternatives. However, a clear and commonly accepted answer on the question how to explain irreversible phenomena in statistical physics has not been reached.

On a theorem of dynamics and the mechanical heat theory

1896a

A theorem proved in the second chapter of *Poincaré*'s prize essay on the three-body problem[1] suggests that widespread ideas about the heat motion of molecules such as those underlying the kinetic theory of gases need to be fundamentally revised so as to become compatible with the law of thermodynamics concerning the increase of entropy. This theorem by *Poincaré* states *that, in a system of material points under the action of forces depending only on location in space, a state of motion that is characterized by configuration and velocities and that has occurred once must in general occur again one more time, even if not as exactly the same, but at least in arbitrary approximation, and even arbitrarily many times, provided that neither the coordinates nor the velocities grow infinitely large.* In a system of this sort, *irreversible processes cannot possibly occur*, with the exception of singular initial states, and no single-valued and continuous function of the state variables such as entropy can always increase, since to every finite increase there would have to correspond a decrease when the system returns to the initial state. In his essay, *Poincaré* uses his theorem for astronomical considerations on the stability of the solar system. However, it seems that he has failed to notice its applicability to systems of molecules or atoms, and hence to the mechanical theory of heat, his special interest in the foundational questions

[1] *Poincaré 1890*; the relevant theorem p. 67–72.

den Grundfragen der Thermodynamik besonderes Interesse zugewandt und auf einem anderen Wege den Nachweis versucht hat, dass die irreversiblen Vorgänge aus der v. *Helmholtz*'schen Theorie der „monocyclischen Systeme“
486 nicht immer | erklärt werden können.[1] Um nun das Studium der umfangreichen und vielen Physikern schwerer zugänglichen *Poincaré*'schen Arbeit nicht voraussetzen zu müssen, schicke ich einen möglichst einfachen Beweis des angeführten Satzes voraus.

Sei N die Anzahl der materiellen Punkte und werden die $n = 6\,N$ Zustandsgrössen, d. h. die $3\,N$ Coordinaten und $3\,N$ Geschwindigkeitscomponenten mit $x_1, x_2, \ldots x_n$ bezeichnet, so sind die nach der Zeit genommenen Differentialquotienten der ersteren identisch mit den entsprechenden Geschwindigkeitscomponenten, die Ableitungen der letzteren aber, d. h. die Beschleunigungscomponenten, die Kräfte, nach unserer Annahme eindeutige und stetige Functionen der Coordinaten. Jene sind also von den Coordinaten, diese von den Geschwindigkeiten unabhängig, und die Differentialgleichungen der Bewegung sind von der Form

$$\left.\begin{array}{l} \dfrac{dx_\mu}{dt} = X_\mu\,(x_1, x_2, \ldots x_n) \\ (\mu = 1, 2, \ldots n)\ , \end{array}\right\} \tag{1}$$

wo keine der Functionen X_μ die entsprechende Variable x_μ selbst enthält, sodass die Beziehung besteht:

$$\frac{\partial X_1}{\partial x_1} + \frac{\partial X_2}{\partial x_2} + \cdots + \frac{\partial X_n}{\partial x_n} = 0\ . \tag{2}$$

In einem solchen Systeme (1) von Differentialgleichungen erster Ordnung entspricht einem beliebigen Anfangszustand P_0:

$$x_1 = \xi_1\ , \quad x_2 = \xi_2 \ldots x_n = \xi_n\ , \quad (t = t_0)$$

ein bestimmter veränderter Zustand P zur Zeit t, ausgedrückt durch die Integralgleichungen von (1):

$$\left.\begin{array}{l} x_\mu = \varphi_\mu\,(t - t_0, \xi_1, \xi_2, \ldots \xi_n)\ , \\ (\mu = 1, 2, \ldots n)\ , \end{array}\right\} \tag{3}$$

wo die φ_μ eindeutige und stetige Functionen ihrer sämmtlichen Argumente sind, die, unabhängig von der Wahl des Zeitanfanges t_0, durch die Functionen X_μ allein bestimmt werden. Diese Beziehungen gelten eben so gut für vorher-
487 gehende wie für nachfolgende Zeiten, d. h. eben so gut für | negative wie für positive Werthe von $t - t_0$; der Anfangszustand P_0 ist eine beliebige, willkürlich hervorgehobene Phase der Bewegung, die nicht immer zeitlich voranzu-

[1] *Poincaré*, Compt. rend. **108**. p. 550–552. 1889; „Vorles. über Thermodynamik“, p. 294–296.

of thermodynamics notwithstanding, and even though he has tried to show by other means that the irreversible processes cannot always be explained on the basis of *v. Helmholtz*'s theory of "monocyclic systems".[2] To save the reader the trouble of delving into *Poincaré*'s work, which is extensive and not easily accessible to many physicists, I begin by providing as simple a proof as possible of this theorem.

Let N be the number of material points and let the $n = 6\,N$ state variables, that is, the $3\,N$ coordinates and $3\,N$ velocity components, be denoted by $x_1, x_2, \ldots x_n$. Then the derivatives of the former taken with respect to time are identical with the corresponding velocity components, while the derivatives of the latter, i.e., the acceleration components, the forces, are, by our assumption, single-valued and continuous functions of the coordinates. Thus, the former are independent of the coordinates, and the latter are independent of the velocities, and the differential equations of the motion have the form

$$\left.\begin{array}{l} \dfrac{dx_\mu}{dt} = X_\mu\,(x_1, x_2, \ldots x_n) \\ (\mu = 1, 2, \ldots n)\ , \end{array}\right\} \tag{1}$$

where none of the functions X_μ contains the corresponding variable x_μ itself so that the following relations obtains:

$$\frac{\partial X_1}{\partial x_1} + \frac{\partial X_2}{\partial x_2} + \cdots + \frac{\partial X_n}{\partial x_n} = 0\ . \tag{2}$$

In such a system (1) consisting of differential equations of first order to an arbitrary initial state P_0:

$$x_1 = \xi_1\ , \quad x_2 = \xi_2 \ldots x_n = \xi_n\ , \quad (t = t_0)$$

there corresponds a particular altered state P at time t, which is expressed by means of the integral equation of (1):

$$\left.\begin{array}{l} x_\mu = \varphi_\mu\,(t - t_0, \xi_1, \xi_2, \ldots \xi_n)\ , \\ (\mu = 1, 2, \ldots n)\ , \end{array}\right\} \tag{3}$$

where the ϕ_μ's are single-valued and continuous functions of all their arguments which are determined solely by the functions X_μ, independently of the choice of the starting time t_0. These relations hold equally for both earlier and later instants of time, i.e., for both negative and positive values of $t - t_0$; the initial state P_0 is any arbitrarily specified phase of motion, which need not be the temporal predecessor. Likewise, to a continuously ex-

[2] *Poincaré 1889*; *Poincaré 1893b*, p. 294–296.

gehen braucht. Ebenso entspricht auch einem stetig ausgedehnten *Gebiet* g_0 von Anfangszuständen, ausdrückbar durch Beziehungen der Form:

$$F(\xi_1, \ldots \xi_n) < 0 \,,$$

ein bestimmtes verändertes Gebiet $g = g_t$ zur Zeit t und somit auch dem über g_0 erstreckten $n$fachen Integrale

$$\gamma_0 = \int d\xi_1 d\xi_2 \ldots d\xi_n \,,$$

das wir als die „*Ausdehnung*“ von g_0 bezeichnen wollen, im allgemeinen eine andere Ausdehnung von g

$$\gamma = \int dx_1 dx_2 \ldots dx_n \,.$$

In dem besonderen Falle aber, wo die Functionen X_μ der Bedingung (2) genügen, ist nach dem Satz von *Liouville*[1] das zweite Integral gleich dem ersten und damit von der Zeit unabhängig, wie auch das Gebiet g_0 oder g, deren jedes durch das andere bestimmt ist, gewählt sein möge, sodass man abgekürzt schreiben kann:

$$d\gamma = dx_1 dx_2 \ldots dx_n = d\gamma_0 = \text{const.} \tag{4}$$

„Die Folgezustände, die den Anfangszuständen eines beliebigen Gebietes entsprechen, erfüllen in jedem Augenblick Gebiete von der gleichen Ausdehnung.“

Ein beliebiges Gebiet g_0 von Zuständen geht also mit der Zeit stetig in immer neue Gebiete $g = g_t$, die „Phasen“ seiner Veränderung, über, welche sämmtlich die gleiche Ausdehnung γ besitzen. Alle diese „*späteren*“ Phasen g_t ($t \geqq 0$) bilden zusammen genommen wieder ein stetiges Gebiet G_0, die „*Zukunft*“ von g_0, d. h. den Inbegriff aller Zustände, welche *künftig irgend einmal* in endlicher Zeit aus solchen von g_0 hervorgehen. Dieses Gebiet $G = G_0$ wird ganz im endlichen liegen und eine endliche Ausdehnung $\Gamma \geqq \gamma$ besitzen, wenn wir voraussetzen, dass die Grössen $x_1, x_2, \ldots x_n$ für alle Anfangs-
488 | zustände von g_0 gewisse endliche Grenzen niemals überschreiten. Während sich nun das Gebiet g von g_0 ausgehend mit der Zeit verändert, zugleich mit allen seinen „späteren Phasen“, deren jede immer in die folgende übergeht, so ändert sich auch ihre Gesammtheit G wie jedes andere Gebiet und stellt dabei in jedem Augenblicke t die „Zukunft“ der entsprechenden Phase g_t dar. Nach der Definition der Zukunft erfolgt diese Veränderung in der Weise, dass immer nur frühere Zustände austreten, niemals neue eintreten können: jede Phase von G enthält alle späteren in sich, und die Ausdehnung Γ kann immer *nur abnehmen*. Da aber nach (4) diese Ausdehnung *constant* bleiben muss,

[1] *Jacobi*, Dynamik, p. 93; *Kirchhoff*, Theorie der Wärme, p. 142–144.

tended *area* g_0 of initial states, which may be expressed by relations of the form:

$$F(\xi_1, \ldots \xi_n) < 0 \,,$$

there corresponds a particular altered area $g = g_t$ at time t, and hence also to the n-fold integral extended over g_0

$$\gamma_0 = \int d\xi_1 d\xi_2 \ldots d\xi_n \,,$$

which we shall call the "*extension*" of g_0, there corresponds in general another extension of g

$$\gamma = \int dx_1 dx_2 \ldots dx_n \,.$$

In the special case, however, where the functions X_μ satisfy condition (2), the second integral is equal to the first one, by *Liouville*'s theorem,[3] and hence independent of time, irrespective of how the area g_0, or g, each of which is determined by the other, has been chosen, so that we may write, using abbreviations,

$$d\gamma = dx_1 dx_2 \ldots dx_n = d\gamma_0 = \text{const.} \tag{4}$$

"The succeeding states corresponding to the initial states of an arbitrary area fill areas of the same extension at all instants of time."

An arbitrary area g_0 of states is therefore continuously transformed into ever new areas $g = g_t$, the "phases" of its change, as time passes, all of which have the same extension γ. Taken together, all these "*later*" phases g_t $(t \geq 0)$ in turn form a continuous area G_0, the "*future*" of g_0, i.e., the epitome of all states arising from ⟦states⟧ of g_0 *at some arbitrary point in the future* over a finite period of time. This area $G = G_0$ will be completely finite and will have a finite extension $\Gamma \geq \gamma$, if we assume that, for all initial states of g_0, the quantities $x_1, x_2, \ldots x_n$ never exceed certain finite boundaries. As the area g starting with g_0 changes with time, along with all its "later phases", each of which is transformed into its respective successor, so does their totality G, just like any other area, and, at every instant t, it represents the "future" of the corresponding phase g_t. According to the definition of future, this change is due always to the disappearance of earlier states but never to the appearance of new states: each phase of G contains in itself all later ones, and the extension Γ may *decrease only*. But since, by (4), this extension

[3] *Jacobi 1866*, p. 98; *Kirchhoff 1894*, p. 142–144.

so können die austretenden Zustände niemals Gebiete von endlicher Ausdehnung erfüllen, ihre Anzahl verschwindet gegen die der bleibenden, sodass wir sie als *singuläre* bezeichnen können. Nun ist g_0 in G_0 enthalten, also zum überwiegenden Theil auch in jeder folgenden Phase G_τ, der Zukunft von g_τ, für ein beliebig grosses Zeitintervall τ. Das bedeutet aber: es gibt immer Zustände innerhalb g_τ, die später einmal in Zustände von g_0 übergehen, und, ihnen rückwärts entsprechend, Zustände von g_0, die auch nach Ablauf der Zeit τ irgend einmal wieder nach g_0 zurückkehren. Diese letzteren finden sich in allen noch so kleinen Theilen des Gebietes, von denen ja dasselbe wie von g_0 selbst gilt, und hängen stetig zusammen, da mit jedem einzelnen zurückkehrenden Zustand auch seine nächste Umgebung zurückkehren muss, d. h. sie erfüllen das *ganze* Gebiet g_0 mit Ausnahme singulärer Zustände von der Gesammtausdehnung 0. Schliesst man daher alle diese singulären Zustände aus, die zu *irgend welchen* endlichen Zeiten τ gehören, so verbleibt ein Restgebiet g', das nun nicht mehr nothwendig stetig zu sein braucht, aber immer noch die überwiegende Mehrzahl der Zustände von g_0 umfasst. Diese Zustände von g' werden nun nach beliebiger Zeit immer noch einmal, also unendlich oft nach g_0 zurückkehren und damit ihren Anfangszuständen beliebig nahe kommen, wenn man g_0 genügend klein angenommen hat.

Damit ist der Satz von *Poincaré* in voller Ausdehnung bewiesen; für den vorliegenden Zweck genügt aber schon der Nachweis, dass die Zustände von g_0 im Allgemeinen wenigstens noch *einmal* nach g_0 zurückkehren. Schon
489 hieraus folgt | unmittelbar, *dass es keine eindeutige und stetige Function* $S = S(x_1, x_2, \ldots x_n)$ *des Zustandes geben kann, die für alle Anfangszustände eines noch so kleinen Gebietes beständig zunähme.* Denn wäre S für einen Anfangszustand P_0 während der Zeit τ von einem Werthe $< R$ gewachsen auf einen anderen $> R$, so müsste das gleiche gelten von allen Zuständen einer gewissen Umgebung g von P_0, und für die nach g zurückkehrenden Zustände dieses Gebietes müsste die Function nachher wieder abnehmen.

Dasselbe lässt sich aber auch sehr einfach direct beweisen. Würde die Function S für alle Anfangszustände von g beständig zunehmen, so würde sie es auch für alle Zustände des grösseren Gebietes G, der Zukunft von g, und wegen (4) müsste dann auch das über G erstreckte $n$fache Integral

$$\int S dx_1 dx_2 \ldots dx_n$$

beständig zunehmen. Das ist aber unmöglich, weil sich das Integrationsgebiet G immer nur um singulare Zustände ohne endliche Ausdehnung verändert, wobei der Werth des Integrales *constant* bleibt.

Sehr anschaulich wird Bedeutung und Beweis des entwickelten Satzes für den Fall $n = 3$, wenn man die Variablen x_1, x_2, x_3 als die Coordinaten eines materiellen Punktes im Raume auffasst. Dann bestimmen die Gleichungen (1) in Verbindung mit (2) oder mit (4) eine *stationäre Strömung einer incompressiblen Flüssigkeit* und zwar in einem *geschlossenen* Gefässe, wenn die Grössen x_μ nicht ins Unendliche wachsen sollen. Einem be-

must remain *constant*, the disappearing states can never fill areas of finite extension, and their number vanishes with respect to those of the remaining states. Thus, we may refer to them as *singular* states. Now, g_0 is contained in G_0, and hence, for the most part, also in every succeeding phase G_τ, the future of g_τ, for an arbitrarily large interval of time τ. But this means that there are always states within g_τ that, at some later time, are transformed into states of g_0, and, in reverse correspondence to them, states of g_0 that return to g_0 at some point after the expiration of time τ. These latter ones can be found in even the smallest parts of the area, which, after all, are subject to the same conditions as g_0 itself, and continuously hang together since for each returning state its immediate neighborhood must return as well. In other words, they fill the *entire* area with the exception of singular states of a total extension equal to 0. Hence, if we exclude all these singular states belonging to *arbitrary* finite times τ, then what we obtain is a remainder area g' that no longer needs to be continuous but still encompasses the great majority of states of g_0. These states of g' will always return to g_0 one more time after some arbitrary period of time, and hence indefinitely many times, thereby getting arbitrarily close to their initial states, if g_0 is taken sufficiently small.

This concludes the proof of *Poincaré*'s theorem in its full generality; but for the present purposes it already suffices to show that the states of g_0 generally return to g_0 at least *one more time*. From this alone it immediately follows *that there can be no single-valued and continuous function* $S = S(x_1, x_2, \ldots x_n)$ *of the state that steadily increases for all initial states of even the smallest area.* For if S had increased from a value $< R$ to a value $> R$ for an initial state P_0 over a period of time τ, then the same would have to hold true for all states of a certain neighborhood g of P_0, and, later on, the function would have to decrease again for the states of this area that return to g.

The same can, however, easily be shown by a direct proof. If the function S would steadily increase for all initial states g, then it would do so for all states of the greater area G, the future of g, and, on account of (4), the n-fold integral extended over G

$$\int S dx_1 dx_2 \ldots dx_n$$

would *steadily increase* as well. This, however, is impossible since the domain of integration G always changes only by singular states lacking a finite extension, where the value of the integral remains *constant*.

The significance and proof of the developed theorem becomes very clear for the case $n = 3$ when we conceive of the variables x_1, x_2, x_3 as the coordinates of material points in space. In combination with (2) or (4), equations (1) then determine a *stationary flow of an incompressible fluid*, and in particular one held in a *closed* container, if the quantities x_μ are supposed not to grow infinitely large. In this case, to a particular "state" there corresponds

stimmten „Zustand“ entspricht hier ein Punkt im Raum, einem in der Zeit veränderten Zustande ein in Bewegung begriffener materieller Punkt. Die von diesen Flüssigkeits-Punkten beschriebenen Bahnen, die „Stromlinien“, bilden in stetiger Zusammensetzung „Stromröhren“ oder „Stromfäden“, je nachdem sie von geschlossenen Curven oder von Flächenstücken ausgehen, und bleiben bei der stationären Bewegung immer unverändert. Nun lehrt die Anschauung, dass hier alle Stromfäden *in sich selbst zurücklaufen* müssen, weil die durchströmende Flüssigkeit weder die Röhren durchbrechen, noch sich im Inneren irgendwo ansammeln kann. Daraus folgt aber, dass
490 jedes endliche Flüssigkeitstheilchen einem einmal ange- | nommenen Orte immer wieder so nahe kommen muss, als man will, wenn man nur die Flüssigkeitsfäden dünn genug annimmt und genügende Zeit zur Verfügung hat. Daneben gibt es freilich auch nicht zurückkehrende singuläre Stromlinien, z. B. solche, die sich umströmten eingeschlossenen festen Körpern oder Hohlräumen zwischen den nach den verschiedenen Seiten ausweichenden übrigen Stromlinien *asymptotisch nähern*; diese vermögen aber niemals Stromfäden von endlicher Dicke zu bilden. Sollte dagegen die Strömung ein *Geschwindigkeitspotential* besitzen, so müsste dasselbe in dem vollständig geschlossenen Gefässe nothwendig *mehrdeutig* sein, während von der Function S in unserer Betrachtung ausdrücklich *Eindeutigkeit* gefordert wurde. Auch in dem allgemeineren Falle $n > 3$ ist es bei der weitgehenden Analogie oft von heuristischem Werth, die gleiche Ausdrucksweise beizubehalten und die Gleichungen (1) und (2) oder (4) als die einer „stationären Strömung einer incompressiblen Flüssigkeit in einem Raume von n Dimensionen“ zu deuten.

Das Ergebniss unserer Betrachtung ist also das folgende:

In einem System beliebig vieler materieller Punkte, deren Beschleunigungen nur von ihrer Lage im Raum abhängen, gibt es keine „irreversiblen“ Vorgänge für alle Anfangszustände, die ein noch so kleines Gebiet von endlicher Ausdehnung erfüllen, falls sowohl die Coordinaten als die Geschwindigkeiten der Punkte endliche Grenzen niemals überschreiten.

Der Satz gilt aber auch allgemeiner, insbesondere für ein *beliebiges* mechanisches System mit den verallgemeinerten Coordinaten q_μ und ihren Bewegungsmomenten p_μ, dessen Bewegungsgleichungen sich in der Hamilton'schen Form schreiben lassen:

$$\frac{dp_\mu}{dt} = \frac{\partial H}{\partial q_\mu} , \quad \frac{dq_\mu}{dt} = -\frac{\partial H}{\partial p_\mu} ,$$

und das wir als ein „conservatives“ bezeichnen können, weil hier alle Kräfte ein Potential besitzen und daher die mechanische Energie erhalten bleibt. In einem solchen System nämlich ist offenbar immer

$$\frac{\partial\, dp_\mu}{\partial\, p_\mu dt} + \frac{\partial\, dq_\mu}{\partial q\, dt} = 0 ,$$

a point in space, and to a state that changes in time a material point in motion. The paths taken by these fluid-points, the "streamlines", form either "stream tubes" or "stream filaments" as they are continuously compounded, depending on whether they issue forth from closed curves or from plane areas, and they always remain unaltered at stationary motion. Intuition tells us that, in this case, all stream filaments must *run back in themselves* since the passing fluid can neither break through the tubes nor accumulate somewhere in the interior. From this, however, it follows that, time and again, every finite particle of the fluid has to get arbitrarily close to any position once occupied by it, provided only that the fluid filaments are taken sufficiently thin and that sufficient time is available. There are of course also non-recurring singular streamlines, such as those *asymptotically approximating* solid bodies embedded in the flow or hollow spaces between the other streamlines, which are diverted in different directions; but those can never form stream filaments of finite thickness. In contrast, if the flow possesses a *velocity potential*, then the latter would necessarily have to be *many-valued* in the completely closed container, while, in our considerations, we expressly demanded that the function S be *single-valued*. In the more general and largely analogous case $n > 3$, too, it is of heuristic value to retain the same terminology and to interpret equations (1) and (2) or (4) as those of a "stationary flow of an incompressible fluid in a space of n dimensions".

Our considerations lead to the following result:

In a system of arbitrarily many material points whose acceleration only depends on their location in space there are no "irreversible" processes for all initial states that fill an arbitrarily small area of finite extension, if both the coordinates and the velocities of the points never exceed finite boundaries.

The theorem is, however, also valid in a more general form, and in particular for an *arbitrary* mechanical system with the generalized coordinates q_μ and their momenta of motion p_μ whose equations of motion can be written in *Hamiltonian* form:

$$\frac{dp_\mu}{dt} = \frac{\partial H}{\partial q_\mu} , \quad \frac{dq_\mu}{dt} = -\frac{\partial H}{\partial p_\mu} ,$$

and which we may call a "conservative" system since, in this case, all forces possess a potential, and hence the mechanical energy is preserved. For in such a system we apparently always have

$$\frac{\partial\, dp_\mu}{\partial\, p_\mu dt} + \frac{\partial\, dq_\mu}{\partial q\, dt} = 0 ,$$

491 | und mit dem Analogon der Beziehung (2) müssen auch alle aus ihr fliessenden Folgerungen ihre Gültigkeit behalten.

Nach der mechanischen Theorie in ihrer gewöhnlichen atomistischen Darstellung wäre nun die ganze Natur als ein System der betrachteten Art aufzufassen: alle Naturvorgänge sind nichts als Bewegungen der Atome oder Molecüle, die entweder selbst als ausdehnungslose Punkte oder als Aggregate solcher Punkte behandelt werden können und ausschliesslich „Centralkräften“, die ein Potential haben, und von den Geschwindigkeiten unabhängig sind, unterliegen. Eben diese Annahme sucht man in der „kinetischen Gastheorie“ durchzuführen, indem man die Molecüle eines „vollkommenen Gases“ als abstossende Centren, als elastische Kugeln oder mit *Boltzmann* als elastische feste Körper anderer Gestalt, jedenfalls aber als „conservative“ Systeme in dem angegebenen Sinne betrachtet, nur dass man sich hier bei der Wirkung zweier Molecüle aufeinander auf „Stosskräfte“ beschränkt, d. h. auf Abstossungen, die erst bei sehr grosser gegenseitiger Annäherung wirksam werden.

Unter diesen Voraussetzungen könnten also auf Grund der vorhergehenden Betrachtungen „irreversible“ Vorgänge für allgemeinere Anfangszustände nur dadurch möglich werden, dass, von einer gleichförmig fortschreitenden Bewegung des Gesammtschwerpunktes natürlich abgesehen, Molecüle sich ins Unendliche zerstreuen oder schliesslich unendlich grosse Geschwindigkeiten gewinnen. Ist aber das erstere durch die besondere Natur des Systems, das wir uns z. B. von einer festen Hülle umgeben denken können, ausgeschlossen, so ist es auch das letztere auf Grund des Princips von der Energie. Denn sonst müsste zur Erreichung einer unendlich grossen lebendigen Kraft erst eine unendlich grosse Arbeit geleistet werden, was nur bei unbegrenzter Annäherung zweier anziehenden Centren eintreten könnte, während wir doch nach unserer Erfahrung bei sehr grosser Annäherung schlechterdings keine anderen als abstossende Kräfte voraussetzen dürfen. Haben wir z. B. ein in ein festes Gefäss mit elastischen und für Wärme undurchdringlichen Wänden eingeschlossenes Gas, so gäbe es zwar im allgemeinen eine unendliche Mannigfaltigkeit von Anfangszuständen der Molecüle, für welche das Gas
492 *bleibenden* | Zustandsänderungen, wie Reibung, Wärmeleitung oder Diffusion entgegenginge. Aber daneben gäbe es noch sehr viel mehr von vornherein ebenso mögliche Anfangszustände, wie man sie schon durch beliebig kleine Verrückungen eines Molecüls aus den früheren erhalten könnte, für welche anstatt solcher irreversiblen Processe alle Zustände sich mit beliebig kleinen Abänderungen in dem oben angegebenen Sinne *periodisch wiederholten*. Das müsste auch gelten, wenn etwa der auf unsere Sinne wirkende *physikalische* Zustand, z. B. die Temperatur, und mit ihm auch der Werth der Entropie, nicht durch den augenblicklichen *Bewegungszustand* definirt wäre, sondern erst durch eine endliche *Folge von Bewegungen*, die aber jedenfalls durch den anfänglichen Bewegungszustand *bestimmt* wäre und mit ihm immer wiederkehren müsste.

Um daher die allgemeine Gültigkeit des zweiten Hauptsatzes festzuhalten, wäre man zu der Annahme genöthigt, dass trotz ihrer geringeren Anzahl

and, given the analogue of the relation (2), all consequences following from it must retain their validity as well.

The common, atomistic account of the mechanical theory suggests that we view nature in its entirety as a system of the sort under consideration: all processes in nature are but motions of atoms or molecules, which themselves can be treated either as points without extension or as aggregates of such points and which are subject only to "central forces" that have a potential and are independent of the velocities. It is this assumption that one seeks to make in the "kinetic theory of gases" by considering the molecules of a "perfect gas" as repulsive centers, elastic spheres, or, with *Boltzmann*, elastic solid bodies of a different form, but always as "conservative" systems in the sense under consideration, except only that the action of any two molecules upon one another is limited to "impact forces", i.e., to repulsions which take effect only at very great mutual approximation.

According to the previous considerations, "irreversible" processes for more general initial states could possibly occur under these conditions only if molecules disperse into the infinite or eventually move at infinitely large velocities, except, of course, for the uniform motion of the total center of mass. But if the first possibility is excluded by the special nature of the system, which can be thought of, e.g., as being enclosed in a solid casing, then so is the latter possibility by virtue of the principle ⟦of the conservation⟧ of energy. For, otherwise, reaching an infinitely great vivid force would require the performance of an infinitely large amount of work, which could only occur in the case of the unlimited approximation of two attracting centers, while our experience teaches us to assume no other than repulsive forces in cases of very great approximation. Consider, e.g., a gas enclosed in a solid container whose walls are elastic and impenetrable to heat. In this case, there would in general be an infinite manifold of initial states of molecules in which the gas would resist *permanent* changes of state such as friction, heat conduction and diffusion. But there would also be many more initial states equally possible from the outset, such as those obtainable from earlier ones through arbitrarily small displacements of a molecule, for which, instead of such irreversible processes, all states would *periodically repeat* themselves in the sense specified above with arbitrarily small variations. This would also have to be true if, e.g., the *physical* state affecting our senses, such as the temperature and, along with it, the value of the entropy, were not defined by the current *state of motion* but only by a *finite sequence of motions*, which, however, would at least be *determined* by the initial state of motion and would always have to recur together with it.

In order to retain the general validity of the second law, we therefore would have to assume that just those initial states leading to irreversible processes

gerade jene zu irreversiblen Vorgängen führenden Anfangszustände in der Natur einmal *verwirklicht* seien, während die anderen, mathematisch betrachtet, wahrscheinlicheren thatsächlich *nicht vorkämen*.

So unwiderleglich eine solche Annahme auch wäre, so wenig entspräche sie unserem Causalitätsbedürfniss und jedenfalls dem Geiste der mechanischen Naturbetrachtung selbst, der uns immer nöthigen wird, alle *denkbaren* mechanischen Anfangszustände, wenigstens innerhalb gewisser Grenzen auch als physikalisch *möglich* vorauszusetzen, zumal solche, die eine überwiegende Mehrheit ausmachen und von wirklich vorkommenden nur beliebig wenig abweichen. Beziehen sich doch, streng genommen, alle unsere Naturgesetze nicht auf *bestimmte* Grössen oder Vorgänge, die sich genau ja niemals beobachten lassen, sondern immer nur auf gewisse Spielräume, Annäherungen und Wahrscheinlichkeiten, während Singularitäten ausschliesslich als Grenzfälle in der Abstraction existiren. Die hier erörterte Annahme stände also einzig da in der Physik, und ich glaube daher nicht, dass sie irgend jemand würde dauernd befriedigen können.

Dass *nicht alle denkbaren* Anfangszustände dem zweiten Hauptsatz ent-
493 sprechen können, geht schon daraus hervor, | dass bei einer *Umkehrung der Geschwindigkeitsrichtungen* aller Molecüle zu einem beliebigen Zeitpunkt sich auch der ganze zeitliche Verlauf eines Vorganges *umkehren* müsste. In der That ist auch dieses Bedenken schon längst gegen die mechanische Ableitung irreversibler Processe geltend gemacht worden und hat noch im Winter 1894/95, angeregt durch eine Aeusserung *Culverwell*'s, zu einer ausgedehnten Discussion dieser Fragen in der *„Nature“* Veranlassung gegeben, ohne indess, wie mir scheint, zu einer befriedigenden Lösung geführt zu haben. Es liess sich eben nicht beweisen, dass der physikalische Zustand eines Gases, auf den es allein ankommt, für gleiche und entgegengesetzte Geschwindigkeiten aller Molecüle immer derselbe sein müsse, in welchem Falle allein hier von einer wirklichen Umkehrung des Vorganges gesprochen werden dürfte, und es blieb ferner noch die Möglichkeit offen, dass wenigstens für ein ausgedehntes *Gebiet* von Anfangszuständen beständige Vermehrung der Entropie stattfinden könne. Beides sind Einwände gegen die angegebene Argumentation, die erst durch die Anwendung des *Poincaré*'schen Satzes beseitigt werden.

Nach alledem bestände also die Nothwendigkeit, entweder dem *Carnot-Clausius*'schen Princip oder aber der mechanischen Grundansicht eine principiell andere Fassung zu geben, sofern man sich immer noch nicht entschliessen kann, die letztere überhaupt endlich aufzugeben. Geringere Abänderungen würden hier, wie mir scheint, kaum zum Ziele führen. Wollte man beispielsweise versuchen, die zwischen den Molecülen oder Atomen wirkenden Kräfte statt allein von ihrer gegenseitigen Lage auch von ihren Geschwindigkeiten abhängig zu machen, womit allerdings die Anwendbarkeit unseres Satzes vermieden würde, so müsste man, um nicht gegen das Princip der Energie zu verstossen, Zusatzkräfte einführen, deren Arbeit beständig verschwindet, deren *Richtung* also durch die Geschwindigkeiten mit bestimmt wird. Dann aber könnten die Kräfte nicht mehr unabhängig voneinander nach Wirkung und

are *realized* in nature, their small number notwithstanding, while the other ones, whose probability of existence is higher, mathematically speaking, *do not* actually *occur*.

While such an assumption would certainly be irrefutable, it would hardly accord with our need for causality or even with the spirit of the mechanical approach to nature itself, which will always compel us, at least within certain limits, to consider all those mechanical initial states as physically *possible* that are *conceivable*, and in particular those that constitute a great majority and deviate from those actually occurring in reality by an arbitrarily small amount only. For, strictly speaking, all of our laws of nature refer not to *specific* magnitudes or processes, which, after all, defy precise observation, but always only to certain margins, approximations and probabilities, while singularities only exist as limiting cases in abstraction. The assumption discussed here would thus be unique in physics, and I therefore doubt that anyone could find lasting satisfaction in it.

That *not every conceivable* initial state is capable of satisfying the second law already follows from the fact that the *reversal of the directions of the velocities* of all molecules at an arbitrary instant of time would necessarily lead to a *reversal* of the entire time evolution of a process. In fact, this concern has already been raised against the mechanical derivation of irreversible processes a long time ago, and, as recently as in the winter of 1894/95, has led to an extensive discussion of these questions in *Nature*, which was instigated by a comment by Culverwell but which, as far as I can see, failed to reach a satisfactory resolution. It was simply not possible to prove that the physical state of a gas, which is all that counts, must always be the same for identical and opposing velocities of all molecules, which is the only case in which it would be legitimate to speak of an actual reversal of the process. Furthermore, the possibility was not excluded that at least for some extended *area* of initial states there could be a constant increase of entropy. Both are objections to the line of argument under discussion and can only be removed by applying *Poincaré*'s theorem.

All of this seems to suggest that it is imperative to provide an altogether different version of either the *Carnot-Clausius* principle or of the basic mechanical approach, unless one decides to entirely abandon the latter at last. I do not believe that minor modifications would be effectual. For instance, if one were to try to make the forces acting between molecules or atoms dependent also on their velocities in addition to their relative locations, which would, however, avoid the applicability of our theorem, then, so as not to violate the principle of ⟦the conservation of⟧ energy, it would be necessary to introduce additional forces whose work constantly vanishes, and hence whose *direction* is also determined by the velocities. But then, the forces could no

Gegenwirkung von Punkt zu Punkt wirken, wie doch der ganzen Atomtheorie wesentlich ist.

Aber mag es auch gelingen, durch geeignete Abänderung der Voraussetz-
494 zungen, z. B. unter Zugrundelegung der *Hertz*'-| schen „Principien der Mechanik“[1], dem dargelegten Widerspruche zu entgehen, so ist es doch jedenfalls *unmöglich*, auf Grund der *bisherigen* Theorie ohne Specialisirung der Anfangszustände eine mechanische Ableitung des zweiten Hauptsatzes durchzuführen, und es ist ebenso unmöglich, unter den gleichen Voraussetzungen das bekannte Gesetz der Geschwindigkeitsvertheilung unter den Gasmolecülen, wie seine Entdecker *Maxwell* und *Boltzmann* wollten, als den nach einiger Zeit sich regelmässig einstellenden stationären Endzustand zu erweisen. Von einer eingehenden Prüfung der bisherigen Versuche einer solchen Ableitung im einzelnen, namentlich der von *Boltzmann* und *Lorentz* (in den Berichten der Wiener Akademie)[2], habe ich bei der Schwierigkeit des Gegenstandes vorläufig Abstand genommen, um lieber mit möglichster Klarheit darzulegen, was mir hier als streng beweisbar und principiell wichtig erscheint, und dadurch zu einer erneuten Erörterung und schliesslichen Lösung der vorliegenden Frage beizutragen.

Berlin, im December 1895.

Entgegnung auf die wärmetheoretischen Betrachtungen des Hrn. E. Zermelo

Boltzmann 1896

Schon *Clausius*, *Maxwell* u. A. haben wiederholt darauf hingewiesen, dass die Lehrsätze der Gastheorie den Charakter statistischer Wahrheiten haben. Ich habe besonders oft und so deutlich als mir möglich war betont[1], dass das *Maxwell*'sche Gesetz der Geschwindigkeitsvertheilung unter Gasmolecü-

[1] Die v. *Helmholtz*'sche Theorie der „cyclischen Systeme“ in ihrer ursprünglichen Form dagegen würde von den Folgerungen des *Poincaré*'schen Satzes mit betroffen werden, da sie in letzter Linie gleichfalls, wenn auch in anderer Form, auf die *Hamilton*'schen Gleichungen zurückgeht.

[2] Neuerdings zusammengestellt in *Boltzmann*'s „Vorlesungen über Gastheorie“ 1. 1896.

[1] *L. Boltzmann,* Wien. Sitzungsber. II **75.** p. 67. 1877; **76.** p. 373. 1877; **78.** p. 740. 1878. „Der zweite Hauptsatz der Wärmetheorie“, Rede gehalten am 29. Mai 1886, Almanach d. Wien. Akad. Nature **51.** p. 413, 28. Febr. 1895; Vorlesung über Gastheorie p. 42. 1895, Leipzig bei J. A. Barth.

longer act independently of one another in action and reaction from point to point, as is essential to the entire atomistic theory.

But even if it is possible to escape the contradiction considered here by making suitable changes to the underlying assumptions, e.g., by building on *H. Hertz 1894*,[4] it is certainly *impossible* to carry out a mechanical derivation of the second law on the basis of the *existing* theory without specializing the initial states, and it is equally impossible to show, under the same assumptions, the well-known law of velocity distribution among gas molecules to be the stationary final state regularly reached after some time, as its discoverers, *Maxwell* and *Boltzmann* intended. Faced with the difficulty of the subject matter, I have for now refrained from a thorough review of the previous attempts at such a derivation, in particular the one undertaken by *Boltzmann* and *Lorentz* (published in the *Berichte der Wiener Akademie*).[5] Instead, I chose to present here as clearly as possible what I consider strictly provable and essentially important in order to contribute to a renewed discussion of the question at hand and to its eventual resolution.

Berlin, December 1895.

Rejoinder to the heat-theoretic considerations of Mr. E. Zermelo

Boltzmann 1896

⟦The introductory note just before *1896a* also addresses *Boltzmann 1896*.⟧

As *Clausius* and *Maxwell*, among others, have already repeatedly pointed out, the principles of the theory of gases have the character of statistical truths. I have pointed out particularly often, and as clearly as I possibly could,[1] that *Maxwell*'s law of velocity distribution among gas molecules cer-

[4] By contrast, the original version of *v. Helmholtz*'s theory of "cyclical systems" would also be affected by the consequences of *Poincaré*'s theorem since it, too, albeit in a different form, is, in the end, a descendant of the Hamiltonian equations.

[5] Recently assembled in *Boltzmann 1896a*.

[1] *Boltzmann 1877a*, p. 67; *Boltzmann 1877b*, p. 373; *Boltzmann 1878*, p. 740; "Der zweite Hauptsatz der Wärmetheorie", lecture delivered on May 29, 1886, *Boltzmann 1886*, *Boltzmann 1895b*, p. 413; *Boltzmann 1896a*, p. 42.

len keineswegs wie ein Lehrsatz der gewöhnlichen Mechanik aus den Bewegungsgleichungen allein bewiesen werden kann, dass man vielmehr nur beweisen kann, dass dasselbe weitaus die grösste Wahrscheinlichkeit hat und bei einer grossen Anzahl von Molecülen alle übrigen Zustände damit verglichen so unwahrscheinlich sind, dass sie praktisch nicht in Betracht kommen. An derselben Stelle habe ich auch betont, dass der zweite Hauptsatz vom moleculartheoretischen Standpunkte ein blosser Wahrscheinlichkeitssatz ist. Die Abhandlung des Hrn. *Zermelo* „Ueber einen Satz der Dynamik und die mechanische Wärmetheorie“[2] zeigt nun zwar, dass meine betreffenden Arbeiten trotzdem nicht verstanden worden sind, demungeachtet muss ich mich über diese Abhandlung freuen, als über den ersten Beweis, dass diesen Arbeiten in Deutschland überhaupt Aufmerksamkeit geschenkt wird.

Der von Hrn. Zermelo zu Anfang auseinander gesetzte Satz *Poincaré*'s ist selbstverständlich richtig, aber dessen Anwendung auf die Wärmetheorie ist es nicht.

Ich habe den Beweis des *Maxwell*'schen Geschwindigkeitsvertheilungsgesetzes aus dem Satze abgeleitet, dass nach den Wahrscheinlichkeitsgesetzen eine gewisse Grösse H (gewissermaassen das Maass des Grades der Abweichung des herrschenden Zustandes vom *Maxwell*'schen) für ein in einem
774 | ruhenden Gefässe ruhendes Gas nur abnehmen kann. Die Art und Weise dieser Abnahme wird am besten klar, wenn man sich, wie ich es[1] that, die Zeit als Abscisse und die dazu gehörigen Werthe der Grösse H, vermindert um deren kleinsten Werth $H_{\min}$, als Ordinate aufgetragen denkt, wodurch man die sogenannte H-Curve erhält.

Setzt man, wie es bei dem in meiner Gastheorie § 5 auseinandergesetzten Beweise ausdrücklich geschieht, zuerst die Anzahl der Gasmolecüle unendlich und lässt erst nachher die Zeit der Bewegung sehr gross werden, so erhält man in der weitaus überwiegenden Mehrheit der Fälle eine Curve, welche sich asymptotisch der Abscissenaxe nähert. Dann ist auch, wie man leicht sieht, der *Poincaré*'sche Satz nicht anwendbar.

Nimmt man aber die Zeit der Bewegung unendlich gross, dagegen die Anzahl der Molecüle zwar sehr, aber nicht absolut unendlich gross an, so hat die H-Curve einen anderen Charakter. Sie verläuft, wie ich schon am citirten Orte in der Nature zeigte, fast immer sehr nahe der Abscissenaxe. Nur äusserst selten erhebt sie sich höher über dieselbe, was wir einen Buckel nennen wollen, und zwar nimmt die Wahrscheinlichkeit eines Buckels mit wachsender Höhe desselben rapid ab. Für jede Zeit, für welche die Ordinate der H-Curve sehr klein ist, herrscht fast genau die *Maxwell*'sche Geschwindigkeitsvertheilung; bedeutende Abweichung von derselben aber finden an den hohen Buckeln der H-Curve statt. Hr. *Zermelo* glaubt nun aus dem *Poincaré*'schen Satze schliessen zu können, dass sich das Gas nur bei gewissen singulären Anfangszuständen, deren Anzahl unendlich klein ist gegen die aller möglichen An-

[2] *Zermelo*, Wied. Ann. **57.** p. 485. 1896.

[1] *L. Boltzmann,* Nature l. c.

tainly cannot be proved from the equations of motion alone like a principle of ordinary mechanics and that, in fact, it is only possible to prove that it has by far the highest probability and that, given a large number of molecules, all other states are so improbable in comparison that they are practically irrelevant. At the same time, I also stressed that the second law is but a principle of probability theory as far as the molecular-theoretic point of view is concerned. Although *Zermelo*'s essay "Ueber einen Satz der Dynamik und die mechanische Wärmetheorie"[2] shows that my works on these issues are still not understood, I cannot but take delight in it for this essay constitutes the first proof that these works of mine receive any attention at all in Germany.

While the theorem by *Poincaré* that *Zermelo* discusses in the beginning of his paper is of course correct, its application to heat theory is not.

I derived the proof of *Maxwell*'s law of velocity distribution from the theorem which states that, according to the laws of probability theory, a certain magnitude H (the measure of the degree of deviation of the prevailing state from the *Maxwell* state, as it were) can only decrease for a gas at rest in a container at rest. The way in which this decrease occurs can best be understood by plotting time along the abscissa axis, as I did,[3] and the corresponding values of the magnitude H, decreased by their least value $H_{\min}$, along the ordinate axis, which yields the so-called H-curve.

If we first take the number of gas molecules to be infinite, as was clearly done in the proof discussed in §5 of my *1896a*, and only then let the time of the motion grow very large, then, in the vast majority of cases, we obtain a curve asymptotically approximating the abscissa axis. Moreover, as can easily be seen, *Poincaré*'s theorem is not applicable in this case.

If, however, we take the time of the motion to be infinitely great and, in contrast, the number of molecules to be very great but not absolutely infinite, then the H-curve has a different character. As I have already shown in the contribution to Nature quoted above, it almost always runs very close to the abscissa axis. It is only in the rarest cases that it rises higher above it, which we shall call a "hump". In particular, the probability of a hump rapidly decreases as its height increases. For every moment in time for which the ordinate of the H-curve is very small, the *Maxwell* velocity distribution obtains almost exactly; significant deviations from it occur, however, at high humps of the H-curves. Now, Mr. *Zermelo* believes that he can conclude from *Poincaré*'s theorem that it is only for certain singular initial states whose number is infinitesimal compared to that of all possible initial states that

[2] *Zermelo 1896a.*
[3] *Boltzmann 1895b*, l. c.

fangszustände, dem *Maxwell*'schen Geschwindigkeitsvertheilungsgesetze immer mehr nähert, während bei den meisten Anfangszuständen dieses Gesetz nicht Platz greift. Dies scheint mir nicht richtig zu sein. Gerade für gewisse singuläre Anfangszustände tritt das *Maxwell*'sche Geschwindigkeitsvertheilungsgesetz niemals ein, z. B. wenn alle Molecüle anfangs in einer an beiden Enden auf der Gefässwand senkrechten Geraden lagen. In der weitaus (unendlich) überwiegenden Mehrzahl von Anfangsbedingungen dagegen hat die H-Curve den soeben geschilderten Charakter.

775 | Liegt der Anfangszustand des Gases auf einem enorm hohen Buckel, d. h. weicht er gänzlich von der *Maxwell*'schen Geschwindigkeitsvertheilung ab, so wird sich der Zustand mit enormer Wahrscheinlichkeit dieser Geschwindigkeitsvertheilung nähern und während enorm langer Zeit nur verschwindend wenig davon abweichen. Allerdings kann man, wenn die Zeit der Bewegung noch mehr verlängert wird, wieder zu einem grösseren Buckel der H-Curve gelangen, ja, wenn diese Verlängerung nur genügend fortgesetzt wird (also selbstverständlich für in mathematischem Sinn unendlich lange Bewegungsdauer unendlich oft), muss sogar der alte Zustand wiederkehren.

Hr. *Zermelo* hat daher vollständig recht, wenn er behauptet, dass die Bewegung im mathematischen Sinne eine periodische ist, aber, weit entfernt meine Sätze zu widerlegen, ist diese Periodicität vielmehr in vollster Harmonie mit denselben. Man vergesse nicht, dass die *Maxwell*'sche Geschwindigkeitsvertheilung kein Zustand ist, wobei jedem Molecül ein bestimmter Ort und eine bestimmte Geschwindigkeit angewiesen wird und welcher etwa dadurch erreicht wird, dass sich der Ort und die Geschwindigkeit jedes Molecüls diesem bestimmten Orte und dieser bestimmten Geschwindigkeit assymptotisch nähern. Unter einer endlichen Zahl von Molecülen kann überhaupt niemals exact, sondern nur mit grosser Annäherung die *Maxwell*'sche Geschwindigkeitsvertheilung bestehen. Diese ist keineswegs eine ausgezeichnete singuläre Geschwindigkeitsvertheilung, welcher unendlich vielmal mehr Nicht-*Maxwell*'sche Geschwindigkeitsvertheilungen gegenüber stehen; sondern sie ist im Gegentheile dadurch charakterisirt, dass die weitaus grösste Zahl der überhaupt möglichen Geschwindigkeitsvertheilungen die charakteristischen Eigenschaften der *Maxwell*'schen haben und gegenüber dieser Zahl die Anzahl derjenigen möglichen Geschwindigkeitsvertheilungen, welche bedeutend von der *Maxwell*'schen abweichen, verschwindend klein ist. Während daher Hr. *Zermelo* sagt, die Anzahl derjenigen Zustände, welche schliesslich zum *Maxwell*'schen führen, sei verschwindend gegenüber der aller möglichen Zustände, so behaupte ich dagegen, dass überhaupt die weitaus grösste Zahl der gleich möglichen Zustände „*Maxwell*'sche" sind und dagegen die Zahl der we-
776 sentlich von der *Maxwell*'schen | Geschwindigkeitsvertheilung abweichenden nur verschwindend klein ist.[1]

[1] Ueber das, was hierbei unter gleich möglichen Zuständen zu verstehen ist, vgl. meine eingangs citirten Abhandlungen.

the gas comes ever closer to satisfying *Maxwell*'s law of velocity distribution, while this law does not apply to most initial states. This does not seem right to me. It is particularly for certain singular initial states that *Maxwell*'s law of velocity distribution never holds, such as when all molecules were at first lying in a straight line perpendicular at both ends to the wall of the container. By contrast, for the vast (infinite) majority of initial conditions, the H-curve has the character just described.

If the initial state of the gas lies on an enormously high hump, i.e., if it entirely deviates from the *Maxwell* velocity distribution, then the state will approximate this velocity distribution with enormously high probability and deviate from it only by a vanishingly small amount over an enormously long period of time. It is, however, possible to reach again a greater hump of the H-curve by further extending the time of motion. In fact, it is even the case that the original state must return, provided only that we continue to sufficiently extend the time of motion (that is, of course, infinitely many times for a duration of motion that is infinitely long in the mathematical sense of the word).

Mr. *Zermelo* is therefore right in claiming that, mathematically speaking, the motion is periodic. He has by no means succeeded, however, in refuting my theorems, which, in fact, are entirely consistent with this periodicity. One should not forget that *Maxwell*'s velocity distribution is not a state where a specific location and a specific velocity are assigned to each molecule and which is reached when, say, the location and velocity of each molecule asymptotically approximate this specific location and velocity. Generally, among a finite number of molecules, the *Maxwell* velocity distribution can never obtain exactly but only with great approximation. It is by no means a distinguished, singular velocity distribution, which is pitted against infinitely many more non-*Maxwell* velocity distributions. Quite the contrary. It is characterized by the fact that the by far greatest number of all possible velocity distributions have the characteristic properties of the *Maxwell* distribution and that, in comparison, the number of possible velocity distributions significantly deviating from the *Maxwell* distribution is vanishingly small. Thus, contrary to Mr. *Zermelo*'s assertion that the number of states eventually leading to the *Maxwell* distribution vanishes compared to that of all possible states, I hold that the by far greatest number of equally possible states are "*Maxwell*" and that, by contrast, the number of states essentially deviating from the *Maxwell* velocity distribution is only vanishingly small.[4]

[4] As for what I mean by equally possible states, see the works referred to at the beginning.

Für das erste Molecül ist jeder Ort im Raume und für Geschwindigkeitscomponente dessen erste jede mit dem Energieprincipe verträgliche Grösse gleich wahrscheinlich.

Combinirt man aber alle Zustände aller Molecüle, so erhält man in den weitaus meisten Fällen mit grosser Annäherung das *Maxwell*'sche Geschwindigkeitsvertheilungsgesetz. Nur ganz wenige Combinationen geben eine total davon abweichende Zustandsvertheilung.

Ein Analogon hierfür bietet die Theorie der Methode der kleinsten Quadrate, wo für jeden Elementarfehler ein positiver oder ein gleicher negativer Werth als gleich wahrscheinlich angenommen und dann bewiesen wird, dass wenn man alle möglichen Werthe der Elementarfehler in allen möglichen Weisen combinirt, bei der grössten Mehrzahl der Combinationen das *Gauss*'sche Fehlergesetz herauskommt und nur bei verhältnissmässig verschwindend wenigen Combinationen bedeutende Abweichungen davon eintreten, welche also nicht unmöglich, aber unendlich unwahrscheinlich sind.

Ein noch einfacheres Beispiel bietet das Würfelspiel. Bei 6000 Würfen mit demselben Würfel wird man beiläufig 1000 Einser-, 1000 Zweierwürfe etc. machen; aber keineswegs deshalb, weil die gerade zufällig eingetretene Reihenfolge der Würfe wahrscheinlicher wäre, als eine Reihe von 6000 Einserwürfen, sondern bloss, weil weit mehr mögliche Combinationen auf eine nahe gleiche Zahl von Einserwürfen, Zweierwürfen etc., als auf lauter Einserwürfe führen.

Die Wahrscheinlichkeitsrechnung führt daher, wie längst bekannt, ebenfalls zu dem Resultate, dass eine Wiederkehr des ursprünglichen Zustandes durchaus nicht mathematisch ausgeschlossen ist, ja dass dieselbe sogar zu erwarten ist, wenn die Zeit der Bewegung genügend lange ausgedehnt wird, da die Wahrscheinlichkeit eines dem Anfangszustande sehr nahe liegenden Zustandes sehr klein, aber nicht unendlich klein ist. Die Consequenz des *Poin-*
777 *caré*'schen Satzes, dass | abgesehen von wenigen singulären Zustandsvertheilungen ein dem Anfangszustande sehr naher Zustand nach einer, wenn auch sehr langen Zeit immer wiederkehren muss, steht daher in vollstem Einklange mit meinen Lehrsätzen.

Nun der Schluss, dass an den mechanischen Grundanschauungen irgend etwas geändert oder diese gar aufgegeben werden müssten, darf daraus nicht gezogen werden. Dieser Schluss wäre nur berechtigt, wenn sich aus den mechanischen Grundanschauungen irgend eine mit der Erfahrung in Widerspruch stehende Consequenz ergäbe. Dies wäre aber nur der Fall, wenn Hr. *Zermelo* beweisen könnte, dass die Zeitdauer dieser Periode, innerhalb welcher der alte Zustand des Gases nach dem *Poincaré*'schen Satze eintreten muss, eine beobachtbare Länge hat. Es dürfte nun zwar schon a priori evident sein, dass, wenn etwa eine Trillion winziger Kugeln, jede mit einer grossen Geschwindigkeit begabt, zu Anfang der Zeit in einer Ecke eines Gefässes mit absolut elastischen Wänden beisammen waren, sich dieselben in kurzer Zeit ziemlich gleichmässig im Gefässe vertheilen werden, und dass die Zeit, wo sich alle ihre Stösse so compensirt haben, dass sie alle wieder in derselben Ecke zu-

Every location in space is equally probable for the first molecule, and so is every magnitude compatible with the principle of energy conservation for its first velocity component.

However, by combining all states of all molecules we obtain *Maxwell*'s law of velocity distribution in by far the greatest number of cases with great approximation. Only very few combinations yield a distribution of states entirely deviating from it.

The theory of the method of least squares provides the following analogue, where for every basic error some positive or identical negative value is taken to be equally probable, and it is then proved that by combining all possible values of the basic errors in every way possible we obtain the *Gauss*ian law of error for the greatest majority of combinations and that we find significant deviations from it only for a few combinations, whose number is vanishingly small by comparison and which are hence not impossible but infinitely improbable.

An even simpler example is provided by the game of dice. If you roll the same die 6000 times, then you will randomly get 1000 ones, 1000 twos, etc.; but this is not because the sequence of rolls occurring incidentally is more probable than a sequence of 6000 ones, but only because a far greater number of possible combinations lead to an approximately equal number of ones, twos, etc., rather than only to ones.

Hence, as has long been known, the calculus of probabilities also leads to the result that, from a mathematical point of view, a return to the original state is certainly not excluded and that it is even to be expected, assuming that the time of motion is sufficiently extended, since the probability of a state lying very close to the initial state is very small but not infinitely small. The consequence of *Poincaré*'s theorem, which states that, apart from a few singular state distributions, a state very close to the initial state must always return after some, albeit very long, time has elapsed, is therefore fully consistent with my principles.

Now, from this we must not conclude that the mechanical approach has to be modified in any way or even abandoned. This conclusion would be justified only if the mechanical approach had a consequence that runs contrary to experience. But this would be the case only if Mr. *Zermelo* were able to prove that the duration of this period within which the old state of the gas must occur in accordance with *Poincaré*'s theorem has an observable length. It should be a priori evident by now that if, say, a quintillion tiny spheres each of which is endowed with great velocity were assembled in one corner of the container with elastic walls at the beginning, they would be distributed fairly equally in the container after a short time, and that the period of time within which all of its collisions offset one another so that they all return

sammenkommen, so gross sein muss, dass sie niemand zu erleben im Stande ist. Zum Ueberflusse ergiebt die im Anhange beigefügte Rechnung für diese Zeit einen Betrag, dessen enorme Grösse wahrhaft beruhigend ist. So wenig nun die im Anhange gegebene Rechnung irgend einen Anspruch auf Genauigkeit machen kann, so zeigt dieselbe doch, dass aus dem *Poincaré*'schen Satze jedenfalls nicht bewiesen werden kann, dass die theoretische Existenz einer Periode, nach welcher derselbe Zustand des Gases wiederkehrt, irgend einen Widerspruch mit der Erfahrung involvire, da die Länge dieser Periode jeder Beobachtbarkeit spottet. Die Zustände, die wir beobachten, aber fallen ja alle in die Zwischenzeit zwischen den Anfang und das Ende der Periode, wo der *Poincaré*'sche Satz Zustände, die sich im beliebigen Grade den *Maxwell*'schen nähern, nicht ausschliesst.

Der *Zermelo*'sche Fall ist daher nur einer jener vielen Fälle (und zwar ein gegen die Gastheorie besonders wenig beweisender), wo ein theoretisch nur
778 sehr unwahrscheinlicher | Zustand praktisch als niemals eintretend betrachtet werden muss. So müssen z. B. selbst bei gewöhnlicher Temperatur im Knallgase einzelne Molecüle mit grosser Geschwindigkeit zu Zweien und selbst zu Dreien aufeinander stossen. Dasselbe muss sich also auch bei gewöhnlicher Temperatur in Wasser verwandeln.

Um ein anderes Beispiel zu geben, ist der Fall, dass in einem Gase während einer Secunde kein Molecül auf einen Stempel von bestimmter Grösse stösst, nur sehr unwahrscheinlich, nicht unmöglich.

Die Zeit, wie lange man warten müsste, bis im Knallgase bei gewöhnlicher Temperatur eine messbare Wassermenge entsteht oder bis ein nicht allzu kleiner Stempel während einer Secunde einen messbar kleinern Druck als den durchschnittlichen Gasdruck erfährt, sind bei weitem nicht so lange, als die *Zermelo*'sche Periode, aber doch ausreichend lang, um jede Beobachtbarkeit auszuschliessen. Ein Argument gegen die Gastheorie könnte aus solchen Betrachtungen nur dann abgeleitet werden, wenn derartige Erscheinungen in Fällen ausblieben, wo sie nach der Rechnung in beobachtbaren Zeiten eintreten müssten. Dies scheint aber nicht der Fall zu sein, im Gegentheil: bei einer Temperatur, die tiefer als die allgemeine Umsetzungstemperatur ist, wurden wirklich Spuren chemischer Umsetzungen gefunden; ebenso wurden an ganz kleinen, in einem Gase befindlichen Körperchen Bewegungen wahrgegenommen, welche davon herrühren können, dass in solchen Fällen in der That auf einem gegen ihre ganze Oberfläche nicht mehr verschwindenden Theil derselben bald ein etwas grösserer, bald ein etwas kleinerer Druck wirkt.

Wenn daher Hr. *Zermelo* aus der theoretischen Notwendigkeit, dass in einem Gase der Anfangszustand wiederkehren muss, ohne zu berechnen, nach wie langer Zeit dies geschehen muss, den Schluss zieht, dass die Hypothesen der Gastheorie verlassen oder im Fundamente verändert werden müssen, so gleicht er einem Würfelspieler, welcher berechnet hat, dass die Wahrscheinlichkeit 1000mal hintereinander ein Auge zu werfen nicht gleich Null ist und nun schliesst, dass seine Würfel falsch sein müssen, weil ihm dieser Fall noch nie vorgekommen ist.

to the same corner must be so great that nobody can live to see it happen. Moreover, the length of this time period, which is obtained by calculations in the appendix, is truly staggering in its enormity. With no pretense to exactitude, the appended calculations show that we cannot conclude from *Poincaré*'s theorem that the theoretical existence of a period after which the same state of the gas returns involves any contradiction with experience since the length of this period defies any observation. But the states we observe all fall in the interim between the beginning and the end of the period, in which *Poincaré*'s theorem does not exclude states arbitrarily closely approximating the *Maxwell* states.

Zermelo's case is therefore but one among a multitude of those cases (and one, at that, which does particularly little to refute the theory of gases) in which a state that is, theoretically speaking, highly improbable must, for all practical purposes, be taken never to occur at all. Thus, for instance, even at ordinary temperatures individual molecules in an oxyhydrogen gas must collide at great velocity in groups of two and even three. The gas must therefore turn into water even at ordinary temperatures.

As another example, consider that the case where not a single gas molecule strikes a piston of a certain size within one second is only highly improbable but not impossible.

The time it takes for a measurable amount of water to form in an oxyhydrogen gas at ordinary temperatures or for a not excessively small piston to encounter a pressure measurably smaller than the average gas pressure within one second is by far not as long as the *Zermelo* period, yet long enough to exclude any possibility of observation. From such considerations we could derive an argument against the theory of gases only if phenomena of this sort failed to occur in cases where, according to the calculations, they would have to occur within measurable periods of time. But this does not seem to be the case here. On the contrary. Traces of chemical transformations were actually found at a temperature lower than the general transformation temperature; likewise, observations were made of movements of very small corpuscles, which may be due to the fact that in such cases a pressure, which is sometimes a little greater, sometimes a little smaller, really acts on a part of their surface that no longer vanishes compared to their entire surface.

Hence, when Mr. *Zermelo* cites the theoretical necessity that, in a gas, the initial state must return, without calculating the time it takes for this to occur, in order to draw the conclusion that we must either abandon the hypotheses of the theory of gases or fundamentally change them, he is like a dice player who has calculated that the probability of rolling a one 1000 times in a row does not equal zero and now concludes that his dices are loaded on the ground that he has not yet come across this case.

779 | Mit dem Vorgebrachten hängt nach meinen Ausführungen an den Eingangs citirten Stellen der 2. Hauptsatz aufs Innigste zusammen. Auch er ist nach den molecular-theoretischen Anschauungen lediglich ein Wahrscheinlichkeitssatz. Nach diesen Anschauungen kann nicht aus den Bewegungsgleichungen bewiesen werden, dass sich alle Erscheinungen immer in einem bestimmten Sinne abspielen müssten. Bei allen Erscheinungen, wo nur sichtbare Bewegungen vorkommen, wo sich also die Körper bloss als Ganzes bewegen, muss jeder Bewegungssinn gleichberechtigt sein. Wo dagegen die Bewegung auf eine sehr grosse Anzahl sehr kleiner Molecüle übergeht, dürfen wir, abgesehen von verschwindend wenigen Fällen, die um so weniger zur Beobachtung gelangen können, je mehr Molecüle in's Spiel kommen, den Uebergang von einem unwahrscheinlichen zu einem wahrscheinlicheren Zustande, also immerwährende Veränderungen in einem bestimmten Sinne erwarten, wie in einem Gase den Eintritt der *Maxwell*'schen Zustandsvertheilung. Wenn dagegen die Bewegungen einzelner Molecüle in Frage kämen, wäre dies nicht mehr zu erwarten.

Der erste und zweite Fall bestätigen sich in der Erfahrung: der dritte Fall wurde noch niemals realisirt. Seine Möglichkeit ist daher nicht bewiesen, aber auch nicht widerlegt. Namhafte Forscher, z. B. *Helmholtz*[1], glaubten an dieselbe und wie ich in meinem Buche über Gastheorie nachzuweisen suchte[2], wird die Ansicht, dass der zweite Hauptsatz ein blosser Wahrscheinlichkeitssatz sei, durch die Thatsachen nicht nur nicht widerlegt, sondern dieselben schliessen sich dieser Ansicht sogar besonders gut an. Auch *Gibbs*[3] gelangt aus rein empirischen Thatsachen zu folgendem Schlusse: „The impossibility of an incompensated decrease of entropy seems to be reduced to an improbability."

Wir kommen also zu folgendem Resultate: Wenn man die Wärme als eine Bewegung von Molecülen auffasst, welche gemäss den allgemeinen Gleichungen der Mechanik stattfindet | und annimmt, dass sich der Complex von 780 Körpern, den wir wahrnehmen, jetzt gerade in einem sehr unwahrscheinlichen Zustande befindet, so ergiebt sich ein Satz, welcher für alle bisher beobachteten Erscheinungen mit dem zweiten Hauptsatze übereinstimmt.

Freilich sobald man Körper von so kleinen Dimensionen beobachtet, dass dieselben nur mehr wenige Molecüle enthalten, muss die Gültigkeit dieses Satzes aufhören. Da aber über das Verhalten so kleiner Körper keinerlei Versuche vorliegen, so widerspricht diese Annahme keiner bisherigen Erfahrung; ja, gewisse mit sehr kleinen, in Gasen befindlichen Körpern vorgenommene Versuche sprechen eher zu ihren Gunsten, wenn man auch noch weit davon entfernt ist, von einem experimentellen Beweise ihre⟦r⟧ Richtigkeit sprechen zu können.

[1] Berl. Ber. **17.** p. 172. 1884; ebend. p. 34. Febr. 1882.
[2] l. c. p. 61.
[3] *Gibbs*, Conn. acad. trans. **3.** p. 229. 1875; *Ostwald*'s deutsche Ausgabe p. 198.

According to considerations of mine to which I referred at the beginning, the foregoing remarks are intimately connected with the second law. On the molecular-theoretic approach, it, too, is but a principle of probability theory. According to this approach, it is impossible to prove on the basis of the equations of motion that all phenomena must always unfold in one direction. For all phenomena that only involve visible motions, and hence where the bodies only move as a whole, the directions of motion must all be equivalent. In contrast, when the motion is transferred to a very large number of very small molecules, we may expect the transition from an improbable state to a more probable state to take place, and hence perennial changes in a certain direction, such as the onset of the *Maxwell* state distribution in a gas, with the exception of vanishingly few cases, which become ever less observable as the number of molecules involved grows. On the other hand, this would no longer have to be expected if the motions of individual molecules were relevant.

The first and the second case are confirmed by experience: the third case has not yet been realized. Hence, its possibility has not been proved. However, it has not been refuted either. Notable scientists such as *Helmholtz*[5] believed in it. Moreover, not only do the facts fail to refute the view that the second law is but a principle of probability theory, but they even line up with it particularly well, as I have tried to show in my book on the theory of gases.[6] *Gibbs*,[7] too, infers from purely empirical facts the following conclusion: "The impossibility of an incompensated decrease of entropy seems to be reduced to an improbability."

We thus reach the following result: If we consider heat as a motion of molecules that occurs in accordance with the general equations of mechanics and assume that the complex of bodies that we perceive is currently in a highly improbable state, then a theorem follows that is in agreement with the second law for all phenomena so far observed.

Of course, this theorem can no longer hold once we observe bodies of so small a scale that they only contain a few molecules. Since, however, we do not have at hand any experimental results on the behavior of bodies so small, this assumption does not run counter to previous experience. In fact, certain experiments conducted on very small bodies in gases seem rather to support the assumption, although we are still far from being able to assert its correctness on the basis of experimental proof.

[5] *von Helmholtz 1884*, p. 172; *von Helmholtz 1882*, p. 34.
[6] l. c. ⟦i.e., *Boltzmann 1896a*⟧ p. 61.
[7] *Gibbs 1875*, p. 229; *Gibbs 1892*, p. 198.

Auch wenn die in Frage kommenden Körper sehr viele Molecüle enthalten, müssen noch immer enorm kleine Abweichungen von diesem Satze eintreten, da die Zahl der Molecüle nicht unendlich ist. Allein diese Abweichungen könnten nur in so langen Zeiträumen sich bis zu einem beobachtbaren Werthe summiren, dass auch diese Consequenz der Atomistik nicht durch die Erfahrung widerlegt wird. Dies gilt um so mehr, da ja die Gastheorie nur beansprucht, ein angenähertes Bild der Wirklichkeit zu sein. Störungen, welche die Molecularbewegung durch den Lichtäther, durch electrische Eigenschaften der Molecüle etc. erfährt, muss sie wegen unserer völligen Unbekanntschaft mit der Natur dieser Agentien vernachlässigen, absolut glatte Wände kommen niemals vor, vielmehr steht jedes Gas mit dem ganzen Universum in Wechselwirkung und die Zulässigkeit der Gastheorie im grossen und ganzen wird daher durch kleine Abweichungen von der Erfahrung nicht widerlegt.

Eine Antwort auf die Frage, woher es komme, dass sich gegenwärtig die uns umgebenden Körper gerade in einem sehr unwahrscheinlichen Zustande befinden, kann man natürlich von der Naturwissenschaft ebenso wenig erwarten, wie etwa auf die Frage, woher es komme, dass es überhaupt Erscheinungen gibt und dass sich dieselben nach gewissen gegebenen Gesetzen abspielen.

781 | Die Gastheorie ist nicht zu verwechseln mit der Kraftcentratheorie, d. h. mit der Hypothese, dass sich alle Naturerscheinungen durch Centralkräfte zwischen materiellen Punkten erklären lassen, da die Gastheorie weder die Voraussetzung macht, dass sich das Verhalten des Lichtäthers, noch dass sich die innere Beschaffenheit der Molecüle durch Kraftcentra erklären lässt, sondern bloss, dass für die Wechselwirkung zweier Molecüle während der Zusammenstösse mit einer für die Erklärung der Wärmeerscheinungen genügenden Annäherung die *Lagrange*'schen Bewegungsgleichungen gelten.

Gegen diese letztere Kraftcentratheorie könnte noch eine Consequenz des *Poincaré*'schen Satzes bezüglich des Verhaltens des ganzen Universums ins Feld geführt werden. Man könnte sagen, dass nach dem *Poincaré*'schen Satze auch das ganze Universum nach genügend langer Zeit in seinen Anfangszustand zurückkehren müsste und daher Zeiten kommen müssten, wo sich alle Vorgänge im entgegengesetzten Sinne wie jetzt abspielen. Allein derartige Schlüsse scheinen mir jeder Berechtigung zu entbehren. Wie sollen wir, sobald wir die Sphäre des Beobachtbaren verlassen, entscheiden, ob die Existenzdauer des Universums oder die Anzahl der Kraftcentra, welche es enthält, unendlich gross höherer Ordnung ist? Auch wird dann die Annahme, dass der Bewegungsraum und der gesammte Energieinhalt endlich sind, fraglich. Es führt ja auch die Annahme der unbedingten Gültigkeit des Irreversibilitätsprincips bei Anwendung auf das Universum unter Voraussetzung einer unendlich langen Dauer desselben bekanntlich zu der kaum mehr verlockenden Consequenz, dass, wenn sich alle irreversiblen Processe abgespielt haben, das Universum noch unendlich lange Zeit ohne jedes Geschehen fortexistiren oder wegen Mangels an Geschehen allmählich verschwinden muss.

Even if the bodies under consideration contain a great number of molecules, enormously small deviations from the principle must still occur, since the number of molecules is not infinite. These deviations, however, could only add up to observable values in periods of time so long that experience fails to refute this consequence of the atomistic theory as well. This is all the more true as the theory of gases is only supposed to be an approximate picture of reality after all. It is bound to neglect disturbances of molecular motions brought about by the ether, or by the electrical properties of the molecules etc. owing to our complete ignorance of the nature of these agents. Perfectly smooth walls do not exist. Rather, every gas interacts with the entire universe, and, by and large, the legitimacy of the theory of gases is therefore not refuted by minor deviations from experience.

Of course, we cannot expect natural science to answer the question as to why the bodies surrounding us currently exist in a highly improbable state, just as we cannot expect it to answer the question as to why there are any phenomena at all and why they adhere to certain given principles.

The theory of gases is not to be confused with the theory of force centers, i.e., with the hypothesis that all phenomena in nature can be explained by means of central forces acting between material points, since the theory of gases assumes neither that the behavior of the ether can be explained by means of force centers nor that the inner constitution of molecules can be so explained. It only assumes that *Lagrange*'s equations of motion hold for the interaction of two molecules during the collisions with an approximation sufficient for the explanation of the thermal phenomena.

We can also use a consequence of *Poincaré*'s theorem as applied to the behavior of the entire universe as an argument against the theory of force centers. One could say that, according to *Poincaré*'s theorem, the entire universe, too, must return to its initial state after a sufficient amount of time and that hence there will have to be times when all processes unfold in the direction opposite to the current direction. Yet such conclusions are, in my view, completely lacking in justification. Once we have left the sphere of the observable behind us, how are we to decide then whether the duration of the universe, or the number of force centers contained in it, is infinitely large of higher order? Moreover, the assumption that the space of motion and the entire energy content are finite then becomes questionable. Also, it is well-known that the assumption of the strict validity of the principle of irreversibility, when applied to the universe on the supposition of its infinite duration, leads to the hardly more appealing consequence that once all irreversible processes have completely unfolded, the universe must continue to exist for an infinite amount of time without any events occurring in it or that it will gradually fade away for want of events. While it would be illegitimate

So wenig es nun berechtigt wäre, hieraus auf die Unrichtigkeit des Irreversibilitätsprincips Schlüsse zu ziehen, so wenig beweist der gleiche Fall etwas gegen die Atomistik.

Alle gegen die mechanische Naturanschauung erhobenen Einwände sind daher gegenstandslos und beruhen auf Irrthümern. Wer aber die Schwierigkeiten, welche die klare Erfassung der gastheoretischen Sätze bietet, nicht zu
782 überwinden | vermag, der sollte in der That dem Rathe Hrn. *Zermelo*'s folgen und sich entschliessen, dieselbe ganz aufzugeben.

Anhang.

Wir setzen ein Gefäss von 1 ccm Rauminhalt voraus. Darin soll sich Luft von gewöhnlicher Dichte, also rund eine Trillion (n) Molecüle befinden. Die Geschwindigkeit eines jeden sei anfangs 500 m pro Secunde. Der mittlere Abstand der Centra zweier Nachbarmolecüle ist also etwa 10^{-6} cm.

Wir construiren nun um den Mittelpunkt jedes Molecüls einen Würfel von 10^{-7} cm Seitenlänge, welchen wir den Anfangsraum des betreffenden Molecüls nennen. Wir zeichnen ferner das Geschwindigkeitsdiagramm, indem wir die Geschwindigkeit jedes Molecüls vom Coordinatenursprunge aus in Grösse und Richtung auftragen. Der Endpunkt dieser Geraden heisse der Geschwindigkeitspunkt des betreffenden Molecüls. Hierauf theilen wir den ganzen unendlichen Raum in lauter Würfel von 1 m Seitenlänge, welche wir die Elementarwürfel nennen. Denjenigen Elementarwürfel, in welchem sich der Geschwindigkeitspunkt eines Molecüls zu Anfang der Zeit befindet, nennen wir den Anfangsraum seines Geschwindigkeitspunktes.

Wir fragen nun zunächst, nach wie langer Zeit gemäss des *Poincaré*'schen Satzes die Centra sowie die Geschwindigkeitspunkte aller Molecüle wieder gleichzeitig in die betreffenden Anfangsräume zurückkehren müssen, wobei wir, wie man sieht, den Spielraum für das, was wir Rückkehr zu einem gleichen Zustande nennen, gewiss nicht enge gezogen haben, da wir den Geschwindigkeitszustand eines Molecüls als den alten bezeichnen, wenn jede seiner Geschwindigkeitscomponenten zu einem Werthe zurückgekehrt ist, der sich um nicht mehr als 1 m von seinem ursprünglichen Werthe unterscheidet.

Wir nehmen an, dass jedes Molecül in der Secunde $4 \cdot 10^9$ Zusammenstösse erfährt. Es erfolgen also im ganzen in der Secunde etwa $b = 2 \cdot 10^{27}$ Zusammenstösse im Gase. Bei jedem solchen Zusammenstosse werden im allgemeinen die Geschwindigkeitspunkte zweier Molecüle in andere Elementarwürfel versetzt. Nach dem *Poincaré*'schen Satze braucht der ursprüngliche
783 Zustand nicht früher wiederzukehren, bis | die Geschwindigkeitspunkte alle möglichen (N) Combinationen von Elementarwürfeln durchlaufen haben.

Das erste Molecül kann alle Geschwindigkeiten von Null bis ($500 \cdot 10^9 = a$) m/sec annehmen. Hat es die Geschwindigkeit $v_1 \times$ m/sec, so kann das zweite noch alle Geschwindigkeiten von Null bis $\sqrt{a^2 - v_1^2}$ m/sec annehmen etc.

to infer from this case that the principle of irreversibility is incorrect, it would be equally wrong to assume that it refutes the atomistic theory in any way whatsoever.

All objections against the mechanical approach to nature are therefore unfounded and based on mistakes. Anyone unable to overcome the difficulties attendant on a clear understanding of the principles of the theory of gases really ought to heed Mr. *Zermelo*'s advice and resolve to abandon the theory altogether.

Appendix.

Consider a container of 1 cc capacity which holds air at ordinary density, and hence contains about one quintillion (n) molecules. Set the velocity to 500 m per second at the beginning. Hence, the mean distance between the centers of two neighboring molecules is about 10^{-6} cm.

We now imagine a cube of edge length 10^{-7} cm which is constructed around the center of each molecule and to which we refer as the initial space of the molecule. Furthermore, we draw the velocity diagram by plotting the velocity for each molecule by magnitude and direction from the point of origin. The endpoint of this straight line is said to be the velocity point of the molecule under consideration. We then divide the entire infinite space into cubes of edge length 1 m, which are called the "elementary cubes". The elementary cube containing the velocity point of a molecule at the beginning of the time interval is said to be the initial space of its velocity point.

Now, we first ask how long, according to *Poincaré*'s principle, it takes for both the centers and the velocity points of all molecules to simultaneously return to their respective initial spaces, where the scope of the term "return to the same state" is evidently not taken too narrowly, since we recognize a velocity state of a molecule as its old velocity state when each of its velocity components returns to a value that differs from the original value by not more than 1 m.

Let us assume that every molecule experiences $4 \cdot 19^9$ collisions per second. Hence, the total number of collisions in the gas per second is about $b = 2 \cdot 10^{27}$. As a rule, at each collision, the velocity points of two molecules are displaced into another elementary cube. According to *Poincaré*'s principle, the initial state need not recur before the velocity points have passed through all possible (N) combinations of elementary cubes.

The first molecule can assume any velocity from zero to ($500 \cdot 10^9 = a$) m/s. If its velocity is $v_1 \times$ m/s, then the second molecule can assume any velocity from zero to $\sqrt{a^2 - v_1^2}$ m/s, etc.

Die Anzahl aller möglichen Combinationen, aller Geschwindigkeitspunkte in die verschiedenen Elementarwürfel ist also:

$$N = (4\pi)^{n-1} \int_0^a v_1^2 dv_1 \int_0^{\sqrt{a^2-v_1^2}} v_2^2 dv_2 \cdots \int_0^{\sqrt{a^2-v_1^2\cdots v_{n-2}^2}} v_{n-1}^2 d \cdot v_{n-1}$$
$$= \frac{\pi^{\frac{3n-3}{2}} a^{3(n-1)}}{2 \cdot 3 \cdot 4 \cdots [3(n-1)/2]},$$

oder

$$\frac{2 \cdot (2\pi)^{\frac{3n-4}{2}} a^{3(n-1)}}{3 \cdot 5 \cdot 7 \cdots 3(n-1)},$$

je nachdem n ungerade oder gerade ist.

Da jede dieser Combinationen durchschnittlich nach $1/b$ Secunde wechselt, so werden sie alle in N/b Secunden durchlaufen sein. Nach dieser Zeit also müssten alle Molecüle bis auf eines das erlangt haben, was wir ihren ursprünglichen Geschwindigkeitszustand nannten. Dabei ist noch die Geschwindigkeitsrichtung dieses letzten Molecüls gar nicht beschränkt, ebenso wenig die Lage des Mittelpunktes irgend eines Molecüls. Damit aber der Zustand wieder der alte würde, müsste auch der Mittelpunkt jedes Molecüls wieder in seinen Anfangsraum zurückkehren, also die obige Zahl noch mit einer zweiten von ähnlicher Grössenordnung multiplicirt werden.

Wie gross aber schon die Zahl N/b ist, davon erhält man einen Begriff, wenn man bedenkt, dass sie viele Trillionen Stellen hat. Wenn dagegen um jeden mit dem besten Fernrohr sichtbaren Fixstern so viele Planeten, wie um die Sonne kreisten, wenn auf jedem dieser Planeten so viele Menschen wie auf der Erde wären und jeder dieser Menschen eine Trillion Jahre lebte, so hätte die Zahl der Secunden, welche alle zusammen erleben, noch lange nicht fünfzig Stellen.

784 | Wären hingegen die Gasmolecüle anfangs im Gefässe ziemlich gleichmässig vertheilt, hätten aber alle genau dieselbe Geschwindigkeit, so würde sich schon nach einhundertmilliontel Secunde sehr nahe die *Maxwell*'sche Geschwindigkeitsvertheilung hergestellt haben. Die Vergleichung dieser Zahlen zeigt einerseits, einen wie kleinen Bruchtheil der Zahl aller möglichen Zustandsvertheilungen diejenigen bilden, welche von der *Maxwell*'schen erheblich abweichen, andererseits wie zweifellos solche Sätze, welche theoretisch nur den Charakter von Wahrscheinlichkeitssätzen haben, praktisch mit Naturgesetzen gleichbedeutend sind.

Wien, den 20. März 1896.

The number of all possible combinations, of all velocity points in the various elementary cubes is therefore:

$$N = (4\pi)^{n-1} \int_0^a v_1^2 dv_1 \int_0^{\sqrt{a^2-v_1^2}} v_2^2 dv_2 \cdots \int_0^{\sqrt{a^2-v_1^2\cdots v_{n-2}^2}} v_{n-1}^2 d \cdot v_{n-1}$$
$$= \frac{\pi^{\frac{3n-3}{2}} a^{3(n-1)}}{2 \cdot 3 \cdot 4 \cdots [3(n-1)/2]},$$

or

$$\frac{2 \cdot (2\pi)^{\frac{3n-4}{2}} a^{3(n-1)}}{3 \cdot 5 \cdot 7 \cdots 3(n-1)},$$

depending on whether n is even or odd.

Since each of these combinations is replaced by another one after $1/b$ seconds on average, all of them will have had their turn after N/b seconds. Hence, after this time, all molecules, except one, will have reached what we called its initial velocity state. Here, the velocity direction of this last molecule has not even been restricted yet; nor has the location of the center of any molecule. But for the state to turn into its former self the center of each molecule would have to return to its initial space as well. Hence, the number above would have to be multiplied by another one of similar magnitude.

In order to get an idea of how great even the number N/b is, just consider that it has many quintillions of digits. On the other hand, if we assume that every fixed star visible through the best telescope is orbited by as many planets as the sun, and that there are as many people on each of these planets as there are on earth, and that each of them has a life span of one quintillion years, then the number of seconds of all their life spans combined is far from reaching 50 digits.

If, in contrast, the gas molecules were fairly equally distributed in the container at the beginning but all had exactly the same velocity, then the *Maxwell* velocity distribution would be almost completely established after merely one one-hundred-millionth part of one second. The comparison of these numbers suggests, on the one hand, that the number of state distributions deviating from the *Maxwell* state distribution is but a small fraction of the number of all possible state distributions, and, on the other hand, that those laws which, theoretically speaking, only have the character of principles of probability theory are, without any doubt, practically equivalent to laws of nature.

Vienna, March 20, 1896.

Ueber mechanische Erklärungen irreversibler Vorgänge. Eine Antwort auf Hrn. Boltzmann's „Entgegnung"

1896b

Durch meine Abhandlung im letzten Märzheft dieser Zeitschrift „Ueber einen Satz der Dynamik und die mechanische Wärmetheorie"[1] hat sich Hr. *Boltzmann* zu einer sofortigen „Entgegnung"[2] veranlasst gefunden, einer Entgegnung freilich, in der ich eher eine Bestätigung als eine Widerlegung meiner Ausführungen erblicken kann. Nicht nur den zu Grunde liegenden Satz *Poincaré*'s erkennt Hr. *Boltzmann* als „selbstverständlich richtig" an, sondern auch seine Anwendbarkeit auf ein abgeschlossenes System von Gasmolecülen im Sinne der kinetischen Theorie. In der That seien in einem solchen Systeme die Vorgänge, mathematisch betrachtet, *periodischer* Natur, also *nicht-irreversibel* im strengen Sinne, sodass eine wirklich *fortdauernde* Vermehrung der Entropie, wie der zweite Hauptsatz in seiner gewöhnlichen Fassung sie fordert, *nicht* angenommen werden dürfe. Eben dies zu beweisen und damit eine gesicherte Grundlage für die Erörterung der principiellen Frage zu gewinnen, war das Ziel meiner Arbeit; Hrn. *Boltzmann*'s gastheoretische Untersuchungen wurden mir erst später bekannt, schienen mir aber diese allgemeine Aufklärung noch keineswegs, wie er selbst annimmt, überflüssig zu machen.

Die von mir (p. 493) behauptete „Nothwendigkeit, entweder dem *Carnot-Clausius*'schen Princip oder aber der mechanischen Grundansicht eine principiell andere Fassung zu geben", wäre also zugestanden, und nur die Entscheidung zwischen den beiden Möglichkeiten bliebe zunächst noch dem Einzelgeschmack überlassen. Hier würde ich allerdings, und ich wohl nicht allein, die einfache Zusammenfassung einer Fülle gesicherter *Erfahrungen* zu einem einzi-
794 gen allgemein | gültigen Satze nach den Regeln der Induction für zuverlässiger halten, als eine ihrer Natur nach niemals direct beweisbare *Theorie* und wäre daher *diese* eher aufzugeben oder abzuändern bereit wie *jenen*, wenn doch einmal beide nicht zu vereinigen sind. — Hr. *Boltzmann* dagegen, der an der herkömmlichen mechanischen Auffassung gar nichts geändert wissen will, verwandelt den zweiten Hauptsatz in einen „blossen Wahrscheinlichkeitssatz" von zeitlich beschränkter Gültigkeit; er hält aber diese Abänderung, deren *principielle* Bedeutung er selbst nicht verkennen wird, für durchaus unbedenklich, ja unwesentlich, denn „praktisch" seien beide Fassungen „völlig gleichbedeutend". Sehen wir nun, wie weit ihm dies nachzuweisen gelungen ist.

Das eine ist zweifellos richtig, was Hr. *Boltzmann* betont: bei der ungeheuren Anzahl der Molecüle in einem endlichen Gasvolumen wird auch die mittlere Dauer der *Poincaré*'schen Perioden, in denen sich jeder Zustand

[1] E. *Zermelo*, Wied. Ann. **57**. p. 485. 1896.
[2] H. *Boltzmann*, Wied. Ann. **57**. p. 773. 1896.

On mechanical explanations of irreversible processes. An answer to Mr. Boltzmann's "rejoinder"

1896b

⟦The introductory note just before *1896a* also addresses *1896b*.⟧

My essay "On a theorem of dynamics and the mechanical heat theory"[1] published in the most recent, March edition of this journal immediately elicited a critical "rejoinder"[2] from Mr. *Boltzmann*, which, however, as I read, confirms rather than refutes my line of argument. Not only does Mr. *Boltzmann* acknowledge that the underlying theorem by *Poincaré* is "of course correct" but also that it can be applied to a closed system of gas molecules in the sense of the kinetic theory. He acknowledges that, in fact, the processes in a system of this kind are essentially *periodic*, mathematically speaking, and hence strictly *non-irreversible*, so that we must *not* assume a really *continual* increase of entropy as demanded by the common version of the second law. It was the goal of my paper to prove just this and to thereby place the discussion of the basic question on a firm footing ; it was only later that Mr. *Boltzmann*'s gas-theoretic investigations came to my attention. Unlike him, however, I did not believe that they render superfluous this general clarification in any way.

My contention (p. 493) that it is "imperative to provide an altogether different version of either the *Carnot-Clausius* principle or of the basic mechanical approach" would thus be conceded, and only the choice between the two possibilities would at first depend on individual taste. In this case, however, I would not be the only one to consider the simple summation of an abundance of trusted experiences into a single generally valid principle according to the rules of induction more reliable than a *theory*, which, by its very nature, can never be proved directly and to thus be more inclined to abandon or modify the *latter* rather than the *former*, if it turns out that they cannot be combined.—Mr. *Boltzmann*, on the other hand, who does not want to see anything changed with the common mechanical approach, converts the second law into a "mere principle of probability theory" whose validity is limited by constraints of time; but he is not alarmed by this modification, whose *fundamental* significance he will not fail to acknowledge, and even considers it marginal, since "from a practical point of view" the two versions are "completely equivalent". Let us now consider how far he succeeded in showing this.

There is one point stressed by Mr. *Boltzmann* that is undoubtedly true: given the enormous number of molecules in a finite gas volume, the mean duration of *Poincaré* periods within which every state recurs is too great to

[1] *Zermelo 1896a.*
[2] *Boltzmann 1896.*

wiederholt, allzu gross werden, als dass eine directe *Beobachtung* der theoretisch nachgewiesenen Periodicität zu erwarten wäre. Nur seine numerische Abschätzung (p. 782–784), der er einen einzelnen exceptionellen Anfangszustand, eine ganz bestimmte Combination der Molecüle zu Grunde legt, kann ich hier nicht für maassgebend halten; handelt es sich doch in der Praxis immer nur um den „physikalischen Zustand", der durch sehr viel verschiedene Combinationen verwirklicht werden und darum auch sehr viel früher wiederkehren kann. Auch genügte es für meinen Zweck bereits, die Wiederkehr irgend eines anderen Zustandes von gleichem oder kleinerem Entropiewerthe nachzuweisen, und die Perioden einer solchen einzelnen Function S werden je nach ihrer Natur verschieden, im Ganzen aber durchaus nicht mehr so „beruhigend" gross ausfallen. Immerhin wird es auch solche Functionen geben, deren Periodicität sich gleichwohl noch jeder Beobachtung entzieht, und zu ihnen mag auch gerade die Entropiefunction gehören.

Für eine solche Function kann nun allerdings der Fall eintreten, dass sie beständig zuzunehmen *scheint*, weil der theoretisch immer vorhandene absteigende Ast der periodischen Curve erst so spät beginnt, dass er praktisch nicht
795 mehr in Betracht kommt. Daraus folgt aber noch keineswegs, dass es | auch Functionen gebe, für welche *immer* nur der aufsteigende, niemals der absteigende zur Erscheinung gelangte, was doch von dem mechanischen Analogon der Entropiefunction gelten müsste. Diese Eigenschaft für die in unserer Zeit nun einmal geltenden Anfangszustände einfach als Thatsache hinzunehmen, geht doch wohl nicht an; denn es handelt sich ja nicht um eine bestimmte, nur einmal vorhandene Variable, wie z. B. die Excentricität der Erdbahn, die jetzt gerade auf noch sehr lange Zeit im Abnehmen begriffen ist, sondern um die Entropie *jedes beliebigen* Systemes, solange es der Einwirkung äusserer Kräfte entzogen ist. Woher kommt es also, dass in einem solchen System immer nur *Zunahme* der Entropie, *Ausgleichung* der Temperatur- und Concentrationsunterschiede, niemals aber von selbst das Gegentheil eintritt? Und welches Recht haben wir, dieselbe Erscheinung auch wenigstens für die nächste Zukunft zu erwarten? Darauf muss durchaus eine befriedigende Antwort gegeben werden, wenn von einer wirklichen mechanischen Analogie zum zweiten Hauptsatze die Rede sein soll.

Auch die Wahrscheinlichkeitsrechnung, richtig angewendet, kann hier, wie mir scheint, nicht helfen, weil eben jeder Zunahme eine (später irgend einmal eintretende) Abnahme entspricht und daher beide gleich wahrscheinlich, ihre Wahrscheinlichkeiten wenigstens von gleicher Ordnung sein müssten. Meines Erachtens in Uebereinstimmung mit der *Poincaré*'schen Definition in der angeführten Abhandlung[1] kann die Wahrscheinlichkeit für das Bestehen einer bestimmten Eigenschaft des molecularen Zustandes, z. B. auch für einen bestimmten Werth der Function S, nur gemessen werden durch die „Ausdehnung"[2] γ des „Gebietes" g aller möglichen Zustände, welche diese Eigenschaft

[1] *Poincaré*, Act. Math. **13**. p. 71. 1890.
[2] *Zermelo*, Wied. Ann. **57**. p. 487. 1896.

allow for a direct *observation* of the theoretically demonstrated periodicity. I would only object to his numerical estimate (pp. 782–784), which he bases on one exceptional initial state, one particular combination of molecules; for, in practice, we are after all always concerned only with the "physical state", which can be realized by many different combinations, and hence can recur much sooner. Also, it already suffices for my purposes to show the return of any other state of the same or smaller entropy value, and, depending on their nature, the periods of an individual function S of this sort will turn out to be different but, by and large, no longer as "alarmingly" great. After all, there will also be functions whose periodicity defies observation nonetheless, and in particular the entropy function may be one of those of functions.

It is now, however, possible that a function of this sort *appears* to steadily increase, since the downward branch of the periodic curve, which, theoretically speaking, always exists, begins too late to be of any practical relevance. From this, however, it by no means follows that there are also functions for which it is *always* only the upward branch but never the downward branch that makes its appearance, which, however, should be the case for the mechanical analogue of the entropy function. Yet it will not do to simply accept this property as a fact for the initial states currently in existence; for, after all, we are not concerned with a particular, unique variable, such as the eccentricity of the earth's orbit, which is currently bound to decrease for a very long time still. Rather, we are concerned with the entropy of *any system whatsoever*, as long as it is removed from the action of external forces. So, why is it that in such a system the entropy only *increases* and that differences in temperature and concentration are *compensated*? Why does the opposite case never occur by itself, and are we really justified in expecting this phenomenon to occur for at least the proximate future? These are questions that require fairly exhaustive answers if we are to speak of a genuine mechanical analogy with the second law.

It seems to me that probability theory, when applied correctly, cannot be of any help in this case either, since to every increase there corresponds a decrease (which occurs at some later time), and hence both would have to be equally probable, or at least have a probability of the same order. According to *Poincaré*'s definition in the cited essay,[3] it is, in my view, possible to measure the probability for a molecular state to have a particular property, and, e.g., also for a particular value of the function S, only by means of the "extension"[4] γ of the "domain" g of all possible states possessing this property,

[3] *Poincaré 1890*, p. 71.
[4] *Zermelo 1896a*, p. 487.

besitzen, dividirt natürlich durch die Gesammtausdehnung Γ des ein für allemal gegebenen Gebietes G aller überhaupt möglichen Zustände. Da nun aber nach dem *Liouville*'schen Satze jede Ausdehnung γ in der Zeit unveränderlich ist, so muss auch die betreffende Eigenschaft, der betreffende Werth
796 der Function, für jede | spätere Zeit ebenso wahrscheinlich sein wie für den Anfangszustand, den Beginn der Bewegung, sodass eine überwiegende Zu- oder Abnahme auch aus Wahrscheinlichkeitsgründen nicht zu erwarten wäre.

Hr. *Boltzmann* verfährt anders. Er nimmt eine Function H an, deren Curve, bezogen auf die Zeit t als Abscisse, im allgemeinen sehr nahe der t-Axe verläuft und nur selten einzelne Erhebungen, „Buckel", besitzt, die um so seltener, um so unwahrscheinlicher sein sollen, je grösser sie sind.[1] Dass er diese Beschaffenheit wirklich von seiner anders definirten Function H *nachgewiesen* habe, kann ich nicht finden, da nach meiner Auffassung Wahrscheinlichkeit und Dauer eines Zustandes nicht identisch sind; indessen Functionen von der angegebenen Eigenschaft mag es ja geben. Von dieser Function H nimmt er weiter an, dass sie anfangs einen ungewöhnlich grossen Werth H_0 besitze, also einem Buckel angehöre, und folgert daraus, dass die Curve diesen Buckel bald überschreiten und fast bis auf Null abnehmen werde, um schliesslich wieder ausserordentlich lange dicht an der Abscissenaxe zu verlaufen. Entspricht nun diesem Grenzwerthe Null der Function H eine durch das *Maxwell*'sche Gesetz ausgedrückte Vertheilung der Geschwindigkeiten, so kann allerdings das Verhalten dieser H-Curve als Erläuterung für die wahrscheinlichkeitstheoretische Bedeutung des Vertheilungsgesetzes, die ich meinerseits aber auch gar nicht bestritten habe, aufgefasst werden. Nur einen „stationären *End*zustand" im strengen Sinne stellt das Gesetz eben nicht dar, weil sich die Curve, wenn auch nach langer Zeit, schliesslich doch wieder zu neuen Buckeln erhebt, und wenn auch Hr. *Boltzmann* selbst diesen *Maxwell*'schen Zustand nur empirisch angenähert als „Endzustand" gelten lassen will, so schien mir doch *diese* Auffassung aus seinen früheren Schriften nicht mit hinreichender Deutlichkeit hervorzugehen.

Hier handelt es sich aber nicht um das *Maxwell*'sche Gesetz, sondern um die Analogie, die zwischen den Eigenschaften der H-Curve und dem zweiten Hauptsatze der Wärmetheorie bestehen soll, und diese Analogie ist es
797 eben, die ich | bestreite. Es genügt doch nicht zu zeigen, dass alle Störungen *schliesslich* wieder zu einem lange dauernden Gleichgewichtszustande zurückführen, sondern es wäre nachzuweisen, dass die Veränderungen *beständig im selben Sinne*, im Sinne des Ausgleiches erfolgen, dass die Function H während beobachtbarer Zeit *immer nur* abnimmt, oder dass wenigstens geringe, praktisch unmerkbare Zunahmen durch unmittelbar folgende stärkere Abnahmen aufgehoben werden; dieser Nachweis aber lässt sich meines Erachtens für die H-Function so wenig führen wie für irgend eine andere. Offenbar kann der Anfangszustand, dessen Wahrscheinlichkeit ja nur vom Anfangswerthe H_0 abhängen soll, ebenso gut vor wie hinter dem Maximum, ebenso gut im

[1] *Boltzmann*, Wied. Ann. **57**. p. 774. 1896.

divided, of course, by the total extension Γ of the domain G of all possible states whatsoever, which is given once and for all. But since now, according to *Liouville*'s theorem, every extension γ is invariant in time, the respective property, the respective value of the function, must be as probable for every later point in time as for the initial state, the beginning of the movement, so that a predominant increase, or decrease, would not have to be expected even for probabilistic reasons.

Mr. *Boltzmann* takes a different route. He assumes a function H whose curve usually runs very close to the t-axis, where the time t is plotted along the abscissa, and only rarely features individual elevations, or "humps". The bigger these "humps" are, the rarer, or less probable, they are supposed to be.[5] That he has actually *demonstrated* this property for his function H, which is differently defined, I fail to see, since, in my view, probability and duration of a state are not identical; nevertheless, functions of the stated property may of course exist. Furthermore, he assumes that the function H has an unusually large value, H_0, at first, and hence that it belongs to a hump, and concludes from this that the curve soon exceeds this hump and almost decreases to zero only to eventually run closely again along the abscissa for an extraordinarily long period of time. But if now to this limit zero of the function H there corresponds a velocity distribution that is expressed by *Maxwell*'s law, then the behavior of this H-curve may as well serve to illustrate the meaning of the law of distribution in terms of probability theory, which I, for my part, have by no means contested. It is only the "stationary *final* state" in the strict sense that is not represented by the law since the curve still rises to new humps eventually, even if only after a long time. And although Mr. *Boltzmann* wants to accept this *Maxwell* state only as an empirical approximation of a "final state", I did not believe that *this* conception emerges from his earlier writings with sufficient clarity.

It is not, however, *Maxwell*'s law with which we are concerned here, but the analogy that is supposed to hold between the properties of the H-curve and the second law of heat theory, and it is this analogy that I contest. It does not suffice after all to show that all perturbations *eventually* return to a long-lasting state of equilibrium. Instead, it would be necessary to demonstrate that the changes take place *continually in the same direction*, in the direction of the balance, that the function H *always only* decreases in observable periods of time, or that at least insignificant, practically indiscernible increases are offset by immediately succeeding, more significant decreases. In my view, however, this cannot be demonstrated for the H-function, nor for any other function. The initial state, whose probability is supposed to depend only on the initial value H_0 after all, may obviously lie before as well as after

[5] *Boltzmann 1896*, p. 774.

aufsteigenden wie im absteigenden Aste gelegen sein, und im ersteren Falle müsste zunächst eine *Zunahme* erfolgen, die ebenso lange dauern könnte wie die Abnahme nachher, solange noch $H > H_0$ ist. Jeder beobachteten Abnahme $H_1 \ldots H_2$ im absteigenden Aste entspräche eine gleich grosse Zunahme $H_2 \ldots H_1$ im vorhergehenden aufsteigenden, mit der einen könnte der Vorgang um nichts leichter beginnen als mit der anderen. Auch wenn die Aufstiege immer in kürzerer Zeit erfolgten und daher unwahrscheinlicher wären als die Abstiege, wozu aber nach den gemachten Voraussetzungen gar kein Grund vorliegt, so würden sie doch zugleich um so *steiler* sein und dadurch, wie ich glaube, ebenso gut ins Gewicht fallen.

Hrn. *Boltzmann*'s Meinung scheint nun freilich, wenn ich ihn richtig verstanden habe[1], dahin zu gehen, dass die Anfangszustände mit erheblichen H-Werthen, sagen wir $H_0 > H'$, damit der doch einmal vorhandene Buckel nicht unnöthig gross und zugleich unnöthig unwahrscheinlich angenommen zu werden brauchte, in der Regel *Maxima* darstellen müssten, und dann allerdings würde immer nur der aufsteigende Ast beobachtet werden können. Wie ich mir das aber vorstellen soll, weiss ich nicht. Da sollen also die Schnittpunkte der Curve mit einer Parallelen zur Abscissenaxe $H = H_0$ meistens Maxima sein und zwar für alle beträchtlicheren Werthe $H_0 > H'$; wo bleiben dann aber die
798 übrigen Punkte dieser Buckel ($H > H'$), | welche *keine* Maxima sind? Sollten die in der That gegen die Maxima in der Minderheit sein? Es ist klar, dass die Betrachtung nur Sinn haben kann, wenn man die Maxima nicht als mathematische Punkte betrachtet, sondern ihnen eine gewisse Breite, eine gewisse Zeitdauer zugesteht. Dann müsste eben für jeden Anfangszustand der Werth der Function eine längere oder kürzere Zeit nahezu unverändert *andauern*, eine Art labilen Gleichgewichtes darstellen, während doch der *Erfahrung* zufolge, z. B. bei der Wärmeleitung, der Ausgleichungsvorgang um so *schneller* beginnt, je grösser die anfänglichen Temperaturdifferenzen sind, d. h. je weiter der Anfangszustand vom stabilen Gleichgewichte entfernt ist. Aber auch davon abgesehen, begreife ich nicht, was denn überhaupt in der ganzen Betrachtung der *Anfangs*zustand ausser seiner geringen Wahrscheinlichkeit, die er doch mindestens mit den benachbarten theilt, vor den übrigen voraus haben soll. Hr. *Boltzmann* nimmt die ganze H-Curve, also wohl die sämmtlichen vom System zu durchlaufenden Zustände, als *gegeben* an und fragt nun nach der Wahrscheinlichkeit eines bestimmten Anfangszustandes, d. h. desjenigen in der Curve enthaltenen Zustandes, von welchem aus das System nach Ausschliessung äusserer Kräfte sich thatsächlich zu bewegen beginnt. Nun kann aber, wie die Erfahrung lehrt, auch ist hier für das Gegentheil kein Grund gegeben, durch geeignete Einwirkung *jeder beliebige* mögliche Zustand hervorgerufen und dann das System isolirt, sich selbst überlassen, d. h. aber, jeder beliebige Zustand P_0 zum Anfangszustande gemacht werden. Alsdann wird das System, solange es isolirt bleibt, alle auf P_0 *folgenden* Zustände P der Reihe nach wirklich durchlaufen, während die vorhergehenden nur mathematisch zu ergänzen blieben. Wäre

[1] *Boltzmann*, Vorl. über Gastheorie p. 44.

the maximum, on the ascending as well as on the descending branch. In the first case, there would initially have to be an *increase*, which may last as long as the subsequent decrease, as long as we still have $H > H_0$. To every observable decrease $H_1 \ldots H_2$ in the descending branch there would correspond an equally great increase $H_2 \ldots H_1$ in the preceding ascending branch, and the process could begin with the one just as easily as with the other. Even if the ascents would always take less time than the descents, and hence would be more improbable, which, however, is not entailed by the given assumptions, the ascents would, at the same time, still be all the steeper and would thus carry the same weight, or so I believe.

Mr. *Boltzmann*, however, if I understood him correctly,[6] tends towards the view that the initial states with considerable H-values, say, $H_0 > H'$ must usually represent *maxima* lest we have to take the inevitable hump to be both needlessly large and needlessly improbable, in which case we would only be able to observe the ascending branch. Yet I do not know how to conceive of this. The intersection points of the curve and a line parallel to the abscissa axis $H = H_0$ are hence supposed to be mostly maxima, namely for all more considerable values $H_0 > H'$; but then what about the remaining points of these humps ($H > H'$) that are *not* maxima? Are the maxima really supposed to outnumber them? It is obvious that the consideration makes sense only if we do not consider maxima as mathematical points but accord to them a certain breadth, a certain duration in time. Then for every initial state the value of the function would have to *last* almost unchanged for a longer or shorter period of time. It would have to represent a sort of unstable equilibrium, while *experience* teaches that, in the case of, say, thermal conduction, the process of equalization begins all the *more quickly* the greater the initial differences in temperature, i.e., the greater the distance of the initial state from the stable equilibrium. But even apart from this, I fail to understand wherein the advantage lies that the entire consideration accords to the *initial* state over all others, except its low probability, which it has in common with its neighboring states at the very least. Mr. *Boltzmann* assumes the entire H-curve as *given*, and hence, presumably, all states through which the system must pass, and he then asks about the probability of a particular initial state, i.e., of that state contained in the curve from which the system actually sets out to move when all external forces are removed. But now, as experiences shows (and moreover there is no reason for the opposite), one can produce *any* possible state by means of suitable action and thereafter isolate the system, leave it alone, i.e., however, turn any state P_0 into the initial state. Then, so long as it remains isolated, the system will successively pass through all states P *following* P_0, while the preceding ones would have to be supplied

[6] *Boltzmann 1896a*, p. 44.

nun die obige Argumentation richtig, stellten in der That die Anfangszustände meistens Maxima der H-Function dar, so müsste dasselbe auch von *allen übrigen* Zuständen gelten, für welche H über H' hinausgeht, da eben jeder nach Belieben zum Anfangszustande gemacht werden könnte, vor allem aber, weil der ganze Wahrscheinlichkeitsschluss auf jeden anderen Zustand mit demselben Recht wie auf den Anfangszustand anwendbar wäre. Alle diese Zustände müs-
799 sten ebenfalls Maxima darstellen, | die Curve müsste von einer gewissen Höhe *ab aus lauter Maximis bestehen*, was widersinnig ist, da die Function doch keinesfalls constant sein soll. Um also auch nur ein empirisch angenähertes Analogon des Entropiesatzes zu gewinnen, genügte es keineswegs, den Anfangszustand als äusserst unwahrscheinlich vorauszusetzen, man müsste vielmehr immer noch die *neue Annahme* hinzufügen, dass im Anfange die H-Curve gerade ein Maximum besitze oder ein solches eben überschritten habe, d. h. aber, solange man eine solche Annahme nicht aus der *physikalischen Entstehung* des Anfangszustandes begreiflich machen kann, man müsste eben das voraussetzen, was man beweisen will; anstatt einer Erklärung wäre das ein Verzicht auf jede Erklärung.

Ich habe mich also noch keineswegs davon überzeugen können, dass Hrn. *Boltzmann*'s Wahrscheinlichkeitsbetrachtungen, auf denen „die klare Erfassung der gastheoretischen Sätze“[1] beruhen soll, in der That im Stande seien, die aus dem *Poincaré*'schen Satze fliessenden Bedenken gegen die Möglichkeit einer mechanischen Erklärung irreversibler Vorgänge zu zerstreuen, selbst wenn man auf die strenge Irreversibilität zu Gunsten einer blos empirischen verzichtet. Ist es doch schon a priori klar, dass der Wahrscheinlichkeitsbegriff gar nichts Zeitliches enthält und daher aus ihm auch nichts auf die *Richtung*, in der die Vorgänge sich abspielen, geschlossen werden kann; vielmehr würde jede solche Herleitung sich mit genau demselben Rechte, indem man Anfangs- und Endzustand vertauscht, auch auf den *umgekehrten* Vorgang, den Verlauf in entgegengesetzter Richtung anwenden lassen. Zutreffender daher auf die vorliegende Streitfrage als das von Hrn. *Boltzmann* angeführte Beispiel des Würfelspielers scheint mir das folgende zu sein. Zwei Spieler, nehmen wir an, hätten die Beobachtung gemacht, dass die von ihnen benutzten Würfel aus einer bestimmten Bezugsquelle beim Beginn ihrer Benutzung immer eine bestimmte Augenzahl, sagen wir, die Eins bevorzugten, dass sie also bei den ersten 600 Würfen nicht 100mal, sondern etwa 200mal die Eins aufzuweisen pflegten, bei den nächsten 600 aber schon weniger oft
800 und nach | längerer Fortsetzung des Spieles zuletzt unter je 600 Würfen die Eins ebenso wie jede andere Zahl durchschnittlich 100mal. Der eine Spieler nun würde diese Erscheinung ganz in Ordnung finden, weil sich doch die Gesetze der Wahrscheinlichkeitslehre immer erst nach längerem Spiele geltend machen könnten, der andere aber erklärte: Nein! Diese Würfel müssen jedenfalls gefälscht sein, und erst bei längerer Benutzung nehmen sie durch Abschleifung oder dergleichen ihren normalen Zustand an. — Dieser letzteren Meinung wäre ich nun auch.

[1] *Boltzmann*, Wied. Ann. **57**. p. 778. 1896.

only mathematically. If the reasoning above were correct, if the initial states really represented maxima of the H-function for the most part, then the same would have to be true of *all remaining* states for which H exceeds H', since any state could arbitrarily be turned into the initial state, but first and foremost since the probabilistic conclusion in its entirety would be just as applicable to any other state as it is to the initial state. All these states would have to represent maxima as well, and from a certain height on, the curve would have to *entirely consist of maxima*, which is paradoxical, since the function is by no means supposed to be constant. Hence, in order to obtain a mere empirical approximation of the analogue of the law of entropy it would by no means suffice to assume that the initial state is highly improbable. Rather, it would always also be necessary to add the *new assumption* that the H-curve possesses a maximum at the beginning or has just exceeded a maximum. But this means that, so long as we are unable to render intelligible an assumption of this kind in terms of the *physical origin* of the initial state, we would have to assume just what we want to prove; instead of providing an explanation, we would renounce any explanation.

Thus, I have not yet been able to convince myself that *Boltzmann*'s probabilistic considerations on which "the clear understanding of the principles of the theory of gases"[7] is supposed to rest are really capable of allaying those concerns against the possibility of a mechanical explanation of irreversible processes that arise from *Poincaré*'s theorem, even if strict irreversibility is replaced by merely empirical irreversibility. Is it really a priori clear that the concept of probability has no temporal element whatsoever, and hence that nothing can be inferred from it about the *direction* in which the processes unfold; rather, we could just as well apply any such derivation to the *reverse* process, the process unfolding in the opposite direction, by interchanging initial and final state. It seems to me that the following example is more relevant to the issue at hand than *Boltzmann*'s example of a dice player. Let us assume that two players have made the observation that the dice they use, which are obtained from a certain source, always show a bias toward a particular face at the beginning of the game, say, the one, so that while the one comes up 200 instead of 100 times in the first 600 rolls, it comes up less often in the following 600 rolls and, as the game goes on, it eventually comes up 100 times out of 600 rolls, just like any other number on the dice on average. Now assume that one of the two players sees nothing wrong with this situation and, in fact, argues that the laws of probability can get a foothold only after the game has gone on for a while. The other player, however, declares: No! The dices must be loaded and they only return to their normal condition through, say, abrasion due to extended use.—And I would agree with the latter player.

[7] *Boltzmann 1896*, p. 781. ⟦Zermelo erroneously writes "778".⟧

Ebenso wenig aber wie das allgemeine Irreversibilitäts*princip* werden sich die einzelnen irreversiblen *Processe* selbst ohne neue physikalische Annahmen aus den mechanischen Voraussetzungen erklären lassen, wenigstens nicht in Bezug auf ihren zeitlichen Verlauf. Namentlich gilt dies auch von der Differentialgleichung der Wärmeleitung und Diffusion: $\partial u/\partial t = a^2\partial^2 u/\partial x^2$, die ihrer Natur nach ausschliesslich nicht umkehrbare Vorgänge darstellt. Die Versuche, auch diese Differentialgleichung allein aus den mechanischen Grundgleichungen in Verbindung mit Wahrscheinlichkeitsbetrachtungen abzuleiten, wie sie u. a. von *Clausius*, *Maxwell* und *Boltzmann* angestellt wurden, können daher nicht zum Ziele führen, weil sie Unmögliches unternehmen, und ein scheinbares Gelingen könnte nur auf Fehlschlüssen beruhen. Zu den Hauptfehlern der hierzu verwendeten Methoden scheint mir vor allem die unbeweisbare, weil unrichtige Annahme zu gehören, dass der moleculare Zustand eines Gases nach Hrn. *Boltzmann*'s Ausdruck[1] jedesmal ein „ungeordneter" sei und alle möglichen Richtungen und Combinationen überall gleichmässig vertreten, wenn man über den wahren Zustand, der doch immer vom „geordneten" Anfangszustand abhängig sein wird, nichts Bestimmtes aussagen kann. Die Wahrscheinlichkeitstheorie, meine ich, berechtigt solche Annahmen in gewissem Umfange höchstens für den *Anfangszustand*; die Wahrscheinlichkeit späterer Zustände aber und damit der Vorgänge selbst müsste immer erst durch die der zugehörigen Anfangszustände ausgedrückt werden, und erst dann liesse sich über Zulässigkeit solcher Durchschnittsannahmen entscheiden. Die | Schwierigkeiten, auf einem ähnlichen Wege die Untersuchungen 801
wenigstens im wahrscheinlichkeitstheoretischen Sinne *streng* durchzuführen, mögen gewiss sehr gross sein, doch scheinen sie mir nicht durchaus unüberwindlich; jedenfalls würden sie allein die Mängel der bisherigen „statistischen Methode" offenbar noch nicht rechtfertigen können, für principielle Fragen wie die vorliegende aber dürften meines Erachtens immer nur solche Entwickelungen in Betracht kommen, deren mathematische Berechtigung ausser Zweifel steht. Auf diese Andeutung muss ich mich vorläufig beschränken, hoffe aber, bei späterer Gelegenheit noch ausführlicher auf diese methodologischen Fragen zurückzukommen.

Aus den grossen Erfolgen der kinetischen Gastheorie in der Erklärung von *Zustands*beziehungen darf ihre völlige Durchführbarkeit, ihre Anwendbarkeit auch auf *zeitliche Vorgänge* nicht gefolgert werden, denn beides sind getrennte Gebiete; gibt sie uns auch auf dem einen ein in vielen Beziehungen zutreffendes und darum werthvolles Bild, so muss sie doch auf dem anderen, wo es sich vor allem um die Erklärung irreversibler Vorgänge handelt, ohne ganz neue Annahmen, das ist noch jetzt meine Ueberzeugung, nothwendig versagen.

Berlin, den 15. September 1896.

[1] *Boltzmann*, Gastheorie, p. 21. 1896. ⟦In the original main text, the footnote mark is missing.⟧

Like the general *principle* of irreversibility, the individual irreversible *processes* themselves cannot be explained on the basis of mechanical suppositions in the absence of new physical assumptions, at least as far as their time evolution is concerned. In particular, this is true for the differential equation of heat conduction and diffusion: $\partial u/\partial t = a^2 \partial^2 u/\partial x^2$, which, by its very nature, exclusively represents non-reversible processes. Hence, endeavors, such as those undertaken by *Clausius*, *Maxwell* and *Boltzmann*, among others, to derive this differential equation, too, solely from the basic equations of mechanics in conjunction with probabilistic considerations cannot succeed since they seek the impossible, and apparent success could only be due to faulty inference. One of the main flaws of the methods used seems to me to consist in the unprovable (because false) assumption that the molecular state of a gas is always "disordered", to use *Boltzmann*'s term,[8] and that all possible combinations and directions are everywhere equally represented, if we have nothing specific to say about the true state, which always depends on the "ordered" initial state after all. In my view, probability theory licenses assumptions of this kind to some extent only if they concern the *initial state*; the probability of later states, however, and hence that of the processes themselves, would always have to be expressed first in terms of those of the corresponding initial states. Only then would it be possible to decide on the legitimacy of such averaging assumptions. It may well prove very difficult to carry out an investigation along similar lines that is also *rigorous*, at least by probabilistic standards. But these difficulties do not seem to me to be insurmountable; in any event, they alone would obviously be insufficient to justify the shortcomings of the current "statistical" method. But I believe that in order to address basic questions such as the one considered here, only those developments may be taken into account whose mathematical legitimacy is beyond doubt. While I must limit myself to these suggestive remarks for now, I hope to be able to further elaborate on these methodological questions on a later occasion.

From the great successes of the kinetic theory of gases in explaining the relationships among *states* we must not deduce its general practicability, ⟦i.e.,⟧ its applicability also to *temporal processes*, since these are two separate domains; although, in the former domain, it offers an account that, in many respects, is accurate, and hence valuable, I am still convinced that it necessarily fails in the absence of entirely new assumptions in the latter domain, where mostly explanations of irreversible processes are concerned.

Berlin, September 15, 1896.

[8] *Boltzmann 1896a*, p. 21. ⟦In the original main text, the footnote mark is missing.⟧

Zu Hrn. Zermelo's Abhandlung „Ueber die mechanische Erklärung irreversibler Vorgänge"*

Boltzmann 1897

Ich will mich in der Duplik so kurz fassen, als es ohne Gefährdung der Klarheit möglich ist.

§ 1. Der 2. Hauptsatz wird mechanisch durch die natürlich unbeweisbare Annahme A erklärt, dass das Universum, wenn man es als mechanisches System auffasst, oder wenigstens ein sehr ausgedehnter, uns umgebender Theil desselben von einem sehr unwahrscheinlichen Zustande ausging und sich noch in einem solchen befindet. Wenn man daher ein kleineres System von Körpern in dem Zustande, in dem es sich gerade befindet, plötzlich von der übrigen Welt abschliesst, so befindet sich dasselbe vermöge der Annahme über den Zustand des Universums anfangs oft in einem ganz unwahrscheinlichen Zustande und dieser geht dann, solange das System abgeschlossen ist, in immer wahrscheinlichere über. Dagegen hat es eine an Unmöglichkeit grenzende Unwahrscheinlichkeit, dass das abgeschlossene System sich anfangs im Wärmegleichgewichte befand, und sich, während es abgeschlossen ist, soweit davon entfernt, dass seine Entropieverminderung wahrnehmbar wäre.

Es handelt sich also nicht um das Verhalten eines ganz beliebigen, sondern eines gerade dem jetzigen Weltzustande entnommenen Systems (l. c. p. 795). Dies hat der Anfangszustand vor den späteren Zuständen voraus (l. c. p. 798), wodurch Hrn. *Zermelo*'s Schluss entfällt, dass alle Punkte der H-Curve Maxima sein müssten (l. c. p. 798). Daher kommt es, dass die Entropie jedesmal zunimmt, sich Temperatur- und Concentrationsunterschiede ausgleichen (l. c. p. 795), dass der Anfangswerth des H ein solcher ist, der in beobachtbarer Zeit fast ausnahmslos abnimmt (l. c. p. 797), dass Anfangs- und Endzustand nicht vertauschbar sind (l. c. p. 799). Die Annahme A ist die nach den Gesetzen der Mechanik begreifliche physikalische Erklärung der Besonderheit der Anfangszustände (l. c. p. 799) oder besser ein einheitlicher, diesen Gesetzen
393 | entsprechender Gesichtspunkt, der die Art der Besonderheit des Anfangszustandes in jedem speciellen Falle voraussagen lässt; denn niemand wird verlangen, dass man das letzte Erklärungsprincip selbst wieder erkläre.

Würden wir dagegen über den gegenwärtigen Zustand des Universums keine Voraussetzung machen, so könnten wir natürlich nicht erwarten, dass sich das vom Universum abgetrennte System, dessen Anfangszustand dann ein ganz beliebiger wäre, eher anfangs als später in einem unwahrscheinlichen Zustande befinde. Dann wäre vielmehr zu erwarten, dass es sich schon im Momente der Abtrennung im Wärmegleichgewichte befindet. Unter den wenigen Fällen, wo dies nicht eintreffen würde, wären solche am häufigsten, wo der Zustand des Systems, wenn man ihn in der Zeit (immer im abgetrenn-

* *Zermelo,* Wied. Ann. **59**, p. 793. 1896.

On Mr. Zermelo's paper "On the mechanical explanation of irreversible processes"*

Boltzmann 1897

⟦The introductory note just before *1896a* also addresses *Boltzmann 1897*.⟧

I will keep this rejoinder as brief as is possible without compromising clarity.

§ 1. The second law receives a mechanical explanation by virtue of the assumption A, which is of course unprovable, that the universe, when considered as a mechanical system, or at least a very extensive part thereof surrounding us, started out in a highly improbable state and still is in such a state. Thus, if a smaller system of bodies in its current state is suddenly being closed off from the rest of the world, then, by virtue of the assumption about the state of the universe, it is often in a highly improbable state at first, which subsequently passes into ever more probable ones so long as the system remains closed off. In contrast, it is highly improbable, if not impossible, that the closed system was initially in a thermal equilibrium from which it moved away sufficiently, while still remaining closed off, so that its entropy decrease would be noticeable.

Hence, this is not a matter of the behavior of an entirely arbitrary system but of that of a system that has just been removed from the current state of the world (l. c. p. 795). This is the advantage that the initial state has over the later states (l. c. p. 798), whereby Mr. *Zermelo*'s conclusion that all points of the H-curve have to be maxima (l. c. p. 798) fails to apply. It is for this reason that entropy always increases, that differences in temperature and concentration balance out (l. c. p. 795), that the initial value of H is one that almost invariably decreases in observable time (l. c. p. 797), and that the initial and final states are not interchangeable (l. c. p. 799). Assumption A is the physical explanation, comprehensible in accordance with the laws of mechanics, of the peculiarity of the initial states (l. c. p. 799). Or, more aptly, it is a particular coherent stance corresponding to these laws from which it is possible to predict the kind of peculiarity of the initial state in each special case; for nobody demands that the ultimate principle of explanation itself be explained as well.

If, in contrast, we made no assumption about the current state of the universe, then we would of course not be able to expect that the system which is closed off from the universe and whose initial state would then be an entirely arbitrary one, would be in an improbable state at the beginning rather than at a later stage. In fact, in this case, we would have to expect that it is already in a thermal equilibrium at the moment of separation. Among the few cases in which this would not be true, the most frequent ones

* *Zermelo 1896b.*

ten Zustande) vor- oder rückwärts verfolgt, sich fast augenblicklich einem wahrscheinlicheren nähert. Noch weit seltener wären Fälle, wo der Zustand während längerer Zeit noch unwahrscheinlicher wird; diese aber wären ebenso häufig wie die, wo er die gleiche Zeit nach rückwärts verfolgt, noch unwahrscheinlicher wird.

§ 2. Die Anwendbarkeit der Wahrscheinlichkeitsrechnung auf einen bestimmten Fall kann natürlich niemals exact bewiesen werden. Wenn von 100 000 Objecten einer bestimmten Gattung jährlich etwa 100 durch Brand zerstört werden, so können wir nicht sicher schliessen, dass dies auch im nächsten Jahre eintreffen wird. Im Gegentheile, wenn die gleichen Bedingungen durch $10^{10^{10}}$ Jahre andauern würden, so würde es während dieser Zeit oft vorkommen, dass an einem Tage alle 100 000 Objecte gleichzeitig abbrennen, und auch dass während eines ganzen Jahres nicht ein einziges Object Schaden leidet. Trotzdem vertraut jede Versicherungsgesellschaft der Wahrscheinlichkeitsrechnung.

Um wie viel mehr scheint wegen der grossen Zahl der Molecüle in einem Cubikmillimeter die freilich für keinen einzigen speciellen Fall mathematisch beweisbare Annahme gerechtfertigt und allen unseren Erfahrungen entsprechend, dass, wenn zwei materiell verschiedene oder ungleich warme Gase in Berührung gebracht werden, jedes Molekül nicht nur im ersten Momente,
394 sondern während langer Zeit Molekülen | von den verschiedensten Zuständen begegnet, entsprechend den Wahrscheinlichkeitsgesetzen, welche durch die an der betreffenden Stelle herrschenden Mittelwerthe bestimmt sind. Diese Warscheinlichkeitsbetrachtungen können die directe Verfolgung der Bewegung jedes Moleküls zwar nicht ersetzen, aber wenn man von den verschiedensten, gleichen Mittelwerthen entsprechenden (also für die Beobachtung gleichen) Anfangsbedingungen ausgeht, so ist man berechtigt zu erwarten, dass die nach beiden Methoden erhaltenen Resultate bis auf einzelne Ausnahmen genügend übereinstimmen werden, welche relativ noch viel seltener sind, als im obigen Beispiele der Fall, dass alle 100 000 Objecte am selben Tage verbrennen. Die Annahme, dass diese seltenen Fälle in der Natur nicht zur Beobachtung kommen, ist nicht strenge beweisbar (strenge beweisbar ist das ganze mechanische Bild nicht), aber sie ist nach dem Gesagten so natürlich und naheliegend, so allen Erfahrungen über Wahrscheinlichkeiten von der Methode der kleinsten Quadrate bis zum Würfelspiel entsprechend, dass der Zweifel daran gewiss nicht die Berechtigung des Bildes, wenn es sonst brauchbar ist, in Frage stellen wird.

Ganz unbegreiflich aber ist es mir, wie man darin eine Widerlegung der Anwendbarkeit der Wahrscheinlichkeitsrechnung sehen kann, wenn irgendwelche andere Betrachtungen zeigen, dass innerhalb Aeonen hin und wieder Ausnahmen eintreten müssen; denn gerade das lehrt ja die Wahrscheinlichkeitsrechnung ebenfalls.

§ 3. Denken wir uns speciell plötzlich eine Scheidewand, welche zwei mit verschiedenartigen Gasen erfüllte Räume trennte, hinweggezogen. Man dürfte kaum bei irgend einer andern Gelegenheit (am wenigsten in allen Fällen, wo

would be those in which the state of the system approaches a more probable state almost instantaneously when it is traced forward or backward in time (always in the separated condition). Cases in which the state would become even more improbable during an extended period of time would be much rarer still; but these would be as frequent as those in which the state would become even more improbable when it is traced backward by the same period of time.

§ 2. The applicability of the calculus of probabilities to a particular case can of course never be proved with precision. If 100 out of 100,000 objects of a particular sort are consumed by fire per year, then we cannot infer with certainty that this will also be the case next year. On the contrary! If the same conditions continue to obtain for $10^{10^{10}}$ years, then it will often be the case during this period that the 100,000 objects are all consumed by fire at once on a single day, and even that not a single object suffers damage over the course of an entire year. Nevertheless, every insurance company places its faith in the calculus of probabilities.

Thus, on account of the great number of molecules in a cubic millimeter, it seems even more justified and consistent with our entire experience to assume that if two materially different, or unequally warm, gases are brought into contact, then, in accordance with the laws of probability, each molecule encounters molecules of the most varied states, which are determined by the average values at the place in question, not only in the first moment but also over a long period of time, even though this assumption cannot be proved mathematically for any special case. While these probabilistic considerations cannot replace the immediate tracking of the motion of each molecule, it is legitimate to expect, assuming that we start with initial conditions that correspond to the various equal average values (and hence are observationally equal), that the results obtained by using the two methods match one another with few exceptions, which are much rarer still than the case of the example above in which all 100,000 objects burn up on the same day. The assumption that these rare cases are not observable in nature cannot be rigorously proved (the entire mechanical picture cannot be rigorously proved). Yet from what has been said, it is so natural and obvious, so consistent with all experiences involving probabilities ranging from the method of least squares to the game of dice, that doubts concerning it certainly will not undermine the justification of the picture, provided it has other uses.

I find it incomprehensible, however, how it is possible to interpret some other observations according to which exceptions must occur once in a while within aeons as refuting the applicability of the calculus of probabilities; for this is precisely what the calculus of probabilities shows us as well.

§ 3. Let us assume, in particular, that a dividing wall separating two spaces each of which is filled with a different gas is suddenly lifted. Hardly any other occasion (and certainly not any of all those cases in which the method of

sich die Methode der kleinsten Quadrate bewährt) so viele voneinander unabhängige, in der verschiedensten Weise wirkende Ursachen haben, welche die Anwendung der Wahrscheinlichkeitsgesetze rechtfertigen. Die Ansicht, dass sich gerade hier die Wahrscheinlichkeitsgesetze nicht bewähren, dass in der Mehrzahl der Fälle die Moleküle nicht diffundiren, dass vielmehr fortwährend grosse Theile des Gefässes bedeutend mehr Sauerstoff-, andere wieder mehr Stickstoffmoleküle enthalten werden, kann und will ich nicht dadurch widerlegen, | dass ich die Bewegung von Trillionen von Molekülen in Millionen von 395
verschiedenen speciellen Fällen exact rechnend verfolge; so viel Berechtigung dürfte diese Ansicht sicher nicht haben, dass dadurch die Brauchbarkeit des Bildes, welches von der Annahme der Anwendbarkeit der Wahrscheinlichkeitsgesetze ausgeht und daraus die logischen Consequenzen zieht, in Frage gestellt würde.

Der *Poincaré*'sche Satz aber spricht nicht gegen, sondern insofern sogar für die Anwendbarkeit der Wahrscheinlichkeitsrechnung, da auch diese lehrt, dass in Aeonen wieder verhältnissmässig kurz dauernde Zeiten eintreten werden, während welcher die Zustandswahrscheinlichkeit, die Entropie des Gasgemisches wieder erheblich abnimmt, wo also wieder mehr geordnete, ja hie und da sogar dem Anfangszustande sehr ähnliche Zustände eintreten. In diesen enorm viel späteren Zeiten ist natürlich fortwährend jede bemerkbare Abweichung der Entropie von ihrem Maximalwerthe äusserst unwahrscheinlich, aber eine augenblickliche Zu- oder Abnahme derselben gleich wahrscheinlich.

Es ist auch in diesem Beispiele wieder klar, dass sich der Process in beobachtbarer Zeit deshalb in nicht umkehrbarer Weise abspielt, weil man absichtlich von einem ganz unwahrscheinlichen Zustand ausging. Bei den Naturvorgängen wird dies durch die Annahme erklärt, dass man das Körpersystem aus dem Universum ausscheidet, welches augenblicklich einen sehr unwahrscheinlichen Gesammtzustand hat.

Dieses Beispiel zweier anfangs vermischter Gase giebt uns sogar ein beiläufiges Bild, wie man sich den Anfangszustand der Welt zu denken hat. Denn wenn wir in dem Beispiele eine in einem kleineren Raume befindliche Gasmasse bald nach begonnener Diffusion von der übrigen Gasmasse isoliren, so wird sie bezüglich des Vor- und Rückschritts in der Zeit ganz dieselbe Einseitigkeit zeigen, wie das in § 1 isolirte Körpersystem.

§ 4. Ich habe selbst wiederholt gewarnt, einer Ausdehnung unserer Gedankenbilder über die Erfahrung hinaus zu sehr zu vertrauen und erinnert, dass man darauf gefasst sein muss, dass sich die Bilder der heutigen Mechanik und besonders die Auffassung der kleinsten Theilchen der Körper als materielle
396 | Punkte, als provisorisch herausstellen werden. Unter allen diesen Reserven aber kann derjenige, welcher dazu Lust hat, dem Drange nachgeben, sich specielle Vorstellungen über das Universum zu machen.

Man hat dann die Wahl zwischen zweierlei Vorstellungen. Man kann annehmen, dass sich das gesammte Universum gegenwärtig in einem sehr unwahrscheinlichen Zustande befindet. Man kann sich aber auch die Aeonen, innerhalb deren wieder unwahrscheinliche Zustände eintreten, winzig gegen

least squares has proved its worth) features so many mutually independent causes acting in manifold ways which justify the application of the laws of probability. The view that it is particularly here that the laws of probability do not prove their worth, that, in the majority of cases, the molecules will not diffuse and that, instead, some large parts of the container will always hold a significantly higher number of, say, oxygen molecules while others will hold a higher number of, say, nitrogen molecules, I neither can nor wish to refute by tracing through precise calculations the motions of quintillions of molecules in millions of different special cases; this view is certainly not sufficiently justified to undermine the usefulness of the picture, which proceeds from the assumption that the laws of probability are applicable and draws the logical conclusions from it.

In this respect, however, *Poincaré*'s theorem does not count against but even in favor of the applicability of the calculus of probabilities, since also the latter shows that within eons relatively short periods of time will recur during which the state probability, the entropy of the gas mixture, again significantly decreases, and hence when a greater number of ordered states occur again, some of which are even very similar to the initial state. During those very much later times any noticeable deviation of the entropy from its maximal value is of course always highly improbable while its instantaneous increase, or decrease, is equally probable.

As is obvious in this example, too, the reason why in observable time the process unfolds in an irreversible way is that we chose to assume an entirely improbable state. As for natural processes, this is explained by the assumption that that system of bodies is separated from the universe whose total state is highly improbable at the time.

This example of two initially mixed gases even suggests how to conceive of the initial state of the world. For if, in our example, we isolate a gas mass in a smaller space from the remaining gas mass soon after the start of the diffusion, then it will exhibit the very same partiality regarding the forward and backward movement in time as the isolated system of bodies of § 1.

§ 4. I myself have repeatedly cautioned against placing excessive trust in the extension of our mental pictures beyond experience and issued reminders that the pictures of contemporary mechanics, and in particular the conception of the smallest particles of bodies as material points, will turn out to be provisional. Given all these constraints, however, anyone feeling the urge to form particular ideas about the universe may surrender to it.

He then has the choice between two ideas. He may assume that the entire universe is currently in a highly improbable state. But he may also assume that the eons during which improbable states occur again are tiny compared

die Dauer, die Siriusfernen winzig gegen die Dimensionen des Universums denken. Es müssen dann im Universum, das sonst überall im Wärmegleichgewichte, also todt ist, hier und da solche verhältnissmässig kleine Bezirke von der Ausdehnung unseres Sternenraums, (nennen wir sie Einzelwelten) vorkommen, die während der verhältnissmässig kurzen Zeit von Aeonen erheblich vom Wärmegleichgewichte abweichen, und zwar ebenso häufig solche, in denen die Zustandswahrscheinlichkeit gerade zu- als abnimmt. Für das Universum sind also beide Richtungen der Zeit ununterscheidbar, wie es im Raum kein Oben oder Unten giebt. Aber wie wir an einer bestimmten Stelle der Erdoberfläche die Richtung gegen den Erdmittelpunkt als nach unten bezeichnen, so wird ein Lebewesen, das sich in einer bestimmten Zeitphase einer solchen Einzelwelt befindet, die Zeitrichtung gegen die unwahrscheinlicheren Zustände anders als die entgegengesetzte (erstere als die Vergangenheit, den Anfang, letztere als die Zukunft, das Ende) bezeichnen und vermöge dieser Benennung werden sich für dasselbe kleine Gebiete, die es aus dem Universum isolirt, „anfangs“ immer in einem unwahrscheinlichen Zustande befinden. Diese Methode scheint mir die einzige, wonach man den 2. Hauptsatz, den Wärmetod jeder Einzelwelt ohne eine einseitige Aenderung des ganzen Universums von einem bestimmten Anfangs- gegen einen schliesslichen Endzustand denken kann. Die Einwendung, dass ein Gedankenbild, welches so viel todte Theile des Universums zur Erklärung von so wenig belebten braucht, unökonomisch und daher unzweckmässig sei, lasse ich nicht gelten. Ich erinnere mich noch zu gut einer Person, welche absolut nicht glaubte, dass die Sonne 20 Millionen Meilen von der Erde entfernt sei, denn die Annahme
397 von so viel | nur Lichtäther enthaltenden Raum neben so wenig mit Leben erfülltem, sei einfach einfältig.

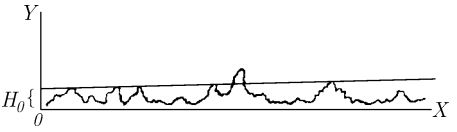
§ 5. Ob man sich in solchen Speculationen ergehen will, ist natürlich Geschmackssache. Von einer Wahl nach Geschmack zwischen der *Carnot-Clausius*'schen Fassung und dem mechanischen Bilde aber kann sicher nicht die Rede sein (l. c. p. 791). Die Wichtigkeit der ersteren als des einfachsten Ausdrucks der bisher beobachteten Thatsachen bestreitet niemand. Ich behaupte nur, dass das mechanische Bild in allem wirklich Beobachteten damit übereinstimmt. Dass es auf die Möglichkeit gewisser neuer Beobachtungen, z. B. über die Bewegung kleiner Körperchen in tropfbaren und gasförmigen Flüssigkeiten, über Reibung und Wärmeleitung in äusserst verdünnten Gasen etc. hinweist, dass es in uncontrollirbaren Fragen (z. B. über das Verhalten des Universums oder eines ganz abgeschlossenen Systems während unendlich langer Zeit) nicht mit der *Carnot-Clausius*'schen Fassung stimmt, mag man einen principiellen Unterschied nennen, jedenfalls scheint es kein Grund, das mechanische Bild aufzugeben, wie Hr. *Zermelo* (l. c. p. 794) meint, wenn es sich nicht, was nicht zu erwarten, principiell abändern lässt. Gerade dieser Unterschied scheint mir dafür zu sprechen, dass es die Allseitigkeit unserer Gedankenbilder fördern muss, neben den Consequenzen des Princips in der *Carnot-Clausius*'schen Fassung auch die des mechanischen Bildes zu studiren.

to the duration of the universe, that the Sirius distance is tiny compared to the scale of the universe. Here and there in the universe, which is in thermal equilibrium anywhere else, and hence dead, there must then exist such relatively small regions whose extension is that of our stellar world (let us call them solitary worlds), which considerably deviate from the thermal equilibrium during the relatively brief period of eons, and in particular as many of those with increasing state probability as of those with decreasing state probability. As for the universe, the two directions of time are thus indistinguishable, just as there is neither above nor below in space. However, just as we refer to the direction from a particular position on the earth's surface to its center as "downward", so a creature living in a certain time period in such a solitary world will refer to the direction of time towards the more improbable states as different from the opposite direction (by calling the former the past, or the beginning, the latter the future, or the end), and in this nomenclature, it will always consider small regions that it isolates from the universe to be in a highly improbable state "initially". This method seems to me the only method according to which it is possible to conceive of the second law, the heat death of each solitary world, without a one-sided change of the entire universe from a particular initial state towards an eventual final state. I refuse to grant the objection that a mental picture requiring so great a number of dead parts of the universe for the explanation of so small a number of animated parts is wasteful, and hence inexpedient. I still vividly remember someone who adamantly refused to believe that the sun's distance from the earth is 20 million miles on the ground that it would simply be foolish to assume so vast a space only containing light ether alongside so small a space filled with life.

§ 5. Whether one wishes to pursue such speculations is of course a matter of taste. But it is certainly not a matter of taste whether one opts for the *Carnot-Clausius* version or for the mechanical picture (l. c. p. 791). The significance of the former as the simplest expression of the facts hitherto observed is beyond dispute. My claim is only that the mechanical picture corresponds to it with respect to everything that can actually be observed. That it suggests the possibility of certain new observations about, e.g., the movements of smaller corpuscles in liquid and gaseous fluids, the friction and heat conduction in extremely dilute gases etc., or that it is not consistent with the *Carnot-Clausius* version regarding uncontrollable questions (e.g., about the behavior of the universe or of an entirely closed system during an infinitely long period of time) may be seen as a fundamental difference. But there is no reason to relinquish the mechanical picture, as Mr. *Zermelo* suggests, (l. c. p. 794) if it cannot be changed fundamentally, which is not to be expected. Precisely this difference seems to me to suggest that the full versatility of our mental pictures is only increased by also examining the implications of the mental picture alongside those of the principle in the *Carnot-Clausius* version.

Anhang.

§ 6. Die Wahrscheinlichkeit eines Zustandes maass ich seit jeher unabhängig vom zeitlichen Verlaufe durch die „Ausdehnung γ“ (l. c. p. 795) des ihm entsprechenden Gebietes, wozu ich schon seit 30 Jahren den *Liouville*'schen Satz benutze.[1] Der *Maxwell*'sche Zustand ist blos deshalb der wahrscheinlichste, weil er in der verschiedensten Weise realisirt sein kann. Die Gesammtausdehnung γ des Gebietes aller jener Zustände, für welche die Geschwindigkeitsvertheilung angenähert durch die *Maxwell*'sche Formel gegeben ist,
398 ist | also viel grösser, als die Gesammtausdehnung des Gebietes aller übrigen Zustände. Nur zur Versinnlichung der in den früheren Paragraphen geschilderten Beziehung zwischen dem zeitlichen Verlaufe der Zustände und deren Wahrscheinlichkeit stellte ich die reciproken Werthe der so gemessenen Wahrscheinlichkeit für die verschiedenen zeitlich sich folgenden Zustände durch die H-Curve dar, falls es sich um eine grosse, endliche Zahl von unendlich wenig deformirbaren Gasmolekülen handelt. Verschwindend wenige specielle Anfangszustände ausgenommen wird dann allerdings auch der wahrscheinlichste Zustand am häufigsten vorkommen (wenigstens bei einer sehr grossen Zahl von Molekülen). Die Ordinaten dieser Curve sind fast ausnahmslos sehr klein und diese kleinen Ordinaten sind natürlich nicht mit Vorliebe Maxima. Lediglich die Ordinaten von ganz ungewöhnlicher Grösse sind es meist und zwar um so wahrscheinlicher, je grösser sie sind. Dass eine sehr grosse Ordinate H_0 öfter einem Maximum als dem Durchschnittspunkte der Geraden $y = H_0$ mit einem noch grösseren Buckel entspricht (l. c. p. 797), kommt von der enormen Zunahme der Seltenheit der Buckel mit wachsender Höhe. Vgl. die nebenstehende Figur, die freilich sehr cum grano zu nehmen ist. Ei-

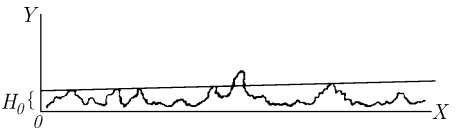


ne richtige Figur hätte der Zinkograph nicht herstellen können, da das, was wir die H-Curve nannten, auf jeder endlichen Strecke sehr viele Maxima und Minima hat und daher nicht durch einen Strich von continuirlich wechselnder Richtung darstellbar ist. Sie hiesse besser: Aggregat vieler sehr naher Punkte oder kleiner horizontaler Striche. Bezüglich des Näheren muss ich auf die „Nature“ 1894 und 1895 verweisen.

[1] Vgl. besonders Wien. Sitz. Ber. **58**, II. p. 517, 1868; **63**, II. p. 679 und p. 712, 1871; **66**, II, 1872; **76**, II. p. 373, 1877. Dort habe ich auch für die oben ohne Beweis angeführten Sätze die Beweise geliefert, die alle zu wiederholen hier natürlich der Platz fehlt.

Appendix.

§ 6. I have always measured the probability of a state independently of the time evolution by means of the "extension γ" (l. c. p. 795) of the area corresponding to it. To this end, I have been using *Liouville*'s theorem for 30 years.[1] The *Maxwell* state is the most probable one only because it can be realized in the greatest variety of ways. Hence, the total extension γ of the area of all those states for which the velocity distribution is approximately given by *Maxwell*'s formula is far greater than the total extension of the area of all remaining states. In order to illustrate the relationship between the time evolution of the states and their probability, which has been described in the previous paragraphs, I took the H-curve to represent the reciprocal values of the thus measured probability for the various states following one another in time, provided we are concerned with a large finite number of infinitely little deformable gas molecules. However, apart from vanishingly few special initial states, the most probable state will then occur most frequently (at least in the case of a very large number of molecules). With very few exceptions, the ordinates of this curve are very small, and, of course, these small ordinates tend not to be maxima. Only extraordinarily large ordinates are usually maxima, and in particular the larger they are, the more probable they become. That a very large ordinate H_0 corresponds to a maximum more often than to the average point of the line $y = H_0$ with an even larger hump (l. c. p. 797) is due to the enormous increase in rarity of humps as the height increases. Consider the adjacent figure, which, however, should be taken with



a large grain of salt. A zincographer would not have been able to produce a real figure since what we called the H-curve has a very large number of maxima and minima on each finite segment, and hence defies representation as a line of continuously changing direction. It would be more apt to call it an aggregate of many points very close to one another, or one of small horizontal lines. For further details, I refer to the 1894 and 1895 issues of "Nature".

[1] Cf. particularly *Boltzmann 1868*; *Boltzmann 1871a* and *1871b*; *Boltzmann 1872*; *Boltzmann 1877b*, where I have also supplied the proofs of the theorems that are stated above without proof, all of which cannot of course be reproduced here due to restrictions of space.

Auf einen in unseren Händen befindlichen irdischen Körper ist natürlich der *Poincaré*'sche Satz niemals anwendbar, weil kein solcher ganz abgeschlossen ist, ebensowenig auf ein ganz abgeschlossenes Gas der kinetischen Theorie, wenn man zuerst die Anzahl der Moleküle, dann erst den Quotienten der Zwischenzeit zwischen zwei benachbarten Zusammenstössen in die Beobachtungszeit unendlich werden lässt.

Wien, den 16. December 1896.

Poincaré's theorem is of course never applicable to terrestrial bodies which we can hold in our hands as none of them is entirely closed. Nor it is applicable to an entirely closed gas of the sort considered by the kinetic theory if first the number of molecules and only then the quotients of the intervals between two neighboring collisions in the observation time is allowed to become infinite.

Vienna, December 16, 1896.

Introductory note to 1899a

Rüdiger Thiele

The physical problem treated here is to find the equations of motion for a point system that is subject to auxiliary conditions (constraints). For an unconstrained point system, the expression P in (1) vanishes, since inside the brackets are the Newtonian equations of motion. In a system whose motion is constrained, however, P is a nonnegative function of location and time. Carl Friedrich Gauss referred to P as *constraint* (*Zwang*) in his paper "Über ein neues allgemeines Grundgesetz der Mechanik" ("On a new fundamental law of mechanics") (*1829*) and required, whenever free motion is impossible, that the constraint P be minimal. It is from the method of least squares that Gauss arrived at this requirement, as he himself noted in the paper.

Gauss's differential principle thus requires that for the actual motion, one have

$$P \equiv \sum_{1}^{n}{}_i \, m_i \Big(x_i'' - \frac{X_i}{m_i} \Big)^2 \equiv P(x_i''). \tag{1}$$

The Gaussian variations do not concern location, time, and velocity (whence $\partial X_i = 0$). This principle makes it possible to write down the equations of motion also for general constraints for which d'Alembert's principle fails.

In his analysis, Zermelo refers to a paper of Adolph Mayer (*1899*), in which the latter used Gauss's principle to obtain solutions for several cases. However, it is advantageous to know beforehand the existence and uniqueness of the solutions in order to apply the Gaussian principle to derive a solution that is known to be the unique solution. Mayer gave uniqueness proofs for some simple cases (two inequalities as constraints). To this end, he exploited his insight that at a given moment in time, only certain ("active") inequalities

H.-D. Ebbinghaus, A. Kanamori (Eds.), *Ernst Zermelo - Collected Works/ Gesammelte Werke II*, DOI 10.1007/978-3-540-70856-8_4, Schriften der Mathematisch-naturwissenschaftlichen Klasse der Heidelberger Akademie der Wissenschaften 23,

need be considered (see Zermelo's equations (3) and (4)). On the one hand, Zermelo generalizes these observations of Mayer, and on the other, gives his own proof, which he prefaces with a geometric interpretation. After a suitable transformation of P, these simple geometric considerations amount to obtaining the distance of a point from a convex surface (see equations $(1)'$ and $(4)'$). From a geometric point of view, uniqueness and minimality are obvious; these considerations can also be applied to function spaces.

Historically, it is worth noting that already in 1889, Mayer allowed inequalities $g_i(x) \geq 0$ $(i = 1, \ldots, r)$ as constraints in extremal problems, and through the introduction of new variables z_i $(i = 1, \ldots, r)$ and the device $z_i^2 - g_i = 0$, he converted them into equations. Frederick A. Valentine made use of this trick in his dissertation *1937*, and to this work can be traced the central requirements of optimization theory, known as the Kuhn-Tucker conditions, from the middle of the previous century (see *Klötzler 1994*, 314). The Gaussian principle of least constraint, that is, equation $(*)$, can be construed in connection with the constraints as a so-called Mayer problem (see, for example, *Giaquinta and Hildebrandt 1996*, vol. 1, 136f.), whereby one can obtain a connection with variational problems with constraints (Lagrange problems; see also *Giaquinta and Hildebrandt 1996*, vol. 1, 137).

Such problems were formulated by Lazar Aronovich Lyusternik (*1934*) in a functional-analytic form; there followed research in Banach spaces on optimization problems with inequalities as constraints (see, for example, *Ioffe and Tichomirov 1974*). For multiple integrals, there remain several open questions in this area, but for single integrals (that is, the case considered here), a satisfactory theory has meanwhile developed.

The treatment of problems with constraints goes hand in hand with a corresponding multiplier theorem. The first multiplier theorems can be traced to Leonhard Euler (*1744*) and Joseph Louis de Lagrange (*1788*). A significant gap in Lagrange's proof was filled by Mayer (*1895*). David Hilbert dealt with the topic in *1905*, and rigorous proofs were given by Constantin Carathéodory (*1935*) and Gilbert Ames Bliss (*1946*); see *Klötzler 1994*, 308–317.

Ueber die Bewegung eines Punktsystemes bei Bedingungsungleichungen

1899a

Vorgelegt von *D. Hilbert* durch den vorsitzenden Secretar in der Sitzung vom 3. Februar.

In seinem Vortrage vom 3. Juli v. J. behandelt Herr *A. Mayer* (Berichte d. k. Sächs. Ges. d. Wissensch, z. Leipzig 1899, S. 224–244) im Anschluß an eine ältere Arbeit von *Ostrogradski* (1834) das Problem, die Differentialgleichungen für die Bewegung eines reibungslosen Punktsystemes aufzustellen, das Bedingungsungleichungen unterworfen ist. Hier giebt er eine Methode, in jedem solchen Falle mit Hilfe des *Gauß*'schen „Princips des kleinsten Zwanges" die Beschleunigungen aller Punkte zu finden, wenn ihre augenblicklichen Lagen und Geschwindigkeiten gegeben sind. Doch ist diese Methode nur eine indirekte: es sind immer eine Anzahl von Fällen zu unterscheiden, je nach deren Eintreten die Schlußformeln verschieden ausfallen; welcher dieser Fälle aber gerade verwirklicht ist, kann im Allgemeinen erst aus dem Resultate entschieden werden, sodaß in der Regel immer erst eine Reihe von Ansätzen durchprobirt werden muß, bevor man einen findet, der allen Bedingungen genügt. Bei diesem Verfahren bedarf es indessen noch des Nachweises, daß eine solche Lösung immer existirt und die einzige ist, damit man überzeugt sein kann, daß das Gaußsche Prinzip zur Bestimmung der wirklichen Bewegung auch ausreicht. Dieser Eindeutigkeitsbeweis, der von Herrn *A. Mayer* nur für die einfachsten Fälle gegeben wird, in denen nicht mehr als zwei Bedingungsungleichungen in Betracht kommen (a. a. O. S. 237), ist der Gegenstand
307 meiner vorliegenden Mitteilung, und gelingt | sehr leicht auf Grund der Bemerkung, auf deren Wichtigkeit ich durch Herrn Prof. Hilbert aufmerksam gemacht wurde, daß die Nebenbedingungen des Minimumsproblems in Bezug auf die Unbekannten, die Beschleunigungen, sämtlich *linear* sind.

Die Coordinaten aller Punkte des Systemes bezeichnen wir der Reihe nach mit $x_1, x_2, \ldots x_n$, die zugehörigen Massen mit $m_1, m_2, \ldots m_n$ und die entsprechenden Componenten der äußeren Kräfte, die nur von den Coordinaten x_i abhängen sollen, mit $X_1, X_2, \ldots X_n$. Dann erfordert das Gaußsche Princip des kleinsten Zwanges, daß für jedes gegebene System (x_i, x_i') der Coordinaten und Geschwindigkeiten der Ausdruck

$$P \equiv \sum_1^n{}_i \, m_i \Big(x_i'' - \frac{X_i}{m_i} \Big)^2 \equiv P(x_i'') \tag{1}$$

durch Wahl der Beschleunigungen x_i'' zu einem Minimum gemacht werden soll, während die Coordinaten x_i zu *allen* Zeiten einer Anzahl von Ungleichungen genügen:

$$f_1(x_i) \leqq 0 \,, \quad f_2(x_i) \leqq 0 \,, \quad \ldots \tag{2}$$

On the motion of a point system with constraint inequalities

1899a

Presented by *D. Hilbert* through the presiding secretary at the meeting on February 3rd.

In his talk given on July 3rd of last year Mr. *A. Mayer* (*1899*), referring to an older paper by *Ostrogradski* (1834), examines the problem of establishing the differential equations for the motion of a frictionless system of points that is subject to constraint inequalities. He provides a method for finding the accelerations of all points by means of *Gauss*'s "principle of least constraint" in every such case where their momentary positions and velocities are given. This method is, however, merely indirect: there are always a number of cases to be distinguished, and the final formula will turn out differently in each case; but which case is currently realized can usually only be determined on the basis of the result so that normally we must always go through a series of ansatze before finding one satisfying all conditions. But when using this procedure we still need proof that such a solution always exists and that it is unique in order to convince ourselves that *Gauss*'s principle really suffices to determine the actual motion. This uniqueness proof, which Mr. *A. Mayer* only gives for the simplest cases with no more than two constraint inequalities (op. cit. p. 237), is the topic of my present note and goes through very easily because the auxiliary conditions of the minimum problem with respect to the unknowns, the accelerations, are all *linear*, an observation whose significance was pointed out to me by Prof. *Hilbert*.

We denote the coordinates of all points of the system successively by $x_1, x_2, \dots x_n$, the associated masses by $m_1, m_2, \dots m_n$ and the corresponding components of the external forces, which are supposed to only depend on the coordinates x_i, by $X_1, X_2, \dots X_n$. Then *Gauss*'s principle of least constraint requires that, for each given system (x_i, x_i') of the coordinates and velocities, the expression

$$P \equiv \sum_1^n{}_i \, m_i \Big(x_i'' - \frac{X_i}{m_i} \Big)^2 \equiv P(x_i'') \tag{1}$$

be turned into a minimum by choice of the accelerations x_i'', whereas the coordinates x_i *always* satisfy a number of inequalities:

$$f_1(x_i) \leqq 0 \,, \quad f_2(x_i) \leqq 0 \,, \quad \dots \tag{2}$$

Wie nun Herr Mayer a. a. O. zeigt, brauchen wir für einen gegebenen Bewegungszustand (x_i, x_i') nur diejenigen r Ungleichungen $f_\mu \leqq 0$ in Betracht zu ziehen, für welche im Augenblicke nicht nur die Functionen f_μ selbst sondern auch ihre nach der Zeit genommenen totalen ersten Ableitungen f_μ' gerade verschwinden, also die $2r$ Gleichungen bestehen:

$$f_\mu(x_i) = 0\ ,\quad f_\mu'(x_i, x_i') = 0 \qquad\qquad (\mu = 1, 2, \ldots r)\ , \tag{3}$$

und die Bedingungen (2) erfordern dann, wenn sie allgemein gültig bleiben sollen, daß auch

$$f_\mu''(x_i'') \equiv f_\mu''(x_i, x_i', x_i'') \equiv \sum_1^n{}_i\ \frac{\partial f_\mu}{\partial x_i} x_i'' + F_\mu(x_i, x_i') \leqq 0\ . \tag{4}$$

Ist uns also ein System von Größen (x_i, x_i') gegeben, das den Bedingungen (3) genügt, so sollen die Unbekannten x_i'' unter den Nebenbedingungen (4) den Ausdruck P zu einem Minimum machen.

Nach der Methode von Mayer geschieht die Bestimmung dieses Minimums folgendermaßen. Wir nehmen an, für einen *Teil* der Functionen f_μ möge in den Bedingungen (4) das Gleichheits-, für die übrigen das Ungleichheitszeichen gelten, so daß wir haben:

$$f_\mu''(x_i'') = 0 \qquad\qquad (\mu = 1, 2, \ldots k) \tag{4a}$$

$$f_\mu''(x_i'') < 0 \qquad\qquad (\mu = k+1, k+2, \ldots r)\ . \tag{4b}$$

308 Dann müssen im Falle des Minimums die Beschleunigungen von | der Form sein

$$m_i x_i'' = X_i - \sum_1^k{}_\mu\ \lambda_\mu \frac{\partial f_\mu}{\partial x_i} \qquad\qquad (i = 1, 2, \ldots n)\,, \tag{5}$$

und die k Factoren λ_μ lassen sich aus den k Gleichungen (4a) welche in Bezug auf die x_i'' und nach der Substitution (5) auch in Bezug auf die λ_μ *linear* sind, eindeutig bestimmen, vorausgesetzt, daß ihre Determinante Δ_k nicht verschwindet (vgl. Mayer a. a. O. S. 234). Diese Werte λ_μ müssen nun, wenn die gemachte Annahme über die Verteilung der Functionen f_μ in die Classen (4a) und (4b) richtig sein soll, selbst sämtlich $\geqq 0$ sein (Mayer S. 236) und müssen außerdem vermöge der Substitution (5) noch die $r-k$ Bedingungen (4b) befriedigen. Ist aber eine solche Lösung gefunden, so ist sie auch die einzige und das Problem ist vollständig gelöst. Denn dann nimmt der Ausdruck P, wie ich zeigen werde, in der That einen *kleineren* Wert an als für *alle anderen* Wertsysteme x_i'', die den Bedingungen (4) genügen, womit die Möglichkeit weiterer secundärer Minima ausgeschlossen und die *Eindeutigkeit* des Problemes bewiesen ist.

Now, as Mr. *Mayer* op. cit. shows, for a given state of motion (x_i, x_i'), we only need to consider those r inequalities $f_\mu \leqq 0$ for which not only the functions f_μ themselves but also their total first derivatives f_μ' taken with respect to time just vanish, and hence for which the following $2r$ equations hold:

$$f_\mu(x_i) = 0\ , \quad f_\mu'(x_i, x_i') = 0 \qquad\qquad (\mu = 1, 2, \ldots r)\ , \tag{3}$$

and conditions (2) then require, if they are to remain generally valid, that also

$$f_\mu''(x_i'') \equiv f_\mu''(x_i, x_i', x_i'') \equiv \sum_1^n{}_i \frac{\partial f_\mu}{\partial x_i} x_i'' + F_\mu(x_i, x_i') \leqq 0\ . \tag{4}$$

Hence, given a system of magnitudes (x_i, x_i') satisfying conditions (3), the unknowns x_i'' are supposed to turn the expression P into a minimum under auxiliary conditions (4).

Following *Mayer*'s method we determine this minimum as follows. We assume that, in conditions (4), the equality sign holds for a *part* of the functions f_μ, while the inequality sign holds for the remaining functions so that we have:

$$f_\mu''(x_i'') = 0 \qquad\qquad (\mu = 1, 2, \ldots k) \tag{4a}$$

$$f_\mu''(x_i'') < 0 \qquad\qquad (\mu = k+1, k+2, \ldots r)\ . \tag{4b}$$

For the minimum, the accelerations must then have the form

$$m_i x_i'' = X_i - \sum_1^k{}_\mu\ \lambda_\mu \frac{\partial f_\mu}{\partial x_i} \qquad\qquad (i = 1, 2, \ldots n)\,, \tag{5}$$

and the k factors λ_μ can be uniquely determined from the k equations (4a), which are *linear* with respect to the x_i'', and subsequent to substitution (5) also with respect to the λ_μ, provided that their determinant Δ_k does not vanish (cf. *Mayer* op. cit. p. 234). Now, these values λ_μ must all be $\geqq 0$ themselves, if the assumption concerning the grouping of the functions f_μ into classes (4a) and (4b) is to be correct (*Mayer 1899*, p. 236) and also satisfy the $r - k$ conditions (4b) by virtue of substitution (5). But when a solution of this kind is found, it is unique, and the problem is completely solved. For then, as I shall show, the expression P indeed assumes a *smaller* value than for *all other* value systems x_i'' satisfying conditions (4), whereby the possibility of further secondary minima is excluded and the *uniqueness* of the problem demonstrated.

Ersetzt man nämlich im Ausdrucke P jedes x_i'' durch $x_i'' + \omega_i$, so erhält man

$$\begin{aligned} P(x_i'' + \omega_i) &\equiv \sum_1^n{}_i\, m_i\Big(x_i'' - \frac{X_i}{m_i}\Big)^2 + 2\sum_1^n{}_i\, m_i\omega_i\Big(x_i'' - \frac{X_i}{m_i}\Big) + \sum_1^n{}_i\, m_i\omega_1^2 \\ &\equiv P(x_i'') + 2P_1(x_i'', \omega_1) + P_2(\omega_2)\ . \end{aligned} \tag{6}$$

Nun ist aber auf Grund von (5)

$$\begin{aligned} P_1(x_i'', \omega_i) &= -\sum_1^n{}_i\, m_i\omega_i \sum_1^k{}_\mu\, \frac{\lambda_\mu}{m_i}\frac{\partial f_\mu}{\partial x_i} \\ &= -\sum_1^k{}_\mu\, \lambda_\mu\left[f_\mu''(x_i + \omega_1) - f_\mu''(x_i)\right] \geqq 0\ , \end{aligned}$$

weil wegen (4)

$$f_\mu''(x'' + \omega_i) = f_\mu''(x_i'') + \sum_1^n{}_i\, \omega_i\frac{\partial f_\mu}{\partial x_i} \leqq 0$$

und nach (4a) für $\mu \leqq k$ $f_\mu''(x_i'') = 0$ sein soll, während alle λ_μ positiv sind. Daher ist, wie behauptet,

$$P(x_i'' + \omega_1) \geqq P(x_i'') + P_2(\omega_1) > P(x_i'')\ ,$$

sofern nicht alle Größen ω_i gleichzeitig verschwinden.

Ich füge noch einen zweiten Beweis für die Existenz eines *einzigen* Mi-
309 nimums hinzu, der eine besonders einfache geome- | trische Deutung zuläßt, wenn wir zunächst den Ausdruck P durch die lineare Substitution

$$\sqrt{m_i}x_i'' - \frac{X_i}{\sqrt{m_i}} = \xi_i \qquad\qquad (i = 1, 2, \ldots n)$$

auf die Form bringen

$$P \equiv \xi_1^2 + \xi_2^2 + \cdots + \xi_n^2 \equiv P(\xi_i)\ , \tag{1}'$$

während die Bedingungen (4) in die folgenden übergehen

$$\varphi_\mu(\xi_i) \equiv a_{\mu 1}\xi_1 + a_{\mu 2}\xi_2 + \cdots + a_{\mu n}\xi_n + A_\mu \leqq 0\ . \qquad (\mu = 1, 2, \ldots r)\ . \tag{4}'$$

Hier sind die r Functionen φ_μ wieder *linear* in Bezug auf die neuen Unbekannten ξ_i, repräsentiren also ein System von $(n-1)$dimensionalen Ebenen $\varphi_\mu = 0$, die aus dem $n$dimensionalen Raume der Coordinaten ξ_i einen Raumteil R' ausschneiden, der den Bedingungen (4)′ genügt. Die Aufgabe ist nun, die quadratische Form $P = \varrho^2$ unter den Bedingungen (4)′ oder den *Abstand* ϱ

For if we replace every x_i'' in the expression P by $x_i'' + \omega_i$, then we obtain

$$P(x_i'' + \omega_i) \equiv \sum_1^n{}_i\, m_i\Big(x_i'' - \frac{X_i}{m_i}\Big)^2 + 2\sum_1^n{}_i\, m_i\omega_i\Big(x_i'' - \frac{X_i}{m_i}\Big) + \sum_1^n{}_i\, m_i\omega_1^2$$
$$\equiv P(x_i'') + 2P_1(x_i'', \omega_1) + P_2(\omega_2)\ . \tag{6}$$

But now, by virtue of (5),

$$P_1(x_i'', \omega_i) = -\sum_1^n{}_i\, m_i\omega_i \sum_1^k{}_\mu\, \frac{\lambda_\mu}{m_i}\frac{\partial f_\mu}{\partial x_i}$$
$$= -\sum_1^k{}_\mu\, \lambda_\mu\left[f_\mu''(x_i + \omega_1) - f_\mu''(x_i)\right] \leqq 0\ ,$$

for, on account of (4),

$$f_\mu''(x'' + \omega_i) = f_\mu''(x_i'') + \sum_1^n{}_i\, \omega_i\frac{\partial f_\mu}{\partial x_i} \leqq 0$$

and, by (4a), we are supposed to have $f_\mu''(x_i'') = 0$ for $\mu \leqq k$, while all λ_μ are positive. Thus, as stated,

$$P(x_i'' + \omega_1) \geqq P(x_i'') + P_2(\omega_1) > P(x_i'')\ ,$$

provided not all magnitudes ω_i vanish at the same time.

As for the existence of a *unique* minimum, I add a second proof, which permits of a particularly simple geometric interpretation if, first, we bring the expression P by the linear substitution

$$\sqrt{m_i}x_i'' - \frac{X_i}{\sqrt{m_i}} = \xi_i \qquad\qquad (i = 1, 2, \ldots n)$$

into the form

$$P \equiv \xi_1^2 + \xi_2^2 + \cdots + \xi_n^2 \equiv P(\xi_i)\ , \tag{1}'$$

while conditions (4) are transformed as follows

$$\varphi_\mu(\xi_i) \equiv a_{\mu1}\xi_1 + a_{\mu2}\xi_2 + \cdots + a_{\mu n}\xi_n + A_\mu \leqq 0\ . \qquad (\mu = 1, 2, \ldots r)\ . \tag{4}'$$

Here, the r functions φ_μ are again *linear* with respect to the new unknowns ξ_i, and hence represent a system of $(n-1)$-dimensional planes $\varphi_\mu = 0$ that cut out a partial space R' satisfying conditions (4)′ from the n-dimensional space of the coordinates ξ_i. The problem is now to turn the quadratic form $P = \varrho^2$ under conditions (4)′, or the *distance ϱ of the point of*

des Coordinatenanfangspunktes von einem Punkte (ξ_i) *des Raumteiles* R' zu einem Minimum zu machen. *Einen* kürzesten Abstand ϱ_0 vom Punkte O muß R' jedenfalls besitzen, weil die „untere Grenze" einer stetigen Function in einem continuirlichen Gebiete immer zugleich auch ein Minimum darstellt. Es ist also nur noch zu zeigen, daß dieses Minimum $P_0 = \varrho_0^2$ nur an *einer* Stelle von R' angenommen wird, und dies folgt einfach aus dem Umstande, daß der Raumteil R' *einfach zusammenhängend* und *überall convex* ist. Liegen nämlich zwei Punkte (ξ_i') und (ξ_i'') auf derselben Seite einer der begrenzenden Ebenen, sodaß sie einer der Bedingungen $(4)'$ genügen, so gilt das gleiche auch von allen zwischen ihnen liegenden Punkten ihrer Verbindungsgeraden, und ebenso verhält es sich auch mit der Gesamtheit dieser r Bedingungen. Gehören also beide Punkte A', A'' dem Gebiete R' an, so liegt auch diese Verbindungsstrecke $A'A''$ ganz im Innern, und wenn beide den *gleichen* Abstand $\varrho' = \varrho''$ von O haben, so haben *alle* Punkte zwischen A' und A'' *kleinere* Abstände und ϱ' kann nicht der kürzeste sein. Dasselbe gilt umsomehr, wenn der Abstand ϱ'' von A'' schon kürzer ist als der von A', und dann giebt es auf der genannten Verbindungslinie in *beliebiger Nähe* von A' wieder Punkte mit Abständen $< \varrho'$. Es kann also auch keine *secundären* Minima geben, sondern nur ein einziges absolutes Minimum $P_0 = \varrho_0^2$, das nur an einer einzigen Stelle angenommen wird.

Analytisch gestaltet sich dieser Nachweis folgendermaßen. Genügen beide
310 Wertsysteme (ξ_i') und (ξ_i'') den Bedingungen $(4)'$, so | ist auch für alle Systeme

$$\xi_i = (1-\vartheta)\xi_i' + \vartheta\xi_i'' \qquad (i = 1, 2, .. n; 0 < \vartheta < 1)$$
$$\varphi_\mu(\xi_i) = (1-\vartheta)\varphi_\mu(\xi_i') + \vartheta\varphi_\mu(\xi_i'') \leqq 0 \,.$$

Ferner wird dann

$$P(\xi_i) = (1-\vartheta)^2 P(\xi_i') + \vartheta^2 P(\xi_i'') + 2\vartheta(1-\vartheta)P_1(\xi_i', \xi_i'') \,,$$

wo $P_1(\xi_i', \xi_i'')$ ähnlich wie in (6) eine bilineare Form (die Polarform von P) ist, also, wenn

$$\varrho''^2 = P(\xi_i'') \leqq P(\xi_i') = \varrho'^2$$

angenommen wird,

$$P(\xi_i) = P(\xi_i') - \vartheta\,[P(\xi_i') - P(\xi_i'')] - \vartheta(1-\vartheta)\,[P(\xi_i') + P(\xi_i'') - 2P_1(\xi_i', \xi_i'')]$$
$$< P(\xi_i) \,,$$

weil

$$P(\xi_i') + P(\xi_i'') - 2P_1(\xi_i', \xi_i'') \equiv P(\xi_i' - \xi_i'')$$

als definite quadratische Form der n Variablen $\xi_i' - \xi_i''$ positiv ist.

origin of the coordinates from a point (ξ_i) *of the partial space* R' into a minimum. In any case, R' must have *one* short distance ϱ_0 from the point O since the "lower limit" of a continuous function in a continuous region also always represents a minimum. It therefore only remains to show that this minimum $P_0 = \varrho_0^2$ only occurs at *one* location of R', and this simply follows from the fact that the partial space R' is *simply connected* and *everywhere convex*. For if two points (ξ_i') and (ξ_i'') lie on the same side of one of the bounding planes so that they satisfy one of the conditions $(4)'$, then the same holds true of all points lying between them along their connecting straight line, and likewise also for the totality of these r conditions. Hence, if the two points A', A'' both belong to the region R', then the entire connecting line segment $A'A''$ lies in the interior, and if both are at the *same* distance $\varrho' = \varrho''$ from O, then *all* points between A' and A'' are at *smaller* distances, and ϱ' cannot be the shortest one. The same is, in particular, true in the case where the distance ϱ'' of A'' is already shorter than that of A', and then along the said connecting straight line there exist points with distances $< \varrho'$ that are *arbitrarily close* to A'. Hence, there cannot be any *secondary* minima either, but only a single, absolute minimum $P_0 = \varrho_0^2$ that is realized only at a single location.

In analytic terms, this proof proceeds as follows. If the two systems of values (ξ_i') and (ξ_i'') satisfy conditions $(4)'$, then we also have for all systems

$$\xi_i = (1-\vartheta)\xi_i' + \vartheta\xi_i'' \qquad (i = 1, 2, .. n; 0 < \vartheta < 1)$$
$$\varphi_\mu(\xi_i) = (1-\vartheta)\varphi_\mu(\xi_i') + \vartheta\varphi_\mu(\xi_i'') \leqq 0 \, .$$

Moreover, we then get

$$P(\xi_i) = (1-\vartheta)^2 P(\xi_i') + \vartheta^2 P(\xi_i'') + 2\vartheta(1-\vartheta)P_1(\xi_i', \xi_i'') \, ,$$

where $P_1(\xi_i', \xi_i'')$ is a bilinear form (the polar form of P), like in (6), and hence, if we assume

$$\varrho''^2 = P(\xi_i'') \leqq P(\xi_i') = \varrho'^2 \, ,$$

also

$$P(\xi_i) = P(\xi_i') - \vartheta\,[P(\xi_i') - P(\xi_i'')] - \vartheta(1-\vartheta)\,[P(\xi_i') + P(\xi_i'') - 2P_1(\xi_i', \xi_i'')]$$
$$< P(\xi_i) \, ,$$

for

$$P(\xi_i') + P(\xi_i'') - 2P_1(\xi_i', \xi_i'') \equiv P(\xi_i' - \xi_i'')$$

is positive, being a definite quadratic form of the n variables $\xi_i' - \xi_i''$.

Introductory note to s1899b

Rüdiger Thiele

The manuscript division of the Göttingen State and University Library has in its possession, in David Hilbert's *Nachlass*, two sheets of paper written by Zermelo, item *s1899b*, bearing the title "Wie bewegt sich ein unausdehnbarer materieller Faden unter dem Einfluß von Kräften mit dem Potentiale $W(x, y, z)$?" ("How does an inextensible material string move under the action of forces with potential $W(x, y, z)$?") (Cod. Ms. D. Hilbert 722). Hilbert's *Nachlass* contains manuscripts of varying lengths by nearly sixty mathematicians. The pages of *s1899b* are in Zermelo's hand, and aside from the title, the text is in German handwriting. There are at least two additions by Hilbert: On the first page, Hilbert notes at the top, "Zermelo 1899", and in the bottom margin, Hilbert adds a sketch, writing, "Klein-example, that"; however, the last addition was again crossed out by Hilbert. Perhaps the entry "$\lambda = \lambda(t, s)$ Lagrange Factor" on the second page was made by Hilbert as well. This is supported by Hilbert's interest in the multiplier rule, which is revealed already in his correspondence with Adolph Mayer in the year 1886 (Cod. Ms. D. Hilbert 246), in which Mayer improves some associated proofs of Hilbert. A similar comment "κ line density" on the first page seems, however, not to be in Hilbert's hand.

How did this manuscript come into Hilbert's possession? In 1895, Hilbert had arrived in Göttingen as a professor; Zermelo arrived two years later for his *Habilitation*, which was completed in 1899. In Göttingen, Hilbert and Zermelo maintained an intensive scientific exchange, so described by Hilbert, for example in his May 1903 assessment of Zermelo concerning a professorship at the University of Breslau (cf. *Ebbinghaus 2007*, 36, 277):

> Before Minkowski came here [...], Zermelo was my main mathematical company, and I have learnt a lot from him, for example, the Weierstrassian calculus of variations.[1]

On the occasion of Zermelo's *Habilitation*, Hilbert had the opportunity to refer to Zermelo's activities in the Göttingen Mathematical Society. For example, shortly after his arrival in Göttingen in the winter semester of 1897, Zermelo had given a lecture on the principles of mechanics (which at the time was also a topic of interest to Hilbert) and had lectured on Adolf Kneser's paper "Zur Variationsrechnung" ("On the calculus of variations") (*1898*). (By

[1] Bevor Mink⟦owski⟧ herkam [...], war Zer⟦melo⟧ mein hauptsächl⟦icher⟧ math⟦ematischer⟧ Verkehr und ich habe viel von ihm gelernt, z.B. die Weierstraßsche Variationsrechung.

H.-D. Ebbinghaus, A. Kanamori (Eds.), *Ernst Zermelo - Collected Works/ Gesammelte Werke II*, DOI 10.1007/978-3-540-70856-8_5, Schriften der Mathematisch-naturwissenschaftlichen Klasse der Heidelberger Akademie der Wissenschaften 23,

the way, Kneser also published on the equilibrium of hanging heavy strings (*1903b*).)

In the summer semester 1898, Hilbert directed the two-hour advanced seminar on selected topics in mechanics, conducted jointly with Felix Klein. Thus it is possible that the two pages written by Zermelo originated in this seminar; the sketch with the mention of Klein would support such an assumption.

Both Zermelo and Hilbert doubtless had applications for the topic treated on the two pages in their lectures on the calculus of variations before the First World War (Hilbert in the summer semester 1898 and the winter semester 1904/1905; Zermelo in the summer semester 1902 and the winter semester 1907/1908); Hilbert also treated this topic in his lectures on mechanics (winter semester 1898/1899 and winter semester 1905/1906), and moreover, it was a subject in his lecture "Mechanics of continua" (winter semester 1902/1903) and in the "Exercises on mechanics" conducted jointly with Hermann Minkowski (winter semester 1904/1905). Finally, in his 1900 Paris lecture, Hilbert provided a prominent place for the calculus of variations (introduction and Problem 23), an interest that Zermelo's extensive knowledge of the subject and of mathematical physics had stimulated (cf. *Ebbinghaus 2007*, 276f.).

Such problems in the calculus of variations were designated Lagrange problems, in which extremes are sought under constraints, wherein traditionally, finite equations, differential/integral equations, and mixed conditions are distinguished; for inequalities as constraints, see the introductory note to *Zermelo 1899a*. The Euler–Lagrange multiplier theorem in the calculus of variations appeared first in the work of Euler, in the second chapter of his *Methodus inveniendi* (*1744*). Joseph Louis de Lagrange spoke very generally of the multiplier rule in his *Méchanique analytique* (*1788*).

Zermelo treats the problem of motion of a string in a potential field W with the help of Hamilton's principle, (1). A string is an elastic physical body with small cross section. It can therefore be represented by a continuous curve $r = r(t) = (x(t), y(t), z(t))$ that can assume every possible position. A string is therefore completely flexible, but by assumption inextensible. The kinetic and potential energies for a line segment ds are respectively

$$\frac{1}{2}\kappa v^2 ds = \frac{1}{2}\kappa \left(\frac{\partial r}{\partial t}\right)^2 ds \quad \text{and} \quad W ds\,;$$

the constraints of the string's length, assumed to be fixed, are $S = 1$ and $\partial S = 0$, respectively. For inextensible strings that are fixed at both ends, the equations of motion (the Euler equations) become, by the multiplier method for the Lagrange function L, $L = (T - U) + \lambda(S - 1)$.

The Lagrange multiplier λ can be physically interpreted as the tension of the string. In the calculations, integration, differentiation, and variation are interchangeable.

Wie bewegt sich ein unausdehnbarer materieller Faden unter dem Einfluss von Kräften mit dem Potentiale $W(x, y, z)$?*

s1899b

Nach dem Hamiltonschen Prinzipe ist:

$$\delta \int_{t_1}^{t_2} (T - U)\, dt = 0, \tag{1}$$

wo, wenn wir die Zeit t und die Bogenlänge s als unabhängige Variable einführen,

$$T = \int_{s_1}^{s_2} \frac{\kappa}{2}\Big[\Big(\frac{\partial x}{\partial t}\Big)^2 + \Big(\frac{\partial y}{\partial t}\Big)^2 + \Big(\frac{\partial z}{\partial t}\Big)^2\Big] ds \quad (\kappa \text{ Liniendichte})$$

$$U = \int_{s_1}^{s_2} W ds,$$

$$\delta s = 0$$

und

$$S = \Big(\frac{\partial x}{\partial s}\Big)^2 + \Big(\frac{\partial y}{\partial s}\Big)^2 + \Big(\frac{\partial z}{\partial s}\Big)^2 = 1. \tag{2}$$

Wir haben also:

21 | $$\partial \int_{t_1}^{t_2} (T - U)\, dt = \int_{t_1}^{t_2} dt \int_{s_1}^{s_2} ds \Big[\kappa \frac{\partial x}{\partial t}\frac{\partial \delta x}{\partial t} + \kappa \frac{\partial y}{\partial t}\frac{\partial \delta y}{\partial t} + \kappa \frac{\partial z}{\partial t}\frac{\partial \delta z}{\partial t} - \frac{\partial W}{\partial x}\delta x - \frac{\partial W}{\partial y}\delta y - \frac{\partial W}{\partial z}\delta z\Big] = 0$$

oder, wenn man nach t partiell integriert:

$$0 = \Big[\frac{\kappa}{2}\int_{s_1}^{s_2} ds \Big(\frac{\partial x}{\partial t}\delta x + \frac{\partial y}{\partial t}\delta y + \frac{\partial z}{\partial t}\delta z\Big)\Big]_{t_1}^{t_2}$$

$$- \int_{t_1}^{t_2} dt \int_{s_1}^{s_2} ds \Big[\kappa\Big(\frac{\partial^2 x}{\partial t^2}\delta x + \frac{\partial^2 y}{\partial t^2}\delta y + \frac{\partial^2 z}{\partial t^2}\delta z\Big) + \frac{\partial W}{\partial x}\delta x + \frac{\partial W}{\partial y}\delta y + \frac{\partial W}{\partial z}\delta z\Big].$$

When the potential W results only from the effect of gravity, then the Euler equations yield for the curve of a homogeneous heavy string, the so-called catenary (hyperbolic cosine).

How does an inextensible material string move under the action of forces with potential $W(x, y, z)$?

s1899b

According to Hamilton's principle we have:

$$\delta \int_{t_1}^{t_2} (T - U)\, dt = 0, \tag{1}$$

where, if we introduce time t and arch length s as independent variables,

$$T = \int_{s_1}^{s_2} \frac{\kappa}{2}\Big[\Big(\frac{\partial x}{\partial t}\Big)^2 + \Big(\frac{\partial y}{\partial t}\Big)^2 + \Big(\frac{\partial z}{\partial t}\Big)^2\Big] ds \quad (\kappa \text{ line density})$$

$$U = \int_{s_1}^{s_2} W ds,$$

$$\delta s = 0$$

and

$$S = \Big(\frac{\partial x}{\partial s}\Big)^2 + \Big(\frac{\partial y}{\partial s}\Big)^2 + \Big(\frac{\partial z}{\partial s}\Big)^2 = 1. \tag{2}$$

Hence, we have:

$$\partial \int_{t_1}^{t_2} (T - U)\, dt = \int_{t_1}^{t_2} dt \int_{s_1}^{s_2} ds \Big[\kappa \frac{\partial x}{\partial t}\frac{\partial \delta x}{\partial t} + \kappa \frac{\partial y}{\partial t}\frac{\partial \delta y}{\partial t} + \kappa \frac{\partial z}{\partial t}\frac{\partial \delta z}{\partial t} - \frac{\partial W}{\partial x}\delta x - \frac{\partial W}{\partial y}\delta y - \frac{\partial W}{\partial z}\delta z\Big] = 0$$

or, performing partial integration with respect to t:

$$0 = \Big[\frac{\kappa}{2}\int_{s_1}^{s_2} ds \Big(\frac{\partial x}{\partial t}\delta x + \frac{\partial y}{\partial t}\delta y + \frac{\partial z}{\partial t}\delta z\Big)\Big]_{t_1}^{t_2}$$

$$- \int_{t_1}^{t_2} dt \int_{s_1}^{s_2} ds \Big[\kappa\Big(\frac{\partial^2 x}{\partial t^2}\delta x + \frac{\partial^2 y}{\partial t^2}\delta y + \frac{\partial^2 z}{\partial t^2}\delta z\Big) + \frac{\partial W}{\partial x}\delta x + \frac{\partial W}{\partial y}\delta y + \frac{\partial W}{\partial z}\delta z\Big].$$

Hier verschwinden die Grenzzylinder ($t = t_1$, $t = t_2$) ohne weiteres nach der Variationsmethode des Ham. Princ., das Integral aber muß verschwinden unter der Nebenbedingung $S = 1$ oder $\delta S = 0$, oder

$$0 = \delta \int_{t_1}^{t_2} dt \int_{s_1}^{s_2} \lambda\, ds\, S = \delta \int_{t_1}^{t_2} dt \int_{s_1}^{s_2} \lambda \Big[\Big(\frac{\partial x}{\partial s}\Big)^2 + \Big(\frac{\partial y}{\partial s}\Big)^2 + \Big(\frac{\partial z}{\partial s}\Big)^2 \Big] ds$$

$$= \int_{t_1}^{t_2} dt \int_{s_1}^{s_2} \lambda \Big[\frac{\partial x}{\partial s}\frac{\partial \delta x}{\partial s} + \frac{\partial y}{\partial s}\frac{\partial \delta y}{\partial s} + \frac{\partial z}{\partial s}\frac{\partial \delta z}{\partial s} \Big] ds$$

$$= \int_{t_1}^{t_2} dt \Big[\lambda \Big(\frac{\partial x}{\partial s}\delta x + \frac{\partial y}{\partial s}\delta y + \frac{\partial z}{\partial s}\delta z \Big) \Big]_{s_1}^{s_2}$$

$$- \int_{t_1}^{t_2} dt \int_{s_1}^{s_2} ds \Big[\partial x \frac{\partial}{\partial s}\Big(\lambda \frac{\partial x}{\partial s}\Big) + \partial y \frac{\partial}{\partial s}\Big(\lambda \frac{\partial y}{\partial s}\Big) + \partial z \frac{\partial}{\partial s}\Big(\lambda \frac{\partial z}{\partial s}\Big) \Big]$$

$\lambda = \lambda(t, s)$ (Lagranges Factor)

wo ebenfalls die Grenzzylinder ($s = s_1$, $s = s_2$) in der Regel verschwinden (d. h. wenn der Faden an den beiden Enden *befestigt* ist).

2r | So erhalten wir schließlich nach der Multiplicatoren-Methode

$$0 = \int_{t_1}^{t_2} dt \int_{s_1}^{s_2} ds \Big\{ \partial x \Big[\kappa \frac{\partial^2 x}{\partial t^2} + \frac{\partial W}{\partial x} + \frac{\partial}{\partial s}\Big(\lambda \frac{\partial x}{\partial s}\Big) \Big]$$

$$+ \partial y \Big[\kappa \frac{\partial^2 y}{\partial t^2} + \frac{\partial W}{\partial y} + \frac{\partial}{\partial s}\Big(\lambda \frac{\partial y}{\partial s}\Big) \Big]$$

$$+ \partial z \Big[\kappa \frac{\partial^2 z}{\partial t^2} + \frac{\partial W}{\partial z} + \frac{\partial}{\partial s}\Big(\lambda \frac{\partial z}{\partial s}\Big) \Big] \Big\}$$

für unabhängig willkürliche Variationen $\partial x, \partial y, \partial z$ (als Funktionen von s und t), und die Bewegungsgleichungen werden:

$$\kappa \frac{\partial^2 x}{\partial t^2} + \frac{\partial}{\partial s}\Big(\lambda \frac{\partial x}{\partial s}\Big) + \frac{\partial W}{\partial x} = 0$$

$$\kappa \frac{\partial^2 y}{\partial t^2} + \frac{\partial}{\partial s}\Big(\lambda \frac{\partial y}{\partial s}\Big) + \frac{\partial W}{\partial y} = 0$$

$$\kappa \frac{\partial^2 z}{\partial t^2} + \frac{\partial}{\partial s}\Big(\lambda \frac{\partial z}{\partial s}\Big) + \frac{\partial W}{\partial z} = 0$$

Hier ist κ die Masse des Fadens in der Längeneinheit, die "Längendichte" und der Lagrangesche Factor $\lambda = \lambda(t, s)$ ein Maß für die Spannung des Fadens.

* ⟦The second page of the manuscript has double the breadth and contains two columns which are indicated on the margin by "2l" and "2r".⟧

Here the boundary cylinders ($t = t_1$, $t = t_2$) simply vanish according to the method of variations of Hamilton's principle. The integral, however, must vanish under the auxiliary condition $S = 1$ or $\delta S = 0$, or

$$0 = \delta \int_{t_1}^{t_2} dt \int_{s_1}^{s_2} \lambda \, ds \, S = \delta \int_{t_1}^{t_2} dt \int_{s_1}^{s_2} \lambda \left[\left(\frac{\partial x}{\partial s} \right)^2 + \left(\frac{\partial y}{\partial s} \right)^2 + \left(\frac{\partial z}{\partial s} \right)^2 \right] ds$$

$$= \int_{t_1}^{t_2} dt \int_{s_1}^{s_2} \lambda \left[\frac{\partial x}{\partial s} \frac{\partial \delta x}{\partial s} + \frac{\partial y}{\partial s} \frac{\partial \delta y}{\partial s} + \frac{\partial z}{\partial s} \frac{\partial \delta z}{\partial s} \right] ds$$

$$= \int_{t_1}^{t_2} dt \left[\lambda \left(\frac{\partial x}{\partial s} \delta x + \frac{\partial y}{\partial s} \delta y + \frac{\partial z}{\partial s} \delta z \right) \right]_{s_1}^{s_2}$$

$$- \int_{t_1}^{t_2} dt \int_{s_1}^{s_2} ds \left[\partial x \frac{\partial}{\partial s} \left(\lambda \frac{\partial x}{\partial s} \right) + \partial y \frac{\partial}{\partial s} \left(\lambda \frac{\partial y}{\partial s} \right) + \partial z \frac{\partial}{\partial s} \left(\lambda \frac{\partial z}{\partial s} \right) \right]$$

$\lambda = \lambda(t, s)$ (Lagrange's factor)

where the boundary cylinders ($s = s_1$, $s = s_2$), too, usually vanish (i. e. when the string is *fixed* at both ends).

Using the multiplier method, we eventually get

$$0 = \int_{t_1}^{t_2} dt \int_{s_1}^{s_2} ds \Bigg\{ \partial x \left[\kappa \frac{\partial^2 x}{\partial t^2} + \frac{\partial W}{\partial x} + \frac{\partial}{\partial s} \left(\lambda \frac{\partial x}{\partial s} \right) \right]$$

$$+ \partial y \left[\kappa \frac{\partial^2 y}{\partial t^2} + \frac{\partial W}{\partial y} + \frac{\partial}{\partial s} \left(\lambda \frac{\partial y}{\partial s} \right) \right]$$

$$+ \partial z \left[\kappa \frac{\partial^2 z}{\partial t^2} + \frac{\partial W}{\partial z} + \frac{\partial}{\partial s} \left(\lambda \frac{\partial z}{\partial s} \right) \right] \Bigg\}$$

for independently arbitrary variations $\partial x, \partial y, \partial z$ (as functions of s and t), and the equations of motion become:

$$\kappa \frac{\partial^2 x}{\partial t^2} + \frac{\partial}{\partial s} \left(\lambda \frac{\partial x}{\partial s} \right) + \frac{\partial W}{\partial x} = 0$$

$$\kappa \frac{\partial^2 y}{\partial t^2} + \frac{\partial}{\partial s} \left(\lambda \frac{\partial y}{\partial s} \right) + \frac{\partial W}{\partial y} = 0$$

$$\kappa \frac{\partial^2 z}{\partial t^2} + \frac{\partial}{\partial s} \left(\lambda \frac{\partial z}{\partial s} \right) + \frac{\partial W}{\partial z} = 0$$

Here κ is the mass of the string per unit of length, the "length density", and the Lagrange factor $\lambda = \lambda(t, s)$ is a measure of the tension of the string.

Introductory note to 1900

Jos Uffink

Zermelo returned to the foundations of statistical physics in his *Habilitation* lecture for the University of Göttingen. Here, he refers to Boltzmann's assumption of molecular disorder, but without attempting to describe what Boltzmann understood by this notion. Zermelo understood this and similar assumptions as implying that various stages of molecular motion can be seen as probabilistically independent. At least, this understanding seems to be suggested by the concern he expressed, namely, that due to the assumed equations of motion in the theory of gases, or dynamical systems in general, these various stages are *not* mutually independent but functionally related.

Zermelo proposes a definition of a probability measure (say ρ on phase space) by taking the probability of any region or open set A (which Zermelo calls g) to be the Lebesgue measure μ of that set divided by the total measure of the phase space: $\rho(A) = \mu(A)/\mu(\Gamma)$.

Zermelo then presents three theorems. The first is actually just the well-known Liouville theorem, and does not need comment. The other two are more interesting since they deal again with the question whether any entropy-like function could be defined on phase space which might have the property of changing monotonically in time.

Zermelo's second theorem applies to the case when we take A to be invariant under time evolution, i.e.: $\forall t \in \mathbb{R} : T_t A = A$. For example, A might be the energy hypersurface $\{x \in \Gamma \mid H(x) = \text{const.}\}$ (in which case the probability measure becomes the microcanonical one) or an energy hypershell $\{x \in \Gamma \mid a < H(x) < b\}$ etc. The theorem states: For every differentiable function S on phase space, the phase average value of $\dfrac{dS}{dt}$ over such an invariant region A is zero. In other words:

$$\frac{1}{\mu(\Gamma)} \int_A \frac{dS(x_t)}{dt} d\mu(x) = 0\,. \tag{1}$$

H.-D. Ebbinghaus, A. Kanamori (Eds.), *Ernst Zermelo - Collected Works/ Gesammelte Werke II*, DOI 10.1007/978-3-540-70856-8_6, Schriften der Mathematisch-naturwissenschaftlichen Klasse der Heidelberger Akademie der Wissenschaften 23,

Zermelo's third theorem extends this result to arbitrary open regions A that need not be invariant.

These last two theorems are analogous to the lemma (Theorem 3.2 on p. 206 of this volume) he presented in his *1896a*. Of course, the difference is that in the present manuscript the result is restricted to differentiable functions, but otherwise both theorems provided here imply the same conclusion: for any such function S, there can be no open set A such that for all $x \in A$, $S(T_t x)$ would be monotonically increasing in t. The present results however seem to be stronger than the lemma based on the recurrence theorem, since the mere non-existence of monotonically increasing functions on an open set A does not by itself imply the vanishing of the phase-average of their derivative.

Zermelo refrained from drawing any conclusion that was critical of Boltzmann. He only stated that these results "may serve as a basis for further considerations by showing us which probabilities and mean values we are permitted to generally replace by other, simpler, ones." Of course, it is hard (at least for me) to resist the temptation of thinking that his reference to other but unspecified "simpler" notions of probability refers to Boltzmann's equivocation of phase-average probabilities and time-average probabilities that he questioned in his response to Boltzmann. But he does not go into details.

Zermelo ends by announcing an intention to attempt a new derivation of Maxwell's distribution law "solely on the basis of the definition of probability given here without any further hypothesis." This intention does not seem to have materialized.

All in all, I would say that this paper contains little of crucial importance to the current debate on the foundations of statistical physics. What it shows is that Zermelo was not in the least bit moved by Boltzmann's replies. He however refrained from mounting or reiterating explicit criticisms of Boltzmann. Zermelo may have chosen to avoid explicitly raising controversy in his *Habilitation* lecture.

Über die Anwendung der Wahrscheinlichkeitsrechnung auf dynamische Systeme*

1900

Auf verschiedenen Gebieten der mathematischen Physik, insbesondere in der kinetischen Gastheorie, entsteht das Problem, die Bewegung solcher mechanischen Systeme zu untersuchen, welche eine zwar endliche, aber sehr grosse Anzahl von Freiheitsgraden besitzen und dementsprechend einer komplicierten, turbulenten Bewegung unterworfen sind. In einem solchen Systeme gemäss den bekannten Principien der Mechanik die Bewegung eines jeden Punktes vollständig zu beschreiben, wäre zwar wünschenswert als eine durchaus sichere Grundlage für alle weiteren Betrachtungen, ist aber fast immer unausführbar und für den vorliegenden Zweck meist auch unnötig, weil sich diese individuelle Bewegung der Beobachtung ebenso entzieht wie der Rechnung und nur der allgemeine, durchschnittliche Grundcharakter des Vorganges eine anschauliche physikalische Bedeutung besitzt. Man wird sich also auf die Betrachtung gewisser Durchschnittsgrössen beschränken, welche physikalisch messbaren Zuständen entsprechen, wie z. B. die mittlere lebendige Kraft in der Gastheorie der Temperatur; man wird aber auch diese Durchschnittswerte nicht genau als Funktionen der Zeit bestimmen können, weil uns der wahre, die ganze Bewegung bestimmende Anfangszustand des mechanischen Systemes, gegeben durch die sämtlichen Koordinaten und Geschwindigkeiten, nicht vollständig bekannt ist, sondern eben nur die anfänglichen Durchschnittswerte, welche, streng genommen, auch zur Bestimmung des sichtbaren Vorganges nicht ausreichen. Vielmehr wird man sich auch hier mit Annäherungen und Wahrscheinlichkeiten begnügen müssen, und als praktisch gewiss wird man schon solche Veränderungen betrachten dürfen, deren Wahrscheinlichkeit bei gewissen, durch die Natur des Systemes nahe gelegten Grenzübergängen, in
318 der Gastheorie z. B. bei unbegrenzt wachsender Anzahl der Moleküle, nach | 1 konvergiert. Wie soll man aber diese Wahrscheinlichkeiten, diese Mittelwerte definieren, wie soll man mit ihnen operieren, um wenigstens im Sinne der Wahrscheinlichkeitsrechnung zu mathematisch zuverlässigen Ergebnissen zu gelangen? Hier hat man sich in verschiedener Weise zu helfen gesucht, man hat (wie z. B. *Helmholtz*) die turbulente Bewegung als eine cyklische aufgefasst, man hat das wirkliche kinetische Potential durch vereinfachte Ausdrücke ersetzt (wie *J. J. Thomson*), die nur von den physikalischen Durchschnittswerten abhängen, oder man ist auch von der Annahme, dass die Bewegung der Moleküle eine „ungeordnete“ sei, ausgegangen (wie *Boltzmann*). Aber alle solche Auskunftsmittel, so plausibel sie auch scheinen, so förderlich sie sich auch für den gerade vorliegenden Zweck erweisen mögen, scheinen mir

* Göttinger Habilitationsvorlesung, gehalten am 4. März 1899.

On the application of the calculus of probabilities to dynamical systems*

1900

In different areas of mathematical physics, and in particular in the kinetic theory of gases, we face the problem of investigating the motions of mechanical systems that possess a finite, albeit very great, number of degrees of freedom, and that, accordingly, are subject to complicated, turbulent motions. While it would certainly be desirable to describe completely the motion of each point in such a system according to the well-known principles of mechanics in order to place all further considerations on an absolutely firm footing, it is almost always infeasible to do so and, for the present purpose, also mostly unnecessary since this individual motion defies both observation and computation and only the general, average basic character of the process has a concrete physical meaning. For this reason, one will limit oneself to the consideration of certain average magnitudes that correspond to physically measurable states such as the mean living force in the gas theory of temperature; but it will not be possible to determine precisely these average values as functions of time either since the true initial state of the mechanical system that determines the entire motion and that is given by all of the coordinates and velocities is not completely known to us, but only the initial average values, that, strictly speaking, are not sufficient either to determine the visible process. Rather, one will have to make do with approximations and probabilities also in this case, and one may already consider such changes as practically certain whose probability converges to 1 for certain transitions to limits suggested by the nature of the system, such as for an indefinitely increasing number of molecules in the theory of gases. But how is one to define these probabilities, these mean values? How is one to operate with them in order to obtain results that are mathematically reliable, at least by the standards of the calculus of probabilities? Several approaches have been taken to answer these questions. *Helmholtz*, e.g., conceived of the turbulent motion as a cyclic motion, and *J.J. Thomson* replaced the real kinetic potential by simplified expressions dependent solely on the physical average values, while *Boltzmann*, e.g., also proceeded from the assumption that the motion of the molecules is an "unordered" motion. But it seems to me that such devices, as plausible as they may appear and as advantageous as they may turn out to be for the particular purpose at hand, are

* Habilitation lecture at Göttingen, delivered on March 4, 1899.

doch einem prinzipiellen Bedenken zu unterliegen: es werden hier Hypothesen zu Grunde gelegt, die sich auf den ganzen Verlauf der Bewegung beziehen sollen, während doch die verschiedenen Phasen derselben nicht von einander unabhängig sind, sondern durch die Gesetze der Mechanik (in uns freilich noch unbekannter Weise) notwendig zusammenhängen, sodass die Frage berechtigt scheint, ob denn eine Annahme, die für einen gegebenen Zeitpunkt im Sinne der Wahrscheinlichkeitsrechnung zulässig sein mag, es auch für spätere Zeiten bleiben werde oder ob nicht vielmehr das System vermöge seiner eigenen Konstitution mit Notwendigkeit anderen Zuständen zustreben werde, in denen diese Voraussetzung nicht mehr erfüllt ist? In der That scheinen mir auch gewisse Widersprüche, die sich bei manchen gastheoretischen Betrachtungen ergeben, der Nichtbeachtung solcher Bedenken zuzuschreiben zu sein.

Um also eine zuverlässigere Grundlage für die Behandlung solcher Probleme zu gewinnen, scheint mir die Forderung unabweisbar, eine bestimmte Definition der Wahrscheinlichkeit an die Spitze zu stellen, die, wenn auch in gewissem Grade willkürlich, doch im Verlaufe der Untersuchung nicht mehr geändert oder durch neue Annahmen ergänzt werden darf, und man wird ferner unbedingt festhalten müssen an dem *Laplace*schen Wahrscheinlichkeitssatze, nach welchem zwei notwendig wie Ursache und Wirkung mit einander verbundene Ereignisse a und b, sodass das Eintreten des einen von ihnen das des anderen bedingt, auch immer gleich wahrscheinlich sein müssen. Diesen Anforderungen werden wir genügen, wenn wir die Wahrscheinlichkeit irgend eines dynamischen Zustandes, welcher bei der Bewegung unseres Systemes zu irgend einer Zeit t eintreten soll, definieren durch die Wahrscheinlichkeit desjenigen Anfangszustandes zu einer Zeit $t = 0$, aus dem der betrachtete selbst hervorgegangen sein muss.

Nehmen wir an, dass auf unser System ausschliesslich Potentialkräfte wirken, welche allein von der augenblicklichen Konfiguration, d. h. von der *Lage* der Punkte, abhängen sollen, so ist der Bewegungszustand zu einer gegebenen Zeit t und damit auch die ganze Bewegung vollständig bestimmt durch das System aller Koordinaten $q_1, q_2, \ldots q_n$ und der zugehörigen Bewegungsmomente (oder Impulskoordinaten) $p_1, p_2, \ldots p_n$, die wir alle zusammen abgekürzt durch (q, p) bezeichnen wollen. Ebenso ist der Anfangszustand zur Zeit $t = 0$ bestimmt durch das System der entsprechenden Grössen $\overline{q}_\lambda$ und $\overline{p}_\lambda$ oder abgekürzt durch $(\overline{q}, \overline{p})$, während die Bewegungsgleichungen sich unter der gemachten Voraussetzung in der *Hamilton*'schen Form schreiben lassen:

$$\frac{dq_\lambda}{dt} = \frac{\partial H}{\partial p_\lambda}\,, \quad \frac{dp_\lambda}{dt} = -\frac{\partial H}{\partial q_\lambda} \qquad (\lambda = 1, 2, \ldots n)\,. \tag{1}$$

Wird nun nach der „Wahrscheinlichkeit“ eines Anfangszustandes $(\overline{q}, \overline{p})$ gefragt, so kann dies wie immer bei „geometrischen Wahrscheinlichkeiten“ nur in folgendem Sinne verstanden werden. Wir suchen die Wahrscheinlichkeit dafür, dass unser Anfangszustand in gegebenen Grenzen liegt, d. h. einem vorgeschriebenen „Gebiete“ g_0 von „möglichen Anfangszuständen“ angehört,

still vulnerable to the one basic concern: the hypotheses forming the basis here are supposed to refer to the entire course of motion, while the various stages of the motion are not mutually independent but necessarily connected with one another by virtue of the laws of mechanics (in a way that is, however, not yet known to us), so that it seems legitimate to ask whether an assumption that may be permissible by the standards of the calculus of probabilities for a given instant in time is also permissible for later instants, or whether the system will rather, by necessity and due to its own constitution, tend towards other states in which this presupposition no longer holds. In fact, it seems to me that certain contradictions arising in many a gas-theoretic consideration must also be attributed to a failure to heed such concerns.

Therefore, in order to place the treatment of such problems on a firmer footing, it seems to me indispensable to first establish a particular definition of probability that, while to a certain extent being arbitrary, must neither be altered over the course of the investigation nor amended by new assumptions. Furthermore, it will be necessary to always hold on to *Laplace*'s probability theorem, according to which two events a and b that are necessarily connected with one another so that the occurrence of one of them brings about the occurrence of the other like cause and effect are also always equally probable. We will meet these requirements by defining the probability of some dynamic state that is supposed to occur when the system moves at some time t in terms of the probability of that initial state at time $t = 0$ from which the state under consideration itself must have arisen.

Let us assume that only those potential forces are acting on our system that are supposed to depend solely on the momentary configuration, i. e., on the *position* of the points. Then the state of motion at a given time t, and hence also the entire motion, is completely determined by the system of all coordinates $q_1, q_2, \ldots q_n$ and the associated momenta of motion (or impulse coordinates) $p_1, p_2, \ldots p_n$, summarily referred to here as (q, p) for short. In the same way, the initial state at time $t = 0$ is determined by the system of the corresponding magnitudes $\overline{q}_\lambda$ and $\overline{p}_\lambda$, or $(\overline{q}, \overline{p})$ for short, while, given the assumptions made, we can write the equations of motion in *Hamilton*ian form:

$$\frac{dq_\lambda}{dt} = \frac{\partial H}{\partial p_\lambda}\,, \quad \frac{dp_\lambda}{dt} = -\frac{\partial H}{\partial q_\lambda} \qquad (\lambda = 1, 2, \ldots n)\,. \tag{1}$$

As is invariably the case with "geometric probabilities", any question about the "probability" of an initial state $(\overline{q}, \overline{p})$ can only be understood in the following way. We are trying to find the probability that our initial state lies within specified limits, i. e., belongs to a prescribed "region" g_0 of "possible

welche einem Systeme von Ungleichungen genügen

$$\overline{g}_\mu\,(\overline{q}_1, \overline{q}_2, \ldots \overline{q}_n; \overline{p}_1, \overline{p}_2, \ldots \overline{p}_n) \equiv \overline{g}_\mu\,(\overline{q}, \overline{p}) < 0\ . \tag{2}$$

Diese „Wahrscheinlichkeit für das Gebiet g_0“ werden wir nun, da alle Anfangszustände als von einander unabhängig anzusehen sind, zweckmässigerweise, wenn auch nicht ganz ohne Willkür, proportional setzen dem über den Bereich g_0 erstreckten $2n$-fachen Integrale

$$\gamma_0 = \int^{(g_0)} d\overline{q}_1 \ldots d\overline{q}_n d\overline{p}_1 \ldots d\overline{p}_n \equiv \int^{(g_0)} dq\, dp\ ,$$

das wir als „die Ausdehnung des Gebietes g_0“ bezeichnen wollen, und den absoluten Wert der Wahrscheinlichkeit erhalten wir dann durch Division in die Ausdehnung Γ_0 des Bereiches G_0 aller überhaupt in Betracht kommenden Anfangszustände, indem wir es als gewiss ansehen, dass unser Zustand $(\overline{q}, \overline{p})$ dem Gebiete G_0 angehöre. Nun entspricht aber jedem Anfangszustande $(\overline{q}, \overline{p})$ zur Zeit $t = 0$ eine ganze Bewegung, also auch ein ganz bestimmter Zustand $(q, p)_t$ zu einer beliebig vorgeschriebenen Zeit t und somit der Gesamtheit aller Anfangszustände von g_0 ebenfalls eine Gesamtheit von Zuständen $(q, p)_t$,
319 welche wieder ein kontinuierliches Gebiet g_t erfüllen werden, von dem | wir sagen wollen, dass es „in der Zeit t aus dem Gebiete g_0 hervorgehe“, und auch die „Ausdehnung“ γ_t dieses Gebietes g_t wird sich als $2n$-faches Integral bestimmen lassen. In dem hier betrachteten Falle nun, wo die Bewegungsgleichungen sich in der Form (1) schreiben lassen, gilt der Satz von *Liouville*, dass die Ausdehnung des Gebietes g_t der Ausdehnung des entsprechenden Gebietes g_0 gleich, mithin von der Zeit t unabhängig ist:

$$\gamma_t = \int^{(g_t)} dq\, dp = \int^{(g_0)} dq\, dp = \gamma_0\ . \tag{3}$$

Ebenso entspricht auch dem Bereiche G_0 zur Zeit t wieder ein Bereich G_t von der gleichen Ausdehnung $\Gamma_t = \Gamma_0$. Durch den Bruch

$$w_g = \frac{\gamma_0}{\Gamma_0} = \frac{\gamma_t}{\Gamma_t}$$

soll aber die Wahrscheinlichkeit gemessen werden, dass unser Anfangszustand $(\overline{q}, \overline{p})$ dem Gebiete g_0 angehöre, und nach unserem oben aufgestellten Reduktionsprinzipe zugleich auch die Wahrscheinlichkeit dafür, dass der spätere Zustand $(q, p)_t$ dem Gebiete g_t angehöre, denn beide Ereignisse sind notwendig mit einander verbunden. Diese Wahrscheinlichkeit ist also durch die Gebiete g_t und G_t allein bestimmt, gerade als ob wir es nur mit einem Anfangszustande zu thun hätten, und von der Zeit selbst gänzlich unabhängig, sodass wir den Satz haben:

initial states" satisfying a system of inequalities

$$\overline{g}_\mu\,(\overline{q}_1, \overline{q}_2, \ldots \overline{q}_n; \overline{p}_1, \overline{p}_2, \ldots \overline{p}_n) \equiv \overline{g}_\mu\,(\overline{q}, \overline{p}) < 0\ . \tag{2}$$

As all initial states are mutually independent, it will be advantageous, albeit somewhat arbitrary, to now set the "probability for the region g_0" proportional to the $2n$-fold integral extended over the domain g_0

$$\gamma_0 = \int^{(g_0)} d\overline{q}_1 \ldots d\overline{q}_n d\overline{p}_1 \ldots d\overline{p}_n \equiv \int^{(g_0)} dq\, dp\ ,$$

which we will call "the extension of the region g_0". We then obtain the absolute value of the probability by division by the extension Γ_0 of the domain G_0 of all initial states that are actually possible by taking it for granted that our state $(\overline{q}, \overline{p})$ belongs to the region G_0. But now to every initial state $(\overline{q}, \overline{p})$ at time $t = 0$ there corresponds an entire motion, and hence also a specific state $(q, p)_t$ at an arbitrarily prescribed time t. Consequently, there corresponds to the totality of all initial states of g_0 also a totality of states $(q, p)_t$ which, in turn, will fill a continuous region g_t, which, as we shall say, "arises in time t from the region g_0", and we will also be able to determine the "extension" γ_t of this region g_t as a $2n$-fold integral. Now, in the case considered here, where it is possible to write the equations of motion in form (1), *Liouville*'s theorem holds according to which the extension of the region g_t is equal to the extension of the corresponding region g_0, and hence independent of time t:

$$\gamma_t = \int^{(g_t)} dq\, dp = \int^{(g_0)} dq\, dp = \gamma_0\ . \tag{3}$$

Likewise, to the domain G_0 at time t there also corresponds a domain G_t of identical extension $\Gamma_t = \Gamma_0$. But by means of the fraction

$$w_g = \frac{\gamma_0}{\Gamma_0} = \frac{\gamma_t}{\Gamma_t}$$

we want to measure the probability that our initial state $(\overline{q}, \overline{p})$ belongs to the region g_0, and, by the reduction principle presented above, also the probability that the later state $(q, p)_t$ belongs to the region g_t, for both events are necessarily connected with one another. Hence, this probability is solely determined by the regions g_t and G_t, just as if we were only dealing with an initial state, and is itself entirely independent of time, so that we have the theorem:

Satz I. Die Wahrscheinlichkeit für eine gegebene Begrenzung eines Bewegungszustandes ist von der Zeit unabhängig, d. h. stets dieselbe, ob der betrachtete Zustand im Anfange oder in irgend einer anderen Phase der Bewegung eintreten soll.

Den Bereich G_0 aller „möglichen“ Anfangszustände, auf den wir unsere Betrachtung beschränken, dessen Wahrscheinlichkeit wir $= 1$ annehmen wollen, wählen wir zweckmässig so, dass er bei der Bewegung unseres Systemes stets *in sich selbst übergeht*, d. h. mit allen seinen späteren Phasen G_t identisch ist. Ein solches Gebiet $G_0 = G_t$, das wir als ein „invariantes Gebiet“ bezeichnen wollen, wird immer begrenzt durch Integrale der Bewegungsgleichungen (1), d. h. durch Gleichungen $G(q, p) = \text{const.}$, welche der Bedingung genügen:

$$\frac{dG}{dt} = \sum_{\lambda=1}^{n} \Big(\frac{\partial G}{\partial q} \frac{\partial H}{\partial p} - \frac{\partial G}{\partial p_\lambda} \frac{\partial H}{\partial q} \Big) \equiv (G, H) = 0 \, . \tag{4}$$

So wird z. B. durch das „Integral der lebendigen Kraft“ $H = \text{const.}$ eine Reihe von „invarianten Gebieten“ $H < c$ oder $a < H < b$ definiert, welche die *Energie* des Systemes nach oben oder unten begrenzen und gelegentlich als Gebiete „aller möglichen“ Zustände interpretiert werden können.

Von der Definition der „Wahrscheinlichkeit“ gehen wir jetzt über zu der des „mittleren“ oder „wahrscheinlichen Wertes“. Es sei nämlich $S \equiv S(q, p)$ eine eindeutige und stetige Funktion des „Zustandes“, d. h. der $2n$ Variablen q_λ, p_λ, und g ein beliebiges Gebiet von der endlichen Ausdehnung γ, so bezeichnen wir als „den mittleren Wert von S im Gebiete g“ den Ausdruck:

$$\overline{S}_g = \frac{1}{\gamma} \overset{(g)}{\int} S(q, p) dq \, dq \quad \left(\gamma = \overset{(g)}{\int} dq \, dp \right) . \tag{5}$$

Dieser Mittelwert ist eindeutig gegeben durch das vorgeschriebene Gebiet g und wird sich im allgemeinen mit der Zeit stetig ändern, wenn g im Verlaufe der Bewegung vom g_0 in g_t übergeht, und es ist dabei immer

$$\frac{d\overline{S}_{g_t}}{dt} = \frac{1}{\gamma} \overset{(g_t)}{\int} \frac{dS}{dt} dq \, dp = \frac{1}{\gamma} \overset{(g_t)}{\int} (S, H) dq \, dp \, , \tag{6}$$

weil nach (3) γ ebenso wie das $2n$-fache Differential $d\gamma = dq \, dp$ von t unabhängig ist. Ist aber $g = g_t$ ein „invariantes“ Gebiet, z. B. das Gebiet G „aller möglichen“ Zustände, so ist auch $\overline{S}_g$ von der Zeit unabhängig, also

$$\gamma \frac{d\overline{S}_g}{dt} = \overset{(G)}{\int} \frac{dS}{dt} dq \, dp = \overset{(G)}{\int} (S, H) dq \, dp = 0 \, , \tag{6a}$$

und wir haben den Satz:

Theorem I. The probability for a given boundary of a state of motion is independent of time, i. e., always the same, whether the considered state is supposed to occur at the beginning or at some other phase of the motion.

It is useful to choose the domain G_0 of all "possible" initial states to which we restrict our consideration and whose probability we shall take to be $= 1$ so that it *is* always *transformed into itself* when our system moves, i. e., so that it is identical with all its later phases G_t. Such a region $G_0 = G_t$, which we shall call an "invariant region", is always bounded by integrals of the equations of motion (1), i. e., by equations $G(q, p) = \text{const.}$ satisfying the condition:[1]

$$\frac{dG}{dt} = \sum_{\lambda=1}^{n} \Big(\frac{\partial G}{\partial q_\lambda} \frac{\partial H}{\partial p_\lambda} - \frac{\partial G}{\partial p_\lambda} \frac{\partial H}{\partial q_\lambda} \Big) \equiv (G, H) = 0 \,. \tag{4}$$

Thus, e. g., the "integral of the living force" $H = \text{const.}$ defines a series of "invariant regions" $H < c$ or $a < H < b$ by which the *energy* of the system is bounded upwards or downwards and which can, on occasion, be interpreted as regions of "all possible" states.

We now move from the definition of the "probability" to that of the "mean" or "probable value". For let $S \equiv S(q, p)$ be a single-valued and continuous function of the "state", i. e., of the $2n$ variables q_λ, p_λ, and let g be any region of finite extension γ. Then by "the mean value of S in the region g" we refer to the expression[2]

$$\overline{S}_g = \frac{1}{\gamma} \int^{(g)} S(q, p) dq\, dp \quad \left(\gamma = \int^{(g)} dq\, dp \right) . \tag{5}$$

This mean value is uniquely given by the prescribed region g and, in general, will change continuously with time, if g is transformed from g_0 into g_t over the course of the motion, and we always have

$$\frac{d\overline{S}_{g_t}}{dt} = \frac{1}{\gamma} \int^{(g_t)} \frac{dS}{dt} dq\, dp = \frac{1}{\gamma} \int^{(g_t)} (S, H) dq\, dp \,, \tag{6}$$

since, by (3), γ as well as the $2n$-fold differential $d\gamma = dq\, dp$ is independent of t. If, however, $g = g_t$ is an "invariant" region, e. g., the region G of "all possible" states, then $\overline{S}_g$, too, is independent of time, and hence

$$\gamma \frac{d\overline{S}_g}{dt} = \int^{(G)} \frac{dS}{dt} dq\, dp = \int^{(G)} (S, H) dq\, dp = 0 \,, \tag{6a}$$

and we have the theorem:

[1] [[On the right side of the following equation (4), Zermelo erroneously omits three lower indices λ.]]

[2] [[In the middle of the following formula (5), Zermelo erroneously writes "*dqdq*" instead of "*dqdp*".]]

Satz II. Ist S eine beliebige eindeutige und differentiierbare Funktion des Bewegungszustandes, so hat der Mittelwert von $\frac{dS}{dt}$ in jedem invarianten Gebiete G den Wert Null.

Es kann also in diesem Gebiete die Grösse S ebenso gut zunehmen wie abnehmen, und es kann keine solche Funktion S existieren, die für alle oder auch nur für die überwiegend meisten Zustände eines invarianten Gebietes vermöge der Bewegungsgleichungen (1) beständig zunehmen oder beständig abnehmen würde.

Ein analoger Satz gilt aber auch für nicht invariante Gebiete, nämlich für solche, welche durch vorgeschriebene numerische Werte von S begrenzt werden, d. h. für Gebiete $a < S < b$. Wir beweisen den Satz zunächst für den einfacheren Fall $S < c$, auf den sich der allgemeine zurückführen lässt. Ist nämlich unser Gebiet $g = g_t$ kein invariantes, so wird es in der Zeit τ in ein anderes Gebiet $g' = g_{t+\tau}$ übergehen. Beide Gebiete haben die gleiche Ausdehnung

320 |
$$\gamma = \int^{(g)} dq\, dp = \int^{(g')} dq\, dp$$

und haben für kleine Werte von τ ein Stück g'' mit einander gemein, während die Ausdehnungen der Restgebiete $g - g''$ und $g' - g''$ mit τ gleichzeitig verschwinden. Daher wird die Differenz der beiden Integrale

$$\begin{aligned}\gamma\,(S_{g'} - S_g) &= \int^{(g')} S\, dq\, dp - \int^{(g)} S\, dq\, dp \\ &= \int^{(g')} (S-c) dq\, dp - \int^{(g)} (S-c) dq\, dp \\ &= \int^{(g'-g'')} (S-c) dq\, dp - \int^{(g-g'')} (S-c) dq\, dp\,,\end{aligned}$$

wo c eine beliebige Konstante sein kann. Ist aber das Gebiet g definiert durch die Ungleichung $S < c$, so werden in den beiden schmalen Randgebieten $g' - g''$ und $g - g''$ die Werte von S nur wenig von c verschieden sein und die beiden Integrale über $S - c$ von der Ordnung τ^2 verschwinden, so dass wir haben, wie behauptet:

$$\lim_{\tau=0} \gamma \frac{\left(S_{g'} - \overline{S}_g\right)}{\tau} = \gamma \frac{dS_g}{dt} = \int^{(g)} \frac{dS}{dt} dq\, dp = 0\,. \tag{7}$$

Ferner ist dasselbe Integral $\int \frac{dS}{dt} dq\, dp$, erstreckt über ein Gebiet $a < S < b$ nichts anderes als die Differenz der beiden analogen Integrale über die Gebiete

Theorem II. If S is any single-valued and differentiable function of the state of motion, then the mean value of $\dfrac{dS}{dt}$ takes the value zero in every invariant region G.

Hence, in this region, the magnitude S may just as well increase as decrease and there can exist no function S that continuously increases or continuously decreases for all or at least the great majority of states of an invariant region by dint of the equations of motion (1).

However, an analogous theorem also holds for regions that are not invariant, namely for those bounded by prescribed numerical values of S, i. e., for regions $a < S < b$. We first prove the theorem for the simpler case $S < c$, to which the general case can be reduced. For if our region $g = g_t$ is not invariant, then it will be transformed into a different region $g' = g_{t+\tau}$ during time τ. Both regions have the same extension

$$\gamma = \int^{(g)} dq\, dp = \int^{(g')} dq\, dp$$

and share a segment g'' for small values of τ, while the extensions of the remainder regions $g - g''$ and $g' - g''$ vanish along with τ. Hence, the difference of the two integrals becomes[3]

$$\begin{aligned}
\gamma\left(\overline{S}_{g'} - \overline{S}_g\right) &= \int^{(g')} S\, dq\, dp - \int^{(g)} S\, dq\, dp \\
&= \int^{(g')} (S-c) dq\, dp - \int^{(g)} (S-c) dq\, dp \\
&= \int^{(g'-g'')} (S-c) dq\, dp - \int^{(g-g'')} (S-c) dq\, dp\,,
\end{aligned}$$

where c may be any constant. But if the region g is defined by the inequality $S < c$, then, in the two narrow boundary regions $g' - g''$ and $g - g''$, the values of S will differ only slightly from c and the two integrals over $S - c$ will vanish with order τ^2 so that we have, as stated,

$$\underset{\tau=0}{Lim}\, \gamma \frac{\left(\overline{S}_{g'} - \overline{S}_g\right)}{\tau} = \gamma \frac{d\overline{S}_g}{dt} = \int^{(g)} \frac{dS}{dt} dq\, dp = 0\,. \tag{7}$$

Furthermore, the same integral $\displaystyle\int \frac{dS}{dt} dq\, dp$, extended over a region $a < S < b$, is but the difference of the two analogous integrals over the regions $S < b$

[3] ⟦In the following formula and in formula (7) below, Zermelo erroneously writes "$S_{g'}$" and "S_g" for "$\overline{S}_{g'}$" and "$\overline{S}_g$", respectively.⟧

$S < b$ und $S < a$ und muss daher gleichfalls verschwinden, auch wenn a und b sich beliebig wenig unterscheiden. So haben wir:

Satz III. Ist uns der Wert S_0 einer eindeutigen Funktion S des Bewegungszustandes mit beliebiger Annäherung σ vorgeschrieben, so hat in dem dadurch definierten Gebiete $S_0 - \sigma < S < S_0 + \sigma$ die mittlere Zunahme $\frac{dS}{dt}$ derselben Funktion den Wert Null.

Die hier entwickelten Sätze könnten als Grundlage für weitere Betrachtungen dienen, indem sie uns lehren, welche Wahrscheinlichkeiten und Mittelwerte wir ganz allgemein durch andere, einfachere ersetzen dürfen. Um nun aber weitere Gleichungen oder Differentialgleichungen zu finden, die zu einer wirklichen Bestimmung der Mittelwerte als Funktionen der Zeit und zu einer Beschreibung des physikalischen Vorganges führen können, wird man diese allgemeinen Betrachtungen verlassen und zu specielleren Annahmen übergehen müssen. Als ein Beispiel hierfür beabsichtige ich, demnächst für das sog. *Maxwell*'sche Gesetz von der Geschwindigkeitsverteilung der Gasmoleküle eine neue Ableitung zu versuchen, die ohne jede weitere Hypothese allein auf der hier gegebenen Definition der Wahrscheinlichkeit beruhen soll.

(Eingegangen 24. März 1900.)

and $S < a$, and hence must also vanish, however small the difference between a and b may be. We thus have:

Theorem III. If the value S_0 of a single-valued function S of the state of motion is prescribed with arbitrary approximation σ, then, in the region $S_0 - \sigma < S < S_0 + \sigma$ thereby defined, the mean increase $\frac{dS}{dt}$ of the same function has value zero.

The theorems developed here may serve as a basis for further considerations by showing us which probabilities and mean values we are permitted to generally replace by other, simpler ones. However, in order to find further equations or differential equations that can lead to an actual determination of the mean values as functions of time and to a description of the physical process we will have to move from these general considerations to more particular assumptions. To illustrate this by an example, I intend to attempt a new derivation of the so-called *Maxwell* law of the velocity distribution of gas molecules soon which is supposed to be based solely on the definition of probability given here without any further hypothesis.

(Received on March 24, 1900.)

Introductory note to 1902a, s1902b, and s1902c

Alexey V. Borisov, Larisa A. Gazizullina, and Sergei M. Ramodanov

It is sometimes the fate of a significant piece of scientific work that it drops out of circulation for quite some time or even passes into oblivion, the results to be rediscovered only by a later generation. When revived and brought out into the open, such work is invariably of important historical interest. Sometimes, it will even be seen as retaining significant scientific value and is able to influence the contemporary reader by its insights and wealth of ideas. It is evident that the author clearly perceived deep concepts which have come to be generally regarded as achievements of contemporary researchers. Perhaps the older work even surpasses the later versions in having obtained and unified what later appears as disparate results found at various times by different scientists.

All of this applies to *1902a, s1902b*, and *s1902c*, the remarkable work by Zermelo which is the subject of the present note. In modern mathematics Zermelo is well known for his decisive contributions towards creating the axiomatic foundation of set theory as well as for his works on statistical mechanics. However, his legacy also includes an important piece of research on hydrodynamics and the theory of vortices: his *Habilitation* thesis *Hydrodynamical investigations of vortex motions in the surface of a sphere* (1899) which was only partially published and afterwards lapsed into obscurity. In it, the author derives and systematically investigates many equations of motion appearing again in various studies beginning from the 1970s. (Of course, Zermelo's analysis lacks the advantages we have nowadays of the availability of techniques such as those of Poincaré and Lie which have become current in mathematical physics). The thesis also contains some material which remains to be assimilated for further clarification and progress in solving many problems of hydrodynamics on surfaces.

With this field, and nonlinear dynamics in general, being subject to intensive research on an international scale, it is not by chance that Zermelo's thesis and other forgotten studies on vortex dynamics are now becoming widely known. While such works did not have a proper influence in the past, they are assuming real importance and gaining the recognition they deserve. This can be said for a number of works of the period from the second half of the 19th century to the early 20th century, a period of fruitful development of vortex theory, which started under the influence of fundamental research by Hermann von Helmholtz (*1858*). Many considerable results in vortex theory achieved then have been obtained independently in the modern era of the second half of the 20th century. Mention should be made of the studies by

H.-D. Ebbinghaus, A. Kanamori (Eds.), *Ernst Zermelo - Collected Works/ Gesammelte Werke II*, DOI 10.1007/978-3-540-70856-8_7, Schriften der Mathematisch-naturwissenschaftlichen Klasse der Heidelberger Akademie der Wissenschaften 23,

such authors as Walter Gröbli,[1] Ippolit S. Gromeka, N. S. Vasilyev, William Hicks, Ernesto Laura, Cataldo Agostinelli and others. Some of these works have been cited in recent reviews but have not yet been subjected to a full scientific and historical analysis.

The biographical data on Zermelo presented here have been drawn from *Ebbinghaus 2007* which contains much interesting documentary material.

Introduction

The *Habilitation* thesis was completed by Zermelo in his second year in Göttingen, to where he had moved in 1897 from Berlin to pursue his academic career. The Berlin period had brought to young Zermelo his first celebrity. In 1894 he obtained his Ph.D. degree with a dissertation (*1894*) on the calculus of variations under the supervision of Hermann Amandus Schwarz. This work, which substantially developed Weierstrass's ideas, was at once most favourably received. The subsequent papers (*1896a, 1896b*) were written by Zermelo when he was already an assistant to Max Planck. They contain criticism of Boltzmann's mechanistic approach to thermodynamics, an approach which had numerous opponents at the time. Zermelo's objections, built on the Poincaré recurrence theorem, were sufficiently strong and serious to initiate a controversy with Boltzmann himself (this interesting discussion is mentioned in many surveys on thermodynamics and statistical mechanics; see also the introductory note to *Zermelo 1896a* in this volume).

In July 1897 Zermelo wrote to Felix Klein requesting assistance toward transferring to another university in a quieter place where he could continue his studies and prepare himself for *Habilitation*. Without specifying the subject of the thesis, he wrote that it would be devoted to issues of mechanics and theoretical physics. Klein's response has been lost, but evidently Zermelo was successful, for by November he had arrived at the university of Göttingen.

Zermelo's thesis was indeed concerned with a problem of applied mathematics, namely of hydrodynamics, and in particular of vortex theory. We quote the first lines of *1902a*: "The work detailed in the present paper seeks to explain the flow of an incompressible, frictionless (two-dimensional) fluid in a spherical surface by the use of a theory as systematic as the one which already exists for planar flows and of which a fairly complete account is given in particular in Poincaré's *Théorie des tourbillions* (*1893a*)."

The defense of the thesis took place on the 9th of February 1899 (by then Zermelo was 27 years old). In his official report, Hilbert gave a highly positive assessment of the thoroughness of the research done and the new results obtained; he also gave a short description of its contents:[2]

[1] Gröbli's work *1877* on the motion of vortices on the plane is dealt with in *Aref, Rott, and Thomann 1992.*

[2] For the German original see *Ebbinghaus 2007*, 275–276.

The thesis submitted by Dr. Zermelo is a purely mathematical investigation, although it is related to the physical-meteorological problem of the motion of cyclones on the surface of the Earth.

The *first* chapter deals with the motion of a fluid on an arbitrary surface in space. Based on the Lagrangian differential equations of the second kind as a basis and introducing two independent Gauss coordinates, the author develops the form of the general laws of fluid motion for the present case. Using the von Helmholtz theorem on the conservation of vorticity in the case of an incompressible fluid, the author obtains a third-order partial differential equation for the stream function ρ, which completely determines the fluid motion in time.

The *second* chapter applies this general theory to the investigation of a flow on a sphere. The partial differential equation $D\psi = 2\rho$, where D is a certain differential operator and ρ is vorticity, is shown to have a special solution which represents a so-called singular "point vortex",[3] while the vorticity on the rest of the spherical surface is constant. Using this solution as a spherical counterpart of the Green function in the planar case, one readily obtains the general solution of the above-mentioned partial differential equation, which is the so-called spherical potential corresponding to the ordinary logarithmic potential on the plane, the vorticity ρ playing the role of mass density. Next, the properties of the spherical potential are examined. If one imagines a mass distributed over the spherical surface with a density differing from the vorticity by a constant value other than zero, the position of the center of gravity of this mass will remain unchanged [...]. In the conclusion of this chapter special classes of continuous stationary and rotational-stationary motions on the sphere are indicated, which are expressed in terms of spherical functions.

In the *third* chapter, the velocity of the point vortex is calculated and the equation of motion is derived for the case where, besides the continuous vortex motion, a finite number of point vortices exist. If the continuous vortex density is constant on the entire sphere, such a flow is called a system of vortices. A system of vortices is completely determined by the number, configuration, and intensity of the point vortices; the goal is to determine the evolution of the configuration in time. Of special interest is the case of equilibrium, which corresponds to the maximum or minimum of the so-called self-potential of vortices. A number of geometric interpretations of this condition are given in which the ordinary polyhedra appear as special figures of equilibrium.

The *last* chapter thoroughly examines the three-vortex problem, which reduces to the investigation of the form of the triangle generated by the vortices. The differential equations of motion are derived and discussed; special attention is paid to the case of equal intensities where integration is possible by elliptic functions. Note that the squares of the sides of the flat triangle formed by straight connections of the point vortices are used as new variables.

The thesis represents a careful and thorough investigation and has led to new and remarkable results; it attests not only to the author's knowledge and abilities but also to his scientific inclination and his ideal pursuit. That is why I vote for admittance to *Habilitation*.

[3] ⟦Zermelo uses the term "Strudelpunkt" which has been translated as "whirl point"; see fn. 2 on page 317.⟧

Zermelo's *Habilitation* thesis appeared as the work entitled *Hydrodynamische Untersuchungen über die Wirbelbewegungen in einer Kugelfläche* (*Hydrodynamical investigations of vortex motions in the surface of a sphere*) in two parts. The first part (*1902a*) was published in the journal *Zeitschrift für Mathematik und Physik*. It systematically develops the general theory of two-dimensional hydrodynamics on the sphere and on surfaces of arbitrary form and introduces the concept of point vortex on the sphere. The second part was not published and has been stored as manuscripts (*s1902b, s1902c*) in the archive of Freiburg University. It examines the system of three point vortices on the sphere, proves its integrability within the framework of the theory of Hamiltonian systems, and studies qualitatively the sequential course of motion.

Unfortunately, several historical circumstances concerning Zermelo's *Habilitation* remain obscure to us. Of interest, first of all, is the question of what caused Zermelo to turn to hydrodynamical investigations. In the introduction to *1902a* he points out that the first inspiration for the work came from geophysical issues related to the understanding of the processes of evolution of cyclones and sea flows. It remains unknown who could have directed him to these issues, which were seemingly far from his main Berlin interests (the calculus of variations and the work vs. Boltzmann). However, as the letter to Klein mentioned above shows, the subject of the *Habilitation* had come to him already in Berlin. An evident stimulus for these investigations would have been the work by Poincaré (*1893a*). Attesting to Zermelo's interest in problems of meteorology was that in 1896 he applied for an appointment as assistant to Deutsche Seewarte (the German sea observatory in Hamburg, which was the central institute of sea meteorology in the country at that time).

One can only guess at the reason why Zermelo did not publish the second part of his work. Already in 1900, the year after the *Habilitation*, he taught a course in set theory, at the same time beginning his active scientific work on the foundations of mathematics with Hilbert. It seems plausible to attribute the failure to publish to these increasingly close relations to Hilbert and to the change in his scientific interests. Both led him to outstanding achievements with which he won the reputation as the founder of modern axiomatic set theory. In a letter to his friend Max Dehn of 25 September 1900, Zermelo complained that he could not complete and publish his "wretched paper on vortices" ("unglückselige Wirbel-Arbeit") and doubted whether he would ever succeed in this. It should also be noted that the *Habilitation* thesis by Zermelo, although it is quite comprehensible, did not gain as much recognition and fame as his Ph.D. thesis. Zermelo never again returned to any issues of vortex theory.

Zermelo's *Habilitation* work in any case is the first analytically completed, systematic investigation of the dynamics of point vortices on the sphere. A little earlier the problem was addressed by the Russian scientist Ippolit S. Gromeka (*1885*), as will be described below. However, although his investigations contain a lot of interesting and profound ideas, Gromeka did not arrive

at the derivation of equations of motion for vortices on the complete surface of a sphere. Zermelo's work by far surpassed most further investigations in this field but unfortunately, owing to its fate, did not have any impact on them. It was not until the 1970s that the equations of motion for vortices on the sphere similar to Zermelo's equations were obtained by Bogomolov (*1977*). Many of Zermelo's results were independently rediscovered and developed in contemporary work. His thesis is still of current interest, and we find here very interesting and suggestive material. Below, we consider its contents more closely with a focus on the second, unpublished, part (Chapters 3 and 4). There will also be some comments made in the context of the present state of the theory of vortex motion on surfaces. This field still contains a great deal of open questions.

Chapters 1 and 2 (*1902a*)

Hydrodynamics on arbitrary surfaces

The first chapter of Zermelo's thesis begins with a systematic construction of the hydrodynamics of an ideal liquid on an arbitrary two-dimensional surface S. To obtain the equations of motion for a liquid analogous to the Euler equations on the plane, he first writes the equations of motion for a material point in local coordinates ξ and η on S and then extends them to the case where the mass is continuously distributed over the surface and thus is well determined by the surface density ρ. Zermelo's equations read:

$$\frac{d}{dt}(u\sqrt{E}) + u^2\sqrt{E}\frac{\partial}{\partial\xi}\frac{1}{\sqrt{E}} + v^2\sqrt{G}\frac{\partial}{\partial\eta}\frac{1}{\sqrt{G}} = -\frac{\partial(P+\Phi)}{\partial\xi}\,,$$
$$\frac{d}{dt}(v\sqrt{E}) + u^2\sqrt{E}\frac{\partial}{\partial\eta}\frac{1}{\sqrt{E}} + v^2\sqrt{G}\frac{\partial}{\partial\xi}\frac{1}{\sqrt{G}} = -\frac{\partial(P+\Phi)}{\partial\xi}\,.$$

Here E, G are the elements of the first quadratic form, u and v are the velocity components of liquid particles in the direction of coordinates ξ and η, and Φ is the potential of body forces; P is the pressure function, and the formula $dP = dp/\rho$ connects the pressure p to the density ρ.

Next, Zermelo develops "from scratch" the surface hydromechanics analogous to that in the planar case. Namely, for the case of an incompressible liquid he introduces the concept of stream function just as it is done on the plane—by examining the flow of liquid through a curve one of the ends of which is fixed, and showing that the rate of fluid flow depends only on the position of the second end. The velocity components u and v are expressed in terms of the stream function by the formulae

$$u = -\frac{1}{\sqrt{G}}\frac{\partial\psi}{\partial\eta}\,, \quad v = \frac{1}{\sqrt{E}}\frac{\partial\psi}{\partial\xi}\,,$$

that is, as in the planar case, $(u, v) = \operatorname{sgrad}\psi$.

Then Zermelo introduces the concept of circulation and proves the theorem of the conservation of circulation along a closed contour, from which he deduces an analogue of the famous von Helmholtz theorem of the conservation of vorticity of a liquid element (*von Helmholtz 1858*). Thus, at least locally, the hydromechanics on surfaces does not at all differ from that in the planar case.

Zermelo was the first to systematically develop two-dimensional hydrodynamics on a sphere and on surfaces of arbitrary form. Although his equations and theorems are of local nature, he created all the essential prerequisites for the further development of hydromechanics on surfaces as a whole. Unfortunately, there has been no substantial advance in this line of research after Zermelo. We can say that the theory of global hydrodynamics on surfaces is still almost in an embryonic state. The literature dealing with this research area only covers a few particular issues concerning basically the motions of point vortices, e.g., *Koiller and Boatto 2008a, 2008b*, *Kidambi and Newton 2000a, 2000b, 1998, 1999, Lamb 1895*, and *Gromeka 1885*; see also the rather "mathematized" study *Arnold and Khesin 1988*. Evidently, the question as to whether the surface on which the flow is examined is compact will be important for this theory. It is well-known that according to the Gauss theorem, the total vorticity on compact surfaces must be equal to zero. It is also clear that in studying the motion of rigid bodies on surfaces it is necessary to solve problems analogous to the Dirichlet and Neumann problems on the plane. In the case of the plane, for these problems to have a unique solution it was necessary to postulate a certain condition at infinity, which is obviously absent on compact surfaces. Thus, for compact surfaces the theorems of existence and uniqueness of solutions in these problems are yet to be refined.

There are some other issues to be elucidated in the hydrodynamics on compact surfaces. In Zermelo's study we find interesting results that are not typical of a planar context. For example, if we consider the flow of a liquid on the sphere and mean by the "mass" of a particle the value of its vorticity, the center of mass of such a distribution will remain fixed in space. Zermelo makes wide use of this and allied theorems in investigating the motion of point vortices.

This and many other issues of the hydrodynamics of an ideal liquid constitute a large body of open problems. One of the important objectives is to extend the methods of planar hydrodynamics to the case of curved surfaces. An example of such an "extension" is *Borisov, Mamaev, and Ramodanov 2010* dealing with the motion of a two-dimensional rigid body interacting dynamically with point vortices on the surface of a two-dimensional sphere. It should be noted that during the preparation of this paper, acqaintance with Zermelo's research was indispensable.

Application of the theory to the sphere. Concept of point vortex

By a point vortex on an arbitrary surface Zermelo means a single-point flow singularity with the following property: the value of velocity circulation along

the contour enclosing the singularity remains different from zero and finite when the contour contracts to a point. Zermelo makes this rather vague definition more precise for the spherical case, assuming that the flow from a point vortex at the pole $\theta = 0$ is given by the stream function $\psi = \frac{m}{\pi} \ln \sin \frac{\theta}{2}$, which is a solution to the Beltrami–Laplace equation on the sphere:

$$D\psi = \frac{\partial^2 \psi}{\partial \theta^2} + \cot\theta \frac{\partial \psi}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2 \psi}{\partial \lambda^2} = -\frac{m\pi}{2}. \tag{1}$$

Here λ and θ are spherical coordinates and $-\frac{m\pi}{2}$ is a background vorticity, which, as stated above, ensures a zero total vorticity on the entire sphere.

Mention should be made here of the little-known work by the Russian scientist Ippolit S. Gromeka who was twenty years older than Zermelo and concerned himself with the same hydrodynamics problems (*Gromeka 1885*). What makes his work interesting is that it contains a deep and original approach to the implementation and analysis of liquid flows on surfaces. In contrast to Zermelo, Gromeka does not construct the hydrodynamics on surfaces "from scratch" but uses the already existing construction of three-dimensional hydrodynamics. Introducing in a special way the orthogonal curvilinear coordinates x_1, x_2, x_3, he examines the flows parallel to some family of parallel surfaces $x_3 =$ constant, e.g. concentric spheres. If the liquid is incompressible, the two-dimensional flow on each of the surfaces will be incompressible as well. Reasoning in this way, Gromeka obtains a new hydrodynamic object—semi-infinite vortex filaments originating from the common center of spheres, whose intensity decreases as $1/R^2$ beginning from their common center. On each sphere $x_3 =$ constant, the filament produces a flow, which is just what Gromeka calls the flow from a point vortex. In this way he derives the general equations of motion of point vortices and applies them to the study of vortex motion of a liquid on the sphere. However, in view of an obstacle created by Gromeka himself (that the total vorticity is zero) he did not manage to derive the equations of motion of point vortices on a full sphere. This problem was completely solved by Zermelo in *s1902b*, which we will discuss below.

Gromeka focused on solving the problem when the vortex moves within a limited region of the spherical layer and, therefore, with zero total vorticity. Thus, he arrives at equation (1) with the right side being zero. In this way he derives the equations governing the path of a point vortex within various contours on the sphere: the vortex moves along the level lines of the stream function. Next, he makes an attempt (although he admits its futility at once) to construct analogous motions on a full sphere assuming that the exterior of the region containing the vortex shrinks to a point. Note that Gromeka derives the equations of vortex motion in rather complicated domains on the sphere. Similar equations have been derived in the contemporary studies *Kidambi and Newton 2000a* and *Crowdy 2006*.

Unfortunately, the remarkable work of Gromeka has lapsed into obscurity even in Russia. In another paper, *1881*, he examined stationary three-

dimensional flows now referred to as ABC-flows (Arnold, Beltrami, Childress). Historically, this designation should also include the name of Gromeka.

Zermelo's results were revisited and further developed to some extent by V. A. Bogomolov in *1977, 1979, 1985* which are cited in the literature as the first rigorous, systematic research on the dynamics of point vortices on a sphere. As we now see, this subject had been completely investigated much earlier though this does not detract from the value of Bogomolov's work. Bogomolov's reasoning is quite interesting and original (although, conceptually, it parallels in a sense Gromeka's approach): He examines the flow from a source placed between concentric spheres. For the impermeable condition to be met on the spherical surfaces, one has to arrange infinitely many auxiliary sources on the line from the sphere's center through the source. Then, letting the distance between the spheres tend to zero, one obtains in the limit a beam consisting of sources whose intensity decays as $1/R^2$. Thus, a flow has been obtained on the spherical surface from a filament source orthogonal to it. The flow from the source is then in a standard way transformed to the flow from a point vortex (on the plane, the stream functions from the source and vortex are conjugate harmonic functions).

As we see, both researchers reach the goal in their own way. Zermelo's approach is more fundamental in that he first systematically builds the general hydrodynamics on a sphere and then postulates the point vortex. At the same time his method is more intuitive and physically natural. Zermelo postulates the vortex in such a way that the constant vorticity is distributed over the entire surface of the sphere, so that the vortex is without the antipodal component. *Borisov, Kilin, and Mamaev 2008* suggest a different method for the physical realization of a point vortex on the sphere, the so-called antipodal model.

Apart from the spherical case, we still have no clear idea of how a vortex can be constructively postulated on other surfaces. Attempts to approach the problem of vortex motions on closed surfaces, specifically on an ellipsoid, are made in *Koiller and Boatto 2008a*. The authors point out that the vortex can move under its own action on a surface of non-constant curvature. This is not surprising, however. In the spherical case this effect is not observed by virtue of symmetry, but, as already pointed out, if a vortex is placed on an arbitrary compact surface, there emerges background vorticity resulting in the vortex motion under the action of the flow caused by the vortex itself.

Stationary motions

In the last section of *1902a* Zermelo studies the so-called stationary flows (when the vorticity remains constant along the streamlines). For this purpose he solves the equation

$$D\psi = f(\psi)\,,$$

where the operator D is determined in (1).

Assuming the function f to be linear of the form $k\psi$ and confining himself to the case $\psi = \psi(\theta)$, Zermelo concludes that the resulting ordinary differen-

tial equation has continuous solutions only for specific values of k. For each admissible k he obtains a solution that divides the entire sphere into parallel zones in each of which the liquid flows in its own direction (Fig. 1). By combining such zonal flows he obtains a vortex structure of cellular type: the surface of the sphere is divided into quadrangles in each of which the sign of vorticity is fixed. In the Mercator projection the resulting image (which, according to Zermelo, resembles a chessboard) is shown in Fig. 2. In conclusion Zermelo points out the practical importance of studying rotational-stationary flows for the solution of meteorological problems.

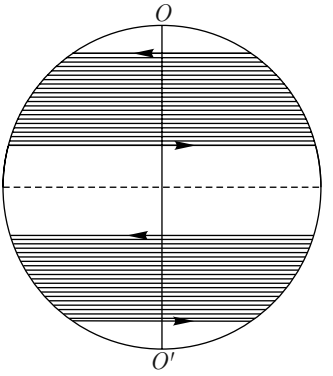


Fig. 1

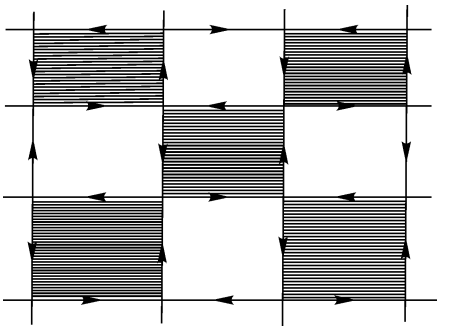


Fig. 2

Chapters 3 and 4 (*s1902b*)

In *s1902b* the dynamics of point vortices on the sphere is systematically developed. In particular, Zermelo carefully studies the problem of three point vortices in relative and absolute space. In fact, he extends to the sphere a number of results which had been obtained about twenty years earlier by Walter Gröbli in his dissertation *Specielle Probleme über die Bewegung geradliniger paralleler Wirbelfäden* (*Special problems on the motion of rectilinear parallel vortices*), *1877*. This work contains a solution to the problem of three vortices on a plane and was defended by Gröbli in Göttingen in 1876; afterwards it lapsed into obscurity for a long time. It is reviewed in *Aref, Rott, and Thomann 1992*. Zermelo might have heard of Gröbli's *1877* from Hermann Amandus Schwarz, their common teacher, who considerably influenced the scientific work of Gröbli and, in particular, attended the defense of his dissertation. However that may be, Zermelo does not cite it.

In *s1902b* Zermelo also obtains quite a few interesting results that are not typical for a planar context.

Equilibrium and motion of a vortex

In the first sections of *s1902b*, Chapter 3, Zermelo formulates some simple propositions concerning a point vortex. For the system of n vortices on the sphere he finds a stream function and shows that as for the planar case, the level lines of the corresponding stream function represent lemniscates.

Further, Zermelo in §3 sets himself the goal of deriving the equations of motion for n point vortices on the sphere. He is guided by the same considerations as for the plane: Having a stream function and assuming that each vortex of the system moves at the velocity induced by other point vortices, he, proceeding from the structure of the stream function, immediately derives the equations of motion in Hamiltonian form:

$$\left.\begin{aligned} \varepsilon_i \sin\theta_i \frac{d\theta}{dt} &= -\frac{\partial H}{\partial \lambda_i} \\ \varepsilon_i \sin\theta_i \frac{d\lambda}{dt} &= -\frac{\partial H}{\partial \theta_i}\,, \qquad i = 1, \dots, n\,. \end{aligned}\right\} \tag{2}$$

Here the (λ_i, θ_i) are the spherical coordinates of the i-vortex of intensity $\Gamma_i = 2\pi\varepsilon_i$, and the Hamiltonian, which Zermelo calls "self-potential", has the form

$$H = \sum_{i,j} \varepsilon_i \varepsilon_j \ln \sin \frac{r_{ij}}{2}\,, \tag{3}$$

where r_{ij} is the chord distance between the i- and j-vortex.

As already noted, Zermelo's equations are identical in form to Bogomolov's equations:

$$\dot\theta_k = \{H, \theta_k\}\,, \qquad \dot\lambda_k = \{H, \lambda_k\}\,,$$

$$\{\lambda_k, \cos\theta_k\} = \frac{\delta_{ik}}{R^2 \Gamma_i}\,,$$

where H coincides with Bogomolov's in *1977* up to a constant multiplier.

Equations (2) look particularly elegant in (redundant) Cartesian coordinates $\boldsymbol{r}_i = (x_{i1}, x_{i2}, x_{i3})$, ${\boldsymbol{r}_i}^2 = R^2$. Paul Newton (*2001*) gave them in the form form:

$$\dot{\boldsymbol{r}}_i = \frac{1}{4\pi R} \sum_{j\neq i}^{N} \Gamma_j \frac{\boldsymbol{r}_j \times \boldsymbol{r}_i}{R^2 - (\boldsymbol{r}_i, \boldsymbol{r}_i)}\,,$$

where $(.\,,\,.)$ is the scalar product in $\mathbb{R}^3$. These equations are also Hamiltonian with the bracket $\{x_{i\alpha}, x_{i\beta}\} = \dfrac{\delta_{i,j}}{R\Gamma_i} \varepsilon_{\alpha\beta\gamma} x_{j\gamma}$.

In *Borisov and Lebedev 1998a, 1998b* another interesting form of the equations is given; this form is based on the use of other variables, namely, the mutual distances between vortices (chords) and the volumes of tetrahedrons and is closely related to equations of *Laura 1902*. It is of interest to note that

the influence of curvature for such equations amounts to the introduction of quadratic terms into the Poisson brackets, which are linear for the motion of vortices on a plane (see *Borisov, Mamaev, and Ramodanov 2008*).

In §4 Zermelo shows that since the equations are Hamiltonian, the self-potential of the vortex system remains unchanged in the course of motion, and so there is an energy-type integral. He also shows that the Hamiltonian depends only on distances and that there exist three integrals of momentum. Further, he proves the theorem of the conservation of the so-called center of gravity of the vortex system (the analogue of the vorticity center on the plane).

It is notable that, to put it in modern terms, these integrals are non-commutative and form the algebra $SO(3)$. That is, to establish integrability one should use the most general, non-commutative variant of the Liouville theorem (see, e.g., *Arnold 1989*, *Borisov and Mamaev 1999*, *Bolsinov and Fomenko 1999*). As is known, Lie's theorem is a local variant of this theorem.

Of course, in Zermelo's time these concepts were poorly assimilated and, therefore, stating the fact that for $2n$ equations of motion of point vortices there are only three (besides energy) first integrals, Zermelo did not notice that this is sufficient for the integrability in quadratures of the problem of motion of three vortices. That is, having obtained the equations in canonical form, Zermelo did not strictly formulate the statement of integrability (in the Liouville sense) of the problem of three vortices on a sphere. Referring to *Poincaré 1893a*, he pointed out that these three integrals of the "center of mass" on the plane give only two independent integrals, and the missing integral is obtained by means of the "law of conservation of the moment of inertia". However, in Poincaré's book we do not find (contrary to what some have said citing it) any exact proof of the Liouville integrability of the three-vortex system. The impression emerges that Poincaré was aware that the three-vortex problem is Liouville integrable, but in his lectures (which may not have been transcribed thoroughly) we do not find a clear formulation of this statement nor any proof. Although Poincaré obtained some integrals, he did not show that they are in involution. Besides, one of the integrals presented by him is not an integral but an identity.

A precise proof of integrability of the problem of three vortices on the plane is obtained in *Laura 1902*. Laura calculates the Poisson bracket of the first integrals, applies Lie's theory, and deduces partial differential equations of the Hamilton-Jacobi type the integration of which is equivalent to the integration of the system of canonical equations. In a subsequent paper (*1905*) Laura obtains the equations of motion in relative distances and indicates a new integral of the equations for n point vortices, which, of course, is expressed in terms of the previously obtained integrals. These equations in relative distances were obtained in modern form in *Borisov and Pavlov 1998*, which allowed new cases of motion to be considered in the four-vortex problem. This form of Laura's equations is—to put it in the modern language of Poisson's structures—a reduced system on the orbit which is determined by the linear Lie-Poisson bracket. As shown in this paper, quadratic terms are

also introduced into the bracket in the spherical case. The issues relating to the application of methods of the theory of Poisson's structures, the theory of Lie's algebra and topology to the study of problems of vortex dynamics, are dealt with in detail in *Borisov and Mamaev 2005b*. A systematic description of Laura's results and some new interesting considerations are given in an extensive work by the little-known Russian scientist N.S. Vasilyev (*1913*). Vasilyev obtained, among other things, a canonical form of the equations of motion for vortex rings.

The last sections of Chapter 3 are of particular interest. Zermelo in §5 concerns himself with the problem of the equilibrium of n vortices on a sphere, for which he uses stereographic and complex coordinates.

§6 considers static configurations of vortices. Zermelo points out the geometric and mechanical interpretations of this problem (Theorem I). The geometric interpretation is to arrange along the circumference n points with masses $\varepsilon_1, \ldots, \varepsilon_n$ (where $\varepsilon_i = \frac{\Gamma_i}{2\pi}$ and Γ_i is the intensity of the i-vortex) in such a way that the product of the $r_{\lambda\mu}^{\varepsilon_\lambda \varepsilon_\mu}$, where $r_{\lambda\mu}$ is the distance from the point λ to the point μ, is maximal or minimal.

The mechanical interpretation deals with the projections of vortices onto the stereographic plane. Theorem II states that the n vortices $\varepsilon_1, \ldots, \varepsilon_n$ are at equilibrium only in the case where for $i = 1, \ldots, n$ the stereographic projection of vortices $\varepsilon_1, \ldots, \varepsilon_{i-1}, \varepsilon_{i+1}, \ldots, \varepsilon_n$ onto the equatorial plane relative to the vortex ε_i has the center of the sphere as its center of gravity (the intensity ε_j is treated as the point's mass).

Further, Zermelo specifies the results obtained for small n. He proves the simple theorem (Theorem III) that two vortices are in equilibrium when they are located at the opposite ends of a diameter; Theorem IV treats the case of the equilibrium of three vortices.

§7 systematically considers the problem of the equilibrium of vortices having equal intensities. Using complex variables, Zermelo obtains an n-th order differential equation with a polynomial solution the roots of which yield the solution of the problem. As far as we know, this equation is not presented in any other work. Of the contemporary papers on this subject, we recommend the papers by O'Neil (*2007, 2008a, 2008b*) which substantially develop Zermelo's results for the cases of equal and different intensities.

Analyzing this differential equation, Zermelo immediately obtains several simple and natural solutions which constitute Theorem III. Its essence is that n equal vortices on a sphere are at equilibrium in the following three cases: 1) if the vortices form a regular polygon inscribed in a great circle; 2) if they form a regular $(n-2)$-gon on a great circle and there is one vortex at each end of the diameter perpendicular to its plane; and 3) if they are the vertices of a regular polyhedron with $n = 4, 6, 8, 12, 20$. All these results are not very difficult and are given in contemporary studies; see, e.g., *Lim, Montaldi, and Roberts 2001*, *Newton 2001*, *Borisov and Mamaev 2005b*.

The final Theorem IV represents an interesting theorem for $n = 4$, which Zermelo proves by methods of elementary geometry: four equal vortices are

at equilibrium if and only if they form either a square inscribed in a great circle or a regular tetrahedron.

Let us formulate an interesting line of research inspired by Zermelo's work. It is known that for $n \leq 4$ all static vortex configurations are symmetric. For $n > 4$ we can assume the existence (not only of symmetric, as partially indicated by Zermelo, but also) of asymmetric static equilibrium configurations. Concerning this assumption, we can refer to an analogy with the case of relative figures of equilibrium (stationary configurations at which the vortex systems rotate as a rigid body), where the existence of asymmetric configurations has already been established (in the case of the plane, examples of such configurations are given in, e.g., *Glass 1997*, *Aref and Vainchtein 1998*.) For contemporary studies of stationary and static vortex configurations on the plane and on the sphere we also recommend *Borisov and Lebedev 1998a* and *Newton 2001*.

A related problem of the stability of rotating regular n-gon vortex configurations, which is due to William Thomson (*1878*) (such configurations now bear his name), has been finally solved in *Kurakin and Yudovich 2002*.

The problem of three vortices on a sphere

Chapter 4 deals with the the integrability problem for three vortices on a sphere. As already noted, without having a general approach, Zermelo virtually reproves for a special case the Liouville theorem of integrability. He introduces mutual distances and studies separately the relative motion of vortices and their absolute motion (i.e. with respect to a fixed in space laboratory frame of reference). He then explicitly integrates the equations of relative motion in mutual distances (§2). He points out that his equations are quite similar to Euler's equations, which describe the rotation of a rigid body around a fixed point, and using this analogy proceeds in the same way as when studying polhodes in the Euler–Poinsot equations (Poinsot's first interpretation).

As was customary in the 19th century, Zermelo understands the integrability problem as one of finding an explicit solution in theta functions (in modern dynamics this seems to be of lesser importance, giving way to more productive methods for the analysis of integrability such as topological and qualitative analysis). He notes that the problem is not solved in the case of arbitrary intensities and considers various special cases. First of all, he specifies the case in which the problem really admits an elliptic quadrature. As is easy to see, it is the case of equal intensities. Analyzing the discriminants of the third-degree equations, Zermelo calculates the periods by inverting elliptic integrals. Selecting the parameters in such a way that the periods tend to infinity, he finds asymptotic solutions and formulates the condition for the stability of relative equilibria. And this, in fact, is how he finds relative equilibria.

These results are all the more remarkable as Zermelo only had at his disposal the technique of theta functions, which is not well suited for obtain-

ing information on the qualitative properties of motion. Of course, from the modern standpoint it seems that his analysis is limited and makes insufficient use of the geometric interpretation. In this sense Zermelo still belongs to the 19th century when the modern methods going back to Poincaré were only emerging and as yet not in wide use among mathematicians.

Further (§3), the general case of various intensities is discussed. In contrast to the case of equal intensities, the solution cannot be obtained in elliptic quadratures. Zermelo points out that the relative motion of three vortices is either periodic or asymptotically tends to the figure of relative equilibrium. In conclusion a classification of these periodic motions and limiting configurations is given.

Absolute motion (*s1902c*)

The main part of Zermelo's *Habilitation* thesis is devoted to the analysis of the relative motion of vortices. But in the concluding section, §5, he also considers the absolute motion for the case of three vortices.[4] He formulates several theorems (rules) some of which admit generalization to the case of an arbitrary number of vortices.

Theorem I is rather simple and asserts that the center of gravity of the vortex system is fixed in space. Theorem II states that during each period of relative motion the entire figure rotates around the central axis (the diameter passing through the center of gravity) by the same angle.

Next, Zermelo introduces the concept of a spherical center, the point A on the sphere such that the radius OA is perpendicular to the plane of the triangle made up of the vortices. Theorem III says that if the spherical center passes through some equatorial plane, its path is symmetric about this plane. The latter two theorems have an immediate analogy with John L. Synge's theorems (*1949*). They state that the general picture of motion of three vortices in absolute space can be made up of separate pieces—paths over the final time interval—by mirroring these pieces according to some symmetry laws. That is to say, generally speaking, it will suffice to describe the path over one period of relative motion and then obtain the entire path already in absolute space.

Theorem IV has no obvious analogy for the case of motion of vortices on a plane. It states that during motion the plane of the vortex triangle sweeps the cone whose vertex is the center of gravity. In a sense, this suggests some parallels with rigid body dynamics. In its geometric and dynamic essence this rule is close to Poinsot's second interpretation of the motion of a free top. As is known, Poinsot imagined the general motion of a rigid body around a fixed point as the motion of some cone that is attached to the rigid body and rolls over another fixed cone (connected with space), the instantaneous

[4] Zermelo's text on absolute motion is rather a tentative draft. It was discovered in the form of a separate, fairly illegible manuscript. See also fn. * on page 465.

axis of this rotation filling some cone both in the body and in space (the so-called axoids). Here Zermelo apparently tries to express analogous ideas for the problem of motion of three vortices. His reasoning is rather interesting, although it lacks the elegant form inherent in Poinsot's reasoning.

Zermelo also endeavours to derive general equations for the angular velocity ω of such motion. However, as he writes himself, he failed to deduce the differential equation by means of which ω is determined in an analytically simple form that would allow a more accurate examination of the cone's shape. Perhaps this interpretation deserves to be studied more closely and made more transparent. In any case, Zermelo's observations are very interesting.

In the conclusion Zermelo considers the simplest motions referred to in the rigid body dynamics as permanent rotations (in Euler's case) or Staude's rotations (in the more general case of a heavy body). (In the general case these are relative equilibria which were systematically examined by Edward John Routh (*1877*); see also *Borisov and Mamaev 2005a*.) Theorem V specifies three cases where the three-vortex system with arbitrary intensities rotates around its central axis as a rigid body: 1) the triangle is equilateral; 2) the main point of this triangle (the center of gravity of the triangle at the vertices of which the masses equal to the cotangents of the corresponding angles are placed) lies on the central axis, and its plane lies in a great circle; and 3) the center of gravity coincides with the center of the sphere. When ω vanishes we come to the case of absolute equilibrium: the vortex triangle lies in a great circle, and its center of gravity coincides with the main point.

In studying the problem of motion of n vortices on a plane, Gröbli (*1877*) considered the absolute motion only for some specific integrable problems, which he managed to solve in quadratures. The general theorems concerning the global properties of motion of the system were obtained much later by Synge (*1949*), which was also devoted to the problem of three vortices on a plane. As we see, Zermelo's statements and theorems are more general than Synge's results in many respects. As far as we know, such results have never been presented in the literature. As a rule, the investigation into the problem of three vortices on a sphere has also been focused on the properties of the relative motion. *Borisov, Kilin, and Mamaev 2005a, 2005b* give a general idea of the absolute motion of a system of three or more vortices on the basis of methods of qualitative analysis. It shows the existence of an extensive class of motions, the so-called choreographies, where in some rotating coordinate system the vortices of equal intensity move one after another along a fixed closed spherical curve. These choreographies can be stable and

play an important role in the analysis of transition to chaos and in gaining a qualitative understanding of other types of motions (*Borisov, Kilin, and Mamaev 2006*).

Epilogue

At Göttingen, under the influence of David Hilbert, Zermelo abandoned his passion for problems of applied nature and focused on abstract issues from set theory and logic. At that time many new branches of mathematics such as quantum mechanics, gravitation theory, etc., were rapidly developing, which, of course, generated new challenges for the mathematical community and required exploration of new avenues and approaches. The leading scientific journals in Europe published a lot of constructive and interesting articles concerned with classical dynamics and dynamical systems (among which Poincaré's contribution was and remains of fundamental unchallenged importance), and most authors were also eminent scientists, real creators of 19th century mathematics, the history of which is thoroughly described and classified in Klein's remarkable lectures.

From about the second quarter of the 20th century the situation changed dramatically. Influenced, perhaps, by the formal aspects of Hilbert's program which aimed at the revision of the foundation of classical mathematics, there appeared a new trend in writing mathematical texts characterized by an unreasonable level of abstraction and formalization and later referred to as "Bourbakism".[5] Journals were flooded with overly formal and obtuse publications often dealing with artificial and hardly relevant constructions. This resulted in a certain "clusterisation" of the scientific community, when even specialists working in the same fields of mathematics became hardly able to understand each other. Perhaps for this reason the 20th century mathematics is still lacking a proper systematization.

The fashion for over-formalization somewhat declined in the 1970's and gave place to a renewed interest (inspired by Poincaré's heritage) in dynamical systems theory and chaos in their connection with mechanics. That is why Zermelo's work on hydrodynamics, now revived and brought into the open, is as never before of current importance and relevance.

[5] "Bourbakism" usually means an unskillful imitation of the Bourbaki style leading to an excessive formalization and cumbersome presentation of trivial things.

Hydrodynamische Untersuchungen über die Wirbelbewegungen in einer Kugelfläche (Erste Mitteilung)

1902a

Die hier vorliegende Arbeit versucht es, die Strömung einer inkompressibelen, reibungslosen (zweidimensionalen) Flüssigkeit in einer Kugelfläche einer ebenso systematischen Theorie zu unterwerfen, wie sie für ebene Strömungen bereits existiert und namentlich in *Poincaré*s „Théorie des tourbillons" (1893) ziemlich vollständig dargestellt ist. Eine solche Untersuchung ist schon an und für sich von geometrischem Interesse, zumal sich auf der Kugel vieles im Endlichen abspielt, was in der Ebene oft im Unendlichen wenigstens für die Anschauung verloren geht. Sodann ist es nicht unmöglich, daß es auf diesem Wege auch gelingen könnte, über manche Vorgänge bei der Fortpflanzung der atmosphärischen Cyklonen, sowie der Meeresströmungen, soweit sie das Erdganze betreffen und soweit die Vertikalkomponente der Strömung gegen die Horizontalkomponenten vernachlässigt werden kann, einigen Aufschluß zu erhalten. Freilich ist dieser geophysikalische Gesichtspunkt, der mir die erste Anregung zu dieser Arbeit gegeben hat, bei der weiteren Durchführung gegenüber den rein geometrisch-analytischen Problemen und Methoden mehr in den Hintergrund getreten. Dabei ist es mein Bestreben gewesen, die Darstellung möglichst einheitlich und unabhängig von fremdartigen Voraussetzungen zu gestalten. Der ganzen Entwickelung liegen ausschließlich die hydrodynamischen Hauptgleichungen in orthogonalen Flächenkoordinaten zu grunde, die gleich im Anfange eingeführt werden, und alle von mir gegebenen litterarischen Citate dienen lediglich als Quellennachweis oder zur Vergleichung.

Unter diesem Gesichtspunkte ist die von *Kirchhoff* zuerst auf das Problem angewandte stereographische Abbildung der Kugel auf die Ebene hier nur beiläufig benutzt worden, obwohl sich mit ihrer Hilfe verschiedene Eigenschaften der Ebene auf die Kugel übertragen lassen, die bei mir direkt hergeleitet werden (cf. *Lamb* Hydrodynamics p. 114, p. 253). Diese Methode der Abbildung ist hier eben keine prinzipielle, sondern nur von beschränkter Anwendbarkeit; sie bezieht sich nur auf das jeweilige momentane Vektorfeld, aber nicht auch auf die Bewegung der Wirbel, auf den zeitlichen Verlauf der Erscheinung. So ist namentlich das Problem der stationären Strömung
202 (II § 4) durchaus nicht durch | Abbildung zu lösen, und vollends das „Gleichgewichtsproblem" der Strudel (K. III §§ 6 u. 7) hat in der Ebene überhaupt kein Analogon. — Aus demselben Grunde ist auch auf die elektromagnetische

Hydrodynamical investigations of vortex motions in the surface of a sphere (First communication)

1902a

The work detailed in the present paper seeks to explain the flow of an incompressible, frictionless (two-dimensional) fluid in a spherical surface[1] by the use of a theory as systematic as the one which already exists for planar flows and of which a fairly complete account is given in particular in *Poincaré*'s "Théorie des tourbillons" (*1893a*). An investigation of this sort is of intrinsic geometric interest in particular because much of what happens in the plane on an infinite scale often defying at least intuition takes place on a sphere on a finite scale. Proceeding along such lines, it should therefore not be impossible to shed light on some of the processes involved both in the propagation of atmospheric cyclones and in the currents of the sea insofar as they affect the entire globe and provided that the vertical component of the current can be disregarded in favor of the horizontal component. In the process, however, this geophysical aspect, which first led me to undertake this work, slowly faded into the background in favor of purely geometric-analytic problems and methods. I have tried to frame an account as uniform as possible and to keep it independent of extraneous assumptions. The entire development rests exclusively on the principal equations of hydrodynamics in orthogonal surface coordinates, which are introduced right at the outset. All quotations from the literature serve but the purpose of reference or comparison.

In following this approach I have made only perfunctory use of the stereographic projection of the sphere onto the plane, which was first applied to the problem by *Kirchhoff*, although it is possible to carry over various properties of the plane, which I directly derive (cf. *Lamb 1895*, p. 114, p. 253), to the sphere by means of it. For, in this case, this method of projection is not comprehensive. Rather, its applicability is only limited; it only refers to the particular current vector field but not to the motion of the vortices,[2] to the temporal evolution of the phenomenon. Thus, in particular, the problem of the stationary flow (II § 4) can definitely not be solved by means of the projection, and the "equilibrium problem" of the whirls (ch. III §§ 6 u. 7) has

[1] ⟦Zermelo's use of 'in' vs. 'auf', i.e., of 'in' vs. 'on', with respect to surfaces is strictly followed.⟧

[2] ⟦When referring to vortices in general (e. g. to vortex patches, vortex filaments, vorticity), Zermelo uses the term "Wirbel". When the vorticity is concentrated at a single point, this point is referred to as "(einfacher) Strudel" or "Strudelpunkt". In order to preserve Zermelo's terminology, but in contrast to the modern use of the term "vortex" also to include the latter case, "Wirbel" is always translated as "vortex"and "Strudel" as "whirl".⟧

Deutung hier keine Rücksicht genommen. Diese Analogie gilt nur in Bezug auf mögliche Strömungszustände, nicht auf ihre zeitliche Veränderung, die vielmehr der Hydrodynamik ganz charakteristisch ist.

Der Kern der hier verwendeten Methode ist in dem Begriffe des „einfachen Strudels“ zu suchen (cf. II § 2): d. h. eines isolierten Strudelpunktes bei konstanter (von 0 verschiedener) Wirbeldichte (Curl) auf der ganzen übrigen Kugel, während die früheren Autoren meines Wissens immer nur Strudelpunkte (unendlich dünne Wirbel) in sonst wirbelfreier Flüssigkeit betrachtet hatten. Auf diesem Begriffe ruhen fast alle weiteren Entwicklungen und verdanken ihm ihre zwanglose Formulierung bei Vermeidung störender Nebenbedingungen.

Das im letzten Kapitel behandelte „Problem der drei Strudel“ als eine spezielle Anwendung der vorausgehenden allgemeinen Theorie dürfte namentlich interessieren durch eine gewisse formale Analogie mit dem Kreiselproblem und eine geometrische mit dem astronomischen „Dreikörperproblem“.

Kapitel I.

Die Flüssigkeitsbewegung auf einer beliebigen Fläche.

§ 1. Die Grundgleichungen in *Gauß*schen Koordinaten.

Sind u und v krummlinige Koordinaten auf einer gegebenen Fläche und E, F, G die bekannten *Gauß*schen Fundamentalgrößen, so ist der Ausdruck für das Quadrat des Linienelementes

$$ds^2 = Edu^2 + 2Fdudv + Gdv^2$$

und daher die lebendige Kraft eines Punktes von der Masse m, welcher sich auf der Fläche bewegt:

$$T = \frac{m}{2}q^2 = \frac{m}{2}\left(Eu'^2 + 2Fu'v' + Gv'^2\right) ,$$

wenn u', v' die nach der Zeit t genommenen Ableitungen von u, v bedeuten. Ist ferner $m\Phi$ das Potential der wirkenden Kräfte, so nehmen die *Lagrange*schen Bewegungsgleichungen (zweiter Art) für unseren Fall die Form an

$$m\frac{d}{dt}\left(Eu' + Fv'\right) - \frac{m}{2}\left(\frac{\partial E}{\partial u}u'^2 + 2\frac{\partial F}{\partial u}u'v' + \frac{\partial G}{\partial u}v'^2\right) = -m\frac{\partial \Phi}{\partial u}$$
$$m\frac{d}{dt}\left(Fu' + Gv'\right) - \frac{m}{2}\left(\frac{\partial E}{\partial v}u'^2 + 2\frac{\partial F}{\partial v}u'v' + \frac{\partial G}{\partial v}v'^2\right) = -m\frac{\partial \Phi}{\partial v} .$$

203 | Wählt man orthogonale Koordinaten u, v, so wird $F = 0$ und man kann setzen:

$$E = \frac{1}{U^2} , \quad G = \frac{1}{V^2} ,$$

absolutely no analog in the plane.—For the same reason, I have also ignored the electromagnetic interpretation. This analogy only holds for possible flow states but not for their changes over time, which properly belong to hydrodynamics.

At its core, the method used here proceeds from the concept of the "simple whirl" (cf. II § 2): i. e., of an isolated whirl point at constant (different from 0) vortex density (curl) on the entire remaining sphere, while, to my knowledge, the previous authors have always only considered whirl points (infinitely thin vortices) in otherwise vortex-free fluids. Nearly all of the further developments are based on this concept, which also makes it possible to frame them in a casual way and avoid troublesome ancillary conditions.

The "problem of three vortices", which is discussed in the last chapter as a special application of the preceding general theory, should be of particular interest on account of a certain, formal analogy with the spinning top problem and a geometric analogy with the "three-body problem" in astronomy.

Chapter I.

Fluid motion on an arbitrary surface.

§ 1. The basic equations in *Gauss*ian coordinates.

If u and v are curvilinear coordinates on a given surface, and E, F, G the well-known *Gauss*ian fundamental quantities, then the expression for the square of the line element is

$$ds^2 = Edu^2 + 2Fdudv + Gdv^2$$

and hence the living force of a point of mass m moving on a surface:

$$T = \frac{m}{2}q^2 = \frac{m}{2}\left(Eu'^2 + 2Fu'v' + Gv'^2\right) ,$$

if u', v' denote the derivatives of u, v taken with respect to time t. Furthermore, if $m\Phi$ is the potential of the acting forces, then, for our case, *Lagrange*'s equations of motion (of the second kind) assume the form

$$m\frac{d}{dt}\left(Eu' + Fv'\right) - \frac{m}{2}\left(\frac{\partial E}{\partial u}u'^2 + 2\frac{\partial F}{\partial u}u'v' + \frac{\partial G}{\partial u}v'^2\right) = -m\frac{\partial \Phi}{\partial u}$$
$$m\frac{d}{dt}\left(Fu' + Gv'\right) - \frac{m}{2}\left(\frac{\partial E}{\partial v}u'^2 + 2\frac{\partial F}{\partial v}u'v' + \frac{\partial G}{\partial v}v'^2\right) = -m\frac{\partial \Phi}{\partial v} .$$

If we choose orthogonal coordinates u, v, then we have $F = 0$ and we can put:

$$E = \frac{1}{U^2} , \quad G = \frac{1}{V^2} ,$$

also

$$ds^2 = \frac{du^2}{U^2} + \frac{dv^2}{V^2} .$$

Dann werden

$$\overline{u} = \frac{u'}{U} \quad \text{und} \quad \overline{v} = \frac{v'}{V}$$

die wahren Geschwindigkeitskomponenten in der Richtung der Koordinaten u, v, also:

$$u' = \frac{du}{dt} = U\overline{u} , \quad v' = \frac{dv}{dt} = V\overline{v} , \tag{a}$$

und unsere Bewegungsgleichungen werden:

$$\begin{aligned} \frac{d}{dt}\left(\frac{\overline{u}}{U}\right) + \frac{\overline{u}^2}{U}\frac{\partial U}{\partial u} + \frac{\overline{v}^2}{V}\frac{\partial V}{\partial u} &= -\frac{\partial \Phi}{\partial u} , \\ \frac{d}{dt}\left(\frac{\overline{v}}{V}\right) + \frac{\overline{u}^2}{U}\frac{\partial U}{\partial v} + \frac{\overline{v}^2}{V}\frac{\partial V}{\partial v} &= -\frac{\partial \Phi}{\partial v} . \end{aligned} \tag{1}$$

Da nun aber

$$\begin{aligned} \frac{d}{dt}\left(\frac{\overline{u}}{U}\right) &= \frac{1}{U}\frac{d\overline{u}}{dt} - \frac{\overline{u}}{U^2}\left(\frac{\partial U}{\partial u}U\overline{u} + \frac{\partial U}{\partial v}V\overline{v}\right) , \\ \frac{d}{dt}\left(\frac{\overline{v}}{V}\right) &= \frac{1}{V}\frac{d\overline{v}}{dt} - \frac{\overline{v}}{V^2}\left(\frac{\partial V}{\partial u}U\overline{u} + \frac{\partial V}{\partial v}V\overline{v}\right) , \end{aligned}$$

so kann man den Gleichungen (1) auch die Form geben:

$$\begin{aligned} \frac{d\overline{u}}{dt} - \overline{v}W &= -U\frac{\partial \Phi}{\partial u} , \\ \frac{d\overline{v}}{dt} + \overline{u}W &= -V\frac{\partial \Phi}{\partial v} , \end{aligned} \tag{2}$$

wenn

$$W = \frac{V}{U}\frac{\partial U}{\partial v}\overline{u} - \frac{U}{V}\frac{\partial V}{\partial u}\overline{v}$$

gesetzt wird.

Diese Differentialgleichungen gelten zunächst nur für einen einzigen materiellen Punkt m. Sie behalten aber ihre Form, wenn man zu einer kontinuierlichen Verteilung von Massenpunkten auf der Fläche, d. h. zu einer reibungslosen zweidimensionalen Flüssigkeit übergeht. Nur hat dann an Stelle von m die variable *Flächendichte* k zu treten, und das auf die Flächeneinheit bezogene Potential der Massenkräfte $k\Phi$ ist noch um eine Funktion $p = p(u, v)$

and hence

$$ds^2 = \frac{du^2}{U^2} + \frac{dv^2}{V^2} .$$

We thus have

$$\overline{u} = \frac{u'}{U} \quad \text{and} \quad \overline{v} = \frac{v'}{V}$$

as the true velocity components in the direction of the coordinates u, v, and hence:

$$u' = \frac{du}{dt} = U\overline{u} , \quad v' = \frac{dv}{dt} = V\overline{v} , \tag{a}$$

and our equations of motion become:

$$\begin{aligned} \frac{d}{dt}\left(\frac{\overline{u}}{U}\right) + \frac{\overline{u}^2}{U}\frac{\partial U}{\partial u} + \frac{\overline{v}^2}{V}\frac{\partial V}{\partial u} &= -\frac{\partial \Phi}{\partial u} , \\ \frac{d}{dt}\left(\frac{\overline{v}}{V}\right) + \frac{\overline{u}^2}{U}\frac{\partial U}{\partial v} + \frac{\overline{v}^2}{V}\frac{\partial V}{\partial v} &= -\frac{\partial \Phi}{\partial v} . \end{aligned} \tag{1}$$

But now since

$$\begin{aligned} \frac{d}{dt}\left(\frac{\overline{u}}{U}\right) &= \frac{1}{U}\frac{d\overline{u}}{dt} - \frac{\overline{u}}{U^2}\left(\frac{\partial U}{\partial u}U\overline{u} + \frac{\partial U}{\partial v}V\overline{v}\right) , \\ \frac{d}{dt}\left(\frac{\overline{v}}{V}\right) &= \frac{1}{V}\frac{d\overline{v}}{dt} - \frac{\overline{v}}{V^2}\left(\frac{\partial V}{\partial u}U\overline{u} + \frac{\partial V}{\partial v}V\overline{v}\right) , \end{aligned}$$

equations (1) can also be given the form:

$$\begin{aligned} \frac{d\overline{u}}{dt} - \overline{v}W &= -U\frac{\partial \Phi}{\partial u} , \\ \frac{d\overline{v}}{dt} + \overline{u}W &= -V\frac{\partial \Phi}{\partial v} , \end{aligned} \tag{2}$$

if we set

$$W = \frac{V}{U}\frac{\partial U}{\partial v}\overline{u} - \frac{U}{V}\frac{\partial V}{\partial u}\overline{v}.$$

At first, these differential equations are only valid for a single material point m. However, they retain their form if we turn to a continuous distribution of mass points on the surface, i. e., to a frictionless two-dimensional fluid. We only need to substitute for m the variable *surface density* k and to increase the potential of mass forces $k\Phi$ that is related to the surface unit by a function $p = p(u, v)$ of the location on the surface, which arises from

des Ortes auf der Fläche zu vermehren, welche von der gegenseitigen Beeinflussung der materiellen Punkte herrührt und der *Druck* der Flüssigkeit genannt wird; *pds* ist dann immer die auf das Linienelement *ds* normal wir-
204 kende Kraft. Be-|merkt sei nur noch, daß wir gleichfalls zu den aufgestellten Grundgleichungen gelangen, wenn wir die Strömung einer dreidimensionalen Flüssigkeit längs eines Systems von Parallelflächen betrachten und dann die Entfernung der beiden Grenzflächen, also die Dicke der Flüssigkeitsschicht an der Grenze null werden lassen. — Ist nun der Druck, wie im Folgenden vorausgesetzt werden soll, eine Funktion der Dichte k allein, so können wir setzen

$$\int \frac{dp}{k} = P \quad \text{(die „Druckfunktion“)} ,$$

also

$$\frac{1}{k}\frac{\partial p}{\partial u} = \frac{\partial P}{\partial u} , \quad \frac{1}{k}\frac{\partial p}{\partial v} = \frac{\partial P}{\partial v} ,$$

und die Differentialgleichungen (1), (2) gelten auch für unsere Flüssigkeit wenn wir setzen:

$$\Phi = P + \Phi_1 ,$$

wo Φ_1 die wahre Potentialfunktion der Massenkräfte, bezogen auf die Masseneinheit, bedeutet.

Bisher haben wir nur die zeitliche Veränderung der Geschwindigkeitskomponenten $\frac{d\overline{u}}{dt}$, $\frac{d\overline{v}}{dt}$ durch die Potentialverteilung Φ und die Geschwindigkeiten $\overline{u}$, $\overline{v}$ selbst ausgedrückt. Wir können aber auch die Geschwindigkeitsveränderung $\frac{\partial\overline{u}}{\partial t}$, $\frac{\partial\overline{v}}{\partial t}$ an einer bestimmten Stelle u, v der Fläche einführen, indem wir die Beziehungen benutzen

$$\left.\begin{aligned} \frac{d\overline{u}}{dt} &= \frac{\partial\overline{u}}{\partial t} + \overline{u}U\frac{\partial\overline{u}}{\partial u} + \overline{v}V\frac{\partial\overline{u}}{\partial v} \\ \frac{d\overline{v}}{dt} &= \frac{\partial\overline{v}}{\partial t} + \overline{u}U\frac{\partial\overline{v}}{\partial u} + \overline{v}V\frac{\partial\overline{v}}{\partial v} \end{aligned}\right\} \quad \text{(cf. (a))} , \tag{b}$$

und erhalten so aus (2)

$$\begin{aligned} \frac{\partial\overline{u}}{\partial t} + U\left(\overline{u}\frac{\partial\overline{u}}{\partial u} + \overline{v}\frac{\partial\overline{v}}{\partial u}\right) - \overline{v}UV\frac{\partial}{\partial u}\left(\frac{\overline{v}}{V}\right) + \overline{v}UV\frac{\partial}{\partial v}\left(\frac{\overline{u}}{U}\right) &= -U\frac{\partial\Phi}{\partial u} \\ \frac{\partial\overline{v}}{\partial t} + V\left(\overline{u}\frac{\partial\overline{u}}{\partial v} + \overline{v}\frac{\partial\overline{v}}{\partial v}\right) + \overline{u}UV\frac{\partial}{\partial u}\left(\frac{\overline{v}}{V}\right) - \overline{u}UV\frac{\partial}{\partial v}\left(\frac{\overline{u}}{U}\right) &= -V\frac{\partial\Phi}{\partial v} \end{aligned}$$

Berücksichtigt man hier, daß

$$\begin{aligned} \overline{u}\frac{\partial\overline{u}}{\partial u} + \overline{v}\frac{\partial\overline{v}}{\partial u} &= \frac{1}{2}\frac{\partial}{\partial u}\left(\overline{u}^2 + \overline{v}^2\right) = \frac{1}{2}\frac{\partial q^2}{\partial u} , \\ \overline{u}\frac{\partial\overline{u}}{\partial v} + \overline{v}\frac{\partial\overline{v}}{\partial v} &= \frac{1}{2}\frac{\partial}{\partial v}\left(\overline{u}^2 + \overline{v}^2\right) = \frac{1}{2}\frac{\partial q^2}{\partial v} , \end{aligned}$$

the mutual interactions among the material points and which is called the *pressure* of the fluid; *pds* is then always the force acting orthogonally on the line element *ds*. It should be noted that we also arrive at the basic equations if we consider the flow of a three-dimensional fluid along a system of parallel surfaces and then allow the distance between the two boundary surfaces, and hence the depth of the fluid stratum, to become zero.—Now if the pressure is a function of the density k alone, as is assumed in what follows, then we may set

$$\int \frac{dp}{k} = P \quad \text{(the "pressure function")} ,$$

and hence

$$\frac{1}{k}\frac{\partial p}{\partial u} = \frac{\partial P}{\partial u} , \quad \frac{1}{k}\frac{\partial p}{\partial v} = \frac{\partial P}{\partial v} ,$$

and the differential equations (1), (2) also hold for our fluid if we set:

$$\Phi = P + \Phi_1 ,$$

where Φ_1 denotes the true potential function of the mass forces, as related to the mass unit.

As yet we have only expressed the temporal change of the velocity components $\frac{d\overline{u}}{dt}$, $\frac{d\overline{v}}{dt}$ by means of the potential distribution Φ and the velocities $\overline{u}$, $\overline{v}$ themselves. We can, however, also introduce the change in velocity $\frac{\partial \overline{u}}{\partial t}$, $\frac{\partial \overline{v}}{\partial t}$ at a particular location u, v of the surface by using the relations

$$\left.\begin{aligned} \frac{d\overline{u}}{dt} &= \frac{\partial \overline{u}}{\partial t} + \overline{u}U\frac{\partial \overline{u}}{\partial u} + \overline{v}V\frac{\partial \overline{u}}{\partial v} \\ \frac{d\overline{v}}{dt} &= \frac{\partial \overline{v}}{\partial t} + \overline{u}U\frac{\partial \overline{v}}{\partial u} + \overline{v}V\frac{\partial \overline{v}}{\partial v} \end{aligned}\right\} \quad \text{(cf. (a))} , \qquad \text{(b)}$$

and, from (2), we thus get

$$\begin{aligned} \frac{\partial \overline{u}}{\partial t} + U\left(\overline{u}\frac{\partial \overline{u}}{\partial u} + \overline{v}\frac{\partial \overline{v}}{\partial u}\right) - \overline{v}UV\frac{\partial}{\partial u}\left(\frac{\overline{v}}{V}\right) + \overline{v}UV\frac{\partial}{\partial v}\left(\frac{\overline{u}}{U}\right) &= -U\frac{\partial \Phi}{\partial u} \\ \frac{\partial \overline{v}}{\partial t} + V\left(\overline{u}\frac{\partial \overline{u}}{\partial v} + \overline{v}\frac{\partial \overline{v}}{\partial v}\right) + \overline{u}UV\frac{\partial}{\partial u}\left(\frac{\overline{v}}{V}\right) - \overline{u}UV\frac{\partial}{\partial v}\left(\frac{\overline{u}}{U}\right) &= -V\frac{\partial \Phi}{\partial v} \end{aligned}$$

If we consider here that

$$\begin{aligned} \overline{u}\frac{\partial \overline{u}}{\partial u} + \overline{v}\frac{\partial \overline{v}}{\partial u} &= \frac{1}{2}\frac{\partial}{\partial u}\left(\overline{u}^2 + \overline{v}^2\right) = \frac{1}{2}\frac{\partial q^2}{\partial u} , \\ \overline{u}\frac{\partial \overline{u}}{\partial v} + \overline{v}\frac{\partial \overline{v}}{\partial v} &= \frac{1}{2}\frac{\partial}{\partial v}\left(\overline{u}^2 + \overline{v}^2\right) = \frac{1}{2}\frac{\partial q^2}{\partial v} , \end{aligned}$$

und setzt weiter:

$$2\varrho = UV\left\{\frac{\partial}{\partial u}\left(\frac{\overline{v}}{V}\right) - \frac{\partial}{\partial v}\left(\frac{\overline{u}}{U}\right)\right\} \tag{c}$$

205 | (ϱ ist die Rotation um die Flächennormale, der „Curl“ oder die „Wirbeldichte“ cf. § 2), so kommt schließlich:

$$\left.\begin{aligned}\frac{\partial \overline{u}}{\partial t} &= \quad 2\overline{v}\varrho - U\frac{\partial}{\partial u}\left(\Phi + \frac{1}{2}q^2\right) \ , \\ \frac{\partial \overline{v}}{\partial t} &= -2\overline{u}\varrho - V\frac{\partial}{\partial v}\left(\Phi + \frac{1}{2}q^2\right) \ .\end{aligned}\right\} \tag{3}$$

Satz. *Die zeitliche Veränderung des Geschwindigkeitsvektors an einer Stelle der Fläche setzt sich zusammen aus zwei Komponenten, deren eine gleich dem doppelten Produkt aus der Wirbeldichte ϱ und der Geschwindigkeit ist und auf der Geschwindigkeitsrichtung senkrecht steht, und deren andere an Richtung und Größe gleich dem Gefälle der Funktion χ ist, die sich additiv zusammensetzt aus dem halben Geschwindigkeitsquadrate der Druckfunktion P und event. der Potentialfunktion Φ_1 der wirkenden Massenkräfte.*

(Den Spezialfall für die Ebene ($U = 1$, $V = 1$) vergl. bei *Lamb*, Hydrodynamics p. 226.)

§ 2. Der Massenfluß und die Inkompressibilität.

Unter dem „Massenfluß“ $K_{\mathfrak{C}}$ durch ein gegebenes Kurvenstück $\mathfrak{C}$ verstehen wir bei stationärer Strömung die Gesamtmasse der Flüssigkeit, welche in der Zeiteinheit das Kurvenstück in einem bestimmten Sinne durchströmt. Ist die Strömung nicht gerade stationär, so haben wir die in der Zeit τ durchströmende Masse durch τ zu dividieren und für $\tau = 0$ zur Grenze überzugehen. So erhalten wir den Ausdruck

$$K_{\mathfrak{C}} = \int^{(\mathfrak{C})} k q_n ds \ , \tag{a}$$

wenn das Integral über das Kurvenstück $\mathfrak{C}$ mit der variablen Bogengänge s erstreckt wird und q_n die Geschwindigkeitskomponente in der positiv gerechneten Normalen n der Kurve bezeichnet. In unseren orthogonalen Koordinaten u, v ist aber

$$q_n = \frac{\overline{v}}{U}\frac{du}{ds} - \frac{\overline{u}}{V}\frac{dv}{ds} \ ; \tag{b}$$

und damit diese Gleichung auch das Vorzeichen von q_n immer richtig bestimmt, wollen wir im Folgenden immer als „positive“ Kurvennormale diejenige Richtung (in der Tangentialebene der Fläche) bezeichnen, welche zur

and also set:

$$2\varrho = UV\left\{\frac{\partial}{\partial u}\left(\frac{\overline{v}}{V}\right) - \frac{\partial}{\partial v}\left(\frac{\overline{u}}{U}\right)\right\} \tag{c}$$

(ϱ is the rotation about the surface normal, the "curl" or the "vortex density" cf. § 2), we eventually get:

$$\left.\begin{aligned}\frac{\partial \overline{u}}{\partial t} &= \quad 2\overline{v}\varrho - U\frac{\partial}{\partial u}\left(\Phi + \frac{1}{2}q^2\right)\ ,\\ \frac{\partial \overline{v}}{\partial t} &= -2\overline{u}\varrho - V\frac{\partial}{\partial v}\left(\Phi + \frac{1}{2}q^2\right)\ .\end{aligned}\right\} \tag{3}$$

Theorem. *The temporal change of the velocity vector at a particular location on the surface is composed of two components one of which equals the double product of the vortex density of ϱ and the velocity and which is perpendicular to the velocity direction, and the other of which is equal with respect to direction and magnitude to the gradient of the function χ, which is additively composed of half the square of the velocity of the pressure function P and possibly $[\![$also$]\!]$ of the potential function Φ_1 of the acting mass forces.*

(As for the special case for the plane ($U = 1$, $V = 1$), see *Lamb 1895*, p. 226.)

§ 2. Mass flux and incompressibility.

By the "mass flux" $K_{\mathfrak{C}}$ through a given curve segment $\mathfrak{C}$ we understand the total mass of a fluid at stationary flow passing through the curve segment in a certain direction per unit time. If the flow is not stationary, then we must divide the mass passing through during time τ by τ and take the limit for $\tau = 0$. We thus get the expression

$$K_{\mathfrak{C}} = \int^{(\mathfrak{C})} k q_n ds\ , \tag{a}$$

if the integral is extended over the curve segment $\mathfrak{C}$ with the variable arc length s and if q_n denotes the velocity component in the normal n of the curve taken as positive. Using our orthogonal coordinates u, v, however, we have

$$q_n = \frac{\overline{v}}{U}\frac{du}{ds} - \frac{\overline{u}}{V}\frac{dv}{ds}\ ; \tag{b}$$

and in order to always ensure that this equation also correctly determines the sign of q_n we shall, in what follows, refer to that direction (in the tangent plane of the surface) as the "positive" normal of the curve which is situated

Fortschreitungsrichtung ds auf der Kurve ebenso liegt, wie die v-Richtung zur u-Richtung, also nach *links*, wenn wir, von einer bestimmten Seite auf die Fläche blickend, das uv-Koordinatensystem ebenso zeichnen wollen wie
206 in der Ebene gewöhnlich das xy-System. | Es wird also immer:

$$K_{\mathfrak{C}} = \overset{(\mathfrak{C})}{\int} k \left(\frac{\overline{v}}{U} du - \frac{\overline{u}}{V} dv \right) , \tag{c}$$

wobei aber auch der Sinn angegeben werden muß, in welchem $\mathfrak{C}$ durchlaufen werden soll.

Besonders wichtig ist der Fall, wo die Kurve $\mathfrak{C}$ *geschlossen* ist, ohne sich selbst zu durchschneiden, und ein endliches Flächenstück C so einschließt, daß die positive (linke) Normale immer nach innen weist; wir sagen dann, die Begrenzung von C werde „im positiven Sinne“ durchlaufen. Dann ist $K = K_{\mathfrak{C}}$ der Gesamtbetrag der in der Zeiteinheit in das Flächenstück C einströmenden Masse, also auf Grund des Prinzips der Konstanz der Materie gleich der gesamten zeitlichen Massenvermehrung in C, d. h. gleich dem Flächenintegrale

$$\overset{(C)}{\int} \frac{\partial k}{\partial t} d\sigma = \overset{(C)}{\int} \frac{\partial k}{\partial t} \frac{du}{U} \frac{dv}{V} ,$$

wenn mit $d\sigma$ das Flächenelement bezeichnet wird.

Nun läßt sich aber das Linienintegral K auch rein formal in ein Flächenintegral verwandeln:

$$K = \overset{(\mathfrak{C})}{\int} k \left(\frac{\overline{v}}{U} du - \frac{\overline{u}}{V} dv \right) = - \overset{(\mathfrak{C})}{\int} du dv \left\{ \frac{\partial}{\partial u} \left(\frac{k\overline{u}}{V} \right) + \frac{\partial}{\partial v} \left(\frac{k\overline{v}}{U} \right) \right\} . \tag{1}$$

Setzen wir diesen Ausdruck für K gleich dem eben gefundenen für die Massenvermehrung, so wird

$$\overset{(\mathfrak{C})}{\int} du dv \left\{ \frac{1}{UV} \frac{\partial k}{\partial t} + \frac{\partial}{\partial u} \left(\frac{k\overline{u}}{V} \right) + \frac{\partial}{\partial v} \left(\frac{k\overline{v}}{U} \right) \right\} = 0 ,$$

und zwar für ein ganz beliebiges Flächenstück C, also muß auch der Integrandus verschwinden, und wir erhalten die sogen. „Kontinuitätsbedingung“:

$$\frac{1}{UV} \frac{\partial k}{\partial t} + \frac{\partial}{\partial u} \left(\frac{k\overline{u}}{V} \right) + \frac{\partial}{\partial v} \left(\frac{k\overline{v}}{U} \right) = 0 , \tag{2}$$

welche als dritte Grundgleichung zu jedem der Gleichungspaare (1), (2) oder (3) in § 1 hinzuzuziehen ist.

on the curve with respect to the direction of progression ds just like the v-direction is situated with respect to the u-direction, and hence towards the *left*, if, looking upon the surface from a particular side, we draw up the uv-coordinate system just as we usually draw up the xy-system. Hence, we always have:

$$K_{\mathfrak{C}} = \overset{(\mathfrak{C})}{\int} k \left(\frac{\overline{v}}{U} du - \frac{\overline{u}}{V} dv \right) , \tag{c}$$

where we also need to specify the direction in which $\mathfrak{C}$ is supposed to be passed through.

Of particular significance is the case where the curve $\mathfrak{C}$ is *closed* without intersecting itself and encloses a finite part C of the surface,[3] so that the positive (left) normal always points inwards; we then say that the boundary of C is passed through "in the positive direction". Then $K = K_{\mathfrak{C}}$ is the total amount of the mass flowing into the part C per time unit, and hence, by virtue of the principle of constant mass, equal to the total temporal mass increase in C, i. e., equal to the surface integral

$$\overset{(C)}{\int} \frac{\partial k}{\partial t} d\sigma = \overset{(C)}{\int} \frac{\partial k}{\partial t} \frac{du}{U} \frac{dv}{V} ,$$

if $d\sigma$ denotes the area element.

Now, however, it is also possible to purely formally transform the line integral K into a surface integral:

$$K = \overset{(\mathfrak{C})}{\int} k \left(\frac{\overline{v}}{U} du - \frac{\overline{u}}{V} dv \right) = - \overset{(\mathfrak{C})}{\int} du dv \left\{ \frac{\partial}{\partial u} \left(\frac{k\overline{u}}{V} \right) + \frac{\partial}{\partial v} \left(\frac{k\overline{v}}{U} \right) \right\} . \tag{1}$$

If we set this expression for K equal to the expression for the mass increase we just found, then we get

$$\overset{(\mathfrak{C})}{\int} du dv \left\{ \frac{1}{UV} \frac{\partial k}{\partial t} + \frac{\partial}{\partial u} \left(\frac{k\overline{u}}{V} \right) + \frac{\partial}{\partial v} \left(\frac{k\overline{v}}{U} \right) \right\} = 0 ,$$

in particular for any part C of the surface, and hence the integrand must disappear as well, and we obtain the so-called "continuity condition":

$$\frac{1}{UV} \frac{\partial k}{\partial t} + \frac{\partial}{\partial u} \left(\frac{k\overline{u}}{V} \right) + \frac{\partial}{\partial v} \left(\frac{k\overline{v}}{U} \right) = 0 , \tag{2}$$

which is to be added as a third fundamental equation to each of the pairs of equations (1), (2) or (3) in § 1.

[3] ⟦Zermelo's terms "Flächenstück", "Flächenelement", and "Gebiet" have been translated as "part of the surface", "area element", and "region", respectively.⟧

Für eine „homogene und inkompressibele“ Flüssigkeit ist nun k der Definition nach in Raum und Zeit *konstant* und zwar, wie wir annehmen wollen, $= 1$, und wir gewinnen aus (2) die „Inkompressibilitätsbedingung“

$$\frac{\partial}{\partial u}\left(\frac{\overline{u}}{V}\right) + \frac{\partial}{\partial v}\left(\frac{\overline{v}}{U}\right) = 0\,. \tag{3}$$

207 | In diesem Falle ist für jede *geschlossene* Kurve $\mathfrak{C}$ das Integral

$$K_{\mathfrak{C}} = \int^{(\mathfrak{C})} \left(\overline{v}\frac{du}{U} - \overline{u}\frac{dv}{V}\right) = 0$$

und daher für jedes *offene* Kurvenstück $\mathfrak{C} = \overset{\frown}{AB}$ der Massenfluß $K_{\mathfrak{C}}$ ganz unabhängig von der Gestalt des Verbindungsweges zwischen A und B, so daß wir einfach von dem Massenfluß $K = K_{\mathrm{AB}}$ „zwischen A und B“ oder „von A nach B“ reden können.

Wählen wir daher einen beliebigen festen Punkt O unserer Fläche zum Anfangspunkt, so hat für jeden anderen Flächenpunkt $P \equiv P(u, v)$ unser Integral

$$K_{OP} = \int_{(O)}^{(P)} \left(\overline{v}\frac{du}{U} - \overline{u}\frac{dv}{V}\right) = \psi(u, v) \tag{1a}$$

einen ganz bestimmten nur von den Koordinaten u, v von P abhängigen Wert, den wir als die „*Stromfunktion*“ $\psi(u, v)$ im Punkte P bezeichnen wollen. Bei einer Veränderung des Anfangspunktes O ändert sich diese Funktion offenbar nur um eine additive Konstante. Die partiellen Ableitungen der Stromfunktion sind dann gegeben durch:

$$\left.\begin{aligned} \frac{\partial\psi}{\partial u} = \frac{\overline{v}}{U}\,, \quad \frac{\partial\psi}{\partial v} = -\frac{\overline{u}}{V} \quad \text{oder} \\ \overline{u} = -V\frac{\partial\psi}{\partial v}\,, \quad \overline{v} = +U\frac{\partial\psi}{\partial u} \end{aligned}\right\} \tag{4}$$

und daher:

$$\left.\begin{aligned} \frac{du}{dt} = U\overline{u} = -UV\frac{\partial\psi}{\partial v}\,, \\ \frac{dv}{dt} = V\overline{v} = +UV\frac{\partial\psi}{\partial u}\,. \end{aligned}\right\} \tag{4a}$$

Die ganze Geschwindigkeitsverteilung ist also vollständig bestimmt durch die einzige Funktion $\psi(u, v)$, wodurch alle Untersuchungen über inkompressible Flüssigkeiten in der Fläche wesentlich vereinfacht werden.

Auf den Kurven $\psi = \mathrm{const}$ ist nach der Definition (1a) von $K = \psi(u, v)$ überall $q_n = 0$, d. h. die Geschwindigkeit tangential gerichtet; sie werden als „Stromlinien“ bezeichnet.

In case of a "homogeneous and incompressible" fluid, k is, by definition, *constant* in space and time, and in particular, as we shall assume, $= 1$, and from (2) we obtain the "incompressibility condition"

$$\frac{\partial}{\partial u}\left(\frac{\overline{u}}{V}\right) + \frac{\partial}{\partial v}\left(\frac{\overline{v}}{U}\right) = 0\,. \tag{3}$$

In this case, for any *closed* curve $\mathfrak{C}$, the integral

$$K_{\mathfrak{C}} = \int^{(\mathfrak{C})} \left(\overline{v}\frac{du}{U} - \overline{u}\frac{dv}{V}\right) = 0$$

and hence, for any *open* curve segment $\mathfrak{C} = \overset{\frown}{AB}$ the mass flux $K_{\mathfrak{C}}$ is entirely independent of the shape of the connecting path between A and B so that we can simply speak of the mass flux $K = K_{\mathrm{AB}}$ "between A and B" or "from A to B".

Hence, if we choose some fixed point O of our surface as a starting point, then, for any other surface point $P \equiv P(u, v)$, our integral

$$K_{OP} = \int_{(O)}^{(P)} \left(\overline{v}\frac{du}{U} - \overline{u}\frac{dv}{V}\right) = \psi(u, v) \tag{1a}$$

assumes a specific value dependent only on the coordinates u, v of P, which is said to be the "*stream function*" $\psi(u, v)$ in point P. In case of a change of the starting point O the function obviously changes only by an additive constant. The partial derivatives of the stream function are then given by:

$$\left.\begin{aligned} \frac{\partial\psi}{\partial u} = \frac{\overline{v}}{U}\,, \quad \frac{\partial\psi}{\partial v} = -\frac{\overline{u}}{V} \quad \text{or} \\ \overline{u} = -V\frac{\partial\psi}{\partial v}\,, \quad \overline{v} = +U\frac{\partial\psi}{\partial u} \end{aligned}\right\} \tag{4}$$

and hence:

$$\left.\begin{aligned} \frac{du}{dt} = U\overline{u} = -UV\frac{\partial\psi}{\partial v}\,, \\ \frac{dv}{dt} = V\overline{v} = +UV\frac{\partial\psi}{\partial u}\,. \end{aligned}\right\} \tag{4a}$$

The entire velocity distribution is therefore completely determined by the unique function $\psi(u, v)$. This significantly simplifies any investigation on incompressible fluids in the plane.

By definition (1a) of $K = \psi(u, v)$, we have $q_n = 0$ everywhere on the curves $\psi = \text{const}$, i. e., the velocity is tangentially directed; the curves are called "streamlines".

Die absolute Eindeutigkeit und Stetigkeit der Stromfunktion ψ ist die notwendige und hinreichende Bedingung dafür, daß die Flüssigkeit auf unserer Fläche inkompressibel und in sich abgeschlossen ist. Die „Inkompressibilitäts-Bedingung“ (3) dagegen bezieht sich nur auf die regulären Punkte und kann an Stellen, wo die Geschwindigkeit unstetig, z. B. unendlich wird, oder eine
208 der Größen U, V verschwindet, ihre | Bedeutung verlieren. Ist dies etwa in einem einzelnen Punkt P_1 der Fall, während sonst überall die Bedingung (3) erfüllt ist, so hätte zwar der Massenfluß K_C, wo C eine beliebige, P_1 einschließende geschlossene Kurve sein kann, einen bestimmten von der Gestalt von C unabhängigen Wert, dieser könnte aber von Null verschieden sein und würde dann die Masseneinströmung in diesen Punkt P_1 angeben; es wäre ein „Quellpunkt“ oder ein „Senkpunkt“ (oder „Abflußpunkt“), je nachdem K_C negativ oder positiv wäre, und ein Verzweigungspunkt der Stromfunktion. Doch wollen wir im Folgenden solche Fälle ausschließen und uns auf den Fall einer absolut eindeutigen Stromfunktion beschränken.

§ 3. Die Zirkulation und das Wirbelmoment. Das Geschwindigkeitspotential.

Ebenso wie im vorigen Paragraphen das Integral $K = \int k q_n ds$ betrachten wir jetzt das Integral

$$2R_{\mathfrak{C}} = \int^{(\mathfrak{C})} q_s ds \,,$$

wo q_s die Geschwindigkeitskomponente in der Richtung ds bezeichnet, und zwar nehmen wir sogleich den Fall, wo $\mathfrak{C}$ eine geschlossene, sich selbst nicht schneidende Kurve ist und in positivem Sinne durchlaufen wird. Dann wird das Integral die „Zirkulation in der Kurve $\mathfrak{C}$“ genannt. Es ist aber

$$q_s = \frac{\overline{u}}{U}\frac{du}{ds} + \frac{\overline{v}}{V}\frac{dv}{ds} \,, \tag{a}$$

also wird:

$$\begin{aligned} 2R_{\mathfrak{C}} &= \int^{(\mathfrak{C})} q_s ds = \int^{(\mathfrak{C})} \left(\frac{\overline{u}}{U}du + \frac{\overline{v}}{V}dv\right) \\ &= \int^{(C)}\int \left\{\frac{\partial}{\partial u}\left(\frac{\overline{v}}{V}\right) - \frac{\partial}{\partial v}\left(\frac{\overline{u}}{U}\right)\right\} du dv = \int^{(C)} 2\varrho d\sigma \,, \end{aligned} \tag{1}$$

wenn C das von $\mathfrak{C}$ eingeschlossene Flächenstück mit dem Flächenelement $d\sigma$ bezeichnet und

$$2\varrho = UV\left\{\frac{\partial}{\partial u}\left(\frac{\overline{v}}{V}\right) - \frac{\partial}{\partial v}\left(\frac{\overline{u}}{U}\right)\right\} \tag{b}$$

gesetzt wird.

The absolute single-valuedness and continuity of the stream function ψ is the necessary and sufficient condition that the fluid on our surface is incompressible and closed in itself. In contrast, the "incompressibility condition" (3) only refers to the regular points and can lose its meaning at places where the velocity becomes discontinuous, e. g., infinite, or where one of the variables U, V vanishes. If this is the case for, say, an individual point P_1, while condition (3) is satisfied everywhere else, then the mass flux K_C would have a particular value independent of the shape of C, where C may be any closed curve enclosing P_1. But this value could be different from zero and would then specify the mass inflow at this point P_1; it would be either a "source point" or a "sink point" (or "outflow point"), depending on whether K_C is negative or positive, and a branching point of the stream function. However, in what follows, we will exclude such cases and restrict ourselves to the case of an absolutely single-valued stream function.

§ 3. Circulation and vorticity moment. Velocity potential.

Just as we considered the integral $K = \int kq_n ds$ in the previous section, so we now consider the integral

$$2R_{\mathfrak{C}} = \overset{(\mathfrak{C})}{\int} q_s ds \,,$$

where q_s denotes the velocity component in the direction ds, and in particular we immediately turn to the case where $\mathfrak{C}$ is a closed curve that does not intersect itself and is traversed in the positive direction. The integral is then called the "circulation in the curve $\mathfrak{C}$". However, we have

$$q_s = \frac{\overline{u}}{U}\frac{du}{ds} + \frac{\overline{v}}{V}\frac{dv}{ds} \,, \tag{a}$$

and hence we get:

$$\begin{aligned} 2R_{\mathfrak{C}} &= \overset{(\mathfrak{C})}{\int} q_s ds = \overset{(\mathfrak{C})}{\int}\left(\frac{\overline{u}}{U}du + \frac{\overline{v}}{V}dv\right) \\ &= \overset{(C)}{\int\int}\left\{\frac{\partial}{\partial u}\left(\frac{\overline{v}}{V}\right) - \frac{\partial}{\partial v}\left(\frac{\overline{u}}{U}\right)\right\} du dv = \overset{(C)}{\int} 2\varrho d\sigma \,, \end{aligned} \tag{1}$$

if C denotes the part of the surface enclosed by $\mathfrak{C}$ with the area element $d\sigma$ and if we set

$$2\varrho = UV\left\{\frac{\partial}{\partial u}\left(\frac{\overline{v}}{V}\right) - \frac{\partial}{\partial v}\left(\frac{\overline{u}}{U}\right)\right\} \,. \tag{b}$$

$R_{\mathfrak{C}} = \overset{(C)}{\int} \varrho d\sigma$ wird auch das „Wirbelmoment von C“ und ϱ die „Wirbeldichte“ im Punkt u, v genannt, die letztere ist nichts anderes als die Rotationskomponente des Flüssigkeitsteilchens um die Flächennormale als Achse. Der Ausdruck ist bereits im ersten Paragraphen ((c) p. 204) zur Vereinfachung der Gleichungen (3) eingeführt worden. Aus der Definition (a) und der Relation (1) ergiebt sich ohne weiteres

209 | Satz I. *Das Wirbelmoment eines Flächenstückes ist gleich der Summe der Wirbelmomente seiner Teile.*

(Vorausgesetzt ist dabei, daß in den Teilkurven selbst die Geschwindigkeit keinen Sprung erleidet. Solche Diskontinuitätskurven müßte man durch Festsetzung entweder ganz zu der einen oder ganz zu der anderen Seite rechnen.)

Satz II. *Das gesamte Wirbelmoment einer geschlossenen Fläche ist null.*

Denn hier kann man die Kurve $\mathfrak{C}$ auf einen einzigen (nicht singulären) Punkt zusammenziehen.

Es kann vorkommen, daß innerhalb eines einfach zusammenhängenden Flächenstückes F die Zirkulation durch *jede* geschlossene Kurve, also das Wirbelmoment *jedes* Flächenstückes in F den Wert Null hat und demgemäß auch die Wirbeldichte ϱ allenthalben verschwindet.

In diesem Falle hat das Integral

$$2R_{OP} = \overset{(P)}{\underset{(O)}{\int}} q_s ds = \overset{(P)}{\underset{(O)}{\int}} \left(\frac{\overline{u} du}{U} + \frac{\overline{v} dv}{V} \right) = \varphi(u, v) \, , \tag{c}$$

welches längs einer beliebigen, doch ganz innerhalb F verlaufenden Kurve $\mathfrak{C}$ von einem festen Anfangspunkt O nach einem beliebigen anderen Flächenpunkte $P = P(u, v)$ erstreckt wird, einen ganz bestimmten von der Gestalt des Weges $\mathfrak{C}$ unabhängigen Wert $\varphi(u, v)$, welcher dem Punkte $P(u, v)$ charakteristisch ist und als das „Geschwindigkeitspotential“ im Punkte P bezeichnet wird. Dann können die Geschwindigkeitskomponenten $\overline{u}$, $\overline{v}$, ebenso wie im Falle der Inkompressibilität in § 2 (4) durch die Stromfunktion ψ, nunmehr durch die partiellen Ableitungen von φ ausgedrückt werden:

$$\overline{u} = U \frac{\partial \varphi}{\partial u} \, , \quad \overline{v} = V \frac{\partial \varphi}{\partial v} \, . \tag{2}$$

Ein solches Geschwindigkeitspotential existiert nach (1) in jedem einfach zusammenhängenden Flächenstücke F als eindeutige Funktion des Ortes, wenn innerhalb F die Geschwindigkeit nirgends Sprünge erleidet und die Wirbeldichte ϱ allenthalben verschwindet:

$$\frac{2\varrho}{UV} = \frac{\partial}{\partial u} \left(\frac{\overline{v}}{V} \right) - \frac{\partial}{\partial v} \left(\frac{\overline{u}}{U} \right) = 0 \, . \tag{3}$$

$R_{\mathfrak{C}} = \int^{(C)} \varrho d\sigma$ is also called the "vorticity moment of C" and ϱ the "vortex density" at point u, v. The latter is but the rotation component of the fluid particle with the surface normal as its axis. We have introduced the expression already in the first section ((c) p. 204) in order to simplify equations (3). From definition (a) and relation (1) we now obtain without further ado

Theorem I. *The vorticity moment of a closed part of the surface is equal to the sum of the vorticity moments of its parts.*

(We assume here that the velocity does not suffer discontinuity jumps in the partial curves themselves. It would be necessary to stipulate that discontinuity curves of this sort belong entirely to either one or the other side.)

Theorem II. *The total vorticity moment of a closed surface is equal to zero.*

For here we can shrink the curve $\mathfrak{C}$ to a single (non-singular) point.

It is possible that, within a simply connected part F of the surface, the circulation through *every* closed curve, and hence the vorticity moment of *every* part of the surface in F, has the value zero, and accordingly that the vortex density ϱ vanishes everywhere.

In this case, the integral

$$2R_{OP} = \int_{(O)}^{(P)} q_s ds = \int_{(O)}^{(P)} \left(\frac{\overline{u} du}{U} + \frac{\overline{v} dv}{V} \right) = \varphi(u, v) \ , \tag{c}$$

which is extended along some curve $\mathfrak{C}$ entirely lying within F from a fixed starting point O to any other surface point $P = P(u, v)$, has a specific value $\varphi(u, v)$ independent of the shape of the path $\mathfrak{C}$ and characteristic of the $P(u, v)$, which is called the "velocity potential" at point P. We can now express the velocity components $\overline{u}$, $\overline{v}$ in terms of the partial derivatives of φ:

$$\overline{u} = U \frac{\partial \varphi}{\partial u} \ , \quad \overline{v} = V \frac{\partial \varphi}{\partial v} \ , \tag{2}$$

just as we expressed them in terms of the stream function ψ in the case of the incompressibility in § 2 (4). By (1), there exists a velocity potential of this sort in every simply connected part F of the surface as a single-valued function of location, if the velocity does not suffer any jump discontinuities within F and if the vortex density ϱ vanishes everywhere:

$$\frac{2\varrho}{UV} = \frac{\partial}{\partial u} \left(\frac{\overline{v}}{V} \right) - \frac{\partial}{\partial v} \left(\frac{\overline{u}}{U} \right) = 0 \ . \tag{3}$$

Ist aber diese Bedingung (3) nur in einem mehrfach zusammenhängenden Flächenstücke F erfüllt oder enthält F singuläre Punkte oder Linien, in denen
210 ϱ seine Bedeutung verliert und deren Ausschließung F | jedenfalls mehrfach zusammenhängend machen würde, so existiert zwar auch ein Geschwindigkeitspotential $\varphi(u, v)$; dasselbe braucht dann aber nicht mehr eindeutig zu sein, sondern ändert sich nach gewissen Umläufen (um die ausgeschlossenen Teile) um additive Perioden.

Die Kurven $\varphi = \text{const.}$, auf denen nach (c) überall $q_s = 0$ ist, stehen in jedem ihrer Punkte auf der dort herrschenden Geschwindigkeitsrichtung senkrecht, schneiden also alle Stromlinien rechtwinklig und werden die *Niveaulinien* genannt.

Ist nun die Flüssigkeit zugleich *inkompressibel* (§ 2), was in den späteren Untersuchungen immer vorausgesetzt werden soll, so kann man nach (4) p. 207 die Stromfunktion ψ einführen und erhält so für die Wirbeldichte aus (b):

$$2\varrho = UV\left\{\frac{\partial}{\partial u}\left(\frac{U}{V}\frac{\partial\psi}{\partial u}\right) + \frac{\partial}{\partial v}\left(\frac{V}{U}\frac{\partial\psi}{\partial v}\right)\right\} \equiv D\psi\,. \tag{4}$$

Diese Formel gestattet die Berechnung der Wirbeldichte aus der Stromfunktion durch zweimalige Differentiation und umgekehrt bei gegebenem ϱ die Berechnung von ψ durch Integration einer partiellen Differentialgleichung zweiter Ordnung.

In einem „wirbelfreien“ Flächenstücke F, in welchem überall $\varrho = 0$ und daher ein Geschwindigkeitspotential φ vorhanden ist, wird demnach überall

$$D\psi = 0, \quad \text{aber zugleich auch:} \quad D\varphi = 0\,,$$

welch letztere Gleichung man erhält, wenn man in der Gleichung (3) § 2 die Ausdrücke (2) für $\overline{u}$ und $\overline{v}$ einführt.

Es ist nun für ein beliebiges Flächenstück C ohne singuläre Stellen im Innern:

$$\int^{(C)} 2\varrho\psi d\sigma = \int^{(C)} \psi D\psi d\sigma = \int\!\!\int^{(C)} \psi\left\{\frac{\partial}{\partial u}\left(\frac{U}{V}\frac{\partial\psi}{\partial u}\right) + \frac{\partial}{\partial v}\left(\frac{V}{U}\frac{\partial\psi}{\partial v}\right)\right\} dudv \tag{5}$$

$$= -\int\!\!\int^{(C)} \left\{\frac{U}{V}\left(\frac{\partial\psi}{\partial u}\right)^2 + \frac{V}{U}\left(\frac{\partial\psi}{\partial v}\right)^2\right\} dudv + \int^{(\mathfrak{C})} \psi\left(\frac{U}{V}\frac{\partial\psi}{\partial u}dv - \frac{V}{U}\frac{\partial\psi}{\partial v}du\right)$$

$$= -\int\!\!\int^{(C)} \left(\overline{u}^2 + \overline{v}^2\right)\frac{dudv}{UV} + \int^{(\mathfrak{C})} \psi\left(\frac{\overline{u}}{U}du + \frac{\overline{v}}{V}dv\right)\,,$$

wenn das Linienintegral rechts über die Begrenzung $\mathfrak{C}$ von C im positiven Sinne erstreckt wird. Hier kann man aber auch die totale Geschwindigkeit q

If, however, this condition (3) is satisfied only in a multiply connected part F of the surface, or if F contains singular points or lines in which ϱ loses its meaning and whose exclusion would render F multiply connected, then, while there also exists a velocity potential $\varphi(u,v)$, it no longer needs to be single-valued but it changes by additive periods after certain revolutions (about the excluded parts).

The curves $\varphi = \text{const.}$, on which, by (c), we have $q_s = 0$ everywhere, are perpendicular in each of their points to the respective velocity direction obtaining there, and hence intersect all streamlines at a right angle, and are called the *contour lines*.

Now if the fluid is also *incompressible* (§ 2), which we shall always assume to be the case in the later investigations, then, by (4) p. 207, we can introduce the stream function ψ, thereby obtaining for the vortex density, from (b):

$$2\varrho = UV\left\{\frac{\partial}{\partial u}\left(\frac{U}{V}\frac{\partial\psi}{\partial u}\right)+\frac{\partial}{\partial v}\left(\frac{V}{U}\frac{\partial\psi}{\partial v}\right)\right\}\equiv D\psi\,. \tag{4}$$

This formula allows for the calculation of the vortex density on the basis of the stream function by twofold differentiation, and, conversely, for given ϱ, the calculation of ψ by integration of a partial differential equation of second order.

In a "vortex-free" part F of the surface in which we have $\varrho = 0$ everywhere, and hence in which there also exists a velocity potential φ, we therefore get everywhere

$$D\psi = 0, \quad \text{but at the same time also:} \quad D\varphi = 0\,,$$

where we obtain the latter equation by introducing into equation (3) § 2 the expressions (2) for $\overline{u}$ and $\overline{v}$.

Given a part C of the surface that has no singular positions in its interior, we now have:

$$\int^{(C)} 2\varrho\psi d\sigma = \int^{(C)} \psi D\psi d\sigma = \int\!\!\int^{(C)} \psi\left\{\frac{\partial}{\partial u}\left(\frac{U}{V}\frac{\partial\psi}{\partial u}\right)+\frac{\partial}{\partial v}\left(\frac{V}{U}\frac{\partial\psi}{\partial v}\right)\right\}dudv \tag{5}$$

$$= -\int\!\!\int^{(C)}\left\{\frac{U}{V}\left(\frac{\partial\psi}{\partial u}\right)^2+\frac{V}{U}\left(\frac{\partial\psi}{\partial v}\right)^2\right\}dudv+\int^{(\mathfrak{C})}\psi\left(\frac{U}{V}\frac{\partial\psi}{\partial u}dv-\frac{V}{U}\frac{\partial\psi}{\partial v}du\right)$$

$$= -\int\!\!\int^{(C)}\left(\overline{u}^2+\overline{v}^2\right)\frac{dudv}{UV}+\int^{(\mathfrak{C})}\psi\left(\frac{\overline{u}}{U}du+\frac{\overline{v}}{V}dv\right)\,,$$

if we extend the line integral on the right side over the boundary $\mathfrak{C}$ of C in the positive direction. Here, however, we can also introduce the total velocity q

und ihre Komponente q_s einführen und erhält:

$$\int^{(C)} 2\varrho\psi d\sigma = -\int^{(C)} q^2 d\sigma + \int^{(\mathfrak{C})} \psi q_s ds \ . \qquad (5')$$

211 | Das Linienintegral über $\mathfrak{C}$ verschwindet aber, wenn C über eine vollständige geschlossene Fläche (ohne Singularitäten) ausgedehnt wird. Ist es dagegen nur ein Flächen*stück*, ein begrenzter Flüssigkeitsbereich, der von einem festen Rande, bestehend aus einer oder mehreren Stromlinien $\psi = \psi_\lambda = \text{const.}$, begrenzt wird, so zerfällt das Randintegral in eine Anzahl von Ausdrücken

$$\int \psi_\lambda q_s ds = \psi_\lambda \int^{(\psi_\lambda)} q_s ds = 2\psi_\lambda R_\lambda$$

und verschwindet wieder für jeden Bereich C mit eindeutigem Geschwindigkeitspotential, in welchem ja *jede* Zirkulation den Wert Null hat, also namentlich für jedes einfach zusammenhängende wirbelfreie Flächenstück ($\varrho = 0$). Dann verschwindet aber gleichzeitig auch die linke Seite von (5), und es bleibt:

$$\int^{(C)} q^2 d\sigma = \int^{(C)} \left(\overline{u}^2 + \overline{v}^2\right) d\sigma = 0 \ .$$

Es muß also überall $q^2 = 0$ und $\psi = \text{const.}$ sein im ganzen Innern von C, d. h. die Flüssigkeit muß hier überall in Ruhe sein. Wir gewinnen also den Satz:

Satz III. *In einer vollständig geschlossenen Fläche sowie in einem einfach zusammenhängenden Flächenstücke von fester Berandung $q_n = 0$ giebt es keine wirbelfreie Bewegung einer inkompressibeln Flüssigkeit, falls im ganzen Innern Unstetigkeiten der Geschwindigkeit ausgeschlossen sind. In einem solchen Bereiche ist daher die gesamte momentane Geschwindigkeitsverteilung, der „Strömungszustand" durch die vorhandenen Wirbel, d. h. durch die Funktion $\varrho = \varrho(u, v)$ eindeutig bestimmt.*

Wären nämlich $\psi = \psi_1$ und $\psi = \psi_2$ zwei Lösungen der Differentialgleichung (4) $D\psi = 2\varrho$, welche auf der vorgeschriebenen Randkurve, die in dem betrachteten Falle höchstens aus einer einzigen Stromlinie bestehen kann, konstante Werte $\overline{\psi}_1$ und $\overline{\psi}_2$ annehmen, so müßte für ihre Differenz $\psi_0 = \psi_1 - \psi_2$ im Innern überall $D\psi_0 = D\psi_1 - D\psi_2 = 0$ sein und auf dem Rande $\overline{\psi}_0 = \overline{\psi}_1 - \overline{\psi}_2 = \text{const.}$, was nach dem oben Bewiesenen nur möglich wäre bei $\psi_0 = \text{const.}$, also $\psi_1 = \psi_2 + \text{const.}$

Bei der Herleitung der Beziehung (1)

$$R = \tfrac{1}{2} \int^{(\mathfrak{C})} q_s ds = \int^{(C)} \varrho d\sigma$$

and its component q_s, obtaining:

$$\int^{(C)} 2\varrho\psi d\sigma = -\int^{(C)} q^2 d\sigma + \int^{(\mathfrak{C})} \psi q_s ds \,. \tag{5'}$$

The line integral over $\mathfrak{C}$ vanishes, however, if we extend C over a complete closed surface (without singularities). If, in contrast, it is only a *part* of a surface, a bounded fluid domain whose fixed boundary consists of one or several streamlines $\psi = \psi_\lambda = \text{const.}$, then the boundary integral decomposes into a number of expressions

$$\int \psi_\lambda q_s ds = \psi_\lambda \int^{(\psi_\lambda)} q_s ds = 2\psi_\lambda R_\lambda$$

and again vanishes for every domain C with single-valued velocity potential, in which, after all, *every* circulation has the value zero, and hence in particular for every simply connected vortex-free region ($\varrho = 0$). But then the left side of (5) too vanishes, and what remains is:

$$\int^{(C)} q^2 d\sigma = \int^{(C)} \left(\overline{u}^2 + \overline{v}^2\right) d\sigma = 0 \,.$$

Hence, we must have $q^2 = 0$ and $\psi = \text{const.}$ everywhere in the entire interior of C, i. e., the fluid must be at rest everywhere here. We thus obtain the theorem:

Theorem III. *In a completely closed surface as well as in a simply connected part of the surface with fixed boundary $q_n = 0$ there exists no vortex-free motion of an incompressible fluid, provided that discontinuities of velocity are excluded in the entire interior. In such a domain the entire current velocity distribution, the "state of the stream", is therefore uniquely determined by the existing vortices, i. e., by the function $\varrho = \varrho(u, v)$.*

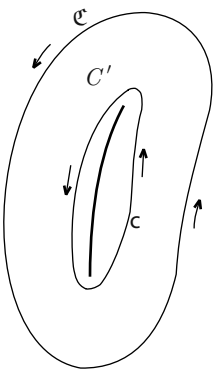
For if $\psi = \psi_1$ and $\psi = \psi_2$ were two solutions of the differential equation (4) $D\psi = 2\varrho$ that assume the constant values $\overline{\psi}_1$ and $\overline{\psi}_2$ on the prescribed boundary curve, which can only consist of at most one single streamline in the case under consideration, then for their difference $\psi_0 = \psi_1 - \psi_2$ we would have to have $D\psi_0 = D\psi_1 - D\psi_2 = 0$ in the interior and $\overline{\psi}_0 = \overline{\psi}_1 - \overline{\psi}_2 = \text{const.}$ on the boundary, which, according to what has been proved above, would only be possible when $\psi_0 = \text{const.}$, and hence $\psi_1 = \psi_2 + \text{const.}$

For the derivation of (1)

$$R = \tfrac{1}{2} \int^{(\mathfrak{C})} q_s ds = \int^{(C)} \varrho d\sigma$$

hatten wir vorausgesetzt, daß das Flächenstück C, über das wir integrierten, von Singularitäten frei sei, die sich auf die Geschwindigkeitskomponenten und ihre ersten Ableitungen beziehen, daß vielmehr überall im Innern ein
212 bestimmter Wert $2\varrho = D\psi$ existiere. Wir wollen | nun die Natur der möglichen Singularitäten untersuchen, indem wir den Fall ausschließen, daß ϱ in einem ganzen endlichen Flächenstück seine Bedeutung verliere. Es sei nun $\mathfrak{L}$ ein singuläres Kurvenstück, das sich event. auch auf einen einzelnen Punkt reduzieren kann, und C mit der Randkurve $\mathfrak{C}$ ein Flächenstück, welches $\mathfrak{L}$ und sonst keine weitere Singularität in seinem Innern enthält. Umgeben wir nun $\mathfrak{L}$ durch eine innerhalb C beliebig verlaufende geschlossene Kurve $\mathfrak{c}$ mit

Fig. 1.



positivem Richtungssinn in Bezug auf $\mathfrak{L}$ als Inneres, so können wir auf das zwischen $\mathfrak{C}$ und $\mathfrak{c}$ liegende Gebiet C' unseren Satz (I) anwenden und erhalten:

$$2R' \equiv \overset{(C')}{\int} 2\varrho d\sigma = \overset{(\mathfrak{C})}{\int} q_s ds - \overset{(\mathfrak{c})}{\int} q_s ds \ .$$

Dies gilt immer, wie eng auch die Kurve $\mathfrak{c}$ die Singularität $\mathfrak{L}$ umgebe, so daß, wenn wir in der Verengung zur Grenze übergehen, auch

$$\lim 2R' \equiv \lim \overset{(C')}{\int} 2\varrho d\sigma = \overset{(\mathfrak{C})}{\int} q_s ds - \lim \overset{(\mathfrak{c})}{\int} q_s ds$$

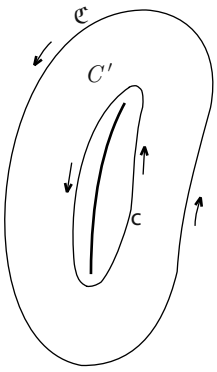
oder

$$2R_{\mathfrak{C}} \equiv \overset{(\mathfrak{C})}{\int} q_s ds = \lim \overset{(C')}{\int} 2\varrho d\sigma + \lim \overset{(\mathfrak{c})}{\int} q_s ds = 2\overline{R} + 2R_{\mathfrak{L}} \ .$$

Unser Satz (I) bleibt also richtig, wenn wir nur auf der rechten Seite zu dem Grenzwerte des Flächenintegrales $\int 2\varrho d\sigma$ bei ausgeschlossener Singularität $\mathfrak{L}$

we had assumed that the part C of the surface over which we integrated is free of singularities related to the velocity components and their first derivations and that, rather, there exists a particular value $2\varrho = D\psi$ everywhere in the interior. We now want to investigate the nature of the possible singularities by excluding the case where ϱ loses its meaning in an entire finite part of the surface. Let $\mathfrak{L}$ be a singular curve segment that may be a single point, and let C with the boundary curve $\mathfrak{C}$ be a part of the surface containing no further singularity in its interior besides $\mathfrak{L}$. If we now surround $\mathfrak{L}$ with any

Fig. 1.



closed curve $\mathfrak{c}$ lying in the interior of C with a positive sense of direction with respect to $\mathfrak{L}$ as the interior, then we can apply our theorem (I) to the region C' lying between $\mathfrak{C}$ and $\mathfrak{c}$ and we obtain:

$$2R' \equiv \int^{(C')} 2\varrho d\sigma = \int^{(\mathfrak{C})} q_s ds - \int^{(\mathfrak{c})} q_s ds \; .$$

This is always true, no matter how closely the curve $\mathfrak{c}$ surrounds the singularity $\mathfrak{L}$, so that, if we take the limit in the narrow part, we also have

$$\lim 2R' \equiv \lim \int^{(C')} 2\varrho d\sigma = \int^{(\mathfrak{C})} q_s ds - \lim \int^{(\mathfrak{c})} q_s ds$$

or

$$2R_{\mathfrak{C}} \equiv \int^{(\mathfrak{C})} q_s ds = \lim \int^{(C')} 2\varrho d\sigma + \lim \int^{(\mathfrak{c})} q_s ds = 2\overline{R} + 2R_{\mathfrak{L}} \; .$$

Thus, our theorem (I) continues to hold when the limit $2R_{\mathfrak{L}}$ of the circulation about $\mathfrak{L}$ itself is only added on the right side to the limit of the surface

noch den Grenzwert $2R_{\mathfrak{L}}$ der Zirkulation um $\mathfrak{L}$ selbst hinzufügen, d. h. wenn wir der Kurve $\mathfrak{L}$ selbst das endliche Wirbelmoment $R_{\mathfrak{L}}$ beilegen. Verfahren wir so bei allen etwa auftretenden Singularitäten, so behalten alle früheren Sätze, z. B. I und II, S. 209, III, S. 211 ihre Giltigkeit. Dabei können wir noch alle die Singularitäten vollständig ignorieren, für welche $2R_{\mathfrak{L}} = \lim \overset{(\mathfrak{c})}{\int} q_s ds$ den Wert 0 hat. Dies ist aber immer der Fall, wenn die Geschwindigkeitskomponenten auch bei $\mathfrak{L}$ stetig bleiben, die Singularität sich also nur auf ihre Ableitungen $\frac{\partial \overline{u}}{\partial u}$, $\frac{\partial \overline{v}}{\partial v}$ etc. bezieht. Denn dann werden bei der Zusammenziehung der Kurve $\mathfrak{c}$ auf $\mathfrak{L}$ selbst die einander gegenüberliegenden Elemente ds, ds' gleiche und entgegengesetze Beiträge $q_s ds$ und $q'_s ds' = -q_s ds' = -q_s ds$ liefern und einander aufheben. Anders dagegen, wenn längs der Kurve $\mathfrak{L}$ die Geschwindigkeit unstetig ist, so daß die Tangentialkomponenten derselben rechts und links von $\mathfrak{L}$ nach zwei verschiedenen Werten $q_{\mathfrak{L}}$ und $q'_{\mathfrak{L}}$ konvergieren, während ihre | Normalkomponenten jedenfalls stetig bleiben müssen, 213
solange unserer Voraussetzung (§ 2 fin.) zufolge auch in $\mathfrak{L}$ keine Quell- oder Senkpunkte vorhanden sind. In diesem Falle wird einfach

$$2R_{\mathfrak{L}} = \overset{(\mathfrak{L})}{\int} (q'_{\mathfrak{L}} - q_{\mathfrak{L}})\, ds \ ,$$

d. h. gleich dem Integrale des Geschwindigkeitssprunges über die Diskontinuitätskurve, falls der Sprung nach der Richtung derjenigen Normalen von $\mathfrak{L}$ gerechnet wird, welche zur Fortschreitungsrichtung ds ebenso liegt, wie die v-Richtung zur u-Richtung. Den erhaltenen Wert bezeichnen wir als das „Wirbelmoment der Diskontinuitätskurve $\mathfrak{L}$“.

Ist aber die Singularität $\mathfrak{L}$ nur ein einzelner Punkt, so verschwindet $R_{\mathfrak{L}}$ immer, vermöge der unbegrenzten Verkürzung der Umlaufskurve $\mathfrak{c}$, so lange die Geschwindigkeit in der Umgebung von $\mathfrak{L}$ eine endliche obere Grenze besitzt. Soll also $R_{\mathfrak{L}}$ einen endlichen Wert annehmen, der von 0 verschieden ist, so muß die Geschwindigkeit bei $\mathfrak{L}$ unendlich groß werden, so aber, daß die Normalkomponente q_n in einem $\mathfrak{L}$ umgebenden kleinen Kreise $\mathfrak{c}$ zugleich mit seinem Radius verschwindet, weil $\mathfrak{L}$ eben kein Quellpunkt sein soll. Mit anderen Worten: die Flüssigkeit muß mit unendlicher Geschwindigkeit um den singulären Punkt herumströmen, und die benachbarten Stromlinien werden kleine ihn umgebende Ovale sein, die, je näher an $\mathfrak{L}$, desto größere Geschwindigkeiten besitzen. Solch einen Punkt bezeichnen wir als einen „*Strudelpunkt*“ und den endlichen Grenzwert $R_{\mathfrak{L}}$ der halben Zirkulation über einen umgebenden kleinen Kreis $\mathfrak{c}$ als das „Strudelmoment“ dieses Strudelpunktes.

Von den etwa auftretenden Singularitäten sind also besonders zu berücksichtigen nur die „Strudelpunkte“ und die „Diskontinuitätslinien“, welch letztere auch „Strudellinien“ genannt werden mögen, da sie sich durch aneinander gereihte Strudelpunkte ersetzen lassen.

integral $\int 2\varrho d\sigma$ with the singularity $\mathfrak{L}$ being excluded, i. e., if we assign the finite vorticity moment $R_{\mathfrak{L}}$ to the curve $\mathfrak{L}$ itself. If we treat all occurring singularities in this fashion, then all previous theorems, e. g. I and II, p. 209, III, p. 211, retain their validity. Moreover, we can ignore all those singularities for which $2R_{\mathfrak{L}} = \lim \overset{(\mathfrak{c})}{\int} q_s ds$ has the value 0. This, however, is always the case when the velocity components remain continuous also at $\mathfrak{L}$, and hence when the singularity is only related to its derivatives $\frac{\partial \overline{u}}{\partial u}$, $\frac{\partial \overline{v}}{\partial v}$ etc. For then the elements ds, ds' facing one another will yield identical and opposite amounts $q_s ds$ and $q'_s ds' = -q_s ds' = -q_s ds$ and cancel each other out when the curve $\mathfrak{c}$ is shrunk to $\mathfrak{L}$ itself. Matters are different, however, when the velocity is not continuous along the curve $\mathfrak{L}$ so that its tangential components on the right and left sides of $\mathfrak{L}$ converge to two different values $q_{\mathfrak{L}}$ and $q'_{\mathfrak{L}}$ while its normal component must remain continuous so long as, according to our assumption (§ 2 fin.), there are no source or sink points in $\mathfrak{L}$ either. In this case, we simply have

$$2R_{\mathfrak{L}} = \overset{(\mathfrak{L})}{\int} (q'_{\mathfrak{L}} - q_{\mathfrak{L}})\, ds\ ,$$

i. e., it is equal to the integral of the velocity jump over the discontinuity curve in case the jump is calculated in the direction of that normal of $\mathfrak{L}$ which lies with respect to the direction of progression ds just as the v-direction lies with respect to the u-direction. The value so obtained is called the "vorticity moment of the discontinuity curve $\mathfrak{L}$".

If, however, the singularity $\mathfrak{L}$ is but a single point, then $R_{\mathfrak{L}}$ always vanishes because of the unbounded shortening of the orbital curve $\mathfrak{c}$ so long as the velocity has a finite upper bound in the neighborhood of $\mathfrak{L}$. Hence, if $R_{\mathfrak{L}}$ is to assume a finite value different from 0, then the velocity must become infinitely large at $\mathfrak{L}$, but so that the normal component q_n in a small circle $\mathfrak{c}$ surrounding $\mathfrak{L}$ vanishes along with its radius precisely because $\mathfrak{L}$ is not supposed to be a source point. In other words: the fluid must flow around the singular point at infinite velocity, and the neighboring streamlines will be small ovals surrounding it whose velocities are the greater, the closer they come to $\mathfrak{L}$. Such a point is called a "*whirl point*", and the finite limit $R_{\mathfrak{L}}$ of half of the circulation about a surrounding small circle $\mathfrak{c}$ is called the "whirl moment" of this whirl point.

Hence, among the possibly occurring singularities, only the "whirl points" and the "discontinuity lines" require special consideration, the latter of which may also be called "whirl lines" as they can be replaced by whirl points that are strung together.

§ 4. Das Helmholtzsche Theorem und die Bewegung der Wirbel.

Bisher hatten wir uns nur mit der momentanen Verteilung der Geschwindigkeit auf unserer Fläche, mit der Lage der Stromlinien u. s. w. in einem einzigen Augenblicke beschäftigt. Nun entsteht aber die Frage: Wie ändert sich dieser Geschwindigkeitszustand, wie deformieren sich die Stromlinien? oder mit anderen Worten: Welches ist der vollständige zeitliche Verlauf der Flüssigkeitsbewegung? Diese Frage beantwortet unter den hier gemachten Voraussetzungen das für unsere zweidimensionale Flüssigkeit auf der Fläche
214 ebenso wie für die | räumliche geltende „*Helmholtzsche* Theorem“, das wir hier so formulieren wollen:

Satz I. *Wenn eine zweidimensionale Flüssigkeit auf einer beliebigen festen Fläche reibungslos unter dem Einflusse von Massenkräften steht, welche ein Potential besitzen, und der Druck eine Funktion der Flächendichte allein ist, so ist die Zirkulation in einer aus bestimmten materiellen Punkten bestehenden geschlossenen Kurve* $\mathfrak{C}$ *oder das Wirbelmoment eines materiellen Flüssigkeitsbereiches* C *bei allen Bewegungen in der Zeit konstant.*

Schreiben wir nämlich nach (1) § 3:

$$2R_{\mathfrak{C}} = \int^{(\mathfrak{C})} \left(\frac{\overline{u}}{U}\delta u + \frac{\overline{v}}{V}\delta v \right) ,$$

wo das Zeichen δ den Übergang von einem materiellen Flüssigkeitspunkte zum anderen, also eine von der Zeit t unabhängige Veränderung ausdrücken soll und daher gegen das Symbol $\frac{d}{dt}$ vertauschbar ist, so wird:

$$\begin{aligned} \frac{d\delta u}{dt} = \delta\frac{du}{dt} = \delta\left(\overline{u}U\right) &= \frac{\partial\left(\overline{u}U\right)}{\partial u}\delta u + \frac{\partial\left(\overline{u}U\right)}{\partial v}\delta v \\ \frac{d\delta v}{dt} = \delta\frac{dv}{dt} = \delta\left(\overline{v}V\right) &= \frac{\partial\left(\overline{v}V\right)}{\partial u}\delta u + \frac{\partial\left(\overline{v}V\right)}{\partial v}\delta v , \end{aligned}$$

und somit

$$\begin{aligned} 2\frac{dR}{dt} &= \int^{(\mathfrak{C})} \left\{ \frac{d}{dt}\left(\frac{\overline{u}}{U}\right)\delta u + \frac{d}{dt}\left(\frac{\overline{v}}{V}\right)\delta v + \frac{\overline{u}}{U}\delta\left(\overline{u}U\right) + \frac{\overline{v}}{V}\delta\left(\overline{v}V\right) \right\} \\ &= \int^{(\mathfrak{C})} \left\{ \frac{d}{dt}\left(\frac{\overline{u}}{U}\right) + \frac{\overline{u}^2\,\partial U}{U\,\partial u} + \frac{\overline{v}^2\,\partial V}{V\,\partial u} \right\}\delta u \\ &\quad + \int^{(\mathfrak{C})} \left\{ \frac{d}{dt}\left(\frac{\overline{v}}{V}\right) + \frac{\overline{u}^2\,\partial U}{U\,\partial v} + \frac{\overline{v}^2\,\partial V}{V\,\partial v} \right\}\delta v + \int \left(\overline{u}\delta\overline{u} + \overline{v}\delta\overline{v}\right) \end{aligned}$$

§ 4. Helmholtz's theorem and the motion of the vortices.

As yet we have only considered the momentary distribution of the velocity on our surface, the position of the streamlines etc., at a single instant in time. Now, however, the question arises: How does this velocity state change? In what way are the streamlines deformed? Or, in other words: What is the complete temporal evolution of the motion of the fluid? It it this question that, under the assumptions made here, is answered by "*Helmholtz*'s theorem", which holds for our two-dimensional fluid on the surface just as well as for the spatial [[fluid]] and which we render as follows:

Theorem I. *If some two-dimensional fluid on any rigid surface is subject without friction to mass forces that possess a potential, and if the pressure is a function of surface density alone, then the circulation in a closed curve* $\mathfrak{C}$ *consisting of particular material points or the vorticity moment of a material fluid domain* C *is constant over time for all motions.*

For if, according to (1) § 3, we write:

$$2R_{\mathfrak{C}} = \overset{(\mathfrak{C})}{\int} \left(\frac{\overline{u}}{U}\delta u + \frac{\overline{v}}{V}\delta v \right) ,$$

where the sign δ is intended to express the transition from one material fluid point to the other, and thus a change independent of time t, and hence can be interchanged with the symbol $\frac{d}{dt}$, then we have:

$$\begin{aligned} \frac{d\delta u}{dt} &= \delta\frac{du}{dt} = \delta\left(\overline{u}U\right) = \frac{\partial\left(\overline{u}U\right)}{\partial u}\delta u + \frac{\partial\left(\overline{u}U\right)}{\partial v}\delta v \\ \frac{d\delta v}{dt} &= \delta\frac{dv}{dt} = \delta\left(\overline{v}V\right) = \frac{\partial\left(\overline{v}V\right)}{\partial u}\delta u + \frac{\partial\left(\overline{v}V\right)}{\partial v}\delta v , \end{aligned}$$

and hence

$$\begin{aligned} 2\frac{dR}{dt} &= \overset{(\mathfrak{C})}{\int} \left\{ \frac{d}{dt}\left(\frac{\overline{u}}{U}\right)\delta u + \frac{d}{dt}\left(\frac{\overline{v}}{V}\right)\delta v + \frac{\overline{u}}{U}\delta\left(\overline{u}U\right) + \frac{\overline{v}}{V}\delta\left(\overline{v}V\right) \right\} \\ &= \overset{(\mathfrak{C})}{\int} \left\{ \frac{d}{dt}\left(\frac{\overline{u}}{U}\right) + \frac{\overline{u}^2}{U}\frac{\partial U}{\partial u} + \frac{\overline{v}^2}{V}\frac{\partial V}{\partial u} \right\}\delta u \\ &+ \overset{(\mathfrak{C})}{\int} \left\{ \frac{d}{dt}\left(\frac{\overline{v}}{V}\right) + \frac{\overline{u}^2}{U}\frac{\partial U}{\partial v} + \frac{\overline{v}^2}{V}\frac{\partial V}{\partial v} \right\}\delta v + \int \left(\overline{u}\delta\overline{u} + \overline{v}\delta\overline{v}\right) \end{aligned}$$

oder auf Grund der Bewegungsgleichungen (1) § 1:

$$2\frac{dR}{dt} = \int^{(\mathfrak{C})} \left\{ -\frac{\partial \Phi}{\partial u}\delta u - \frac{\partial \Phi}{\partial v}\delta v + \overline{u}\delta\overline{u} + \overline{v}\delta\overline{v} \right\}$$
$$= \int^{(\mathfrak{C})} \left(-\delta\Phi + \tfrac{1}{2}\delta\left(q^2\right)\right) = \int \delta\left(\tfrac{1}{2}q^2 - \Phi\right) = \left[\tfrac{1}{2}q^2 - \Phi\right]_A^A = 0 \,,$$

wenn das Integral über eine geschlossene Kurve $\mathfrak{C}$ ausgedehnt wird. Es folgt also in der That, wie behauptet war:

$$R = \tfrac{1}{2}\int^{(\mathfrak{C})} q_s ds = \int^{(C)} \varrho d\sigma = \text{const.} \tag{1}$$

215 | Aus unserem Satze ergeben sich unmittelbar die Folgerungen:

Satz II. *Ist ein Teil der Flüssigkeit zu irgend einer Zeit wirbelfrei* $\varrho = 0$, *so ist er es auch zu allen Zeiten.*

Satz III. *Ist die Flüssigkeit* inkompressibel, *so ist die Wirbeldichte* ϱ *in jedem materiellen Punkte konstant, d. h. an jeder Stelle der Fläche, welche ein bestimmter materieller Punkt* m *einmal erreicht, hat die Wirbeldichte* ϱ *in diesem Augenblicke immer denselben Wert* ϱ_m, *welcher dem Punkte* m *charakteristisch ist.*

Denn nach Satz I ist $\varrho d\sigma = \text{const}$ und der Inkompressibilität wegen zugleich auch $d\sigma = \text{const}$.

Das *Helmholtz*sche Theorem hätten wir auch aus der Gleichung (3) von § 1 ableiten können in folgender Weise: Es ist

$$\begin{aligned} \frac{\partial}{\partial t}\left(\frac{\overline{u}}{U}\right) &= \quad 2\varrho\frac{\overline{v}}{U} - \frac{\partial}{\partial u}\left(\Phi + \tfrac{1}{2}q^2\right) \,, \\ \frac{\partial}{\partial t}\left(\frac{\overline{v}}{V}\right) &= -2\varrho\frac{\overline{u}}{V} - \frac{\partial}{\partial v}\left(\Phi + \tfrac{1}{2}q^2\right) \,, \end{aligned} \tag{a}$$

da U und V von t unabhängig sind. Wenn man hier die zweite Gleichung nach u, die erste nach v differentiiert und subtrahiert, so folgt:

$$\frac{\partial^2}{\partial u \partial t}\left(\frac{\overline{v}}{V}\right) - \frac{\partial^2}{\partial v \partial t}\left(\frac{\overline{u}}{U}\right) = -2\frac{\partial}{\partial u}\left(\frac{\varrho\overline{u}}{V}\right) - 2\frac{\partial}{\partial v}\left(\frac{\varrho\overline{v}}{U}\right) \,.$$

Hier ist aber die linke Seite nach der Definition (c) S. 204 $= \dfrac{\partial}{\partial t}\left(\dfrac{2\varrho}{UV}\right)$, so daß wir die Gleichung gewinnen:

$$\frac{\partial}{\partial t}\left(\frac{\varrho}{UV}\right) + \frac{\partial}{\partial u}\left(\frac{\varrho\overline{u}}{V}\right) + \frac{\partial}{\partial v}\left(\frac{\varrho\overline{v}}{U}\right) = 0 \,. \tag{2}$$

or, by virtue of the equations of motion (1) § 1:

$$2\frac{dR}{dt} = \int^{(\mathfrak{C})} \left\{ -\frac{\partial \Phi}{\partial u}\delta u - \frac{\partial \Phi}{\partial v}\delta v + \overline{u}\delta\overline{u} + \overline{v}\delta\overline{v} \right\}$$
$$= \int^{(\mathfrak{C})} \left(-\delta\Phi + \tfrac{1}{2}\delta\left(q^2\right)\right) = \int \delta\left(\tfrac{1}{2}q^2 - \Phi\right) = \left[\tfrac{1}{2}q^2 - \Phi\right]_A^A = 0\,,$$

if the integral is extended over a closed curve $\mathfrak{C}$. Hence, it indeed follows, as asserted, that:

$$R = \tfrac{1}{2}\int^{(\mathfrak{C})} q_s ds = \int^{(C)} \varrho d\sigma = \text{const.} \tag{1}$$

Our theorem immediately leads to the corollaries:

Theorem II. *If part of the fluid is vortex-free at some point in time ($\varrho = 0$), then so it is at all times.*

Theorem III. *If the fluid is* incompressible*, then the vortex density ϱ is constant in every material point, i. e., at every position on the surface reached by a particular point m the vortex density ϱ has, at this instant in time, always the same value ϱ_m, which is characteristic of the point m.*

For, by Theorem I, we have $\varrho d\sigma = \text{const}$ and, by dint of the incompressibility, also $d\sigma = \text{const}$.

We could have also derived *Helmholtz*'s theorem from equation (3) of § 1 as follows: We have

$$\begin{aligned} \frac{\partial}{\partial t}\left(\frac{\overline{u}}{U}\right) &= \quad 2\varrho\frac{\overline{v}}{U} - \frac{\partial}{\partial u}\left(\Phi + \tfrac{1}{2}q^2\right)\,, \\ \frac{\partial}{\partial t}\left(\frac{\overline{v}}{V}\right) &= -2\varrho\frac{\overline{u}}{V} - \frac{\partial}{\partial v}\left(\Phi + \tfrac{1}{2}q^2\right)\,, \end{aligned} \tag{a}$$

since U and V are independent of t. If we differentiate here the second equation with respect to u, and the first with respect to v and subtract, then it follows that:

$$\frac{\partial^2}{\partial u \partial t}\left(\frac{\overline{v}}{V}\right) - \frac{\partial^2}{\partial v \partial t}\left(\frac{\overline{u}}{U}\right) = -2\frac{\partial}{\partial u}\left(\frac{\varrho\overline{u}}{V}\right) - 2\frac{\partial}{\partial v}\left(\frac{\varrho\overline{v}}{U}\right)\,.$$

Here, however, by definition (c) p. 204, the left side is $= \dfrac{\partial}{\partial t}\left(\dfrac{2\varrho}{UV}\right)$, so that we get the equation:

$$\frac{\partial}{\partial t}\left(\frac{\varrho}{UV}\right) + \frac{\partial}{\partial u}\left(\frac{\varrho\overline{u}}{V}\right) + \frac{\partial}{\partial v}\left(\frac{\varrho\overline{v}}{U}\right) = 0\,. \tag{2}$$

Diese unterscheidet sich von der Gleichung (2) in § 2 nur dadurch, daß hier ϱ an Stelle von k tritt und wie jene, die „Kontinuitätsgleichung“, die Konstanz der *Masse* $kd\sigma$ eines materiellen Teilchens ausdrückt, ebenso diese die Konstanz seines *Wirbelmomentes* $\varrho d\sigma$.

Im Falle der Inkompressibilität ist nun nach (3) § 2

$$\frac{\partial}{\partial u}\left(\frac{\overline{u}}{V}\right)+\frac{\partial}{\partial v}\left(\frac{\overline{v}}{U}\right)=0\,,$$

so daß (2) sich schreiben läßt:

$$\begin{aligned}\frac{\partial\varrho}{\partial t}&=-UV\left(\frac{\overline{u}\,\partial\varrho}{V\,\partial u}-\frac{\overline{v}\,\partial\varrho}{U\,\partial v}\right)\\&=-U\overline{u}\frac{\partial\varrho}{\partial u}-V\overline{v}\frac{\partial\varrho}{\partial v}\end{aligned}\tag{3}$$

oder

$$\frac{\partial\varrho}{dt}\equiv\frac{\partial\varrho}{\partial t}+\frac{\partial\varrho\,du}{\partial u\,dt}+\frac{\partial\varrho\,dv}{\partial v\,dt}=0\,,$$

d. h. $\varrho=\text{const.}$

216 | Hier kann man aber auch $\overline{u}$ und $\overline{v}$ durch die Stromfunktion ψ ausdrücken [(4) § 2] und erhält:

$$\frac{\partial\varrho}{\partial t}=UV\left(\frac{\partial\varrho}{\partial u}\frac{\partial\psi}{\partial v}-\frac{\partial\varrho}{\partial v}\frac{\partial\psi}{\partial u}\right)\tag{4}$$

oder, wenn man für 2ϱ nach (4) § 3 seinen Wert $D\psi$ einsetzt:

$$\frac{\partial D\psi}{\partial t}=UV\left(\frac{\partial\psi}{\partial v}\frac{\partial D\psi}{\partial u}-\frac{\partial\psi}{\partial u}\frac{\partial D\psi}{\partial v}\right)\,,\tag{4$'$}$$

eine (nicht homogene) partielle Differentialgleichung dritter Ordnung für $\psi=\psi(u,v,t)$, welche den ganzen zeitlichen Verlauf der Flüssigkeitsbewegung bestimmt.

Ist nämlich irgend eine Lösung $\psi=\psi(u,v,t)$ dieser Differentialgleichung gefunden, welche den gegebenen Grenz- und Stetigkeitsbedingungen genügt, so können die beiden Gleichungen (a) nur noch zur Bestimmung des *Druckes* dienen. Sie lassen sich nämlich schreiben:

$$\begin{aligned}\frac{\partial}{\partial u}\left(\Phi+\tfrac{1}{2}q^2\right)&=2\varrho\frac{\partial\psi}{\partial u}+\frac{V}{U}\frac{\partial^2\psi}{\partial v\partial t}\\\frac{\partial}{\partial v}\left(\Phi+\tfrac{1}{2}q^2\right)&=2\varrho\frac{\partial\psi}{\partial v}-\frac{U}{V}\frac{\partial^2\psi}{\partial u\partial t}\end{aligned}\tag{5}$$

This equation differs from equation (2) in § 2 only in that, here, ϱ takes the place of k and in that it expresses the constancy of the *vorticity moment* $\varrho d\sigma$ of a material particle, just like equation (2), the "continuity equation", expresses the constancy of its *mass* $kd\sigma$.

In the case of incompressibility we now have, by (3) § 2

$$\frac{\partial}{\partial u}\left(\frac{\overline{u}}{V}\right)+\frac{\partial}{\partial v}\left(\frac{\overline{v}}{U}\right)=0\,,$$

so that we can write (2) as:[4]

$$\begin{aligned}\frac{\partial\varrho}{\partial t}&=-UV\left(\frac{\overline{u}\,\partial\varrho}{V\,\partial u}+\frac{\overline{v}\,\partial\varrho}{U\,\partial v}\right)\\&=-U\overline{u}\frac{\partial\varrho}{\partial u}-V\overline{v}\frac{\partial\varrho}{\partial v}\end{aligned}\tag{3}$$

or

$$\frac{\partial\varrho}{dt}\equiv\frac{\partial\varrho}{\partial t}+\frac{\partial\varrho\,du}{\partial u\,dt}+\frac{\partial\varrho\,dv}{\partial v\,dt}=0\,,$$

i. e., $\varrho=\text{const.}$

But, in this case, we may also express $\overline{u}$ and $\overline{v}$ in terms of the stream function ψ [(4) § 2] and obtain:

$$\frac{\partial\varrho}{\partial t}=UV\left(\frac{\partial\varrho}{\partial u}\frac{\partial\psi}{\partial v}-\frac{\partial\varrho}{\partial v}\frac{\partial\psi}{\partial u}\right)\tag{4}$$

or, if we substitute for 2ϱ its value $D\psi$, by (4) § 3:

$$\frac{\partial D\psi}{\partial t}=UV\left(\frac{\partial\psi}{\partial v}\frac{\partial D\psi}{\partial u}-\frac{\partial\psi}{\partial u}\frac{\partial D\psi}{\partial v}\right)\,,\tag{4$'$}$$

a (non-homogeneous) partial differential equation of third order for $\psi=\psi(u,v,t)$, which determines the entire temporal evolution of the motion of the fluid.

For if we have found any solution $\psi=\psi(u,v,t)$ of this differential equation satisfying the given boundary and continuity conditions, then the two equations (a) can only serve to determine the *pressure*. For they can be written as:

$$\begin{aligned}\frac{\partial}{\partial u}\left(\Phi+\tfrac{1}{2}q^2\right)&=2\varrho\frac{\partial\psi}{\partial u}+\frac{V}{U}\frac{\partial^2\psi}{\partial v\partial t}\\\frac{\partial}{\partial v}\left(\Phi+\tfrac{1}{2}q^2\right)&=2\varrho\frac{\partial\psi}{\partial v}-\frac{U}{V}\frac{\partial^2\psi}{\partial u\partial t}\end{aligned}\tag{5}$$

[4] [[In the brackets of the next formula, Zermelo erroneously writes "−" instead of "+".]]

oder, wenn man die Variationen δu, δv einführt:

$$\delta\left(\Phi + \tfrac{1}{2}q^2\right) = \left(2\varrho\frac{\partial\psi}{\partial u} + \frac{V}{U}\frac{\partial^2\psi}{\partial v\partial t}\right)\delta u + \left(2\varrho\frac{\partial\psi}{\partial v} - \frac{U}{V}\frac{\partial^2\psi}{\partial u\partial t}\right)\delta v\ . \qquad (5)'$$

Hier ist auf Grund von (4) die rechte Seite ein vollständiges Differential (denn es wurde ja (4) gerade durch Elimination von Φ aus den Gleichungen (a) oder (5) abgeleitet). Wir können also die Funktion $\Phi + \frac{1}{2}q^2$ durch Integration finden und brauchen dann nur noch $\frac{1}{2}q^2 = \frac{1}{2}U^2\left(\frac{\partial\psi}{\partial u}\right)^2 + \frac{1}{2}V^2\left(\frac{\partial\psi}{\partial v}\right)^2$ abzuziehen, um Φ, und dann noch das Potential Φ_1 der Massenkräfte (cf. p. 204), um die Druckfunktion $P = \int\frac{dp}{k} = p$, d. h. den Druck selbst als Funktion von u, v und t zu finden.

Besonders wichtig ist der Fall, wo $\frac{\partial\psi}{\partial t} = 0$, d. h. wo die Strömung *stationär* ist. Dann muß auch $\frac{\partial\varrho}{\partial t} = 0$ sein, also nach (4)

$$\frac{\partial\varrho}{\partial u}\frac{\partial\psi}{\partial v} - \frac{\partial\varrho}{\partial v}\frac{\partial\psi}{\partial u} = 0 \qquad (4a)$$

oder integriert:

$$2\varrho = D\psi = f(\psi)\ , \qquad (4b)$$

eine Relation, die sich auch geometrisch so ausdrücken läßt:

217 | Satz IV. *Bei der stationären Strömung einer inkompressiblen Flüssigkeit in einer beliebigen Fläche muß auf jeder Stromlinie* ($\psi = \text{const.}$) *die Wirbeldichte* ϱ *constant sein.*

Dies leuchtet auch unmittelbar ein, weil sich im stationären Falle die materiellen Punkte mit unverändertem ϱ (Satz III) auf den festen Stromlinien bewegen und dabei doch ϱ an jeder Stelle ungeändert bleiben soll.[1]

In diesem Falle ist die Bestimmung des Druckes sehr einfach. Denn hier wird (5)

$$\frac{\partial\left(\Phi + \frac{1}{2}q^2\right)}{\partial u} = 2\varrho\frac{\partial\psi}{\partial u}\ , \quad \frac{\partial\left(\Phi + \frac{1}{2}q^2\right)}{\partial v} = 2\varrho\frac{\partial\psi}{\partial v}\ ,$$

also:

$$\Phi + \tfrac{1}{2}q^2 = \Phi_1 + p + \tfrac{1}{2}q^2 = \int 2\varrho\left(\frac{\partial\psi}{\partial u}du + \frac{\partial\psi}{\partial v}d\sigma\right) = \int 2\varrho d\psi\ ; \qquad (6)$$

[1] Für den Spezialfall der Ebene, wo $D\psi = \Delta\psi$ ist, findet sich die Bedingung $\Delta\psi = f(\psi)$ der Stationarität bereits bei *Lagrange* (Oeuvres t. 4 p. 720) vergl. auch *Stokes* (Math. Phys. Papers v. I p. 264), *Lamb*, (Hydrod. p. 263).

or, if we introduce the variations δu, δv:

$$\delta\left(\Phi + \tfrac{1}{2}q^2\right) = \left(2\varrho\frac{\partial\psi}{\partial u} + \frac{V}{U}\frac{\partial^2\psi}{\partial v\partial t}\right)\delta u + \left(2\varrho\frac{\partial\psi}{\partial v} - \frac{U}{V}\frac{\partial^2\psi}{\partial u\partial t}\right)\delta v\ . \qquad (5)'$$

In this case, the right side is a total differential by virtue of (4) (for (4) has just been derived from equations (a), or (5), by elimination of Φ). Hence, we can find the function $\Phi + \frac{1}{2}q^2$ by means of integration and then only need to subtract $\frac{1}{2}q^2 = \frac{1}{2}U^2\left(\frac{\partial\psi}{\partial u}\right)^2 + \frac{1}{2}V^2\left(\frac{\partial\psi}{\partial v}\right)^2$ in order to find Φ, and also the potential Φ_1 of the mass forces (cf. p. 204), in order to find the pressure function $P = \int\frac{dp}{k} = p$, i. e. the pressure itself as a function of u, v and t.

The case where $\frac{\partial\psi}{\partial t} = 0$, i. e., where the current is *stationary*, is of particular significance. In this case, we must also have $\frac{\partial\varrho}{\partial t} = 0$, and hence, by (4),

$$\frac{\partial\varrho}{\partial u}\frac{\partial\psi}{\partial v} - \frac{\partial\varrho}{\partial v}\frac{\partial\psi}{\partial u} = 0 \qquad (4a)$$

or, after integration:

$$2\varrho = D\psi = f(\psi)\ , \qquad (4b)$$

a relation, which can also be represented geometrically as follows:

Theorem IV. *In case of stationary flow of an incompressible fluid in an arbitrary surface, the vortex density ϱ must be constant on every streamline* ($\psi = \text{const.}$).

This is also immediately evident, since in the stationary case the material points with unaltered ϱ (Theorem III) move on fixed streamlines, while, at the same time, ϱ is supposed to remain unaltered at every position.[5]

In this case, it is very easy to determine the pressure. For, here, (5) becomes

$$\frac{\partial\left(\Phi + \frac{1}{2}q^2\right)}{\partial u} = 2\varrho\frac{\partial\psi}{\partial u}\ , \quad \frac{\partial\left(\Phi + \frac{1}{2}q^2\right)}{\partial v} = 2\varrho\frac{\partial\psi}{\partial v}\ ,$$

and hence:

$$\Phi + \tfrac{1}{2}q^2 = \Phi_1 + p + \tfrac{1}{2}q^2 = \int 2\varrho\left(\frac{\partial\psi}{\partial u}du + \frac{\partial\psi}{\partial v}d\sigma\right) = \int 2\varrho d\psi\ ; \qquad (6)$$

[5] For the special case of the plane where we have $D\psi = \Delta\psi$, the condition $\Delta\psi = f(\psi)$ of the stationarity is already to be found in *Lagrange 1869*, p. 720, cf. also *Stokes 1880*, p. 264, *Lamb 1895*, p. 263.

und auf den Stromlinien $\psi = \text{const}$, $\varrho = \text{const}$ ist gleichzeitig auch

$$\Phi = \text{const}\,,$$

und falls keine Massenkräfte wirken, der Druck konstant:

$$p = \text{const}\,.$$

§ 5. Die Erhaltung der lebendigen Kraft.

Aus der Form (2) unserer hydrodynamischen Grundgleichungen in § 1 folgt unmittelbar:

$$\begin{aligned}\overline{u}\frac{d\overline{u}}{dt} + \overline{v}\frac{d\overline{v}}{dt} &= -U\overline{u}\frac{\partial\Phi}{\partial u} - V\overline{v}\frac{\partial\Phi}{\partial v}\\ &= -\frac{\partial\Phi}{\partial u}\frac{du}{dt} - \frac{\partial\Phi}{\partial v}\frac{dv}{dt}\,,\end{aligned}$$

oder im Falle der Inkompressibilität, wenn man nach § 2 $\overline{u}$ und $\overline{v}$ durch ψ ausdrückt:

$$\overline{u}\frac{d\overline{u}}{dt} + \overline{v}\frac{d\overline{v}}{dt} = UV\left(\frac{\partial\psi}{\partial v}\frac{\partial\Phi}{\partial u} - \frac{\partial\psi}{\partial u}\frac{\partial\Phi}{\partial v}\right)\,, \tag{1}$$

also, wenn man über ein Flächenstück C integriert:

$$\overset{(C)}{\int}\left(\overline{u}\frac{d\overline{u}}{dt} + \overline{v}\frac{d\overline{v}}{dt}\right)d\sigma = \overset{(C)}{\int\int} dudv\left(\frac{\partial\Phi}{\partial u}\frac{\partial\psi}{\partial v} - \frac{\partial\Phi}{\partial u}\frac{\partial\psi}{\partial v}\right)\,.$$

Da im betrachteten Falle auch das Flächenelement $d\sigma$ in der Zeit konstant
218 ist, so ist die linke Seite nichts anderes als der zeitliche | Differentialquotient der gesamten im Bereich C enthaltenen lebendigen Kraft

$$\frac{dT_C}{dt} = \frac{d}{dt}\overset{(C)}{\int}\tfrac{1}{2}\left(\overline{u}^2 + \overline{v}^2\right)d\sigma\,.$$

Die rechte Seite aber läßt sich in ein Randintegral verwandeln, im positiven Sinne erstreckt über die Begrenzung $\mathfrak{C}$ von C:

$$\begin{aligned}\overset{(C)}{\int} dudv\left(\frac{\partial\psi}{\partial v}\frac{\partial\Phi}{\partial u} - \frac{\partial\psi}{\partial u}\frac{\partial\Phi}{\partial v}\right) &= \overset{(\mathfrak{C})}{\int}\Phi\left(\frac{\partial\psi}{\partial u}du + \frac{\partial\psi}{\partial v}dv\right)\\ &= \overset{(\mathfrak{C})}{\int}\Phi\frac{\partial\psi}{\partial s}ds = \overset{(\mathfrak{C})}{\int}\Phi q_n ds\,,\end{aligned}$$

and on the streamlines $\psi = \text{const}$, $\varrho = \text{const}$ we also have

$$\Phi = \text{const}\,,$$

and in case no mass forces act, the pressure is constant:

$$p = \text{const}\,.$$

§ 5. The conservation of the living force.

From the form (2) of our hydrodynamical basic equations in § 1 it immediately follows that:

$$\begin{aligned}\overline{u}\frac{d\overline{u}}{dt} + \overline{v}\frac{d\overline{v}}{dt} &= -U\overline{u}\frac{\partial\Phi}{\partial u} - V\overline{v}\frac{\partial\Phi}{\partial v}\\ &= -\frac{\partial\Phi}{\partial u}\frac{du}{dt} - \frac{\partial\Phi}{\partial v}\frac{dv}{dt}\,,\end{aligned}$$

or, in the case of the incompressibility, if we express $\overline{u}$ and $\overline{v}$ in terms of ψ in accordance with § 2:

$$\overline{u}\frac{d\overline{u}}{dt} + \overline{v}\frac{d\overline{v}}{dt} = UV\left(\frac{\partial\psi}{\partial v}\frac{\partial\Phi}{\partial u} - \frac{\partial\psi}{\partial u}\frac{\partial\Phi}{\partial v}\right)\,, \tag{1}$$

and hence, if we integrate over the part C of the surface:

$$\overset{(C)}{\int}\left(\overline{u}\frac{d\overline{u}}{dt} + \overline{v}\frac{d\overline{v}}{dt}\right)d\sigma = \overset{(C)}{\int}\int dudv\left(\frac{\partial\Phi}{\partial u}\frac{\partial\psi}{\partial v} - \frac{\partial\Phi}{\partial u}\frac{\partial\psi}{\partial v}\right)\,.$$

Since, in the case under consideration, the area element $d\sigma$, too, is constant over time, the left side is but the temporal derivative of the entire living force contained in the domain C

$$\frac{dT_C}{dt} = \frac{d}{dt}\overset{(C)}{\int}\tfrac{1}{2}\left(\overline{u}^2 + \overline{v}^2\right)d\sigma\,.$$

The right side, however, can be transformed into a boundary integral extended in the positive direction over the boundary $\mathfrak{C}$ of C:

$$\begin{aligned}\overset{(C)}{\int} dudv\left(\frac{\partial\psi}{\partial v}\frac{\partial\Phi}{\partial u} - \frac{\partial\psi}{\partial u}\frac{\partial\Phi}{\partial v}\right) &= \overset{(\mathfrak{C})}{\int}\Phi\left(\frac{\partial\psi}{\partial u}du + \frac{\partial\psi}{\partial v}dv\right)\\ &= \overset{(\mathfrak{C})}{\int}\Phi\frac{\partial\psi}{\partial s}ds = \overset{(\mathfrak{C})}{\int}\Phi q_n ds\,,\end{aligned}$$

wenn ds wieder das Bogenelement von $\mathfrak{C}$ und q_n die Normalkomponente der Geschwindigkeit bedeutet. Wir haben also:

$$\frac{dT_C}{dt} = \frac{d}{dt}\int^{(C)} \tfrac{1}{2}q^2 d\sigma = \int^{(\mathfrak{C})} \Phi q_n ds \ . \tag{2}$$

Dieses ist der Ausdruck des Gesetzes von der Erhaltung der Energie, angewendet auf eine inkompressible zweidimensionale Flüssigkeit.

Nehmen wir nun an, daß die Begrenzung $\mathfrak{C}$ unseres Bereiches C durch eine Anzahl von Stromlinien $\psi = \text{const}$ gebildet werden oder daß sie in einen (nicht singulären) Punkt zusammenschrumpft, während C eine geschlossene Fläche vollständig erfüllt, so verschwindet die rechte Seite, und es wird:

$$T_C = \frac{1}{2}\int^{(\mathfrak{C})} \left(\overline{u}^2 + \overline{v}^2\right) d\sigma = \text{const} \ . \tag{3}$$

Also:

Satz I. *Die gesamte lebendige Kraft einer inkompressiblen reibungslosen Flüssigkeit in einer vollständig geschlossenen Fläche oder in einem Flächenstücke von fester aus Stromlinien gebildeten Umrandung ist in der Zeit konstant, vorausgesetzt, daß die Flüssigkeit keine Quellpunkte besitzt und daß die äußeren Kräfte ein Potential haben.*

Für den Fall der *geschlossenen Fläche* können wir aber der gesamten lebendigen Kraft T noch eine andere Form geben durch Einführung der Stromfunktion ψ und der Wirbeldichte ϱ. Es ist nämlich nach (4) S. 207:

$$\begin{aligned}\frac{\overline{u}^2 + \overline{v}^2}{UV} &= -\frac{\overline{u}}{U}\frac{\partial\psi}{\partial v} + \frac{\overline{v}}{V}\frac{\partial\psi}{\partial u} \\ &= \frac{\partial}{\partial u}\left(\frac{\psi\overline{v}}{V}\right) - \frac{\partial}{\partial v}\left(\frac{\psi\overline{u}}{U}\right) - \frac{2\psi\varrho}{UV} \qquad \text{(cf. (c) § [[1]]),}\end{aligned}$$

219 | und daher, wenn man über das Flächenstück C integriert und rechts das entsprechende Randintegral über $\mathfrak{C}$ einführt:

$$2T_C = \int^{(C)}\int \left(\overline{u}^2 + \overline{v}^2\right)\frac{dudv}{UV} = \int^{(\mathfrak{C})} \psi\left(\overline{u}\frac{du}{U} + \overline{v}\frac{dv}{V}\right) - \int^{(C)}\int 2\psi\varrho\frac{dudv}{UV} \tag{4}$$

oder:

$$T_C = \int^{(C)} \tfrac{1}{2}q^2 d\sigma = \tfrac{1}{2}\int^{(\mathfrak{C})} \psi q_s ds - \int^{(C)} \psi\varrho d\sigma \ .$$

if ds again denotes the arc element of $\mathfrak{C}$, and q_n the normal component of the velocity. We thus have:

$$\frac{dT_C}{dt} = \frac{d}{dt}\int^{(C)} \tfrac{1}{2}q^2 d\sigma = \int^{(\mathfrak{C})} \Phi q_n ds \,. \tag{2}$$

This is the expression of the law of the conservation of energy as applied to an incompressible two-dimensional fluid.

If we now assume that the boundary $\mathfrak{C}$ of our domain C is formed by a number of streamlines $\psi = \text{const}$, or that it is shrunk to a (non-singular) point, while C completely fills a closed surface, then the right side disappears, and we get:

$$T_C = \frac{1}{2}\int^{(\mathfrak{C})} \left(\overline{u}^2 + \overline{v}^2\right) d\sigma = \text{const} \,. \tag{3}$$

And hence:

Theorem I. *The entire living force of an incompressible frictionless fluid in a completely closed surface or in a part of the surface with a fixed border formed by streamlines is constant over time, provided that the fluid has no source points and that the extraneous forces have a potential.*

For the case of the *closed surface*, however, we can express the entire living force T in a different form by introducing the stream function ψ and the vortex density ϱ. For, by (4) p. 207, we have:

$$\begin{aligned}\frac{\overline{u}^2 + \overline{v}^2}{UV} &= -\frac{\overline{u}}{U}\frac{\partial\psi}{\partial v} + \frac{\overline{v}}{V}\frac{\partial\psi}{\partial u} \\ &= \frac{\partial}{\partial u}\left(\frac{\psi\overline{v}}{V}\right) - \frac{\partial}{\partial v}\left(\frac{\psi\overline{u}}{U}\right) - \frac{2\psi\varrho}{UV} \qquad \text{(cf. (c) § [[1]]),}\end{aligned}$$

and hence, if we integrate over the part C of the surface and introduce the corresponding boundary integral over $\mathfrak{C}$ on the right side:

$$2T_C = \int^{(C)}\int \left(\overline{u}^2 + \overline{v}^2\right)\frac{dudv}{UV} = \int^{(\mathfrak{C})} \psi\left(\overline{u}\frac{du}{U} + \overline{v}\frac{dv}{V}\right) - \int^{(C)}\int 2\psi\varrho\frac{dudv}{UV} \tag{4}$$

or:

$$T_C = \int^{(C)} \tfrac{1}{2}q^2 d\sigma = \tfrac{1}{2}\int^{(\mathfrak{C})} \psi q_s ds - \int^{(C)} \psi\varrho d\sigma \,.$$

Also, wenn C die ganze geschlossene Fläche F darstellt:

$$T = \int^{(F)} \tfrac{1}{2} q^2 d\sigma = - \int^{(F)} \psi \varrho d\sigma = \text{const} .^1 \tag{4a}$$

Wir haben somit den Satz gewonnen:

Satz II. *Bei der Strömung einer inkompressiblen Flüssigkeit in einer geschlossenen Fläche ist die Summe aller Wirbelelemente $\varrho d\sigma$, jedes multipliziert mit dem jeweiligen Werte der Stromfunktion ψ gleich der negativen lebendigen Kraft der ganzen Strömung und daher unter den Voraussetzungen des Satzes I in der Zeit konstant:*

$$\mathsf{P} = \int \psi \varrho d\sigma = -T = const .$$

Kapitel II

Anwendung der allgemeinen Theorie auf die Kugel.

§ 1. Die Grundformeln in stereographischen und Polarkoordinaten.

Ist die betrachtete Fläche eine Kugel, so empfehlen sich zur Behandlung der hydrodynamischen Probleme insbesondere die folgenden beiden Systeme orthogonaler Koordinaten u, v.

1) *Polarkoordinaten* ϑ, ω, wo ϑ den Winkelabstand eines Punktes P von einem festen Anfangspunkt O und ω den Neigungswinkel des Meridianes OPO' gegen einen festen Anfangsmeridian $\omega = 0$ bedeutet. Wir rechnen ω nach der Richtung als zunehmend, welche einer positiven Drehung um den Durchmesser $O'CO$ entspricht (nach links, wenn der Beobachter sich in die Richtung CO stellt).

2) *Stereographische Koordinaten x, y,* d. h. die rechtwinkligen Koordinaten des Punktes p, den wir aus P durch stereographische Projektion von einem
220 festen Kugelpunkte O' (entgegengesetzt O) aus auf die | Äquatorebene des Durchmessers OO' gewinnen. Auch hier soll der Übergang von der x- zur y-Achse einer positiven Drehung um die Achse $O'O$ entsprechen. Dieses Koordinatensystem ist besonders deswegen für viele Zwecke wichtig, weil die stereographische Abbildung der Kugel auf die Ebene bekanntlich eine kon-
Fig. 2 forme oder winkeltreue ist.

[1] Vergl. *Poincaré*, Théor. d. Tourbillons, § 73. p. 80/81.

And hence, if C represents the entire closed surface F:

$$T = \int^{(F)} \tfrac{1}{2} q^2 d\sigma = - \int^{(F)} \psi \varrho d\sigma = \text{const}\,.^{6} \tag{4a}$$

We have thus obtained the theorem:

Theorem II. *Given a flow of an incompressible fluid in a closed surface, the sum of all vortex elements $\varrho d\sigma$, each of them multiplied with the respective value of the stream function ψ, is equal to the negative living force of the entire flow, and hence constant over time under the assumptions of Theorem I:*

$$\mathsf{P} = \int \psi \varrho d\sigma = -T = const\,.$$

Chapter II

Application of the general theory to the sphere.

§ 1. The basic formulas in stereographic and polar coordinates.

If the surface under consideration is a sphere, then it is advisable to treat the hydrodynamical problems by using in particular the following two systems of orthogonal coordinates u, v.

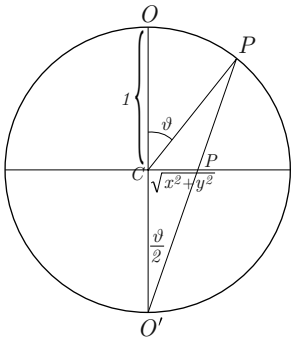
1) *Polar coordinates* ϑ, ω, where ϑ denotes the angular distance of a point P from a fixed starting point O, and ω the inclination angle of the meridian OPO' with respect to a fixed initial meridian $\omega = 0$. We consider ω as increasing in that direction which corresponds to a positive rotation about the diameter $O'CO$ (counterclockwise, assuming the observer faces in the direction CO).

2) *Stereographic coordinates x, y,* i. e., the orthogonal coordinates of the point p we obtain from P by stereographic projection from a fixed point of the sphere O' (antipodal to O) onto the equatorial plane of the diameter OO'. Here, too, the transition from the x- to the y-axis is supposed to correspond to a positive rotation about the axis $O'O$. This coordinate system is important for many purposes in particular because the stereographic projection of the sphere onto the plane is, as is well-known, conformal, or angle-preserving.

Fig. 2

[6] Cf. *Poincaré 1893a*, § 73, p. 80/81.

Fig. 2.



Wählen wir das Projektionszentrum O' dem Punkte O des Polarkoordinatensystems gerade entgegengesetzt, und legen die x-Achse in den Anfangsmeridian $\omega = 0$, so bestehen zwischen ϑ, ω; x, y, wenn wir den Kugelradius $= 1$ annehmen, die Beziehungen:

$$\sqrt{x^2+y^2} = \operatorname{tg}\frac{\vartheta}{2}\,, \quad x = \operatorname{tg}\frac{\vartheta}{2}\cos\omega\,, \quad y = \operatorname{tg}\frac{\vartheta}{2}\sin\omega\,. \tag{1}$$

Also:

$$dx^2 + dy^2 = \frac{d\vartheta^2}{4\cos^4\frac{\vartheta}{2}} + \operatorname{tg}^2\frac{\vartheta}{2}d\omega^2 = \frac{d\vartheta^2 + \sin^2\vartheta d\omega^2}{4\cos^4\frac{\vartheta}{2}}\,.$$

Es wird daher, wenn ds wie früher die Länge des Linienelementes auf der Kugel bezeichnet:

$$ds^2 = d\vartheta^2 + \sin^2\vartheta d\omega^2 = \frac{4}{(1+x^2+y^2)^2}\left(dx^2+dy^2\right)\,, \tag{2}$$

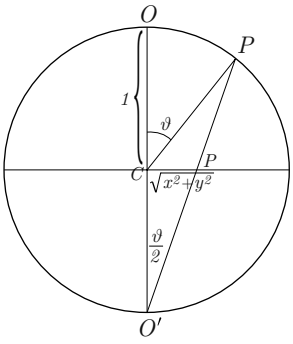
so daß, wenn wir für die Größen U, V in Kap. I bezw. Θ, Ω; X, Y setzen und die Geschwindigkeitskomponenten durch $\overline{\vartheta}$, $\overline{\omega}$; $\overline{x}$, $\overline{y}$ bezeichnen:

$$\left.\begin{array}{l|l} \Theta = 1\,,\ \Omega = \frac{1}{\sin\vartheta}\,. & X = Y = \frac{1}{2}\left(1+x^2+y^2\right)\,. \\ \frac{d\vartheta}{dt} = \overline{\vartheta}\,,\ \frac{d\omega}{dt} = \frac{\overline{\omega}}{\sin\vartheta}\,. & \frac{dx}{dt} = \frac{1}{2}\left(1+x^2+y^2\right)\overline{x}\,,\ \frac{dy}{dt} = \frac{1}{2}\left(1+x^2+y^2\right)\overline{y}\,. \end{array}\right\} \tag{3}$$

Aus den Grundformeln (1), (2) und (3) in Kap. I, § 1 erhalten wir dann die folgenden:

$$\left.\begin{aligned} \frac{d\overline{\vartheta}}{dt} - \cot\vartheta\overline{\omega}^2 &= -\frac{\partial\Phi}{\partial\vartheta} \\ \frac{d\overline{\omega}}{dt} + \cot\vartheta\overline{\vartheta}\overline{\omega}^2 &= -\frac{1}{\sin\vartheta}\frac{\partial\Phi}{\partial\omega} \end{aligned}\right\} \tag{4a}$$

Fig. 2.



If we take the center of projection O' to be exactly antipodal to the point O of the polar coordinate system and place the x-axis in the initial meridian $\omega = 0$, then, assuming we set the radius of the sphere to be $= 1$, the following relations hold between ϑ, ω; x, y:

$$\sqrt{x^2+y^2} = \operatorname{tg}\frac{\vartheta}{2}\,, \quad x = \operatorname{tg}\frac{\vartheta}{2}\cos\omega\,, \quad y = \operatorname{tg}\frac{\vartheta}{2}\sin\omega\,. \tag{1}$$

And hence:

$$dx^2+dy^2 = \frac{d\vartheta^2}{4\cos^4\frac{\vartheta}{2}} + \operatorname{tg}^2\frac{\vartheta}{2}d\omega^2 = \frac{d\vartheta^2+\sin^2\vartheta d\omega^2}{4\cos^4\frac{\vartheta}{2}}\,.$$

We therefore get:

$$ds^2 = d\vartheta^2+\sin^2\vartheta d\omega^2 = \frac{4}{\left(1+x^2+y^2\right)^2}\left(dx^2+dy^2\right)\,, \tag{2}$$

where ds denotes, as previously, the length of the line element on the sphere, so that:

$$\left.\begin{array}{l|c} \Theta = 1\,,\ \Omega = \frac{1}{\sin\vartheta}\,. & X = Y = \frac{1}{2}\left(1+x^2+y^2\right)\,. \\ \frac{d\vartheta}{dt} = \overline{\vartheta}\,,\ \frac{d\omega}{dt} = \frac{\overline{\omega}}{\sin\vartheta}\,. & \frac{dx}{dt} = \frac{1}{2}\left(1+x^2+y^2\right)\overline{x}\,,\ \frac{dy}{dt} = \frac{1}{2}\left(1+x^2+y^2\right)\overline{y}\,, \end{array}\right\} \tag{3}$$

if we substitute for the variables U, V in chap. I [[the variables]] Θ, Ω; X, Y, respectively, and denote the velocity components by $\overline{\vartheta}$, $\overline{\omega}$; $\overline{x}$, $\overline{y}$.

From the basic formula (1), (2) and (3) in chap. I, § 1 we then obtain the following formulas:

$$\left.\begin{aligned} \frac{d\overline{\vartheta}}{dt} - \cot\vartheta\overline{\omega}^2 &= -\frac{\partial\Phi}{\partial\vartheta} \\ \frac{d\overline{\omega}}{dt} + \cot\vartheta\overline{\vartheta}\overline{\omega}^2 &= -\frac{1}{\sin\vartheta}\frac{\partial\Phi}{\partial\omega} \end{aligned}\right\} \tag{4a}$$

$$\left.\begin{aligned}\frac{d}{dt}\left(\frac{\overline{x}}{1+x^2+y^2}\right)+\frac{x}{1+x^2+y^2}\left(\overline{x}^2+\overline{y}^2\right)&=-\tfrac{1}{2}\frac{\partial\Phi}{\partial x}\\ \frac{d}{dt}\left(\frac{\overline{y}}{1+x^2+y^2}\right)+\frac{y}{1+x^2+y^2}\left(\overline{x}^2+\overline{y}^2\right)&=-\tfrac{1}{2}\frac{\partial\Phi}{\partial y}\end{aligned}\right\} \tag{4b}$$

221 |

$$\left.\begin{aligned}\frac{d\overline{x}}{dt}-\overline{y}\left(y\overline{x}-x\overline{y}\right)&=-\tfrac{1}{2}\left(1+x^2+y^2\right)\frac{\partial\Phi}{\partial x}\\ \frac{d\overline{y}}{dt}+\overline{x}\left(y\overline{x}-x\overline{y}\right)&=-\tfrac{1}{2}\left(1+x^2+y^2\right)\frac{\partial\Phi}{\partial y}\end{aligned}\right\} \tag{4b$'$}$$

$$\begin{array}{l|l}\frac{\partial\overline{\vartheta}}{\partial t}=2\varrho\overline{\omega}-\frac{\partial}{\partial\vartheta}\left(\Phi+\frac{1}{2}q^2\right) & \frac{\partial\overline{x}}{\partial t}=2\varrho\overline{y}-\frac{1}{2}\left(1+x^2+y^2\right)\frac{\partial}{\partial x}\left(\Phi+\frac{1}{2}q^2\right)\\ \frac{\partial\overline{\omega}}{\partial t}=-2\varrho\overline{\vartheta}-\frac{1}{\sin\vartheta}\frac{\partial}{\partial\omega}\left(\Phi+\frac{1}{2}q^2\right) & \frac{\partial\overline{y}}{\partial t}=-2\varrho\overline{x}-\frac{1}{2}\left(1+x^2+y^2\right)\frac{\partial}{\partial y}\left(\Phi+\frac{1}{2}q^2\right)\end{array} \tag{5}$$

$$2\varrho=\frac{1}{\sin\vartheta}\left\{\frac{\partial}{\partial\vartheta}\left(\overline{\omega}\sin\vartheta\right)-\frac{\partial\overline{\vartheta}}{\partial\omega}\right\} \tag{6a}$$

$$\begin{aligned}2\varrho&=\tfrac{1}{2}\left(1+x^2+y^2\right)^2\left\{\frac{\partial}{\partial x}\left(\frac{\overline{y}}{1+x^2+y^2}\right)-\frac{\partial}{\partial y}\left(\frac{\overline{x}}{1+x^2+y^2}\right)\right\}\\ &=y\overline{x}-x\overline{y}+\tfrac{1}{2}\left(1+x^2+y^2\right)\left(\frac{\partial\overline{y}}{\partial x}-\frac{\partial\overline{x}}{\partial y}\right).\end{aligned} \tag{6b}$$

Ferner werden die Komponenten der Geschwindigkeit $q=\sqrt{\overline{\vartheta}^2+\overline{\omega}^2}=\sqrt{\overline{x}^2+\overline{y}^2}$ in der Richtung der Tangente und Normale:

$$\begin{aligned}q_n&=\overline{\omega}\frac{d\vartheta}{ds}-\overline{\vartheta}\sin\vartheta\frac{d\omega}{ds} & q_n&=\frac{2}{1+x^2+y^2}\left(\overline{y}\frac{dx}{ds}-\overline{x}\frac{dy}{ds}\right)\\ q_s&=\overline{\vartheta}\frac{d\vartheta}{ds}+\overline{\omega}\sin\vartheta\frac{d\omega}{ds} & q_s&=\frac{2}{1+x^2+y^2}\left(\overline{x}\frac{dx}{ds}+\overline{y}\frac{dy}{ds}\right).\end{aligned} \tag{7}$$

Für eine inkompressible Flüssigkeit kann man noch die Stromfunktion $\psi=\int_0^P q_n ds$ einführen und erhält so nach (4) § 2 von Kap. I:

$$\left.\begin{aligned}\overline{\vartheta}=\frac{d\vartheta}{dt}=-\frac{1}{\sin\vartheta}\frac{\partial\psi}{\partial\omega},\ \overline{x}&=\frac{2}{1+x^2+y^2}\frac{dx}{dt}=-\frac{1}{2}\left(1+x^2+y^2\right)\frac{\partial\psi}{\partial y}\\ \overline{\omega}=\sin\vartheta\frac{d\omega}{dt}=\frac{\partial\psi}{\partial\vartheta},\quad \overline{y}&=\frac{2}{1+x^2+y^2}\frac{dy}{dt}=\ \ \frac{1}{2}\left(1+x^2+y^2\right)\frac{\partial\psi}{\partial x},\end{aligned}\right\} \tag{8}$$

$$\left.\begin{aligned}\frac{d}{dt}\left(\frac{\overline{x}}{1+x^2+y^2}\right)+\frac{x}{1+x^2+y^2}\left(\overline{x}^2+\overline{y}^2\right)&=-\tfrac{1}{2}\frac{\partial\Phi}{\partial x}\\ \frac{d}{dt}\left(\frac{\overline{y}}{1+x^2+y^2}\right)+\frac{y}{1+x^2+y^2}\left(\overline{x}^2+\overline{y}^2\right)&=-\tfrac{1}{2}\frac{\partial\Phi}{\partial y}\end{aligned}\right\} \tag{4b}$$

$$\left.\begin{aligned}\frac{d\overline{x}}{dt}-\overline{y}\left(y\overline{x}-x\overline{y}\right)&=-\tfrac{1}{2}\left(1+x^2+y^2\right)\frac{\partial\Phi}{\partial x}\\ \frac{d\overline{y}}{dt}+\overline{x}\left(y\overline{x}-x\overline{y}\right)&=-\tfrac{1}{2}\left(1+x^2+y^2\right)\frac{\partial\Phi}{\partial y}\end{aligned}\right\} \tag{4b}'$$

$$\left.\begin{aligned}\tfrac{\partial\overline{\vartheta}}{\partial t}&=2\varrho\overline{\omega}-\tfrac{\partial}{\partial\vartheta}\left(\Phi+\tfrac{1}{2}q^2\right)\\ \tfrac{\partial\overline{\omega}}{\partial t}&=-2\varrho\overline{\vartheta}-\tfrac{1}{\sin\vartheta}\tfrac{\partial}{\partial\omega}\left(\Phi+\tfrac{1}{2}q^2\right)\end{aligned}\right| \left.\begin{aligned}\tfrac{\partial\overline{x}}{\partial t}&=2\varrho\overline{y}-\tfrac{1}{2}\left(1+x^2+y^2\right)\tfrac{\partial}{\partial x}\left(\Phi+\tfrac{1}{2}q^2\right)\\ \tfrac{\partial\overline{y}}{\partial t}&=-2\varrho\overline{x}-\tfrac{1}{2}\left(1+x^2+y^2\right)\tfrac{\partial}{\partial y}\left(\Phi+\tfrac{1}{2}q^2\right)\end{aligned}\right. \tag{5}$$

$$2\varrho=\frac{1}{\sin\vartheta}\left\{\frac{\partial}{\partial\vartheta}\left(\overline{\omega}\sin\vartheta\right)-\frac{\partial\overline{\vartheta}}{\partial\omega}\right\} \tag{6a}$$

$$\begin{aligned}2\varrho&=\tfrac{1}{2}\left(1+x^2+y^2\right)^2\left\{\frac{\partial}{\partial x}\left(\frac{\overline{y}}{1+x^2+y^2}\right)-\frac{\partial}{\partial y}\left(\frac{\overline{x}}{1+x^2+y^2}\right)\right\}\\ &=y\overline{x}-x\overline{y}+\tfrac{1}{2}\left(1+x^2+y^2\right)\left(\frac{\partial\overline{y}}{\partial x}-\frac{\partial\overline{x}}{\partial y}\right)\ .\end{aligned} \tag{6b}$$

Furthermore, the components of the velocity $q=\sqrt{\overline{\vartheta}^2+\overline{\omega}^2}=\sqrt{\overline{x}^2+\overline{y}^2}$ in the direction of the tangent and the normal become:

$$\begin{aligned}q_n&=\overline{\omega}\frac{d\vartheta}{ds}-\overline{\vartheta}\sin\vartheta\frac{d\omega}{ds} & q_n&=\frac{2}{1+x^2+y^2}\left(\overline{y}\frac{dx}{ds}-\overline{x}\frac{dy}{ds}\right)\\ q_s&=\overline{\vartheta}\frac{d\vartheta}{ds}+\overline{\omega}\sin\vartheta\frac{d\omega}{ds} & q_s&=\frac{2}{1+x^2+y^2}\left(\overline{x}\frac{dx}{ds}+\overline{y}\frac{dy}{ds}\right)\ .\end{aligned} \tag{7}$$

Moreover, for an incompressible fluid it is possible to introduce the stream function $\psi=\int_0^P q_n ds$ and to thereby obtain, by (4) § 2 of chap. I:

$$\left.\begin{aligned}\overline{\vartheta}=\frac{d\vartheta}{dt}&=-\frac{1}{\sin\vartheta}\frac{\partial\psi}{\partial\omega}\ ,\ \overline{x}=\frac{2}{1+x^2+y^2}\frac{dx}{dt}=-\frac{1}{2}\left(1+x^2+y^2\right)\frac{\partial\psi}{\partial y}\\ \overline{\omega}=\sin\vartheta\frac{d\omega}{dt}&=\frac{\partial\psi}{\partial\vartheta}\ ,\quad \overline{y}=\frac{2}{1+x^2+y^2}\frac{dy}{dt}=\ \ \frac{1}{2}\left(1+x^2+y^2\right)\frac{\partial\psi}{\partial x}\ ,\end{aligned}\right\} \tag{8}$$

$$2\varrho = D\psi = \frac{1}{\sin\vartheta}\left\{\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial\psi}{\partial\vartheta}\right) + \frac{\partial}{\partial\omega}\left(\frac{1}{\sin\vartheta}\frac{\partial\psi}{\partial\omega}\right)\right\}$$
$$= \frac{\partial^2\psi}{\partial\vartheta^2} + \cot\vartheta\frac{\partial\psi}{\partial\vartheta} + \frac{1}{\sin^2\vartheta}\frac{\partial^2\psi}{\partial\omega^2} \qquad \text{nach (4) S. 210.} \tag{9a}$$

$$2\varrho = D\psi = \tfrac{1}{4}\left(1+x^2+y^2\right)^2\left(\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{dy^2}\right) = \tfrac{1}{4}\left(1+x^2+y^2\right)^2\Delta\psi\ . \tag{9b}$$
$$\Delta\psi = \frac{8\varrho}{\left(1+x^2+y^2\right)^2}\ .$$

§ 2. Der einfache Strudel und das sphärische Potential.

Die Gleichung (9a) in § 1 suchen wir zunächst zu befriedigen unter der Bedingung, daß ψ und ϱ *Funktionen von* ϑ *allein* sind, also nach (8)

222 | $$\overline{\vartheta} = 0\ , \quad \overline{\omega} = \frac{d\psi}{d\vartheta} = \psi'(\vartheta)\ ,$$

und die ganze Strömung symmetrisch um die Achse CO in den Parallelkreisen erfolgt. Eine solche Geschwindigkeitsverteilung wollen wir im Folgenden immer als eine „zonale“ oder „axialsymmetrische“ bezeichnen. Die genannte Gleichung nimmt dann die Form an:

$$\frac{d}{d\vartheta}\left(\sin\vartheta\frac{d\psi}{d\vartheta}\right) = 2\varrho\sin\vartheta \tag{1}$$

und ihr wird für den besonderen Fall $\varrho = \text{const.}$ genügt durch den Ansatz:

$$\psi = -4\varrho\ \lg\ \sin\frac{\vartheta}{2}\ , \quad \overline{\omega} = \frac{d\psi}{d\vartheta} = -2\varrho\ \cot\ \frac{\vartheta}{2} \tag{2}$$
$$\overline{\vartheta} = 0\ .$$

Diese spezielle Lösung besitzt im Anfangspunkte O ($\vartheta = 0$) einen singulären Punkt, in welchem ψ und $\overline{\omega}$ unendlich werden. Bilden wir nun die Zirkulation (I § 3) im Parallelkreise ϑ:

$$2R_\vartheta = \int_0^{2\pi}\overline{\omega}\sin\vartheta d\omega = 2\pi\overline{\omega}_\vartheta\sin\vartheta = -4\pi\varrho\left(1+\cos\vartheta\right)\ ,$$

so nimmt diese für $\vartheta = 0$ den Wert an

$$2R_0 = -8\pi\varrho\ ,$$

also ist

$$R_0 = -4\pi\varrho = m$$

$$2\varrho = D\psi = \frac{1}{\sin\vartheta}\left\{\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial\psi}{\partial\vartheta}\right)+\frac{\partial}{\partial\omega}\left(\frac{1}{\sin\vartheta}\frac{\partial\psi}{\partial\omega}\right)\right\}$$
$$= \frac{\partial^2\psi}{\partial\vartheta^2}+\cot\vartheta\frac{\partial\psi}{\partial\vartheta}+\frac{1}{\sin^2\vartheta}\frac{\partial^2\psi}{\partial\omega^2} \qquad \text{by (4) p. 210.} \tag{9a}$$

$$2\varrho = D\psi = \tfrac{1}{4}\left(1+x^2+y^2\right)^2\left(\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{dy^2}\right) = \tfrac{1}{4}\left(1+x^2+y^2\right)^2\Delta\psi\ . \tag{9b}$$
$$\Delta\psi = \frac{8\varrho}{(1+x^2+y^2)^2}\ .$$

§ 2. The simple whirl and the spherical potential.

We first seek to satisfy the equation (9a) in § 1 under the assumption that ψ and ϱ are *functions of* ϑ *alone*, and hence, by (8),

$$\overline{\vartheta} = 0\ , \quad \overline{\omega} = \frac{d\psi}{d\vartheta} = \psi'(\vartheta)\ ,$$

and that the entire flow takes place symmetrically about the axis CO in the parallel circles. In what follows, a velocity distribution of this kind will be called "zonal", or "axially symmetric". The said equation then assumes the form:

$$\frac{d}{d\vartheta}\left(\sin\vartheta\frac{d\psi}{d\vartheta}\right) = 2\varrho\sin\vartheta \tag{1}$$

and for the special case where $\varrho = \text{const.}$ it is satisfied by the ansatz:

$$\psi = -4\varrho\ \lg\ \sin\frac{\vartheta}{2}\ , \quad \overline{\omega} = \frac{d\psi}{d\vartheta} = -2\varrho\ \cot\ \frac{\vartheta}{2} \tag{2}$$
$$\overline{\vartheta} = 0\ .$$

This special solution has a singular point in the initial point O ($\vartheta = 0$) in which ψ and $\overline{\omega}$ become infinite. If we now form the circulation (I § 3) in the parallel circle ϑ:

$$2R_\vartheta = \int_0^{2\pi}\overline{\omega}\sin\vartheta d\omega = 2\pi\overline{\omega}_\vartheta\sin\vartheta = -4\pi\varrho\left(1+\cos\vartheta\right)\ ,$$

then for $\vartheta = 0$ it assumes the value

$$2R_0 = -8\pi\varrho\ ,$$

and hence we have

$$R_0 = -4\pi\varrho = m$$

das Moment des Strudelpunktes O (cf. Kap. I, § 2 S. 213), so daß in der That gemäß Satz II S. 209 die Summe aller Wirbelmomente auf der ganzen Kugel $= 4\pi\varrho - R_0 = 0$ ist.

Satz I. *Jede zonale Strömung in der Kugel, bei welcher, abgesehen von einem einzigen Strudelpunkte m in O die Wirbeldichte auf der ganzen Kugel konstant $= -\frac{m}{4\pi}$ ist, soll als ein „*einfacher Strudel*“ vom Momente m bezeichnet werden und wird dargestellt durch die Stromfunktion*

$$\psi = \frac{m}{\pi} \log \sin \frac{\vartheta}{2} ,$$

wenn ϑ den Bogenabstand vom Strudelpunkte O, dem „Zentrum“ des Strudels, bedeutet.

Hier strömt die Flüssigkeit gleichförmig auf den Parallelkreisen und um so schneller, je näher sie dem Strudelpunkt O sind, um diesen selbst mit unendlicher Geschwindigkeit, während sie in dem Gegenpole O' ganz in Ruhe bleibt.

Einen solchen „Strudel“ können wir uns in einer physikalischen Flüssigkeit,
223 welche keine wirkliche Unstetigkeit zuläßt, angenähert rea- | lisiert denken durch geeignete Zusammensetzung der betrachteten Lösung mit irgend einer anderen Lösung von (1), z. B.

$$\psi = \psi_1 = -\alpha \cos \vartheta , \quad \overline{\omega} = \alpha \sin \vartheta , \quad \varrho = \alpha \cos \vartheta = -\psi_1 ,$$

welche einer starren Rotation der ganzen Flüssigkeit um die Achse $O'O$ mit der Winkelgeschwindigkeit α entspricht. Wir könnten dann auf einem beliebigen Parallelkreise $\vartheta = \vartheta_0$ beide Strömungen ψ_0 und ψ_1 stetig zusammenfügen, indem wir die entsprechenden Geschwindigkeiten gleich setzten:

$$\overline{\omega} = \alpha \sin \vartheta_0 = \frac{m}{2\pi} \cot \frac{\vartheta_0}{2} ,$$

also:

$$\alpha = \frac{m}{4\pi} \frac{1}{\sin^2 \frac{\vartheta_0}{2}} .$$

Dann wäre die zusammengesetzte Strömung gegeben durch:

$$\psi = \psi_1 \quad \left(\vartheta < \vartheta^0\right) , \quad \psi = \psi_0 = \frac{m}{\pi} \operatorname{lg} \sin \frac{\vartheta}{2} \quad (\vartheta > \vartheta_0)$$

und würde mit der des einfachen Strudels umsomehr übereinstimmen, je kleiner ϑ_0 gewählt würde, und doch würde die Geschwindigkeit überall stetig bleiben, solange noch ϑ_0 von 0 verschieden ist.

the momentum of the whirl point O (cf. chap. I, § 2 p. 213) so that the sum of all whirl momenta on the entire sphere is indeed $= 4\pi\varrho - R_0 = 0$ in accordance with Theorem II p. 209.

Theorem I. *Every zonal flow in the sphere for which the vortex density on the entire sphere is constant* $= -\dfrac{m}{4\pi}$ *except for a single whirl point* m *in* O *shall be called a* "simple whirl" *of momentum* m*, and it is represented by the stream function*

$$\psi = \frac{m}{\pi} \log \sin \frac{\vartheta}{2} \ ,$$

where ϑ *denotes the arc distance from the whirl point* O*, the* "*center*" *of the whirl.*

In this case, the fluid flows uniformly on the parallel circles, and it flows more quickly the closer the circles are to the whirl point O. Around the whirl point, it flows at infinite velocity, while it is completely at rest at the antipole O'.

We can imagine such a "whirl" being approximately realized in a physical fluid, which permits no real discontinuity, by a suitable combination of the solution under consideration with some other solution of (1), e. g.,

$$\psi = \psi_1 = -\alpha \cos \vartheta \ , \quad \overline{\omega} = \alpha \sin \vartheta \ , \quad \varrho = \alpha \cos \vartheta = -\psi_1 \ ,$$

which corresponds to a rigid rotation of the entire fluid about the axis $O'O$ at angular velocity α. We then could continuously combine the two flows ψ_0 and ψ_1 on some parallel circle $\vartheta = \vartheta_0$ by setting the corresponding velocities equal to one another:

$$\overline{\omega} = \alpha \sin \vartheta_0 = \frac{m}{2\pi} \cot \frac{\vartheta_0}{2} \ ,$$

and hence:

$$\alpha = \frac{m}{4\pi} \frac{1}{\sin^2 \frac{\vartheta_0}{2}} \ .$$

The composite flow would then be given by:

$$\psi = \psi_1 \quad (\vartheta < \vartheta_0) \ ,^{7} \quad \psi = \psi_0 = \frac{m}{\pi} \lg \sin \frac{\vartheta}{2} \quad (\vartheta > \vartheta_0)$$

and the smaller we choose ϑ_0, the more the composite flow would agree with that of the simple whirl. Nevertheless, the velocity would be continuous everywhere so long as ϑ_0 is still different from 0.

[7] ⟦Zermelo erroneously writes "$\vartheta < \vartheta^0$" instead of "$\vartheta < \vartheta_0$".⟧

Nun war aber der Anfangspunkt O ursprünglich ein beliebiger Punkt auf der Kugel, der keine ausgezeichnete Eigenschaft besitzt. Wir können daher auch jeden beliebigen anderen Punkt P_0 zum Mittelpunkte eines „Strudels“ machen, wenn wir setzen:

$$\psi = \frac{m}{\pi} \lg \sin \frac{\delta_0}{2} = \frac{m}{\pi} \lg \frac{r_0}{2} ,$$

wo unter δ_0 der sphärische und unter r_0 der Sehnenabstand eines beliebigen Kugelpunktes P von P_0 verstanden werden soll. Dabei können wir immer noch ein beliebiges Polar- oder stereographisches Koordinatensystem zu grunde legen, wenn wir nur δ_0 oder r_0 richtig durch die Koordinaten ϑ, ω oder x, y von P ausdrücken. Beachten wir nun weiter, daß unsere Grundgleichung $D\psi = 2\varrho$ in den Variabeln ϱ, ψ und ihren Ableitungen linear und homogen ist, daß wir also aus zwei Lösungen ψ_1, ϱ_1 und ψ_2, ϱ_2 immer neue Lösungen $c_1\psi_1 + c_2\psi_2$, $c_1\varrho_1 + c_2\varrho_2$ linear zusammensetzen können, so gelangen wir zu dem Satze:

Satz II. *Bezeichnen wir mit $\delta_1, \delta_2, .. \delta_n$ die sphärischen und mit $r_1, r_2, .. r_n$ die Sehnenabstände des variablen Punktes P von n festen Kugelpunkten $P_1, P_2, .. P_n$, so stellt die Funktion:*

$$\begin{aligned} \psi &= \frac{m_1}{\pi} \lg \sin \frac{\delta_1}{2} + \frac{m_2}{\pi} \lg \sin \frac{\delta_2}{2} + \cdots + \frac{m_n}{\pi} \lg \sin \frac{\delta_n}{2} \\ &= \frac{m_1}{\pi} \lg \frac{r_1}{2} + \frac{m_2}{\pi} \lg \frac{r_2}{2} + \cdots + \frac{m_n}{\pi} \lg \frac{r_n}{2} , \end{aligned} \tag{4}$$

224 | *welche das „sphärische Potential“ der n Massen $m_1, m_2, .. m_n$ genannt werde, die Stromfunktion einer bestimmten Strömung in der Kugel dar, in welcher die Punkte $P_1, P_2, .. P_n$ Strudelpunkte mit den Wirbelmomenten $m_1, m_2, .. m_n$ sind und welche in jedem anderen Kugelpunkte P eine* konstante *Wirbeldichte*

$$\varrho = \varrho_1 + \varrho_2 + \cdots \varrho_n = -\frac{m_1}{4\pi} - \frac{m_2}{4\pi} \cdots - \frac{m_n}{4\pi} = -\frac{1}{4\pi} \sum m = -\frac{M}{4\pi} \tag{5}$$

besitzt. Jede solche Strömung bezeichnen wir als ein „Strudelsystem“.

Die konstante Wirbeldichte ϱ nimmt den Wert 0 an, wenn die Summe $M = \sum m$ aller Strudelmomente verschwindet.

Besteht beispielsweise das System nur aus zwei gleichen und entgegengesetzten Strudeln $\pm m$, so wird seine Stromfunktion:

$$\psi = \frac{m}{\pi} \lg \frac{\sin \frac{\delta_1}{2}}{\sin \frac{\delta_2}{2}} = \frac{m}{\pi} \lg \frac{r_1}{r_2} , \tag{5a}$$

und die Stromlinien werden gegeben durch:

$$\frac{r_1}{r_2} = \text{const} ,$$

Originally, however, the initial point O was any point on the sphere with no distinguished property. Hence, we can make any other point P_0 the center of a "whirl" by setting:

$$\psi = \frac{m}{\pi} \lg \sin \frac{\delta_0}{2} = \frac{m}{\pi} \lg \frac{r_0}{2} ,$$

where δ_0 and r_0 denote the spherical distance and the chord distance respectively of any circle point P from P_0. Here we can still use any polar or stereographic coordinate system, provided only that we correctly express δ_0 and r_0 in terms of the coordinates ϑ, ω and x, y of P respectively. Furthermore, if we now consider that our basic equation $D\psi = 2\varrho$ is linear and homogeneous in the variables ϱ, ψ and its derivatives, and hence that we can always obtain new solutions $c_1\psi_1 + c_2\psi_2$, $c_1\varrho_1 + c_2\varrho_2$ by linearly combining two solutions ψ_1, ϱ_1 and ψ_2, ϱ_2, then we get the theorem:

Theorem II. *If we denote the spherical distance and the chord distance of the variable point P from n fixed circle points $P_1, P_2, .. P_n$ by $\delta_1, \delta_2, .. \delta_n$ and $r_1, r_2, .. r_n$ respectively, then the function:*

$$\begin{aligned} \psi &= \frac{m_1}{\pi} \lg \sin \frac{\delta_1}{2} + \frac{m_2}{\pi} \lg \sin \frac{\delta_2}{2} + \cdots + \frac{m_n}{\pi} \lg \sin \frac{\delta_n}{2} \\ &= \frac{m_1}{\pi} \lg \frac{r_1}{2} + \frac{m_2}{\pi} \lg \frac{r_2}{2} + \cdots + \frac{m_n}{\pi} \lg \frac{r_n}{2} , \end{aligned} \tag{4}$$

which shall be called the "spherical potential" of the n masses $m_1, m_2, .. m_n$, represents the stream function of a particular flow in the sphere, in which the points $P_1, P_2, .. P_n$ are whirl points with vorticity moments $m_1, m_2, .. m_n$ and which in any other sphere point P possesses a constant *vortex density*

$$\varrho = \varrho_1 + \varrho_2 + \cdots \varrho_n = -\frac{m_1}{4\pi} - \frac{m_2}{4\pi} \cdots - \frac{m_n}{4\pi} = -\frac{1}{4\pi} \sum m = -\frac{M}{4\pi} . \tag{5}$$

Every flow of this kind is called a "whirl system".

The constant vortex density ϱ assumes the value 0 when the sum $M = \sum m$ of all whirl moments vanishes.

If, for example, the system only consists of two equal and antipodal whirls $\pm m$, then its stream function becomes:

$$\psi = \frac{m}{\pi} \lg \frac{\sin \frac{\delta_1}{2}}{\sin \frac{\delta_2}{2}} = \frac{m}{\pi} \lg \frac{r_1}{r_2} , \tag{5a}$$

and the streamlines are given by:

$$\frac{r_1}{r_2} = \text{const} ,$$

d. h. es sind die Kugelkreise, deren Ebenen durch die Schnittlinie der in P_1 und P_2 berührenden Tangentialebenen gehen.[1]

Liegen die beiden Strudel insbesondere in zwei entgegengesetzten Punkten O, O' ($\vartheta = 0$, $\vartheta = \pi$), so wird die Stromfunktion:

$$\psi = \frac{m}{\pi} \operatorname{lg} \operatorname{tg} \frac{\vartheta}{2} \tag{5b}$$

und die Geschwindigkeit:

$$\overline{\vartheta} = 0 \,, \quad \overline{\omega} = \frac{m}{\pi} \frac{1}{\sin \vartheta} \,.$$

Die Strömung erfolgt in den Parallelkreisen mit einer Geschwindigkeit umgekehrt proportional ihren Radien.

Das sphärische Potential, das wir oben für eine endliche Anzahl von Massenpunkten definiert hatten, wollen wir jetzt auf den Fall einer kontinuierlichen Massenverteilung auf der Kugelfläche ausdehnen. Es sei nämlich $k = k(\vartheta, \omega)$ eine wenigstens stückweise stetige Funktion des Ortes auf der Kugel, und es werde gesetzt:

$$M = \int k d\sigma \,, \tag{6}$$

wo die Integration über die ganze Kugel erstreckt wird. Nun zerlegen wir die
225 Kugelfläche in eine Anzahl von Teilbereichen $\sigma_1, \sigma_2, \sigma_3 \ldots$ | und bezeichnen mit $k_1, k_2, k_3 \ldots$ die entsprechenden arithmetischen Mittelwerte der Funktion k, so daß

$$k_1\sigma_1 = \int^{(\sigma_1)} k d\sigma \,, \quad k_2\sigma_2 = \int^{(\sigma_2)} k d\sigma \ldots \tag{7}$$

Ferner nehmen wir in jedem der Bereiche $\sigma_1, \sigma_2, \ldots$ einen festen Punkt $P_1, P_2, \ldots$ an und bezeichnen mit $r_1, r_2, \ldots$ die von diesen Punkten aus gerechneten Sehnenabstände eines variablen Punktes P. Jetzt betrachten wir das Strudelsystem (Satz II, S. 223) mit der Stromfunktion

$$\psi' = \frac{k_1\sigma_1}{\pi} \operatorname{lg} \frac{r_1}{2} + \frac{k_2\sigma_2}{\pi} \operatorname{lg} \frac{r_2}{2} + \cdots \,, \tag{8}$$

für welches die konstante Wirbeldichte, abgesehen von den Strudelpunkten $P_1, P_2, \ldots$, den Wert hat:

$$\varrho' = -\frac{k_1\sigma_1}{4\pi} - \frac{k_2\sigma_2}{4\pi} \cdots = -\frac{M}{4\pi} \,, \tag{9}$$

[1] Vergl. *Lamb*, Hydrodynamics, p. 115 und p. 253.

i. e., they are the circles on the sphere whose planes pass through the line of intersection of the tangent planes making contact at P_1 and P_2.[8]

In particular, if the two whirls lie at two antipodal points O, O' ($\vartheta = 0$, $\vartheta = \pi$), then the stream function becomes:

$$\psi = \frac{m}{\pi} \lg \operatorname{tg} \frac{\vartheta}{2} \tag{5b}$$

and the velocity becomes:

$$\overline{\vartheta} = 0 \,, \quad \overline{\omega} = \frac{m}{\pi} \frac{1}{\sin \vartheta} \,.$$

The flow occurs in the parallel circles at a velocity inversely proportional to their radii.

We shall now extend the spherical potential, which we have defined above for a finite number of mass points, to the case of a continuous mass distribution on the surface of the sphere. For let $k = k\,(\vartheta, \omega)$ be an at least piecewise continuous function of the location on the sphere and set:

$$M = \int k d\sigma \,, \tag{6}$$

where the integration is extended over the entire sphere. We now divide the spherical surface into a number of partial domains $\sigma_1, \sigma_2, \sigma_3 \ldots$ and denote the corresponding arithmetic mean values of the function k by $k_1, k_2, k_3 \ldots$ so that

$$k_1\sigma_1 = \int^{(\sigma_1)} k d\sigma \,, \quad k_2\sigma_2 = \int^{(\sigma_2)} k d\sigma \ldots \tag{7}$$

Furthermore, we take a fixed point $P_1, P_2, \ldots$ in each of the domains $\sigma_1, \sigma_2, \ldots$ and denote the chord distances of a variable point P reckoned from these points by $r_1, r_2, \ldots$. We now consider the system of whirls (Theorem II, p. 223) with the stream function

$$\psi' = \frac{k_1\sigma_1}{\pi} \lg \frac{r_1}{2} + \frac{k_2\sigma_2}{\pi} \lg \frac{r_2}{2} + \cdots \,, \tag{8}$$

for which, except for the whirl points $P_1, P_2, \ldots$, the constant vortex density has the value:

$$\varrho' = -\frac{k_1\sigma_1}{4\pi} - \frac{k_2\sigma_2}{4\pi} \cdots = -\frac{M}{4\pi} \,, \tag{9}$$

[8] Cf. *Lamb 1895*, p. 115 and p. 253.

und bilden die halbe Zirkulation

$$R_\sigma = \frac{1}{2}\int^{(\mathfrak{S})} q_s ds = \frac{1}{2}\int^{(\mathfrak{S})} \frac{\partial \psi'}{\partial n} ds$$

über die Begrenzung $\mathfrak{S}$ eines Aggregates σ solcher Teilgebiete $\sigma_1, \sigma_2, \ldots$ Diese wird dann gleich der Summe der im Innern von σ enthaltenen Wirbelmomente, also

$$\begin{aligned} R_\sigma &= k_1\sigma_1 + k_2\sigma_2 + \cdots + \varrho'\sigma \\ &= \int^{(\sigma_1)} k d\sigma + \int^{(\sigma_2)} k d\sigma + \cdots + \varrho'\sigma \\ &= \int^{(\sigma)} k d\sigma + \varrho'\sigma = \int^{(\sigma)} (k + \varrho') d\sigma = \int^{(\sigma)} \left(k - \frac{M}{4\pi}\right) d\sigma \, . \end{aligned}$$

Lassen wir jetzt die Dimensionen aller Teilgebiete $\sigma_1, \sigma_2 \ldots$ unbegrenzt abnehmen, während ihre Anzahl unbegrenzt wächst, so werden wir schließlich *jedes* Flächenstück C als ein solches Aggregat σ auffassen können und erhalten so:

$$R_C = \int^{(C)} \left(k - \frac{M}{4\pi}\right) d\sigma \, , \tag{10}$$

d. h. die Wirbeldichte der betrachteten Strömung hat an der Grenze den Wert:

$$\varrho = k(\vartheta, \omega) - \frac{M}{4\pi} \, . \tag{11}$$

Gleichzeitig geht aber in (8) die endliche Summe in ein bestimmtes Integral über:

$$\psi = \lim \psi' = \frac{1}{\pi}\int^{(K)} k \, \lg \, \frac{r}{2} d\sigma = \frac{1}{\pi}\int^{(K)} k \, \lg \, \sin\frac{\delta}{2} d\sigma \, , \tag{12}$$

226 | wo die Integration sich auf die ganze Kugelfläche K bezieht und die Buchstaben r und δ den Sehnen- und den Bogenabstand eines variablen Punktes P von dem betreffenden Elemente $d\sigma$ ausdrücken sollen.

So erhalten wir:

Satz III. *Das sphärische Potential*

$$\psi = \frac{1}{\pi}\int^{(K)} k \, \lg \, \sin\frac{\delta}{2} d\sigma$$

and we form the half-circulation

$$R_\sigma = \frac{1}{2}\int^{(\mathfrak{S})} q_s ds = \frac{1}{2}\int^{(\mathfrak{S})} \frac{\partial \psi'}{\partial n} ds$$

over the boundary $\mathfrak{S}$ of an aggregate σ of such partial domains $\sigma_1, \sigma_2, \ldots$. [[The half-circulation]] then becomes equal to sum of the vorticity moments contained in the interior of σ, and hence

$$\begin{aligned} R_\sigma &= k_1\sigma_1 + k_2\sigma_2 + \cdots + \varrho'\sigma \\ &= \int^{(\sigma_1)} k d\sigma + \int^{(\sigma_2)} k d\sigma + \cdots + \varrho'\sigma \\ &= \int^{(\sigma)} k d\sigma + \varrho'\sigma = \int^{(\sigma)} (k + \varrho') d\sigma = \int^{(\sigma)} \left(k - \frac{M}{4\pi}\right) d\sigma \ . \end{aligned}$$

If we now allow the dimensions of all partial domains $\sigma_1, \sigma_2 \ldots$ to decrease indefinitely while its number indefinitely increases, then we can eventually conceive of *every* part C of the surface as such an aggregate σ, thereby obtaining:

$$R_C = \int^{(C)} \left(k - \frac{M}{4\pi}\right) d\sigma \ , \tag{10}$$

i. e., at the limit, the vortex density of the flow under consideration has the value:

$$\varrho = k(\vartheta, \omega) - \frac{M}{4\pi} \ . \tag{11}$$

At the same time, however, the finite sum in (8) is transformed into a definite integral:

$$\psi = \lim \psi' = \frac{1}{\pi}\int^{(K)} k \ \lg \frac{r}{2} d\sigma = \frac{1}{\pi}\int^{(K)} k \ \lg \ \sin \frac{\delta}{2} d\sigma \ , \tag{12}$$

where the integration refers to the entire spherical surface K, and where the letters r and δ denote the chord and the arc distance, respectively, of a variable point P from the respective element $d\sigma$.

We thus obtain:

Theorem III. *The spherical potential*

$$\psi = \frac{1}{\pi}\int^{(K)} k \ \lg \ \sin \frac{\delta}{2} d\sigma$$

einer kontinuierlichen Massenverteilung auf der Kugelfläche mit der variablen Dichte $k = k(\vartheta, \omega)$ und der Gesamtmasse M stellt, als Stromfunktion betrachtet, eine Flüssigkeitsströmung dar, deren Wirbeldichte ϱ an jeder Stelle den Wert hat

$$\varrho = k(\vartheta, \omega) - \frac{M}{4\pi} . \tag{11}$$

Ist daher $M = \overset{(K)}{\int} k d\sigma = 0$, so wird $\varrho = k$ und somit:

$$\psi = \frac{1}{\pi} \overset{(K)}{\int} \varrho \lg \frac{r}{2} d\sigma = \frac{1}{\pi} \overset{(K)}{\int} \varrho \lg \sin \frac{\delta}{2} d\sigma , \tag{13}$$

während gleichzeitig nach (4) S. 210

$$D\psi = 2\varrho \tag{14}$$

sein muß. Die Gleichung (13) stellt also bei vorgeschriebenem $\varrho = \varrho(\vartheta, \omega)$ eine Auflösung der Differentialgleichung (14) dar und zwar (nach Satz III, S. 211) bis auf eine additive Konstante die einzig mögliche, falls ϱ überall endlich bleibt. Sind dagegen außer der kontinuierlichen Wirbelverteilung ϱ noch Strudelpunkte $m_1, m_2, \ldots$ oder Strudellinien mit den Längendichten $\gamma_1, \gamma_2, \ldots$ vorhanden, so hat man noch die entsprechenden sphärischen Potentiale hinzuzufügen:

$$\frac{m_1}{\pi} \lg \sin \frac{\delta_1}{2} + \frac{m_2}{\pi} \lg \sin \frac{\delta_2}{2} + \cdots + \int \gamma_1 \lg \sin \frac{\delta}{2} ds_1 + \cdots$$

und erhält so auf jeden Fall die Stromfunktion eindeutig (bis auf eine additive Konstante), wenn nur, dem Satze II § 3 entsprechend, die Summe *aller* Wirbelmomente verschwindet. Wir haben also den weiteren Satz:

Satz III. *Ist von einer Strömung einer inkompressiblen Flüssigkeit auf der unbegrenzten Kugelfläche die gesamte Wirbelverteilung gegeben, so ist die zugehörige Stromfunktion (bis auf eine additive Konstante) gleich dem sphärischen Potential der entsprechenden Massenverteilung.*

227 Das hier betrachtete sphärische Potential verhält sich auf der | Kugel ganz analog dem gewöhnlichen *Newton*schen Potential im Raume oder dem logarithmischen in der Ebene. Namentlich gilt auch hier das Theorem:

Satz IV. *Das sphärische Potential einer zonalen, d. h. aus homogenen konzentrischen Ringen bestehenden Massenverteilung auf einer Kalotte C bleibt für alle äußeren Punkte (in der restierenden Kalotte C') bis auf eine additive Konstante ungeändert, wenn man alle wirkenden Massen M in ihrem (inneren) Mittelpunkte O vereinigt.*

of a continuous mass distribution on the spherical surface with variable density $k = k(\vartheta, \omega)$ *and the total mass* M *represents, when considered as a stream function, a flow of fluid whose vortex density* ϱ *everywhere has the value*

$$\varrho = k(\vartheta, \omega) - \frac{M}{4\pi} . \tag{11}$$

Therefore, if we have $M = \overset{(K)}{\int} k d\sigma = 0$, then we have $\varrho = k$, and hence:

$$\psi = \frac{1}{\pi} \overset{(K)}{\int} \varrho \lg \frac{r}{2} d\sigma = \frac{1}{\pi} \overset{(K)}{\int} \varrho \lg \sin \frac{\delta}{2} d\sigma , \tag{13}$$

while, by (4) p. 210, we must also have

$$D\psi = 2\varrho . \tag{14}$$

Thus, equation (13) constitutes a solution of the differential equation (14) for prescribed $\varrho = \varrho(\vartheta, \omega)$, and in particular (by Theorem III, p. 211), the only possible one, apart from an additive constant, in case ϱ remains finite everywhere. If, however, there are, besides the continuous vortex distribution ϱ, also whirl points $m_1, m_2, \ldots$ or whirl lines with length densities $\gamma_1, \gamma_2, \ldots$, then we must also add the corresponding spherical potentials:

$$\frac{m_1}{\pi} \lg \sin \frac{\delta_1}{2} + \frac{m_2}{\pi} \lg \sin \frac{\delta_2}{2} + \cdots + \int \gamma_1 \lg \sin \frac{\delta}{2} ds_1 + \cdots$$

and the stream function thus obtained is in any event single-valued (apart from an additive constant), if only, in accordance with Theorem II § 3, the sum of *all* vorticity moments vanishes. We thus have the further theorem:

Theorem III.[9] *If, for a flow of incompressible fluid on the unbounded spherical surface, the entire vortex distribution is given, then the associated stream function* (*apart from an additive constant*) *is equal to the spherical potential of the corresponding mass distribution.*

The spherical potential considered here behaves on the sphere analogously to the common *Newton*'s potential in space, or to the logarithmic potential in the plane. In particular, the following theorem also holds here:

Theorem IV. *The spherical potential of a zonal mass distribution, i. e., one consisting of homogeneous, concentric rings, on a calotte* C *remains unaltered for all outer points* (*in the remaining calotte* C') *apart from an additive constant, if all acting masses* M *are united in its* (*inner*) *center* O.

[9] ⟦Zermelo should have continued with "Theorem IV".⟧

Es sei nämlich $\varrho = \varrho(\vartheta)$ $(\vartheta < \alpha)$ die Dichte der ursprünglichen Massenverteilung und $\psi = \psi(\vartheta)$ ihr sphärisches Potential, und es sei ferner $\psi_1 = \frac{M}{\pi} \lg \sin\frac{\vartheta}{2} = \psi_1(\vartheta)$ das Potential der in O befindlichen Masse M. Dann wird $\psi_0 = \psi - \psi_1$ das sphärische Potential der Massenverteilung mit der vorgeschriebenen Dichte ϱ in C und dem *hinzukommenden* Massenpunkte $-M$ in O. Hier ist natürlich die Summe aller Massen $= 0$, und wir können sie deshalb auch als die Wirbel einer Flüssigkeitsströmung auffassen, welche in O den Strudelpunkt $-M$, innerhalb C die variable Wirbeldichte ϱ und in C' gar keine Wirbel mehr besitzt. Diese Strömung ist ebenfalls zonal und hat die Stromfunktion $\psi_0 = \psi - \psi_1$, und da sie in dem einfach zusammenhängenden und von einer Stromlinie begrenzten Bereiche C' wirbelfrei ist, so muß nach Satz III, S. 211 die Stromfunktion in diesem Bereiche *konstant* sein, d. h.:

$$\psi_0 = \psi - \psi_1 = \text{const.}\,, \quad \psi = \psi_1 + \text{const.} \qquad (\vartheta > \alpha)$$

q. e. d.

Den Wert der Konstanten bestimmt man, indem man das Potential ψ in O', dem Mittelpunkte von C', wo ψ_1 verschwindet, direkt berechnet:

$$\psi_{O'} = \frac{1}{\pi}\int \varrho d\sigma \lg \cos\frac{\vartheta}{2} = 2\int_0^\alpha \varrho \sin\vartheta d\vartheta \lg \cos\frac{\vartheta}{2}\,.$$

Ist z. B. $\varrho = \text{const.}$, d. h. ist unsere Kalotte C *homogen* mit Masse belegt, so wird

$$M = 2\pi\varrho(1-\cos\alpha) = 4\pi\varrho\sin^2\frac{\alpha}{2}$$

und

$$\psi_{O'} = 2\varrho\int_0^\alpha \lg \cos\frac{\vartheta}{2}\sin\vartheta d\vartheta = 2\varrho\left[\cos^2\frac{\alpha}{2}\left(1 - \lg\cos^2\frac{\alpha}{2}\right) - 1\right]\,,$$

also ist das Potential im Äußern (C'):

$$\psi_\alpha = 4\varrho\sin^2\frac{\alpha}{2}\lg\sin\frac{\vartheta}{2} - 2\varrho\left(\sin^2\frac{\alpha}{2} + \cos^2\frac{\alpha}{2}\lg\cos^2\frac{\alpha}{2}\right)\,. \qquad (15)$$

228 | Um aber das Potential im Innern von C $(\vartheta < \alpha)$ zu berechnen, denken wir uns zunächst die *ganze* Kugel homogen mit ϱ belegt und dann die Belegung mit der Dichte $-\varrho$ in der Kalotte C' hinzugefügt. Bei der ersteren Belegung ist natürlich das Potential überall konstant $\overline{\psi} = \psi_0$, also nach (15) für $\alpha = \vartheta = \pi$ $\overline{\psi} = -2\varrho$, und bei der zweiten ist C wieder die *äußere* Kalotte, also das Potential:

$$\psi' = -4\varrho\cos^2\frac{\alpha}{2}\lg\cos\frac{\vartheta}{2} + 2\varrho\left(\cos^2\frac{\alpha}{2} + \sin^2\frac{\alpha}{2}\lg\sin^2\frac{\alpha}{2}\right)\,, \quad (\vartheta < \alpha)$$

For let $\varrho = \varrho(\vartheta)$ $(\vartheta < \alpha)$ be the density of the original mass distribution, and $\psi = \psi(\vartheta)$ its spherical potential. Furthermore, let $\psi_1 = \dfrac{M}{\pi} \lg \sin \dfrac{\vartheta}{2} = \psi_1(\vartheta)$ be the potential of the mass M in O. Then $\psi_0 = \psi - \psi_1$ becomes the spherical potential of the mass distribution with the prescribed density ϱ in C and the *joining* mass point $-M$ in O. Of course, the sum of all the masses is $= 0$ here, and we can therefore also conceive of them as the vortices of a flow of fluid that has the whirl point $-M$ in O and the variable vortex density ϱ in C and that no longer has any vortices in C'. This flow is also zonal and has the stream function $\psi_0 = \psi - \psi_1$, and since it is vortex-free in the simply connected domain C' that is bounded by a stream line, by Theorem III, p. 211, the stream function must be *constant* in this domain, i. e.:

$$\psi_0 = \psi - \psi_1 = \text{const.}\,, \quad \psi = \psi_1 + \text{const.} \qquad (\vartheta > \alpha)$$

q. e. d.

We determine the value of the constant by directly computing the potential ψ in O', the center of C', where ψ_1 vanishes:

$$\psi_{O'} = \frac{1}{\pi} \int \varrho d\sigma \lg \cos \frac{\vartheta}{2} = 2 \int_0^\alpha \varrho \sin \vartheta d\vartheta \lg \cos \frac{\vartheta}{2}\,.$$

If, e. g., we have $\varrho = \text{const.}$, i. e. if our calotte C is *homogeneously* covered with mass, then we have

$$M = 2\pi\varrho(1 - \cos\alpha) = 4\pi\varrho \sin^2 \frac{\alpha}{2}$$

and

$$\psi_{O'} = 2\varrho \int_0^\alpha \lg \cos \frac{\vartheta}{2} \sin\vartheta d\vartheta = 2\varrho \left[\cos^2 \frac{\alpha}{2} \left(1 - \lg \cos^2 \frac{\alpha}{2}\right) - 1\right]\,,$$

and hence the potential in the exterior (C') is:

$$\psi_\alpha = 4\varrho \sin^2 \frac{\alpha}{2} \lg \sin \frac{\vartheta}{2} - 2\varrho \left(\sin^2 \frac{\alpha}{2} + \cos^2 \frac{\alpha}{2} \lg \cos^2 \frac{\alpha}{2}\right)\,. \qquad (15)$$

However, in order to calculate the potential in the interior of C $(\vartheta < \alpha)$, we first imagine that the *entire* sphere is homogeneously covered with ϱ and then that the covering with density $-\varrho$ in the calotte C' is added. For the first covering, the potential is, of course, constant everywhere $\overline{\psi} = \psi_0$, and hence, by (15), for $\alpha = \vartheta = \pi$ $\overline{\psi} = -2\varrho$. For the second covering, C is the *outer* calotte again, and hence the potential is:

$$\psi' = -4\varrho \cos^2 \frac{\alpha}{2} \lg \cos \frac{\vartheta}{2} + 2\varrho \left(\cos^2 \frac{\alpha}{2} + \sin^2 \frac{\alpha}{2} \lg \sin^2 \frac{\alpha}{2}\right)\,, \quad (\vartheta < \alpha)$$

und man erhält im Ganzen für das innere Potential:

$$\psi_i = \overline{\psi} + \psi' = -4\varrho\cos^2\frac{\alpha}{2}\,\lg\,\cos\frac{\vartheta}{2} - 2\varrho\sin^2\frac{\alpha}{2}\left(1 - \lg\,\sin^2\frac{\alpha}{2}\right) . \qquad (16)$$

Am Rande $\vartheta = \alpha$ werden beide Ausdrücke einander gleich:

$$\psi_\alpha = \psi_i = 2\varrho\left(\sin^2\frac{\alpha}{2}\,\lg\,\sin^2\frac{\alpha}{2} - \cos^2\frac{\alpha}{2}\,\lg\,\cos^2\frac{\alpha}{2} - \sin^2\frac{\alpha}{2}\right) .$$

§ 3. Die Erhaltung des Schwerpunktes.

Nach (9a) S. 221 ist für die Kugel:

$$2\varrho = D\psi = \frac{\partial^2\psi}{\partial\vartheta^2} + \cot\vartheta\frac{\partial\psi}{\partial\vartheta} + \frac{1}{\sin^2\vartheta}\frac{\partial^2\psi}{\partial\omega^2} ,$$

also:

$$2\varrho\sin\vartheta\frac{\partial\psi}{\partial\omega} = \frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial\psi}{\partial\vartheta}\frac{\partial\psi}{\partial\omega}\right) - \frac{1}{2}\frac{\partial}{\partial\omega}\left(\sin\vartheta\left(\frac{\partial\psi}{\partial\vartheta}\right)^2 - \frac{1}{\sin\vartheta}\left(\frac{\partial\psi}{\partial\omega}\right)^2\right) \qquad (1)$$

$$= \frac{\partial}{\partial\vartheta}\left(-\sin^2\vartheta\overline{\vartheta}\overline{\omega}\right) - \frac{1}{2}\frac{\partial}{\partial\omega}\left(\sin\vartheta\overline{\omega}^2 - \sin\vartheta\overline{\vartheta}^2\right)$$

und somit, wenn man über den Bereich C nach ϑ und ω integriert[1] und die rechte Seite in ein Randintegral über $\mathfrak{C}$ verwandelt:

$$\int^{(C)} 2\varrho\frac{\partial\psi}{\partial\omega}d\sigma = \int^{(\mathfrak{C})}\left(\frac{1}{2}\left(\overline{\omega}^2 - \overline{\vartheta}^2\right)\sin\vartheta d\vartheta - \overline{\vartheta}\overline{\omega}\sin^2\vartheta d\omega\right) . \qquad (2)$$

Wird hier der Bereich C über die ganze Kugelfläche K ausgedehnt, in welcher sich die Geschwindigkeit überall stetig ändern möge, so verschwindet die rechte Seite und es wird:

$$\int^{(K)} 2\varrho\frac{\partial\psi}{\partial\omega}d\sigma = 0 \qquad (2a)$$

oder, da nach (8) S. 221 $\dfrac{\partial\psi}{\partial\omega} = -\sin\vartheta\dfrac{d\vartheta}{dt} = \dfrac{d\cos\vartheta}{dt}$ ist und gleichzeitig $\dfrac{d}{dt}(\varrho d\sigma) = 0$:

229 |
$$\frac{d}{dt}\int^{(K)}\varrho d\sigma\cos\vartheta = 0$$

[1] Cf. *Poincaré* a. a. O. Nr. 65.

and, in all, we obtain for the potential in the interior:

$$\psi_i = \overline{\psi} + \psi' = -4\varrho \cos^2 \frac{\alpha}{2} \lg \cos \frac{\vartheta}{2} - 2\varrho \sin^2 \frac{\alpha}{2} \left(1 - \lg \sin^2 \frac{\alpha}{2}\right) . \qquad (16)$$

At the boundary $\vartheta = \alpha$ the two expressions become equal to each other:

$$\psi_\alpha = \psi_i = 2\varrho \left(\sin^2 \frac{\alpha}{2} \lg \sin^2 \frac{\alpha}{2} - \cos^2 \frac{\alpha}{2} \lg \cos^2 \frac{\alpha}{2} - \sin^2 \frac{\alpha}{2}\right) .$$

§ 3. The conservation of the center of gravity.

By (9a) p. 221, we have for the sphere:

$$2\varrho = D\psi = \frac{\partial^2 \psi}{\partial \vartheta^2} + \cot \vartheta \frac{\partial \psi}{\partial \vartheta} + \frac{1}{\sin^2 \vartheta} \frac{\partial^2 \psi}{\partial \omega^2} ,$$

therefore:

$$2\varrho \sin \vartheta \frac{\partial \psi}{\partial \omega} = \frac{\partial}{\partial \vartheta} \left(\sin \vartheta \frac{\partial \psi}{\partial \vartheta} \frac{\partial \psi}{\partial \omega}\right) - \frac{1}{2} \frac{\partial}{\partial \omega} \left(\sin \vartheta \left(\frac{\partial \psi}{\partial \vartheta}\right)^2 - \frac{1}{\sin \vartheta} \left(\frac{\partial \psi}{\partial \omega}\right)^2\right) \qquad (1)$$

$$= \frac{\partial}{\partial \vartheta} \left(-\sin^2 \vartheta \overline{\vartheta} \overline{\omega}\right) - \frac{1}{2} \frac{\partial}{\partial \omega} \left(\sin \vartheta \overline{\omega}^2 - \sin \vartheta \overline{\vartheta}^2\right)$$

and hence, if we integrate over the domain C with respect to ϑ and ω[10] and transform the right side into a boundary integral over $\mathfrak{C}$:

$$\int^{(C)} 2\varrho \frac{\partial \psi}{\partial \omega} d\sigma = \int^{(\mathfrak{C})} \left(\frac{1}{2} \left(\overline{\omega}^2 - \overline{\vartheta}^2\right) \sin \vartheta d\vartheta - \overline{\vartheta} \overline{\omega} \sin^2 \vartheta d\omega\right) . \qquad (2)$$

If we extend the domain C over the entire spherical surface K, in which the velocity shall continuously vary everywhere, then the right side vanishes, and we get:

$$\int^{(K)} 2\varrho \frac{\partial \psi}{\partial \omega} d\sigma = 0 \qquad (2a)$$

or, since, by (8) p. 221, we have $\frac{\partial \psi}{\partial \omega} = -\sin \vartheta \frac{d\vartheta}{dt} = \frac{d \cos \vartheta}{dt}$ and also $\frac{d}{dt}(\varrho d\sigma) = 0$:

$$\frac{d}{dt} \int^{(K)} \varrho d\sigma \cos \vartheta = 0$$

[10] Cf. *Poincaré 1893a*, no. 65.

oder

$$L_0 = \int^{(K)} \varrho d\sigma \cos\vartheta = \text{const.}$$

Nun ist $\cos\vartheta$ gleich der Projektion des zum Punkte $P(\vartheta, \omega)$ gehörenden Radius CP auf die Koordinatenachse CO, die aber willkürlich ist, und den Ausdruck $\varrho d\sigma$ hatten wir bereits S. 219 als „Wirbelelement" bezeichnet. Somit haben wir:

Satz I. *Die Summe aller Wirbelelemente auf der Kugel, jedes multipliziert mit der Projektion des zugehörigen Kugelradius auf eine beliebige feste Achse ist in der Zeit konstant.*

Sind also CX, CY, CZ drei auf einander senkrechte Achsen und ξ, η, ζ die entsprechenden Projektionen von OP, so wird:

$$\left.\begin{aligned} L_x &= \int^{(K)} \xi\varrho d\sigma = \text{const} \\ L_y &= \int^{(K)} \eta\varrho d\sigma = \text{const} \\ L_z &= \int^{(K)} \zeta\varrho d\sigma = \text{const}\ , \end{aligned}\right\} \tag{4}$$

und diese drei Relationen sind von einander unabhängig, während jede weitere analoge Gleichung durch lineare Kombination aus ihnen hervorgehen würde. Die Größen L_x, L_y, L_z sind die Komponenten eines Vektors, dessen Endpunkt (vom Kugelmittelpunkt C an gerechnet) ein fester Punkt S' im Raume ist, den wir als den „repräsentierenden Schwerpunkt" der Wirbelelemente bezeichnen. Wollten wir nämlich den wahren Schwerpunkt aller Elemente $\varrho d\sigma$ bestimmen, so wären seine Koordinaten ξ_0, η_0, ζ_0 gegeben durch:

$$\xi_0 = \frac{L_x}{M}\ , \quad \eta_0 = \frac{L_y}{M}\ , \quad \zeta_0 = \frac{L_z}{M}\ ,$$

wo $M = \int^{(K)} \varrho d\sigma$, die Summe aller Wirbelelemente, nach Satz II, S. 209 bekanntlich $= 0$ ist. Der wahre Schwerpunkt fällt daher ins *Unendliche*, doch in eine feste, d. h. in der Zeit unveränderliche Richtung:

$$\xi_0 : \eta_0 : \zeta_0 = L_x : L_y : L_z\ ,$$

und dabei würden von den drei unabhängigen Relationen (4) immer nur zwei zur Geltung kommen, während die Konstanz z. B. der Größe $L_x^2 + L_y^2 + L_z^2$ verloren ginge. Um diesem Übelstande zu entgehen, ersetzen wir die Massen-
230 verteilung mit der Dichte ϱ durch eine andere, | deren Summe nicht mehr $= 0$, sondern $= 1$ ist, indem wir die Dichte $\varrho = \varrho' + \dfrac{1}{4\pi}$ annehmen, also

or

$$L_0 = \int^{(K)} \varrho d\sigma \cos\vartheta = \text{const.}$$

Now, $\cos\vartheta$ is equal to the projection of the radius CP belonging to the point $P(\vartheta, \omega)$ on the coordinate axis CO, which however, is arbitrary, and, already on p. 219, we referred to the expression $\varrho d\sigma$ as the "vortex element". We thus have:

Theorem I. *The sum of all vortex elements on the sphere, each multiplied by the projection of the corresponding sphere radius on some fixed axis, is constant over time.*

Hence, if CX, CY, CZ are three mutually perpendicular axes, and ξ, η, ζ the corresponding projections of OP, then we have:

$$\left.\begin{aligned} L_x &= \int^{(K)} \xi\varrho d\sigma = \text{const} \\ L_y &= \int^{(K)} \eta\varrho d\sigma = \text{const} \\ L_z &= \int^{(K)} \zeta\varrho d\sigma = \text{const}\,, \end{aligned}\right\} \tag{4}$$

and these three relations are mutually independent, while any further analogous equation would be obtained from them by linear combination. The magnitudes L_x, L_y, L_z are the components of a vector whose endpoint (reckoned from the sphere's center C) is a fixed point S' in space, which we call the "representative center of gravity" of the vortex elements. For if we wanted to determine the true center of gravity of all elements $\varrho d\sigma$, then its coordinates ξ_0, η_0, ζ_0 would be given by:

$$\xi_0 = \frac{L_x}{M}\,, \quad \eta_0 = \frac{L_y}{M}\,, \quad \zeta_0 = \frac{L_z}{M}\,,$$

where $M = \int^{(K)} \varrho d\sigma$, the sum of all vortex elements is, by Theorem II, p. 209 $= 0$, as is well-known. The true center of gravity therefore lies at *infinity* but in a direction that is fixed, i. e., invariable over time:

$$\xi_0 : \eta_0 : \zeta_0 = L_x : L_y : L_z\,,$$

and, in this case, only two of the three independent relations (4) would ever come to the fore, while the constancy of, e. g., the magnitude $L_x^2 + L_y^2 + L_z^2$ would be lost. To avert this misfortune, we replace the mass distribution of density ϱ by a different one whose sum no longer is $= 0$ but $= 1$ by assuming the density $\varrho = \varrho' + \dfrac{1}{4\pi}$, and hence by adding a homogeneous mass covering

eine homogene Massenbelegung von der Dichte $\frac{1}{4\pi}$ und der Gesamtmasse 1 hinzufügen. Der Schwerpunkt dieser Belegung ϱ' hat dann die Koordinaten:

$$\begin{aligned}
\xi_0' &= \int^{(K)} \xi\varrho' d\sigma = \int^{(K)} \xi\varrho d\sigma + \frac{1}{4\pi}\int^{(K)} \xi d\sigma = \int^{(K)} \xi\varrho d\sigma = L_x = \text{const.} \\
\eta_0' &= \int^{(K)} \eta\varrho' d\sigma = \int^{(K)} \eta\varrho d\sigma + \frac{1}{4\pi}\int^{(K)} \eta d\sigma = \int^{(K)} \eta\varrho d\sigma = L_y = \text{const.} \qquad (5) \\
\zeta_0' &= \int^{(K)} \zeta\varrho' d\sigma = \int^{(K)} \zeta\varrho d\sigma + \frac{1}{4\pi}\int^{(K)} \zeta d\sigma = \int^{(K)} \zeta\varrho d\sigma = L_z = \text{const.},
\end{aligned}$$

fällt also zusammen mit dem oben definierten „repräsentierenden Schwerpunkt". Also:

Satz II. *Ist ϱ die Wirbeldichte einer kontinuierlichen Strömung einer inkompressiblen Flüssigkeit auf einer Vollkugel, und bildet man in jedem Augenblick den Schwerpunkt S' einer Massenbelegung mit der Dichte $\varrho + \frac{1}{4\pi}$, so ist dieser (immer im Endlichen liegende) „repräsentierende Schwerpunkt" S' ein fester Punkt im Raume bei allen Veränderungen der Geschwindigkeitsverteilung. Der wahre Schwerpunkt S aller Wirbelelemente dagegen fällt in den unendlich fernen Punkt desselben Durchmessers CS'.*

§ 4. Stationäre Strömungen.

Nach Satz IV S. 217 ist eine (von singulären Stellen freie) Strömung auf einer Fläche *stationär*, wenn auf jeder Stromlinie die Wirbeldichte konstant ist, d. h. wenn

$$\varrho = \tfrac{1}{2} D\psi = f(\psi)\ ,$$

eine Funktion der Stromfunktion allein ist. Diese Bedingung schreibt sich für die Kugel a) in Polarkoordinaten ϑ, ω, b) in stereographischen Koordinaten x, y (cf. II § 1) in der Form:

$$\left.\begin{aligned}
&\text{a)} \quad D\psi \equiv \frac{\partial^2\psi}{\partial\vartheta^2} + \cot\vartheta\frac{\partial\psi}{\partial\vartheta} + \frac{1}{\sin^2\vartheta}\frac{\partial^2\psi}{\partial\omega^2} = 2f(\psi) \\
&\text{b)} \quad \Delta\psi \equiv \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} = \frac{8f(\psi)}{(1+x^2+y^2)^2}\ .
\end{aligned}\right\} \qquad (1)$$

Die Bedingung (1a) wird sicher befriedigt, wenn $\frac{\partial\psi}{\partial\omega} = 0$, d. h. wenn $\psi = \psi(\vartheta)$ eine Funktion der Poldistanz allein und demnach auf allen Parallelkreisen konstant ist. Daraus folgt:

of density $\frac{1}{4\pi}$ and of total mass 1. The center of gravity of this covering ϱ' has then the coordinates:

$$\left.\begin{aligned}
\xi_0' &= \int^{(K)} \xi\varrho' d\sigma = \int^{(K)} \xi\varrho d\sigma + \frac{1}{4\pi}\int^{(K)} \xi d\sigma = \int^{(K)} \xi\varrho d\sigma = L_x = \text{const.}\\
\eta_0' &= \int^{(K)} \eta\varrho' d\sigma = \int^{(K)} \eta\varrho d\sigma + \frac{1}{4\pi}\int^{(K)} \eta d\sigma = \int^{(K)} \eta\varrho d\sigma = L_y = \text{const.}\\
\zeta_0' &= \int^{(K)} \zeta\varrho' d\sigma = \int^{(K)} \zeta\varrho d\sigma + \frac{1}{4\pi}\int^{(K)} \zeta d\sigma = \int^{(K)} \zeta\varrho d\sigma = L_z = \text{const.},
\end{aligned}\right\} \quad (5)$$

and hence coincides with the "representative center of gravity" defined above. Thus:

Theorem II. *If ϱ is the vortex density of a continuous flow of an incompressible fluid on a full sphere, and if, at every instant in time, we form the center of gravity S' of a mass covering of density $\varrho + \frac{1}{4\pi}$, then this (always finite) "representative center of gravity" S' is a fixed point in space for all changes of the velocity distribution. In contrast, the true center of gravity S of all vortex elements is the infinitely distant point of the same diameter CS'.*

§ 4. Stationary flows.

By Theorem IV p. 217, a flow (that is free of singular points) is *stationary* on a surface, if the vortex density is constant on every streamline, i. e., if

$$\varrho = \tfrac{1}{2} D\psi = f(\psi)\ ,$$

is a function of the stream function alone. We can express this condition for the sphere a) in polar coordinates ϑ, ω, b) in stereographic coordinates x, y (cf. II § 1) as follows:

$$\left.\begin{aligned}
&\text{a)} \quad D\psi \equiv \frac{\partial^2\psi}{\partial\vartheta^2} + \cot\vartheta\frac{\partial\psi}{\partial\vartheta} + \frac{1}{\sin^2\vartheta}\frac{\partial^2\psi}{\partial\omega^2} = 2f(\psi)\\
&\text{b)} \quad \Delta\psi \equiv \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} = \frac{8f(\psi)}{(1+x^2+y^2)^2}\ .
\end{aligned}\right\} \quad (1)$$

Condition (1a) is, in fact, satisfied if $\frac{\partial\psi}{\partial\omega} = 0$, i. e., if $\psi = \psi(\vartheta)$ is a function of the polar distance alone, and hence if it is constant on all parallel circles. From this follows:

231 | Satz I. *Alle „zonalen" Strömungen* (*welche symmetrisch um einen Durchmesser als Achse in den Parallelkreisen erfolgen*) *sind stationär.*

Bei diesem wie bei den folgenden Sätzen wird auch ohne ausdrückliche Erwähnung vorausgesetzt, daß die Flüssigkeit inkompressibel, auf der Kugel unbegrenzt und von Strudelpunkten etc. frei sei.

Abgesehen von dem zonalen Falle klassifiziert man die stationären Strömungen mit Vorteil nach der Beschaffenheit der Funktion f.

Hier ergiebt sich zunächst, weil nach S. 209

$$\overset{(K)}{\int} \varrho d\sigma = \overset{(K)}{\int} f(\psi) d\sigma = 0$$

ist, daß die Funktion f jedenfalls kein definites Vorzeichen haben kann. So kann *nicht* auf der ganzen Kugel

$$\varrho = \text{const}\,, \qquad \varrho = c\psi^2 \quad \text{oder} \quad \varrho = ce^{g(\psi)}$$

sein.

Der einfachste Fall, der hier in Betracht kommt, wäre der, wo $f(\psi)$ eine *lineare* Funktion von ψ oder, was die Allgemeinheit nicht beschränkt sondern nur die additive Konstante von ψ beeinflußt, wo die Wirbeldichte ϱ der Stromfunktion ψ proportional $= k\psi$ ist. Wir erhalten dann für ψ die partielle Differentialgleichung:

$$D\psi \equiv \frac{\partial^2\psi}{\partial\vartheta^2} + \cot\vartheta \frac{\partial\psi}{\partial\vartheta} + \frac{1}{\sin^2\vartheta} \frac{\partial^2\psi}{\partial\omega^2} = 2k\psi \tag{2a}$$

oder

$$\Delta\psi \equiv \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} = \frac{8k}{(1+x^2+y^2)^2}\psi\,. \tag{2b}$$

Diese Differentialgleichung der „*linear-stationären Strömungen*", wie wir uns zur Abkürzung ausdrücken wollen, ist identisch mit derjenigen, welche z. B. die elastischen Schwingungen der Kugelfläche bestimmt, und spielt auch in der Potentialtheorie eine wichtige Rolle. Ihre Integration erfolgt durch *Kugelfunktionen* („*Laplace*sche Funktionen", „Kugelflächenfunktionen", „Spherical harmonics"), und ihre Theorie ist vielfach ausführlich behandelt, z. B. bei *Lamb* a. a. O., und bei *Maxwell*, Treatise on Elect. a. Magn. Wir können uns hier darauf beschränken, die wichtigsten Ergebnisse dieser Theorie, soweit wir ihrer bedürfen, kurz anzugeben und auf ihre hydrodynamische Bedeutung für das hier vorliegende Problem hinzuweisen.

Zunächst ist zu beachten, daß unsere Differentialgleichung linear und homogen ist, daß sich also alle ihre Lösungen linear superponieren. Also:

Satz II. *Durch additive Übereinanderlagerung* (*d. h. durch Addition der*
232 *entsprechenden Stromfunktionen oder Geschwindigkeitsvektoren*) | *von zwei*

Theorem I. *All "zonal" flows (that run symmetrically about a diameter as axis in the parallel circles) are stationary.*

For this and the following theorems, we assume, even without explicitly mentioning it, that the fluid is incompressible, unbounded on the sphere, and free of whirl points etc.

The zonal case aside, it is advantageous to classify stationary flows according to the constitution of the function f.

Here we first note that, since we have

$$\int^{(K)} \varrho d\sigma = \int^{(K)} f(\psi) d\sigma = 0 \,,$$

by p. 209, the function f cannot, in fact, have a definite sign. Thus we can *not* have

$$\varrho = \text{const}\,, \qquad \varrho = c\psi^2 \quad \text{oder} \quad \varrho = ce^{g(\psi)}$$

on the entire sphere.

The simplest case of relevance here would be the case where $f(\psi)$ is a *linear* function of ψ, or where the vortex density ϱ is proportional to the stream function ψ and $= k\psi$, which does not restrict generality but only affects the additive constant of ψ. We then obtain for ψ the partial differential equation:

$$D\psi \equiv \frac{\partial^2\psi}{\partial\vartheta^2} + \cot\vartheta\frac{\partial\psi}{\partial\vartheta} + \frac{1}{\sin^2\vartheta}\frac{\partial^2\psi}{\partial\omega^2} = 2k\psi \tag{2a}$$

or

$$\Delta\psi \equiv \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} = \frac{8k}{(1+x^2+y^2)^2}\psi \,. \tag{2b}$$

This differential equation of the "*linear-stationary flows*", for short, is identical with the one that, e. g., determines the elastic oscillations of the spherical surface, and also plays a significant role in potential theory. Its integration proceeds via *spherical functions* ("*Laplace*'s functions", "spherical surface functions", "spherical harmonics"), and its theory was discussed in detail on many occasions, e. g., in *Lamb 1895*, and in *Maxwell 1873*. We may confine ourselves here to a brief statement of the most significant results of this theory, as need arises, and to a few remarks on their hydrodynamical significance for the present problem.

First we should note that our differential equation is linear and homogeneous, and hence that all of its solutions linearly superpose. Thus:

Theorem II. *By means of additive superposition (i. e., by adding the corresponding stream functions or velocity vectors) of two linear-stationary flows*

linear-stationären Strömungen, die zu demselben Werte k, wir wollen dafür sagen: zur selben „Klasse", gehören, erhält man immer wieder linear-stationäre Strömungen derselben Klasse.

Unter den Lösungen $\psi = \psi(\vartheta, \omega)$ von (2a) interessieren zunächst die Funktionen $\psi = \psi(\vartheta)$ von ϑ allein, welche zonalen Strömungen entsprechen. Sie müssen der gewöhnlichen Differentialgleichung genügen:

$$\frac{d^2\psi}{d\vartheta^2} + \cot\,\vartheta \frac{d\psi}{d\vartheta} - 2k\psi = 0\ . \tag{3}$$

Diese Differentialgleichung besitzt aber nur dann ein zwischen den Grenzen $\vartheta = 0$ und $\vartheta = \pi$ stetiges partikuläres Integral, wenn

$$-2k = n(n+1) \tag{4}$$

und n eine positive *ganze Zahl* ist, welche die „Klasse" der linear-stationären Strömung angiebt. Die Gleichung muß also von der Form sein

$$\frac{d^2\psi}{d\vartheta^2} + \cot\,\vartheta \frac{d\psi}{d\vartheta} + n(n+1)\psi = 0\ , \tag{3'}$$

und die gesuchte stetige Lösung ist dann die „$n$te Kugelfunktion" (oder „*Legendre*sches Polynom") von $\cos\vartheta$:

$$\psi = P_n(\cos\vartheta)\ , \quad \overline{\omega} = \frac{d\psi}{d\vartheta} = -\sin\vartheta P_n'(\cos\vartheta)\ , \tag{5}$$

wo P_n eine ganze Funktion $n$ten Grades.

$n = 0$ liefert die triviale Lösung $\psi = \text{const}$, d. h. die völlige Ruhe der Flüssigkeit, $n = -k = 1$ dagegen die Lösung:

$$\psi = -\varrho = -\alpha\cos\vartheta\ , \quad \overline{\omega} = \alpha\sin\vartheta\ ,$$

welche einer *starren Rotation* der gesamten Flüssigkeit um die Achse OO' mit der Winkelgeschwindigkeit α entspricht. Erst die höheren Kugelfunktionen P_2, $P_3 \ldots$ liefern eigentliche Strömungen mit wirklicher Deformation der Flüssigkeit.

Da $P_n(x)$ zwischen den Grenzen -1 und $+1$ bekanntlich n mal, $P_n'(x)$ $n-1$ mal verschwindet, so giebt es (unter Ausschluß der Pole O und O') immer n Parallelkreise, auf welchen die Stromfunktion und die Wirbeldichte ϱ, und, von ihnen separiert, $n-1$ Parallelkreise, auf welchen die Geschwindigkeit verschwindet. Beide Kreissysteme verteilen sich symmetrisch um den Äquator $\left(\vartheta = \frac{\pi}{2}\right)$ (auf welchem selbst $\psi = 0$ und $\varrho = 0$ oder $\omega = 0$ ist, je nachdem n ungerade oder gerade ist), und zerlegen die ganze Kugelfläche in $n+1$, bezw. n Zonen, auf welchen ϱ, bezw. ω abwechselnd positiv und negativ ist. Auf
233 zwei | symmetrischen Parallelkreisen oberhalb und unterhalb des Äquators sind entweder die Werte der Wirbeldichte gleich und die der Geschwindigkeit

belonging to the same value k, or, as we shall say: to the same "class", we always obtain linear-stationary flows of the same class.

Among the solutions $\psi = \psi(\vartheta, \omega)$ of (2a), the functions $\psi = \psi(\vartheta)$ that depend only on ϑ and correspond to zonal flows are at first of interest to us here. They must satisfy the usual differential equation:

$$\frac{d^2\psi}{d\vartheta^2} + \cot\,\vartheta \frac{d\psi}{d\vartheta} - 2k\psi = 0\ . \tag{3}$$

This differential equation, however, possesses a continuous particular integral between the limits $\vartheta = 0$ and $\vartheta = \pi$ only if

$$-2k = n(n+1) \tag{4}$$

and if n is a positive *integer*, which indicates the "class" of the linear-stationary flow. Thus the equation must be of the form

$$\frac{d^2\psi}{d\vartheta^2} + \cot\,\vartheta \frac{d\psi}{d\vartheta} + n(n+1)\psi = 0\ , \tag{3'}$$

and the continuous solution sought is then the "nth spherical function" (or "*Legendre*'s polynomial") of $\cos\vartheta$:

$$\psi = P_n(\cos\vartheta)\ , \quad \overline{\omega} = \frac{d\psi}{d\vartheta} = -\sin\vartheta P_n'(\cos\vartheta)\ , \tag{5}$$

where P_n is an entire function of nth order.

$n = 0$ yields the trivial solution $\psi = \text{const}$, i. e., the total rest of the fluid, while $n = -k = 1$ yields the solution:

$$\psi = -\varrho = -\alpha\cos\vartheta\ , \quad \overline{\omega} = \alpha\sin\vartheta\ ,$$

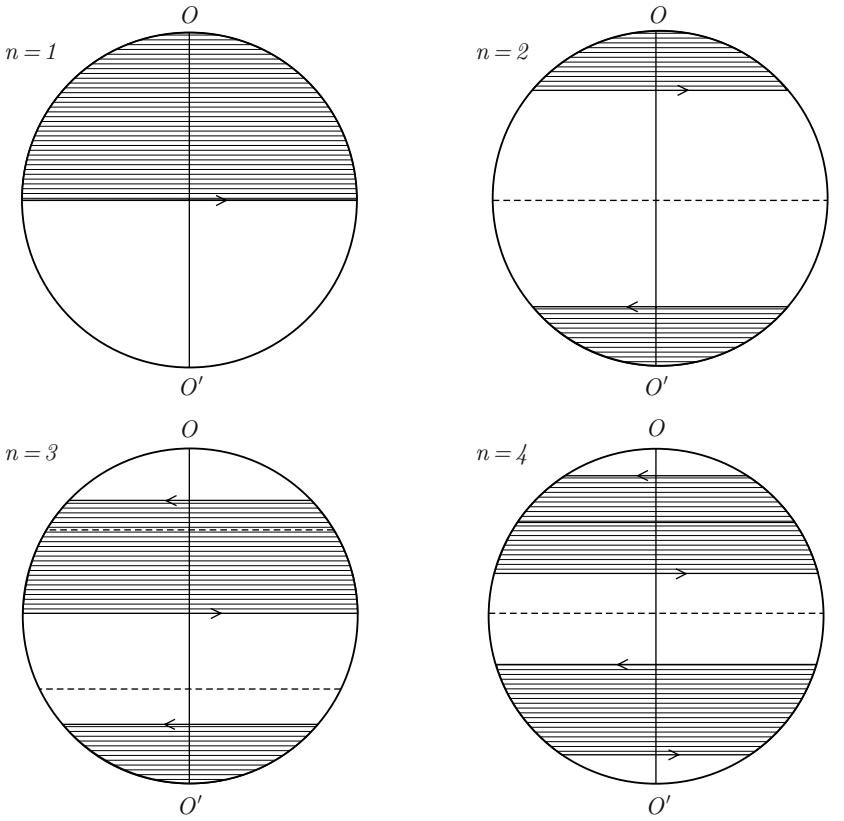
which corresponds to a *rigid rotation* of the entire fluid about the axis OO' at angular velocity α. Only the higher spherical functions P_2, P_3 ... yield actual flows with real deformation of the fluid.

Since $P_n(x)$ vanishes between the limits -1 and $+1$ n times, as is well-known, and $P_n'(x)$ $n-1$ times, there always (the poles O and O' being excluded) exist n parallel circles on which the stream function and the vortex density ϱ vanish, and, separately, $n-1$ parallel circles on which the velocity vanishes. The two systems of circles are arranged symmetrically about the equator $\left(\vartheta = \frac{\pi}{2}\right)$ (on which either $\psi = 0$ and $\varrho = 0$ or $\omega = 0$ depending on whether n is uneven or even), and they divide the entire spherical surface into $n+1$ and n zones respectively, on which ϱ, or ω, is alternately positive and negative. On two symmetric parallel circles above and below the equator, the values of the vortex density are either equal and those of the velocity

entgegengesetzt (wenn n gerade) oder die der Geschwindigkeit gleich und die der Wirbeldichte entgegengesetzt (wenn n ungerade).

[Auf den beistehenden Figuren für $n = 1, 2, 3, 4$ (in orthographischer Projektion) sind die Parallelkreise $\varrho = 0$, $\psi = 0$ ausgezogen und die Kreise $\overline{\omega} = 0$ punktiert gezeichnet, die Geschwindigkeitsrichtung durch Pfeile angedeutet und die Gebiete $\varrho > 0$ schraffiert].

Fig. 3.



Da aber der Pol O auf der Kugel willkürlich angenommen werden kann und die Lösungen ψ von (2) für jedes $2k = -n(n+1)$ sich additiv zusammensetzen lassen, so erhält man neue *nicht* achsial-symmetrische Strömungen, wenn man mehrere zu verschiedenen Polen gehörende zonale zu einander addiert:

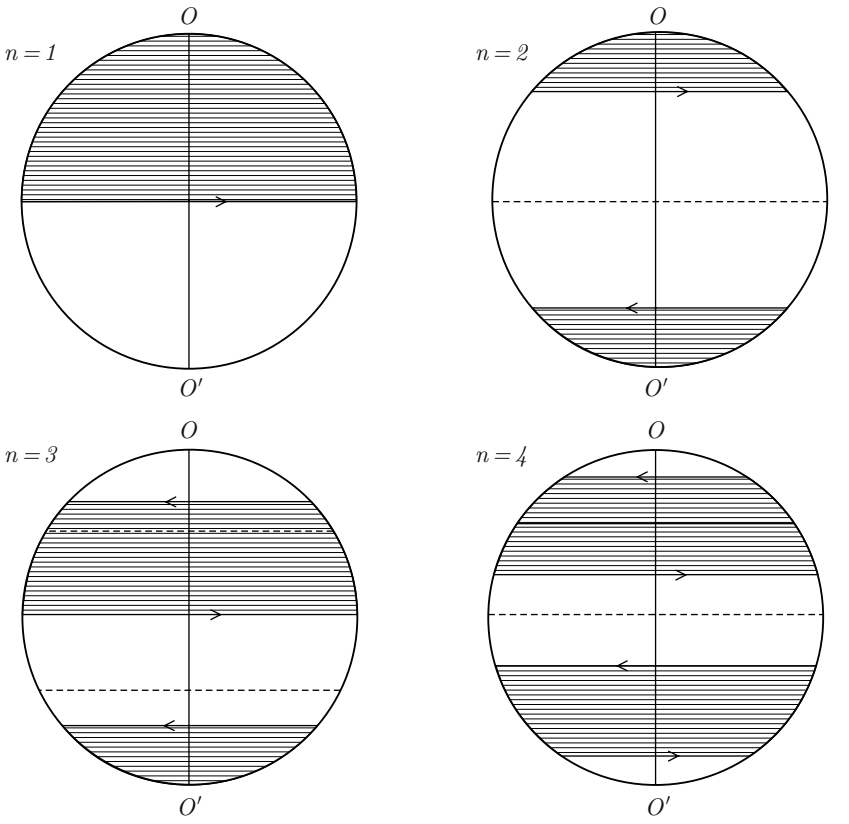
$$\psi_n = c_1 P_n(\cos\delta_1) + c_2 P_n(\cos\delta_2) + \dots \,, \tag{6}$$

234 | wo $\delta_1, \delta_2, \dots$ die sphärischen Abstände von den Polen $P_1, P_2, \dots$ bezeichnen.

opposite (when n is even) or those of the velocity are equal and those of the vortex density opposite (when n is uneven).

[The adjacent figures for $n = 1, 2, 3, 4$ (in orthographic projection) show the parallel circles $\varrho = 0$, $\psi = 0$ as continuous lines and the circles $\overline{\omega} = 0$ as dotted lines, arrows indicate the velocity direction, and the regions $\varrho > 0$ are shaded].

Fig. 3.



Since, however, we can assume any pole O on the sphere, and since we can additively combine the solutions ψ of (2) for any $2k = -n(n+1)$, we obtain new *non* axial-symmetric flows by adding together several zonal flows belonging to different poles:

$$\psi_n = c_1 P_n(\cos\delta_1) + c_2 P_n(\cos\delta_2) + \dots \,, \tag{6}$$

where δ_1, δ_2, ... denote the spherical distances from the poles P_1, P_2, ...

Nun ist aber bekannt, daß auch die partielle Differentialgleichung (2) *nur* in dem Falle (4) $2k = -n(n+1)$ überhaupt eine eindeutige und stetige Lösung auf der Vollkugel besitzt, und daß die allgemeinste Lösung von dieser Beschaffenheit eine ganze rationale Funktion $n$ter Dimension von $\cos\vartheta$, $\cos\omega$ und $\sin\omega$ ist, welche sich mittelst $2n+1$ willkürlicher Konstanten in der Form darstellen läßt:

$$\psi = \psi_n = a_0 P_n(\cos\vartheta) + \sum_{r=1}^{n} (a_r \cos(r\omega) + b_r \sin'(r\omega)) \sin^r\vartheta P_n^{(r)}(\cos\vartheta)\,, \quad (7)$$

wo $a_0, a_1, a_2, \ldots, b_1, b_2, \ldots$ Konstanten sind und $P_n^{(r)}(x)$ die $r$te Ableitung von $P_n(x)$ bedeutet. Eine solche Funktion des Ortes auf der Kugel wird eine „*Laplace*sche Funktion“ oder eine „Kugelflächenfunktion“ genannt.

Wir erkennen somit:

Satz III. *Es giebt keine anderen kontinuierlichen linear-stationären Strömungen auf der Vollkugel als solche n ter Klasse.* ($2k = -n(n+1)$, $n = 1, 2, 3, \ldots$)*, und die allgemeinste n ter Klasse läßt sich aus* $2n+1$ *von einander unabhängigen Grundströmungen linear zusammensetzen, deren Stromfunktionen sämtlich durch die „Kugelflächenfunktionen“ n ten Grades dargestellt werden.*

Die Strömungen 1. Klasse

$$\begin{aligned}\psi = \psi_1 &= a_0\cos\vartheta + a_1\sin\vartheta\cos\omega + b_1\sin\vartheta\sin\omega \\ &= a_0\zeta + a_1\xi + b_1\eta\end{aligned}$$

(wo ξ, η, ζ rechtwinklige Koordinaten mit dem Kugelmittelpunkt als Anfangspunkt bedeuten) sind lediglich starre *Rotationen* um beliebige Durchmesser und lassen sich stets aus drei verschiedenen Rotationen linear zusammensetzen.

Für höhere Werte von n zerfällt die ganze Kugelfläche in eine Anzahl von Teilgebieten, in deren Innerem Stromfunktion und Wirbeldichte abwechselnd positiv und negativ ist und an den durch Stromlinien gebildeten Grenzen verschwindet. Besonders einfach werden diese Gebiete für eine „Grundströmung“

$$\psi = \psi_{n,r} = \sin^r\vartheta P_n^{(r)}(\cos\vartheta)\cos(r\omega) \quad (r = 0, 1, \ldots, n)\,. \qquad (7a)$$

Denn hier verschwinden (abgesehen von dem oben behandelten zonalen Fall $r = 0$, $\psi = P_n(\cos\vartheta)$) ψ und ϱ augenscheinlich

1) auf allen r Meridianen $\omega = \dfrac{\pi}{2r}, \dfrac{3\pi}{2r}, \ldots (2r-1)\dfrac{\pi}{2r}$, für welche $\cos(r\omega) = 0$
235 ist, |
2) auf allen $n-r$ Parallelkreisen $\vartheta = \alpha_1, \alpha_2, \ldots \alpha_{n-r}$, für welche $P_n^{(r)}(\cos\vartheta) = 0$ ist.

Now, however, it is well-known that the partial differential equation (2), too, *only* has a unique and continuous solution on the full sphere in the case (4) $2k = -n(n+1)$ and that the most general solution so constituted is an entire rational function of nth dimension of $\cos\vartheta$, $\cos\omega$ and $\sin\omega$, which can be represented by means of $2n+1$ arbitrary constants as follows:

$$\psi = \psi_n = a_0 P_n(\cos\vartheta) + \sum_{r=1}^{n} (a_r \cos(r\omega) + b_r \sin'(r\omega)) \sin^r\vartheta P_n^{(r)}(\cos\vartheta)\,, \quad (7)$$

where $a_0, a_1, a_2, \ldots, b_1, b_2, \ldots$ are constants, and $P_n^{(r)}(x)$ denotes the rth derivative of $P_n(x)$. Such a function of the location on the sphere is called a "*Laplace* function", or "spherical surface function".

It is thus evident:

Theorem III. *There are no other continuous linear-stationary flows on the full sphere besides those of the n th class.* ($2k = -n(n+1)$, $n = 1, 2, 3, \ldots$), *and the most general one of the n th class can be obtained by linear combination of* $2n+1$ *mutually independent basic flows whose stream functions are all represented by the "spherical surface functions" of n th degree.*

The flows of the 1st class

$$\begin{aligned}\psi = \psi_1 &= a_0 \cos\vartheta + a_1 \sin\vartheta \cos\omega + b_1 \sin\vartheta \sin\omega \\ &= a_0\zeta + a_1\xi + b_1\eta\end{aligned}$$

(where ξ, η, ζ denote the orthogonal coordinates with the sphere's center as the point of origin) are but rigid *rotations* about some diameter and can always be obtained by a linear combination of three different rotations.

For higher values of n, the entire spherical surface breaks down into a number of partial regions in the interior of which stream function and vortex density are alternately positive and negative and vanish at the boundaries formed by the streamlines. These regions are particularly simple for a "basic flow"

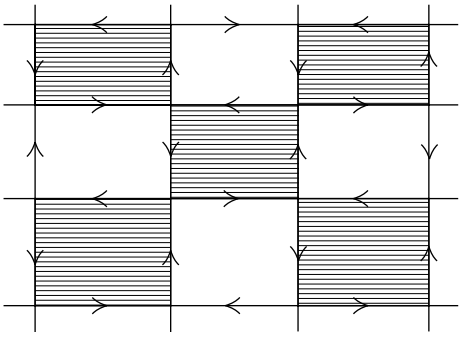
$$\psi = \psi_{n,r} = \sin^r\vartheta P_n^{(r)}(\cos\vartheta)\cos(r\omega) \quad (r = 0, 1, \ldots, n)\,. \qquad (7a)$$

For here (apart from the zonal case discussed above where $r = 0$, $\psi = P_n(\cos\vartheta)$) ψ and ϱ evidently vanish

1) on all r meridians $\omega = \dfrac{\pi}{2r}, \dfrac{3\pi}{2r}, \ldots (2r-1)\dfrac{\pi}{2r}$ for which we have $\cos(r\omega) = 0$,
2) on all $n - r$ parallel circles $\vartheta = \alpha_1, \alpha_2, \ldots \alpha_{n-r}$ for which we have $P_n^{(r)}(\cos\vartheta) = 0$.

Hier zerfällt also die Kugelfläche in ein System von $r(n-r+1)$ rechtwinkligen sphärischen Vierecken (bezw. Dreiecken bei O und O'), deren Grenzen von Stromlinien gebildet sind, welche abwechselnd im positiven und negativen Sinne umkreist werden ($\varrho \gtrless 0$), und in deren Eckpunkten die Flüssigkeit beständig in Ruhe bleibt ($\overline{\vartheta} = 0, \overline{\varpi} = 0$). In Merkatorscher Projektion erhält man demnach eine schachbrettartige Figur wie die beistehende Fig. 4.

Fig. 4.

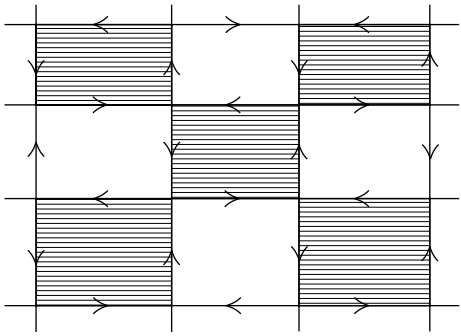


Für andere, zusammengesetzte Kugelflächenfunktionen (6) werden die Teilbereiche unregelmäßiger, aber der allgemeine Charakter des Vorganges bleibt derselbe. Mittelst der *Lamé*schen Funktionen würden wir z. B. Einteilungen durch konfokale sphärische Kegelschnitte erhalten. Doch soll darauf nicht weiter eingegangen werden.

Hätten wir statt der Vollkugel nur ein bestimmt umrandetes *Gebiet C* betrachtet und nach den dort möglichen stationären Strömungen gefragt, so wären wir auf die Randwertprobleme der Differentialgleichung (1) oder (2) gestoßen. Denn da die Berandung immer selbst eine Stromlinie sein muß, so hätten wir unsere partielle Differentialgleichung unter der Randbedingung $\psi = \psi_0 = \text{const}$ zu integrieren gehabt. Bezüglich der Differentialgleichung (2) in der Form (2b) kann dieses Problem als gelöst betrachtet werden durch die Untersuchungen von *Schwarz* (Festschrift „Über ein Problem der Variationsrechnung etc.“ 1885). Die Lösung der Randwertaufgabe für irgend ein $\psi_0 > 0$ ist immer möglich (für beliebiges k), wenn das Gebiet C nicht zu groß gewählt wird. Dagegen existiert für jedes Gebiet C immer ein Wert k derart, daß an dem Rande überall $\psi = \psi_0 = 0$ wird. Unter analogen Bedingungen ist aber nach *Picard* (Liouville Journ., Ser. IV, t. 6, p. 145 ff.) die Randwertaufgabe auch lösbar für die Diffgl. (1b), wenn die Funktion f eine nicht lineare Funktion ist; es giebt also stationäre, aber *nicht*-linear-stationäre Strömungen wenigstens in hinreichend kleinen umrandeten *Bereichen* der Kugelfläche. Ob aber auch auf der Vollkugel solche allgemeineren stationären Strömungen
236 | möglich sind, diese Frage ist als noch ungelöst zu bezeichnen. Jedenfalls bedürfte es zu dieser Untersuchung wohl eines wesentlich anderen Integrationsverfahrens als der *Schwarz-Picard*schen Approximationsmethode.

Hence, in this case, the spherical surface breaks down into a system of $r(n-r+1)$ orthogonal spherical rectangles (or triangles at O and O') whose boundaries are formed by stream lines which are alternately orbited in the positive and negative sense ($\varrho \gtrless 0$) and in whose vertices the fluid always remains at rest ($\overline{\vartheta} = 0$, $\overline{\varpi} = 0$). Thus, in Mercator projection, we obtain a chessboard-like figure like the adjacent Fig. 4.

Fig. 4.



For other, composite spherical surface functions (6) the partial domains turn out to be less regular, but the general character of the procedure remains the same. By means of *Lamé* functions, we would obtain, e. g., divisions by confocal spherical conic sections. We shall not, however, pursue this any further.

If we had not considered the full sphere but only a definitely outlined *region* C and raised the question of stationary flows that are possible there, then we would have encountered the boundary value problem of the differential equation (1) or (2). For since the boundary itself must always be a streamline, we would have had to integrate our partial differential equation under the boundary condition $\psi = \psi_0 = \text{const}$. With respect to the differential equation (2) in the form (2b), this problem can be considered solved by the investigations by *Schwarz* (Festschrift "Über ein Problem der Variationsrechnung etc.", *1885*). The solution of the boundary value problem for any $\psi_0 > 0$ is always possible (for any k), if the region C is not chosen too large. In contrast, for every C there always exists a value k so that we get $\psi = \psi_0 = 0$ everywhere at the boundary. However, the boundary value problem for the diff⟦erential⟧ equ⟦ation⟧ (1b) is also solvable under analogous conditions, according to *Picard* (*1890*), if the function f is a nonlinear function; hence there exist stationary yet *non*-linear-stationary flows at least in sufficiently small outlined *domains* of the spherical surface. But whether such more general stationary flows are also possible on the full sphere is a question we must consider still unsolvable. In any event, this investigation would certainly require a method of integration essentially different from the *Schwarz-Picard* method by approximation.

Den stationären Strömungen am nächsten stehen die „*rotierend-stationären*“, d. h. solche Strömungen, bei denen das ganze Bild der Stromlinien oder die Stromfunktion ψ als Funktion des Ortes auf der Kugel zwar nicht konstant bleibt, aber doch nur eine gleichförmige Rotation um eine feste Achse erfährt, oder m. a. W., die Strömungen, welche, auf ein gleichförmig rotierendes Koordinatensystem bezogen, stationär erscheinen.

Ist $\psi(\vartheta, \omega)$ die wahre Stromfunktion und nach S. 232 $\psi_1 = -\alpha \cos\vartheta$ die einer Rotation, so wird

$$\psi - \psi_1 = \psi + \alpha \cos\vartheta$$

die „scheinbare“ oder „relative Stromfunktion“ in Bezug auf die rotierende Kugel, und auf den „scheinbaren“ Stromlinien $\psi - \psi_1 = \text{const}$ muß jetzt, wenn die „scheinbare Strömung“ stationär sein soll, die wahre Wirbeldichte $\varrho = \frac{1}{2} D\psi$ konstant sein (da eben diese jedem materiellen Teilchen charakteristisch ist), damit die Wirbelverteilung, auf die rotierende Kugel bezogen, ungeändert bleibt, d. h. es muß sein:

$$\varrho = \tfrac{1}{2} D\psi = f(\psi - \psi_1) = f(\psi + \alpha \cos\vartheta) \,. \tag{8}$$

Auch hier betrachten wir ebenso wie bei der Diffgl. (1) vor allem den Spezialfall

$$f(u) = ku = -\frac{n(n+1)}{2} u \,,$$

also die Gleichung:

$$D\psi + n(n+1)\psi = -n(n+1)\alpha \cos\vartheta \,. \tag{9}$$

Eine partikuläre Lösung derselben ist die neue Rotation

$$\psi_0 = c \cos\vartheta = -\frac{n(n+1)}{(n-1)(n+2)} \alpha \cos\vartheta \,, \tag{10}$$

wie man sich mit Hilfe von $D\cos\vartheta = -2\cos\vartheta$ durch Ausrechnung leicht überzeugt, und die allgemeine Lösung von (9) erhält man, wenn man zu dieser partikulären ψ_0 die allgemeine Lösung ψ' der homogenen Diffgl. (2) hinzufügt, also nach (7):

$$\begin{aligned} \psi = \psi_0 + \psi' = &-\frac{n(n+1)}{(n-1)(n+2)} \alpha \cos\vartheta + a_0 P_n(\cos\vartheta) \\ &+ \sum_{r=1}^{n} (a_r \cos r\omega + b_r \sin r\omega) \sin^r\vartheta P_n^{(r)}(\cos\vartheta) \,. \end{aligned} \tag{11}$$

The flows most closely related to the stationary flows are the "*rotating-stationary*" flows, i. e., those for which the entire picture of the streamlines, or the stream function ψ as a function of the position on the sphere, does not remain constant but is only subject to a uniform rotation about a fixed axis, or, in other words, those flows that appear to be stationary relative to a uniformly rotating coordinate system.

If $\psi(\vartheta, \omega)$ is the true stream function, and if, by p. 232, $\psi_1 = -\alpha \cos\vartheta$ is that of a rotation, then

$$\psi - \psi_1 = \psi + \alpha \cos\vartheta$$

becomes the "apparent" or "relative stream function" with regard to the rotating sphere, and if the "apparent flow" is meant to be stationary, then the true vortex density $\varrho = \frac{1}{2} D\psi$ must now be constant on the "apparent" streamlines $\psi - \psi_1 = \text{const}$ (since, after all, it is characteristic of every material particle) for the vortex distribution, in relation to the rotating sphere, to remain unaltered, i. e., we must have:

$$\varrho = \tfrac{1}{2} D\psi = f(\psi - \psi_1) = f(\psi + \alpha \cos\vartheta) \; . \tag{8}$$

Here, just as in the case of the diff⟦erential⟧ equ⟦ation⟧ (1), we consider in particular the special case

$$f(u) = ku = -\frac{n(n+1)}{2} u \; ,$$

and thus the equation:

$$D\psi + n(n+1)\psi = -n(n+1)\alpha \cos\vartheta \; . \tag{9}$$

A particular solution of it is the new rotation

$$\psi_0 = c \cos\vartheta = -\frac{n(n+1)}{(n-1)(n+2)} \alpha \cos\vartheta \; , \tag{10}$$

as is readily evident from $D \cos\vartheta = -2 \cos\vartheta$ by computation, and we obtain the general solution of (9) by adding to this particular ψ_0 the general solution ψ' of the homogeneous diff⟦erential⟧ equ⟦ation⟧ (2), and hence, by (7):

$$\begin{aligned} \psi = \psi_0 + \psi' = -\frac{n(n+1)}{(n-1)(n+2)} \alpha \cos\vartheta + a_0 P_n(\cos\vartheta) \\ + \sum_{r=1}^{n} (a_r \cos r\omega + b_r \sin r\omega) \sin^r \vartheta P_n^{(r)}(\cos\vartheta) \; . \end{aligned} \tag{11}$$

237 | So haben wir den Satz:

Satz IV. *Aus jeder linear-stationären Strömung* $2k = -n(n+1)$ *kann man eine mit der Winkelgeschwindigkeit* α *rotierend-stationäre ableiten durch bloße Hinzufügung einer Rotation von der Winkelgeschwindigkeit* $c = \frac{n(n+1)}{(n-1)(n+2)}\,\alpha$ *um dieselbe Achse. Oder anders ausgedruckt: Eine linear-stationäre Strömung* n *ter Klasse verbunden mit einer beliebigen Rotation* c *rotiert gleichförmig um dieselbe Achse mit der Winkelgeschwindiqkeit* $\alpha = c\,\frac{(n-1)(n+2)}{n(n+1)}$.

Die Theorie der rotierend-stationären Strömungen hat ein physikalisches Interesse, insofern sie sich anwenden läßt auf stationäre Luft- oder Wasserströmungen auf der rotierenden Erde.

Hydrodynamische Untersuchungen über die Wirbelbewegungen in einer Kugelfläche (Zweite Mitteilung)

s1902b

Capitel III
Gleichgewicht und Bewegung der Strudel

§1. Die Bewegung eines Strudelpunktes

Einen Strudelpunkt (I §3 Ende), in welchem die Geschwindigkeit mit unbestimmter Richtung unendlich groß wird, können wir uns nach II §2 durch Grenzübergang entstanden denken aus einem sehr kleinen Flächenstücke σ der Kugel, z. B. einer Calotte mit sehr kleinem Radius, sehr großer mittlerer Wirbeldichte ϱ und endlichem Wirbelmomente m. Da nun nach dem Helmholtzschen Theorem (I §4) bei der Bewegung der incompressibelen Flüssigkeit die Wirbeldichte ϱ jedes materiellen Teilchens ungeändert bleibt, so wird auch das Flächenstückchen σ immer eine sehr große Wirbeldichte und
62 sein endliches Wirbelmoment bewahren, | wird also beständig durch einen Strudelpunkt vom constanten Momente m ersetzt werden können, so daß wir sagen können:

Satz I. *Ein Strudelpunkt bleibt immer Strudelpunkt mit dem constanten Momente* m *und bewegt sich wie ein materieller Punkt.*

We thus have the theorem:

Theorem IV. *From every linear-stationary flow* $2k = -n(n+1)$ *we can derive a rotating-stationary flow of angular velocity* α *by simply adding a rotation of angular velocity* $c = \dfrac{n(n+1)}{(n-1)(n+2)}\,\alpha$ *about the same axis. Or, put differently: A linear-stationary flow of the n th class combined with any rotation c uniformly rotates about the same axis at angular velocity* $\alpha = c\,\dfrac{(n-1)(n+2)}{n(n+1)}$.

The theory of rotating-stationary flows is of physical interest insofar as it is applicable to stationary air and water currents on the rotating Earth.

Hydrodynamical investigations of vortex motions in the surface of a sphere (Second communication)

s1902b

⟦The introductory note just before *1902a* also addresses *s1902b*.⟧

Chapter III
Equilibrium and motion of the whirls

§1. The motion of a whirl point

According to II §2 we may assume that a whirl point (I §3 end) at which the velocity becomes infinitely large in an indeterminate direction has come about in a passage to the limit from a very small part of the surface, σ, of the sphere, e. g., a spherical calotte with a very small radius, very large median vortex density ϱ and finite vorticity moment m. Now since, by *Helmholtz*'s theorem (I §4), for the motion of an incompressible fluid the vortex density ϱ of every material particle remains unaltered, the small part of the surface σ will always retain a very large vortex density and its finite vorticity moment, and hence it will always be possible to replace it by a whirl point of constant momentum m so that we can say:

Theorem I. *A whirl point always remains a whirl point with constant momentum* m *and moves like a material point.*

Wollen wir aber die Geschwindigkeit eines solchen Strudelpunktes bestimmen, so entsteht eine Schwierigkeit dadurch, daß die Flüssigkeit ihn von allen Seiten umströmt und dadurch die Geschwindigkeitsrichtung unbestimmt wird.

Nehmen wir zunächst den Fall eines *einfachen Strudels* (II §2), d. h. eines einzigen Strudelpunktes m bei constanter Wirbeldichte $\varrho_0 = -\frac{m}{4\pi}$ auf der ganzen Kugel. Ein solcher Strudel läßt sich, wie dort gezeigt wird, angenähert realisieren durch eine zonale Strömung ($\psi = \psi(\vartheta)$) auf der ganzen Kugel ohne wahre Singularitäten. Eine zonale Strömung ist aber nach II §4 gleichzeitig immer eine stationäre, d. h. alle Stromlinien und Geschwindigkeiten an jeder Stelle bleiben ungeändert. Wir schließen, daß dies auch im Grenzfalle so sein wird, und erhalten den Satz:

63 | **Satz II.** *Ein einfacher Strudel behält seine anfängliche Lage auf der Kugel bei, er erteilt sich selbst keine Geschwindigkeit.*

Ist jedoch außerhalb des Strudelpunktes ($\vartheta = 0$) die Wirbeldichte ϱ eine beliebige Funktion des Ortes die der Bedingung $m + \int \varrho d\sigma = 0$ genügt, so können wir die geltende Stromfunktion ψ in zwei Teile zerlegen,

$$\psi = \psi_0 + \psi',$$

deren einer

$$\psi_0 = \frac{m}{\pi} \lg \sin \frac{\vartheta}{2}$$

einem einfachen Strudel m in 0, der andere ψ' aber einer continuierlichen Wirbelverteilung mit der Dichte

$$\varrho' = \varrho - \varrho_0 = \varrho + \frac{m}{4\pi}$$

entspricht. Demgemäß zerfällt auch der Geschwindigkeitsvector q des Strudelpunktes in zwei Teile:

$$q = q_0 + q',$$

deren erster q_0 nach Satz II im Punkte 0 verschwindet, während der letztere dort einen bestimmten Grenzwert besitzt. Die zur Stromfunction $\psi' = \psi - \psi_0$ gehörende Geschwindigkeit q' an der betreffenden Stelle ist also die wahre Geschwindigkeit des Strudelpunktes. Dabei brauchen wir in der Berechnung
64 von ψ' durch das sphärische Potential (II §2) auf das konstante | Glied $\varrho_0 = -\frac{m}{4\pi}$ der Wirbeldichte keine Rücksicht zu nehmen, weil es die Stromfunction nur um eine additive Constante verändern würde. Wir sehen also:

Satz III. *Ein Strudelpunkt bewegt sich mit derjenigen Geschwindigkeit q', welche an der betreffenden Stelle gerade herrschen würde, wenn der Strudel selbst nicht vorhanden wäre* (*cf. Poincaré, a. a. O. Nr. 66*).

Dies gilt natürlich auch, wenn die Flüssigkeit mehrere Strudelpunkte enthält.

But if we wish to determine the velocity of such a whirl point, then a difficulty arises from the fact that the liquid flows around it from all sides, thereby rendering the velocity direction indeterminate.

Let us first consider the case of a *simple whirl* (II §2), i. e., of a single whirl point m at constant vortex density $\varrho_0 = -\frac{m}{4\pi}$ on the entire sphere. Such a whirl can, as is shown there, be approximately realized by a zonal current ($\psi = \psi(\vartheta)$) on the entire sphere without true singularities. But, according to II §4, a zonal current is also always a stationary one, i. e., all streamlines and velocities remain unaltered at every point. We conclude that this will also be true in the limit case and obtain the theorem:

Theorem II. *A simple whirl retains its original position on the sphere, it does not impart a velocity upon itself.*

But if outside of the whirl point ($\vartheta = 0$) the vortex density ϱ is any function of the position satisfying the condition $m + \int \varrho d\sigma = 0$, then we can divide the prevalent stream function ψ into two parts,

$$\psi = \psi_0 + \psi',$$

one of which,

$$\psi_0 = \frac{m}{\pi} \lg \sin \frac{\vartheta}{2},$$

corresponds to a simple whirl m at 0, while the other one ψ' corresponds to a continuous vortex distribution of density

$$\varrho' = \varrho - \varrho_0 = \varrho + \frac{m}{4\pi}.$$

Accordingly, the velocity vector q of the whirl point, too, breaks down into two parts:

$$q = q_0 + q',$$

the first of which, q_0, vanishes at point 0 according to Theorem II, while the latter has a certain limit there. Hence, the velocity q' at the point in question that belongs to the stream function $\psi' = \psi - \psi_0$ is the true velocity of the whirl point. Here, we do not need to take into account the constant term $\varrho_0 = -\frac{m}{4\pi}$ of the vortex density for the calculation of ψ' by means of the spherical potential (II §2) since this would change the stream function only by an additive constant. We thus see:

Theorem III. *A whirl point moves at the velocity q' that would currently obtain at the point in question if the whirl itself did not exist* (*cf. Poincaré 1893a, no. 66*).

Of course, this also holds when the liquid contains several whirl points.

§2. Strömungen mit Strudelpunkten

Es seien gegeben n Strudelpunkte

$$P_1(\vartheta_1, \omega_1), \qquad P_2(\vartheta_2, \omega_2), \quad \cdots \quad P_n(\vartheta_n, \omega_n)$$

mit den Momenten

$$m_1, \quad m_2, \quad \cdots \quad m_n,$$

und außerdem eine kontinuierliche Wirbelverteilung auf der Kugel mit der überall endlichen und stetigen Wirbeldichte $\varrho = \varrho(\vartheta, \omega)$, wobei aber immer

$$m_1 + m_2 + \cdots + m_n = -\overset{(K)}{\int} \varrho d\sigma$$

sein muß. Dann wird die Stromfunction $\psi(\vartheta, \omega)$ in einem beliebigen Punkte $P(\vartheta, \omega)$ ausgedrückt durch das sphärische Potential (II §2)

65 |
$$\psi = \psi_1 + \psi_2 + \cdots + \psi_n + \psi' \tag{1}$$
$$= \frac{m_1}{\pi} \lg \frac{r_1}{2} + \frac{m_2}{\pi} \lg \frac{r_2}{2} + \cdots + \frac{m_n}{\pi} \lg \frac{r_n}{2} + \frac{1}{\pi} \overset{(K)}{\int} \varrho d\sigma \lg \frac{r}{2} ,$$

wenn man mit $r_\lambda = 2 \sin \frac{\delta_\lambda}{2}$ den Sehnenabstand der Punkte P und P_λ und mit r den Abstand des Punktes P von dem Flächenelemente $d\sigma$ bezeichnet.

Somit wird die Geschwindigkeit q in einem regulären Punkte $P(\vartheta, \omega)$ gegeben durch

$$\left.\begin{aligned} \overline{\vartheta} &= \frac{d\vartheta}{dt} = -\frac{1}{\sin\vartheta} \frac{\partial\psi}{\partial\omega} = -\frac{1}{\sin\vartheta} \left[\frac{\partial\psi_1}{\partial\omega} + \frac{\partial\psi_2}{\partial\omega} + \cdots + \frac{\partial\psi_n}{\partial\omega} + \frac{\partial\psi}{\partial\omega} \right] \\ \overline{\omega} &= \sin\vartheta \frac{d\omega}{dt} = \frac{\partial\psi}{\partial\vartheta} = \frac{\partial\psi_1}{\partial\vartheta} + \frac{\partial\psi_2}{\partial\vartheta} + \cdots + \frac{\partial\psi_n}{\partial\vartheta} + \frac{\partial\psi'}{\partial\vartheta} , \end{aligned}\right\} \tag{2}$$

die Geschwindigkeit q_λ eines Strudelpunktes P_λ aber durch

$$\left.\begin{aligned} \overline{\vartheta_\lambda} \sin\vartheta_\lambda &= \frac{d\vartheta_\lambda}{dt} \sin\vartheta_\lambda = -\left[\frac{\partial(\psi - \psi_\lambda)}{\partial\omega} \right]_{\vartheta=\vartheta_\lambda, \omega=\omega_\lambda} , \\ \overline{\omega_\lambda} &= \frac{d\omega_\lambda}{dt} \sin\vartheta_\lambda = -\left[\frac{\partial(\psi - \psi_\lambda)}{\partial\vartheta} \right]_{\vartheta=\vartheta_\lambda, \omega=\omega_\lambda} \end{aligned}\right\} \tag{3}$$

oder durch:

$$\left.\begin{aligned} \frac{d\vartheta_\lambda}{dt} \sin\vartheta_\lambda &= -\sum_{\mu\neq\lambda} m_\mu \frac{\partial\psi_{\lambda,\mu}}{\partial\omega_\lambda} - \frac{\partial\psi'(\vartheta_\lambda, \omega_\lambda)}{\partial\vartheta_\lambda} \\ \frac{d\omega_\lambda}{dt} \sin\vartheta_\lambda &= \sum_{\mu\neq\lambda} m_\mu \frac{\partial\psi_{\lambda,\mu}}{\partial\vartheta_\lambda} + \frac{\partial\psi'(\vartheta_\lambda, \omega_\lambda)}{\partial\vartheta_\lambda} , \end{aligned}\right\} \tag{3'}$$

§2. Currents with whirl points

Consider n whirl points

$$P_1(\vartheta_1, \omega_1), \qquad P_2(\vartheta_2, \omega_2), \quad \cdots \quad P_n(\vartheta_n, \omega_n)$$

with momenta

$$m_1, \quad m_2, \quad \cdots \quad m_n,$$

and, furthermore, a continual vortex distribution on the sphere with a vortex density $\varrho = \varrho(\vartheta, \omega)$ that is everywhere finite and continuous, where, however, always necessarily

$$m_1 + m_2 + \cdots + m_n = -\overset{(K)}{\int} \varrho d\sigma \ .$$

Then at any point $P(\vartheta, \omega)$ the stream function $\psi(\vartheta, \omega)$ is expressed by the spherical potential (II §2)

$$\begin{aligned} \psi &= \psi_1 + \psi_2 + \cdots + \psi_n + \psi' \qquad (1) \\ &= \frac{m_1}{\pi} \lg \frac{r_1}{2} + \frac{m_2}{\pi} \lg \frac{r_2}{2} + \cdots + \frac{m_n}{\pi} \lg \frac{r_n}{2} + \frac{1}{\pi} \overset{(K)}{\int} \varrho d\sigma \lg \frac{r}{2} \ , \end{aligned}$$

if one denotes by $r_\lambda = 2 \sin \dfrac{\delta_\lambda}{2}$ the chord distance of the points P and P_λ, and by r the distance of the point P from the area element $d\sigma$.

Thus, at a regular point $P(\vartheta, \omega)$, the velocity q is given by

$$\left.\begin{aligned} \overline{\vartheta} &= \frac{d\vartheta}{dt} = -\frac{1}{\sin\vartheta} \frac{\partial\psi}{\partial\omega} = -\frac{1}{\sin\vartheta} \left[\frac{\partial\psi_1}{\partial\omega} + \frac{\partial\psi_2}{\partial\omega} + \cdots + \frac{\partial\psi_n}{\partial\omega} + \frac{\partial\psi}{\partial\omega} \right] \\ \overline{\omega} &= \sin\vartheta \frac{d\omega}{dt} = \frac{\partial\psi}{\partial\vartheta} = \frac{\partial\psi_1}{\partial\vartheta} + \frac{\partial\psi_2}{\partial\vartheta} + \cdots + \frac{\partial\psi_n}{\partial\vartheta} + \frac{\partial\psi'}{\partial\vartheta} \ , \end{aligned}\right\} \qquad (2)$$

but the velocity q_λ of a whirl point P_λ by

$$\left.\begin{aligned} \overline{\vartheta_\lambda} \sin\vartheta_\lambda &= \frac{d\vartheta_\lambda}{dt} \sin\vartheta_\lambda = -\left[\frac{\partial(\psi - \psi_\lambda)}{\partial\omega} \right]_{\vartheta=\vartheta_\lambda, \omega=\omega_\lambda} , \\ \overline{\omega_\lambda} &= \frac{d\omega_\lambda}{dt} \sin\vartheta_\lambda = -\left[\frac{\partial(\psi - \psi_\lambda)}{\partial\vartheta} \right]_{\vartheta=\vartheta_\lambda, \omega=\omega_\lambda} \end{aligned}\right\} \qquad (3)$$

or by

$$\left.\begin{aligned} \frac{d\vartheta_\lambda}{dt} \sin\vartheta_\lambda &= -\sum_{\mu\neq\lambda} m_\mu \frac{\partial\psi_{\lambda,\mu}}{\partial\omega_\lambda} - \frac{\partial\psi'(\vartheta_\lambda, \omega_\lambda)}{\partial\vartheta_\lambda} \\ \frac{d\omega_\lambda}{dt} \sin\vartheta_\lambda &= \sum_{\mu\neq\lambda} m_\mu \frac{\partial\psi_{\lambda,\mu}}{\partial\vartheta_\lambda} + \frac{\partial\psi'(\vartheta_\lambda, \omega_\lambda)}{\partial\vartheta_\lambda} \ , \end{aligned}\right\} \qquad (3)'$$

wenn

$$\psi_{\lambda,\mu} = \psi_\lambda(\vartheta_\mu, \omega_\mu) = \psi_\lambda(\vartheta_\lambda, \omega_\lambda) = \frac{1}{\pi} \lg \frac{r_{\lambda\mu}}{2} = \frac{1}{\pi} \lg \sin \frac{\delta_{\lambda\mu}}{2} \tag{4}$$

und $\delta_{\lambda\mu} = \delta(\vartheta_\lambda, \omega_\lambda; \vartheta_\mu, \omega_\mu)$ den Bogenabstand der Punkte P_λ und P_μ bezeichnet, so daß

$$\cos \delta_{\lambda,\mu} = \cos \vartheta_\lambda \cos \vartheta_\mu + \sin \vartheta_\lambda \sin \vartheta_\mu \cos(\omega_\lambda - \omega_\mu). \tag{5}$$

Durch die Differentialgleichungen $(3)'$ wird also die Bewegung der Strudel-
66 punkte bestimmt, während die Function ψ' und | mit ihr die gesamte Stromfunction $\psi = \psi' + \psi_1 + \cdots + \psi_n$ und die Wirbeldichte $\varrho = \frac{1}{2} D\psi = \frac{1}{2} D\psi' + \text{const}$ der partiellen Differentialgleichung genügen muß (I §4):

$$\frac{\partial D\psi}{\partial t} \sin \vartheta = \frac{\partial \psi}{\partial \omega} \frac{\partial D\psi}{\partial \vartheta} - \frac{\partial \psi}{\partial \vartheta} \frac{\partial D\psi}{\partial \omega} \tag{6}$$

oder

$$\frac{\partial D\psi'}{\partial t} \sin \vartheta = \left(\frac{\partial \psi'}{\partial \omega} + \frac{\partial \psi_1}{\partial \omega} + \cdots \right) \frac{\partial D\psi'}{\partial \vartheta} - \left(\frac{\partial \psi'}{\partial \vartheta} + \frac{\partial \psi_1}{\partial \vartheta} + \cdots \right) \frac{\partial D\psi'}{\partial \omega} \ .$$

Wir haben also im Ganzen $2n + 1$ Relationen für die $2n$ Functionen der Zeit $\vartheta_\lambda = \vartheta_\lambda(t)$, $\omega_\lambda = \omega_\lambda(t)$ und für die eine Function von Ort und Zeit: $\psi' = \psi'(\vartheta, \omega, t)$ zur Bestimmung des gesamten Bewegungsvorganges.

Die Strömung wird *stationär* (I §4 f., II §4), wenn

$$2\varrho = D\psi = D\psi' + \text{const} = f(\psi) = f(\psi' + \psi_1 + \psi_2 + \cdots + \psi_n) \tag{7}$$

und gleichzeitig beständig

$$\frac{d\vartheta_\lambda}{dt} = 0 \ , \quad \frac{d\omega_\lambda}{dt} = 0 \quad (\lambda = 1, 2, \ldots n) \, , \tag{8}$$

d. h. wenn alle Strudelpunkte fest bleiben, und sie wird „*rotierend stationär*“ (II §4 fin), wenn einmal

$$D\psi' = f(\psi + \alpha \cos \vartheta) \tag{9}$$

und sodann immer

$$\frac{d\vartheta_\lambda}{dt} = 0 \ , \quad \frac{d\omega_\lambda}{dt} = \alpha = \text{const} \qquad (\lambda = 1, 2, \ldots n), \tag{10}$$

d. h. wenn auch die Strudelpunkte mit der vorgeschriebenen Winkelgeschwindigkeit α rotieren.

if

$$\psi_{\lambda,\mu} = \psi_\lambda(\vartheta_\mu, \omega_\mu) = \psi_\lambda(\vartheta_\lambda, \omega_\lambda) = \frac{1}{\pi} \lg \frac{r_{\lambda\mu}}{2} = \frac{1}{\pi} \lg \sin \frac{\delta_{\lambda\mu}}{2} \tag{4}$$

and $\delta_{\lambda\mu} = \delta(\vartheta_\lambda, \omega_\lambda; \vartheta_\mu, \omega_\mu)$ denotes the arc distance of the points P_λ and P_μ so that

$$\cos \delta_{\lambda,\mu} = \cos \vartheta_\lambda \cos \vartheta_\mu + \sin \vartheta_\lambda \sin \vartheta_\mu \cos(\omega_\lambda - \omega_\mu). \tag{5}$$

Thus, the motion of the whirl points is determined by the differential equations $(3)'$, whereas the function ψ', and with it the entire stream function $\psi = \psi' + \psi_1 + \cdots + \psi_n$ and the vortex density $\varrho = \frac{1}{2} D\psi = \frac{1}{2} D\psi' + \text{const}$ must satisfy the partial differential equation (I §4):

$$\frac{\partial D\psi}{\partial t} \sin \vartheta = \frac{\partial \psi}{\partial \omega} \frac{\partial D\psi}{\partial \vartheta} - \frac{\partial \psi}{\partial \vartheta} \frac{\partial D\psi}{\partial \omega} \tag{6}$$

or

$$\frac{\partial D\psi'}{\partial t} \sin \vartheta = \left(\frac{\partial \psi'}{\partial \omega} + \frac{\partial \psi_1}{\partial \omega} + \cdots \right) \frac{\partial D\psi'}{\partial \vartheta} - \left(\frac{\partial \psi'}{\partial \vartheta} + \frac{\partial \psi_1}{\partial \vartheta} + \cdots \right) \frac{\partial D\psi'}{\partial \omega} \, .$$

Hence, all in all, we have $2n + 1$ relations for the $2n$ functions of time $\vartheta_\lambda = \vartheta_\lambda(t)$, $\omega_\lambda = \omega_\lambda(t)$ and for the one function of location and time: $\psi' = \psi'(\vartheta, \omega, t)$ for the determination of the entire motion.

The current becomes *stationary* (I §4 f., II §4) if

$$2\varrho = D\psi = D\psi' + \text{const} = f(\psi) = f(\psi' + \psi_1 + \psi_2 + \cdots + \psi_n) \tag{7}$$

and, at the same time, always

$$\frac{d\vartheta_\lambda}{dt} = 0 \, , \quad \frac{d\omega_\lambda}{dt} = 0 \quad (\lambda = 1, 2, \ldots n) \, , \tag{8}$$

i. e., if all whirl points remain fixed. It becomes "*rotationally stationary*" (II §4 fin) if once

$$D\psi' = f(\psi + \alpha \cos \vartheta) \tag{9}$$

and thereafter always

$$\frac{d\vartheta_\lambda}{dt} = 0 \, , \quad \frac{d\omega_\lambda}{dt} = \alpha = \text{const} \qquad (\lambda = 1, 2, \ldots n), \tag{10}$$

i. e., if the whirl points, too, rotate at the prescribed angle velocity α.

67 |

§3. Systeme von einfachen Strudeln

Wir beschränken uns jetzt auf den einfacheren Fall, wo, von den Strudelpunkten abgesehen, die Wirbeldichte ϱ auf der ganzen Kugel constant ist, $\varrho = -\frac{M}{4\pi}$, wenn mit M die Summe aller Strudelmomente bezeichnet wird. Wir haben dann einfach ein „System von n Strudeln“ in den Punkten $P_1, P_2, \ldots, P_n$ mit den Momenten $m_1, m_2, \ldots, m_n$ und der Stromfunktion

$$\left.\begin{aligned}\psi &= \frac{m_1}{\pi}\lg\sin\frac{\delta_1}{2} + \frac{m_2}{\pi}\lg\sin\frac{\delta_2}{2} + \cdots + \frac{m_n}{\pi}\lg\sin\frac{\delta_n}{2}\\ &= \epsilon_1\lg\frac{r_1}{2} + \epsilon_2\lg\frac{r_2}{2} + \cdots + \epsilon_n\lg\frac{r_n}{2}\ ,\end{aligned}\right\} \qquad (1)$$

wenn wieder unter $\delta_1, \delta_2, \ldots$ die sphärischen und unter $r_1, r_2, \ldots$ die Sehnenabstände von den Punkten $P_1, P_2, \ldots$ verstanden werden und zur Abkürzung gesetzt wird:

$$m_1 = \epsilon_1\,\pi,\ m_2 = \epsilon_2\pi,\ \ldots, m_n = \epsilon_n\pi.$$

Die Stromlinien $\psi = \text{const}$ sind dann *lemniskatische Curven*, deren Gleichung sich schreiben läßt:

$$r_1^{\epsilon_1}\, r_2^{\epsilon_2} \ldots r_n^{\epsilon_n} = \text{const}\,,$$

und lassen sich für jede Strudelverteilung ohne Schwierigkeit ihrem allgemeinen Verlaufe nach angeben, wenn man beachtet, daß zwischen zwei gleichstim-
68 migen Strudelpunkten immer eine | Gleichgewichtslage, in dem eine Stromlinie sich selbst durchschneidet, vorhanden sein und zwischen zwei entgegengesetzten die Flüssigkeit in einem durch ihre Vorzeichen bestimmten Sinne strömen wird.

Durch die jeweilige Configuration der Strudelpunkte oder einfach der „Strudel“ (die wir uns durch ihre Strudelpunkte $P_1, P_2, \ldots$ repräsentiert denken) ist der Strömungszustand in jedem Augenblicke vollständig bestimmt, durch ihre Bewegung also auch der ganze Verlauf des Vorganges. Nach §1 aber bewegt sich jeder Strudelpunkt mit einer solchen Geschwindigkeit, als ob nur die $n-1$ übrigen Strudel vorhanden wären. Wir können uns somit den Vorgang etwa in folgender Weise veranschaulichen:

Wir denken uns die Kugel von n in einander geschobenen starren aber gegen einander und die Kugel drehbaren Schalenflächen $K_1, K_2, \ldots, K_n$ überdeckt. Jede derselben soll zu einem Strudel $P_\mu(\epsilon_\mu)$ gehören und die entsprechende Strömung $\psi_\mu = \epsilon_\mu \lg\sin\frac{\delta_\mu}{2}$ als ein starres Vectorfeld aufgezeichnet
69 enthalten. Die wahre Geschwindigkeit der Flüssigkeit in einem be- | liebigen Kugelpunkte erhält man dann für jeden Augenblick, indem man bei der jeweiligen Lage der Schalen die n Geschwindigkeitsvectoren $q_1, q_2, \ldots, q_n$ an

§3. Systems of simple whirls

We now limit ourselves to the simpler case in which, except for the whirl points, the vortex density ϱ is constant on the entire sphere, $\varrho = -\dfrac{M}{4\pi}$, where M denotes the sum of all whirl momenta. We then simply have a "system of n whirls" in the points $P_1, P_2, \ldots, P_n$ with the momenta $m_1, m_2, \ldots, m_n$ and with the stream function

$$\left.\begin{aligned}\psi &= \frac{m_1}{\pi}\lg\sin\frac{\delta_1}{2} + \frac{m_2}{\pi}\lg\sin\frac{\delta_2}{2} + \cdots + \frac{m_n}{\pi}\lg\sin\frac{\delta_n}{2}\\ &= \epsilon_1\lg\frac{r_1}{2} + \epsilon_2\lg\frac{r_2}{2} + \cdots + \epsilon_n\lg\frac{r_n}{2}\ ,\end{aligned}\right\} \tag{1}$$

if, again, by $\delta_1, \delta_2, \ldots$ we denote the spherical distances and by $r_1, r_2, \ldots$ the chord distances from the points $P_1, P_2, \ldots$ and set for short:

$$m_1 = \epsilon_1\ \pi,\ m_2 = \epsilon_2\pi,\ \ldots, m_n = \epsilon_n\pi.$$

Then the streamlines $\psi = \text{const}$ are *lemniscatic curves*, whose equation can be written:

$$r_1^{\epsilon_1}\, r_2^{\epsilon_2} \ldots r_n^{\epsilon_n} = \text{const}\,,$$

and can be easily specified for every whirl distribution according to its general course, assuming we take into account that between two resonant whirl points there will always be a position of equilibrium in which a streamline intersects itself, and that between two opposite whirl points the fluid will flow in a direction determined by its sign.

The state of the current is completely determined at every moment by the respective configuration of the whirl points, or simply of the "whirls" (which we imagine to be represented by their whirl points $P_1, P_2, \ldots$), and hence their motion also determines the entire course of the process. But, by §1, each whirl point moves at such a velocity as if only the $n-1$ remaining whirls existed. Thus, we can illustrate the process as follows:

We assume that the sphere is covered by n rigid spherical shell surfaces $K_1, K_2, \ldots, K_n$ that are fitted into each other but can be rotated with respect to one another and to the sphere. Each of them shall belong to one whirl $P_\mu(\epsilon_\mu)$ and contain the corresponding current $\psi_\mu = \varepsilon_\mu \lg\sin\dfrac{\delta_\mu}{2}$ drawn upon it as a rigid vector field. Then the true velocity of the fluid at any point of the sphere is obtained for every instant by adding for the particular position of the shells the n velocity vectors $q_1, q_2, \ldots, q_n$ at the particular point according

der betreffenden Stelle nach dem Kräfteparallelogramm addiert. Die Bewegung der Schale K_λ selbst aber erhält man, indem man dem entsprechenden Strudelpunkte P_λ die durch Addition der $n-1$ übrigen Vectoren $q_1, q_2, \ldots$ an derselben Stelle entstehende Geschwindigkeit zuschreibt.

Die gegenseitige Einwirkung der Strudel können wir nach den Formeln (2)' in II §2 auch so ausdrücken:

Satz I. *Jeder Strudel P_λ erteilt jedem anderen P_μ eine Geschwindigkeit, deren Richtung senkrecht steht auf dem beide verbindenden grössten Kugelkreise und deren Grösse dem Momente m_λ des ersteren einerseits und der Cotangente des halben sphärischen Abstandes $\delta_{\lambda\mu}$ der beiden proportional ist. Der Proportionalitätsfactor ist dabei $\frac{1}{2\pi}$ und der Richtungssinn ist immer so zu wählen, wie es dem Drehungssinne des wirkenden Strudels P_λ entspricht, d. h. von P_λ aus betrachtet nach links oder nach rechts, je nachdem m_λ positiv oder negativ ist.*

Bezeichnen wir dafür mit $\vartheta_\lambda, \omega_\lambda$ die Polarkoordinaten des Strudels $m_\lambda =$
70 $\pi\epsilon_\lambda$ | und mit $\psi_\lambda = \epsilon_\lambda \lg \sin \frac{\delta_\lambda}{2} = \epsilon_\lambda \lg \frac{r_\lambda}{2}$ das ihm entsprechende Glied der Stromfunktion, so werden (cf. (3) §2) die Differentialgleichungen der Bewegung

$$\left.\begin{aligned} \frac{d\vartheta_\lambda}{dt}\sin\vartheta_\lambda &= -\left[\frac{\partial(\psi-\psi_\lambda)\psi_\lambda}{\partial\omega_\lambda}\right]_{\vartheta_\lambda,\omega_\lambda} = -\sum_{\mu\neq\lambda}\epsilon_\mu\frac{\partial \lg\frac{r_{\lambda\mu}}{2}}{\partial\omega_\lambda} \\ \frac{d\omega_\lambda}{dt}\sin\vartheta_\lambda &= +\left[\frac{\partial(\psi-\psi_\lambda)}{\partial\vartheta_\lambda}\right]_{\vartheta_\lambda,\omega_\lambda} = \sum_{\mu\neq\lambda}\epsilon_\mu\frac{\partial \lg\frac{r_{\lambda\mu}}{2}}{\partial\vartheta_\lambda}\ , \end{aligned}\right\} \quad (2)$$
$$(\lambda = 1, 2, \ldots, n)$$

wo $\psi = \psi_1 + \cdots + \psi_n$ und $r_{\lambda\mu} = 2\sin\frac{\delta_{\lambda\mu}}{2}$ den Sehnenabstand $P_\lambda P_\mu$ bezeichnet.

Führt man hier die Function ein

$$\mathcal{H} = \mathcal{H}(\vartheta_1, \omega_1; \vartheta_2, \omega_2; \ldots; \vartheta_n, \omega_n) = \sum_{\lambda,\mu}\epsilon_\lambda\epsilon_\mu \lg\frac{r_{\lambda\mu}}{2} = \sum_{\lambda,\mu}\epsilon_\lambda\epsilon_\mu \lg\sin\frac{r_{\delta\mu}}{2}, \quad (3)$$

so nehmen die Differentialgleichungen (2) die Gestalt an

$$\left.\begin{aligned} \epsilon_\lambda\sin\vartheta_\lambda\frac{d\vartheta_\lambda}{dt} &= -\frac{\partial\mathcal{H}}{\partial\omega_\lambda} \\ \epsilon_\lambda\sin\vartheta_\lambda\frac{d\omega_\lambda}{dt} &= +\frac{\partial\mathcal{H}}{\partial\vartheta_\lambda}\ , \quad (\lambda = 1, 2, \ldots, n) \end{aligned}\right\} \quad (4)$$

to the parallelogram of forces. But the motion of the shell K_λ itself is obtained by attributing to the corresponding whirl point P_λ the velocity arising from the addition of the $n-1$ remaining vectors $q_1, q_2, \ldots$ at the same point.

We may also express the mutual interaction of the whirls in accordance with formulas (2)' in II §2 as follows:

Theorem I. *Every whirl P_λ imparts on every other whirl P_μ a velocity whose direction is perpendicular to the great circle of the sphere joining both of them and whose magnitude is proportional to the momentum m_λ of the former on the one hand and to the cotangent of half the spherical distance $\delta_{\lambda\mu}$ between both of them on the other. The factor of proportionality is $\frac{1}{2\pi}$, and the sense of direction is to be chosen always corresponding to the sense of rotation of the acting whirl P_λ, i. e., counterclockwise or clockwise, considered from P_λ, depending on whether m_λ is positive or negative.*

To this end, we denote the polar coordinates of the whirl $m_\lambda = \pi\epsilon_\lambda$ by $\vartheta_\lambda, \omega_\lambda$, and the term of the stream function corresponding to it by $\psi_\lambda = \epsilon_\lambda \lg \sin \frac{\delta_\lambda}{2} = \epsilon_\lambda \lg \frac{r_\lambda}{2}$. Then (comp. (3) §2) the differential equations of the motion become

$$\left.\begin{aligned} \frac{d\vartheta_\lambda}{dt}\sin\vartheta_\lambda &= -\left[\frac{\partial(\psi-\psi_\lambda)\psi_\lambda}{\partial\omega_\lambda}\right]_{\vartheta_\lambda,\omega_\lambda} = -\sum_{\mu\neq\lambda}\epsilon_\mu\frac{\partial \lg \frac{r_{\lambda\mu}}{2}}{\partial\omega_\lambda} \\ \frac{d\omega_\lambda}{dt}\sin\vartheta_\lambda &= +\left[\frac{\partial(\psi-\psi_\lambda)}{\partial\vartheta_\lambda}\right]_{\vartheta_\lambda,\omega_\lambda} = \sum_{\mu\neq\lambda}\epsilon_\mu\frac{\partial \lg \frac{r_{\lambda\mu}}{2}}{\partial\vartheta_\lambda}\ , \end{aligned}\right\} \quad (2)$$

$$(\lambda = 1, 2, \ldots, n)$$

where $\psi = \psi_1 + \cdots + \psi_n$ and $r_{\lambda\mu} = 2\sin\frac{\delta_{\lambda\mu}}{2}$ denote the chord distance $P_\lambda P_\mu$.

If one introduces here the function

$$\mathcal{H} = \mathcal{H}(\vartheta_1, \omega_1; \vartheta_2, \omega_2; \ldots; \vartheta_n, \omega_n) = \sum_{\lambda,\mu}\epsilon_\lambda\epsilon_\mu \lg\frac{r_{\lambda\mu}}{2} = \sum_{\lambda,\mu}\epsilon_\lambda\epsilon_\mu \lg\sin\frac{r_{\delta\mu}}{2}, \quad (3)$$

then the differential equations (2) assume the form

$$\left.\begin{aligned} \epsilon_\lambda\sin\vartheta_\lambda\frac{d\vartheta_\lambda}{dt} &= -\frac{\partial\mathcal{H}}{\partial\omega_\lambda} \\ \epsilon_\lambda\sin\vartheta_\lambda\frac{d\omega_\lambda}{dt} &= +\frac{\partial\mathcal{H}}{\partial\vartheta_\lambda}\ , \quad (\lambda = 1, 2, \ldots, n) \end{aligned}\right\} \quad (4)$$

welche (vergl. Poincaré a. a. O. §67 für ebene Strudel) der *Hamilton*schen Form der dynamischen Differentialgleichungen entspricht und die wir als die „*Normalform*“ der Bewegungsgleichungen bezeichnen wollen. Die Funktion $\mathcal{H}$ nennen wir das „*Selbstpotential*“ des Strudelsystemes.

71 |

§4. Integrale der Bewegungsgleichungen

1) Aus den Gleichungen (4) des vorigen § folgt unmittelbar:

$$\frac{\partial \mathcal{H}}{\partial \vartheta_\lambda}\frac{d\vartheta_\lambda}{dt} + \frac{\partial \mathcal{H}}{\partial \omega_\lambda}\frac{d\omega_\lambda}{dt} = 0$$

und daher durch Summation über $\lambda = 1, 2, \ldots, n$:

$$\frac{d\mathcal{H}}{dt} = \sum_{\lambda=1}^{n}\left(\frac{\partial \mathcal{H}}{\partial \vartheta_\lambda}\frac{d\vartheta_\lambda}{dt} + \frac{\partial \mathcal{H}}{\partial \omega_\lambda}\frac{d\omega_\lambda}{dt}\right) = 0 \quad \text{oder} \quad \mathcal{H} = \text{ const.} \tag{1}$$

Satz I. *Das Selbstpotential eines Strudelsystems bleibt während der ganzen Bewegung constant.*

Dieser Satz ist aber nur eine Folge das in Cap. I §5 aufgestellten Satzes von der Erhaltung der lebendigen Kraft. Es ist nämlich nach (4a) p. 30:

$$\overset{(K)}{\int} \psi \varrho d\sigma = -T = \text{ const.} \tag{2}$$

Nun ist aber nach II §2 die Stromfunction ψ gleich dem sphärischen Potentiale

$$\psi = \frac{1}{\pi}\overset{(K)}{\int} \varrho' d\sigma' \lg \frac{r}{2},$$

und daher

$$-T = \frac{1}{\pi}\overset{(K)}{\int} \varrho d\sigma \overset{(K)}{\int} \varrho' d\sigma' \lg \frac{r}{2} = \frac{1}{\pi}\int\int \varrho d\sigma \varrho' d\sigma' \lg \frac{r}{2} = \text{ const} \tag{3}$$

wo r den Sehnenabstand der Elemente $d\sigma$, $d\sigma'$ bezeichnet.

Diese Formel geht aber für den Fall, daß nur Strudelpunkte m und sonst keine flächenhaften Wirbel vorhanden sind ($\varrho = 0$) durch Grenzübergang in die folgende über

$$-T = \frac{2}{\pi}\sum_{\lambda,\mu} m_\lambda m_\mu \lg \frac{r_{\lambda\mu}}{2} = 2\pi \sum_{\lambda,\mu} \epsilon_\lambda \epsilon_\mu \lg \frac{r_{\lambda\mu}}{2} = -\mathcal{H}, \tag{4}$$

72 | wenn man die Selbstpotentiale der einzelnen Strudel für sich vernachlässigt, welche zwar beim Grenzübergang unendlich werden, aber ihrer Natur nach

which (comp. *Poincaré 1893a*, §67 for whirls in the plane) correspond to the *Hamilton*ian form of the dynamic differential equations and which we call the "*normal form*" of the equations of motion. We call the function $\mathcal{H}$ the "*self-potential*" of the whirl system.

§4. Integrals of the equations of motion

1) From equations (4) of the previous § it immediately follows that

$$\frac{\partial \mathcal{H}}{\partial \vartheta_\lambda}\frac{d\vartheta_\lambda}{dt} + \frac{\partial \mathcal{H}}{\partial \omega_\lambda}\frac{d\omega_\lambda}{dt} = 0$$

and hence, by summation over $\lambda = 1, 2, \ldots, n$,

$$\frac{d\mathcal{H}}{dt} = \sum_{\lambda=1}^{n}\left(\frac{\partial \mathcal{H}}{\partial \vartheta_\lambda}\frac{d\vartheta_\lambda}{dt} + \frac{\partial \mathcal{H}}{\partial \omega_\lambda}\frac{d\omega_\lambda}{dt}\right) = 0 \quad \text{or} \quad \mathcal{H} = \text{ const.} \tag{1}$$

Theorem I. *The self-potential of a whirl system remains constant during the entire motion.*

This theorem is, however, but a corollary of the theorem established in chap. I §5 on the conservation of the living force. For, by (4a) p. 30,

$$\overset{(K)}{\int} \psi \varrho d\sigma = -T = \text{ const.} \tag{2}$$

But now, by II §2, the stream function ψ is equal to the spherical potential

$$\psi = \frac{1}{\pi}\overset{(K)}{\int} \varrho' d\sigma' \lg\frac{r}{2},$$

and therefore

$$-T = \frac{1}{\pi}\overset{(K)}{\int} \varrho d\sigma \overset{(K)}{\int} \varrho' d\sigma' \lg\frac{r}{2} = \frac{1}{\pi}\int\int \varrho d\sigma \varrho' d\sigma' \lg\frac{r}{2} = \text{ const} \tag{3}$$

where r denotes the chord distance of the elements $d\sigma$, $d\sigma'$.

But in the case where there exist only whirl points m and no other surface-like whirls ($\varrho = 0$) this formula is transformed by a passage to the limit into

$$-T = \frac{2}{\pi}\sum_{\lambda,\mu} m_\lambda m_\mu \lg\frac{r_{\lambda\mu}}{2} = 2\pi\sum_{\lambda,\mu}\epsilon_\lambda\epsilon_\mu \lg\frac{r_{\lambda\mu}}{2} = -\mathcal{H}, \tag{4}$$

if the self-potentials of the individual whirls themselves are disregarded, which, as the limit is taken, do become infinite but are constant in nature.

konstant sind. Auch für den Fall $\varrho = -\dfrac{M}{4\pi} \neq 0$ würden, wie leicht zu sehen, lediglich constante Glieder hinzutreten. So ergiebt sich in jedem Falle $\mathcal{H} =$ const.

2) Das Selbstpotential $\mathcal{H}$ des Systemes ist eine Function der Abstände $r_{\lambda,\mu} = 2\sin\dfrac{\delta_{\lambda\mu}}{2}$ *allein*, und es ist

$$\cos\delta_{\lambda\mu} = 1 - \frac{1}{2}r_{\lambda\mu}^2 = \cos\vartheta_\lambda\cos\vartheta_\mu + \sin\vartheta_\lambda\sin\vartheta_\mu\cos(\omega_\lambda - \omega_\mu)\,,$$

so daß

$$\left.\begin{aligned} &\frac{\partial\cos\delta_{\lambda\mu}}{\partial\omega_\lambda} + \frac{\partial\cos\delta_{\lambda\mu}}{\partial\omega_\mu} = 0 \quad \text{und ebenso} \\ &\frac{\partial r_{\lambda\mu}}{\partial\omega_\lambda} + \frac{\partial r_{\lambda\mu}}{\partial\omega_\mu} = 0\,. \end{aligned}\right\} \tag{5}$$

Hieraus folgt aber durch Summation

$$\sum_{\lambda=1}^{n}\frac{\partial\mathcal{H}}{\partial\omega_\lambda} = \sum_{\lambda,\mu}\frac{\partial\mathcal{H}}{\partial r_{\lambda,\mu}}\left(\frac{\partial r_{\lambda\mu}}{\partial\omega_\lambda} + \frac{\partial r_{\lambda\mu}}{\partial\omega_\mu}\right) = 0$$

und daher wegen (4) von §3:

$$\sum_{\lambda=1}^{n}\epsilon_\lambda\frac{d\cos\vartheta_\lambda}{dt} = -\sum_{\lambda=1}^{n}\epsilon_\lambda\sin\vartheta_\lambda\frac{d\vartheta_\lambda}{dt} = \sum_{\lambda=1}^{n}\frac{\partial\mathcal{H}}{\partial\omega_\lambda} = 0$$

oder integriert:

$$L = \sum_{\lambda=1}^{n}\epsilon_\lambda\cos\vartheta_\lambda = \text{ const.} \tag{6}$$

Dies gilt für *jede* Lage des Polarcoordinatensystemes, also ist auch, wenn ξ, η, ζ beliebige rechtwinklige Coordinaten sind:

73 | $$\epsilon\,\xi_0 = \sum_\lambda\epsilon_\lambda\xi_\lambda = \text{const},\ \epsilon\,\eta_0 = \sum_\lambda\epsilon_\lambda\eta_\lambda = \text{const},\ \epsilon\,\zeta_0 = \sum_\lambda\epsilon_\lambda\zeta_\lambda = \text{ const},\tag{6}'$$

und hier sind ξ_0, η_0, ζ_0 die Coordinaten des wahren Schwerpunktes der Massen $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$, wenn $\epsilon = \epsilon_1 + \epsilon_2 + \cdots + \epsilon_n$ gesetzt wird.

Satz II. (*Erhaltung des Schwerpunktes*) *Der wahre Schwerpunkt eines Strudelsystemes, dessen Strudelmomente $m_\lambda = \pi\epsilon_\lambda$ als Massen aufgefasst werden, ist ein* fester *Punkt S im Raume bei allen Bewegungen des Systemes.*

In the case where $\varrho = -\dfrac{M}{4\pi} \neq 0$ there would also appear only additional constant terms, as is readily seen. Thus, in any case, $\mathcal{H} = \text{const}$.

2) The self-potential $\mathcal{H}$ of the system is a function of the distances $r_{\lambda,\mu} = 2\sin\dfrac{\delta_{\lambda\mu}}{2}$ *only*, and it is

$$\cos\delta_{\lambda\mu} = 1 - \frac{1}{2}r_{\lambda\mu}^2 = \cos\vartheta_\lambda\cos\vartheta_\mu + \sin\vartheta_\lambda\sin\vartheta_\mu\cos(\omega_\lambda - \omega_\mu)\,,$$

so that

$$\left.\begin{aligned} &\frac{\partial\cos\delta_{\lambda\mu}}{\partial\omega_\lambda} + \frac{\partial\cos\delta_{\lambda\mu}}{\partial\omega_\mu} = 0 \quad \text{and likewise} \\ &\frac{\partial r_{\lambda\mu}}{\partial\omega_\lambda} + \frac{\partial r_{\lambda\mu}}{\partial\omega_\mu} = 0\,. \end{aligned}\right\} \tag{5}$$

But from this it follows by summation that

$$\sum_{\lambda=1}^{n}\frac{\partial\mathcal{H}}{\partial\omega_\lambda} = \sum_{\lambda,\mu}\frac{\partial\mathcal{H}}{\partial r_{\lambda,\mu}}\left(\frac{\partial r_{\lambda\mu}}{\partial\omega_\lambda} + \frac{\partial r_{\lambda\mu}}{\partial\omega_\mu}\right) = 0$$

and therefore, on account of (4) in §3,

$$\sum_{\lambda=1}^{n}\epsilon_\lambda\frac{d\cos\vartheta_\lambda}{dt} = -\sum_{\lambda=1}^{n}\epsilon_\lambda\sin\vartheta_\lambda\frac{d\vartheta_\lambda}{dt} = \sum_{\lambda=1}^{n}\frac{\partial\mathcal{H}}{\partial\omega_\lambda} = 0$$

or, when integrated,

$$L = \sum_{\lambda=1}^{n}\epsilon_\lambda\cos\vartheta_\lambda = \text{ const.} \tag{6}$$

This holds for *every* position of the polar coordinate system, and hence, assuming ξ, η, ζ are any orthogonal coordinates, we also have

$$\epsilon\,\xi_0 = \sum_\lambda\epsilon_\lambda\xi_\lambda = \text{const},\ \epsilon\,\eta_0 = \sum_\lambda\epsilon_\lambda\eta_\lambda = \text{const},\ \epsilon\,\zeta_0 = \sum_\lambda\epsilon_\lambda\zeta_\lambda = \text{ const}, \tag{6}'$$

where ξ_0, η_0, ζ_0 are the coordinates of the true center of gravity of the masses $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$, if one sets $\epsilon = \epsilon_1 + \epsilon_2 + \cdots + \epsilon_n$.

Theorem II. (*Conservation of the center of gravity*) *The true center of gravity of a whirl system whose whirl momenta $m_\lambda = \pi\epsilon_\lambda$ are conceived of as masses is a* fixed *point S in space for all motions of the system.*

Natürlich ist dieser Satz eine Folge der Schwerpunktssätze C. II §3, wenn man den bekannten Grenzübergang von der kontinuierlichen zur diskontinuierlichen Wirbelverteilung vornimmt. Dabei können die von der konstanten Wirbeldichte $\varrho = -\frac{M}{4\pi} = -\frac{\epsilon}{4}$ herrührenden Beiträge zu L_x, L_y, L_z als so wie so konstant unterdrückt werden. Während aber dort der wahre Schwerpunkt wegen ${}^{(K)}\int \varrho d\sigma = 0$ immer ins Unendliche fiel, kann er hier im Endlichen bleiben, wenn nämlich die Summe $M = \pi\epsilon$ aller Strudelmomente und damit auch die konstante Wirbeldichte ϱ von 0 verschieden ist.

Vermöge dieses „Schwerpunktsatzes“ können wir in allen Fällen unser Problem etwas vereinfachen, indem wir die 2. Axe CO unseres Polarkoordi-
74 natensystems | durch den Schwerpunkt S legen. In diesem Falle wird immer

$$\epsilon\,\xi_0 = \sum_\lambda \sin\vartheta_\lambda \cos\omega_\lambda = 0\,, \quad \epsilon\,\eta_0 = \sum_\lambda \epsilon_\lambda \sin\vartheta_\lambda \sin\omega_\lambda = 0 \qquad (6)''$$

$$\epsilon\,h = \sum_\lambda \epsilon_\lambda \cos\vartheta_\lambda = \text{ const},$$

wo h die Entfernung des Schwerpunktes S vom Kugelmittelpunkte bedeutet.

So haben wir vermittelst unserer Sätze I und II bereits 4 von einander unabhängige Integrale unserer $2n$ Differentialgleichungen (4) p. 70 gefunden. Mehr als diese 4 lassen sich aber auch allgemein nicht aufstellen. In der Ebene, wo der analoge Schwerpunktsatz nur 2, *nicht* wie hier 3 unabhängige Integrale liefert (Poincaré a. a. O. Nr. 70) wird ein 4tes Integral gewonnen durch den Satz von der „Constanz des Trägheitsmomentes“ (Poincaré Nr. 75). Ein entsprechender Satz gilt allerdings auch für die Kugel, ist hier aber lediglich eine *Folge der 3 Schwerpunktsätze.*

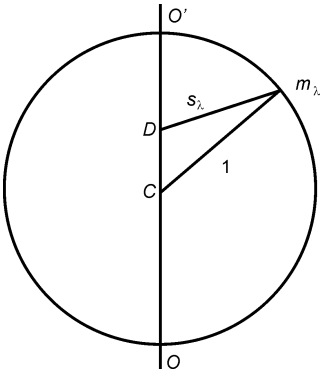


Fig. 1.

This theorem is, of course, a corollary of the theorems on centers of gravity C. II §3 if one takes the well-known limit from the continuous to the discontinuous whirl distribution. In this case, it is possible to suppress the contributions made to L_x, L_y, L_z by the constant whirl density $\varrho = -\frac{M}{4\pi} = -\frac{\epsilon}{4}$ since they are constant anyway. But while, in the former case, the true center of gravity always became infinite on account of ${}^{(K)}\int \varrho d\sigma = 0$, in this case, it can remain finite, namely if the sum $M = \pi\epsilon$ of all whirl momenta, and hence also the constant vortex density ϱ, is different from 0.

By virtue of this "theorem on gravity centers" we can somewhat simplify our problem in all cases by letting the 2nd axis of our polar coordinate system CO pass through the center of gravity S. In this case, always

$$\epsilon\,\xi_0 = \sum_\lambda \sin\vartheta_\lambda \cos\omega_\lambda = 0\,, \quad \epsilon\,\eta_0 = \sum_\lambda \epsilon_\lambda \sin\vartheta_\lambda \sin\omega_\lambda = 0 \qquad (6)''$$

$$\epsilon\,h = \sum_\lambda \epsilon_\lambda \cos\vartheta_\lambda = \text{ const},$$

where h denotes the distance of the center of gravity S from the center of the sphere.

Thus, using our theorems I and II, we have already found 4 mutually independent integrals of our $2n$ differential equations (4) p. 70. It is, however, not possible to establish more than these 4 generally. In the plane, where the analogous theorem on centers of gravity yields only 2 independent integrals and *not* 3, as in the present case (*Poincaré 1893a*, no. 70), a 4th integral is obtained by means of the theorem on the "constancy of the moment of inertia" (*Poincaré 1893a*, no. 75). While a corresponding theorem also holds for the sphere, in this case, it is but a *corollary of the 3 theorems on centers of gravity.*

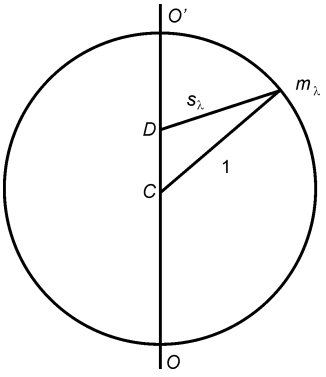


Fig. 1.

Es sei nämlich D ein beliebiger Punkt auf der (willkürlich gewählten) Coordinatenachse $O'CO$ und $CD = a$. Ist nun m_λ ein beliebiger Strudel und $Dm_\lambda = s_\lambda$, so wird in dem Dreieck CDm_λ:

$$s_\lambda^2 = 1 + a^2 - 2a\cos\vartheta_\lambda\,,$$

75 | also durch Summation über *alle* Strudel:

$$\begin{aligned}\sum_\lambda m_\lambda s_\lambda^2 &= (1+a^2)\sum_\lambda m_\lambda - 2a\sum m_\lambda\cos\vartheta_\lambda \\ &= (1+a^2)M - 2aM\zeta_0 = \text{ const}\,,\end{aligned}$$

wenn $M = \sum m_\lambda$ die Summe der Strudelmomente und ζ_0 die Coordinate ihres Schwerpunktes S in der Richtung CD ist. Wir haben also den weiteren Satz:

Satz III. *Die Summe der Momente aller Strudel, jedes multipliziert mit dem Quadrate seines Abstandes von einem beliebigen festen Punkte D im Raume ist bei der ganzen Bewegung constant. Dies ist eine Folge aus dem Satze von der Erhaltung des Schwerpunktes.*

§5. Die Strudelbewegung in stereographischen und conjugierten Coordinaten

In diesem Capitel III haben wir bisher alle Entwicklungen ausschließlich in Polarcoordinaten ϑ, ω gegeben, nachdem wir bereits Cap II §1 auch stereographische Coordinaten x, y eingeführt hatten. Auf die letzteren wollen wir jetzt wieder zurückkommen, doch in der modifizierten Form als „conjugierte
76 Coordinaten“ $(z,\overline{z})$, wie sie in der Sphärik | häufig angewendet werden (Vergl. namentlich Klein, Das Ikosaeder K. II) und sich auch im Folgenden zuweilen als nützlich erweisen werden. Wir setzen nämlich (cf. II §1)

$$\left.\begin{aligned} z = x + iy &= \tan\frac{\vartheta}{2}(\cos\omega + i\sin\omega) = \tan\frac{\vartheta}{2}e^{i\omega} \\ \overline{z} = x - iy &= \tan\frac{\vartheta}{2}(\cos\omega - i\sin\omega) = \tan\frac{\vartheta}{2}e^{-i\omega}\,, \end{aligned}\right\} \qquad (1)$$

wo i die complexe Einheit $\sqrt{-1}$ bedeutet. Es wird dann

$$z\overline{z} = \tan^2\frac{\vartheta}{2} - x^2 + y^2\ ,\ 1 + z\overline{z} = 1 + x^2 + y^2 = \frac{1}{\cos^2\frac{\vartheta}{2}}\,,$$

und

$$\cos\vartheta = \frac{1 - z\overline{z}}{1 + z\overline{z}}\,,\quad \sin\vartheta = \frac{2\sqrt{z\overline{z}}}{1 + z\overline{z}}\,. \qquad (2)$$

For let D be any point on the (arbitrarily chosen) coordinate axis $O'CO$ and $CD = a$. Now if m_λ is any whirl, and $Dm_\lambda = s_\lambda$, then in the triangle CDm_λ

$$s_\lambda^2 = 1 + a^2 - 2a\cos\vartheta_\lambda\,,$$

and hence by summation over *all* whirls

$$\begin{aligned}\sum_\lambda m_\lambda s_\lambda^2 &= (1+a^2)\sum_\lambda m_\lambda - 2a\sum m_\lambda\cos\vartheta_\lambda \\ &= (1+a^2)M - 2aM\zeta_0 = \text{ const}\end{aligned}$$

if $M = \sum m_\lambda$ is the sum of the whirl momenta, and ζ_0 the coordinate of its center of gravity S in the direction CD. We thus have the further theorem:

Theorem III. *The sum of the momenta of all whirls, each multiplied by the square of its distance from any fixed point D in space, is constant during the entire motion. This is a corollary of the theorem on the conservation of the center of gravity.*

§5. The whirl motion in stereographic and conjugate coordinates

In the present chapter III we have so far provided all developments exclusively in polar coordinates ϑ, ω, after also having introduced stereographic coordinates x, y already in chap. II §1. We shall now return to the latter, albeit in the modified form of "conjugate coordinates" $(z,\overline{z})$, as they are often used in the study of spheres (comp. in particular *F. Klein 1884*, ch. II), which will at times prove useful in what follows. For we set (comp. II §1)

$$\left.\begin{aligned} z = x + iy &= \tan\frac{\vartheta}{2}(\cos\omega + i\sin\omega) = \tan\frac{\vartheta}{2}e^{i\omega} \\ \overline{z} = x - iy &= \tan\frac{\vartheta}{2}(\cos\omega - i\sin\omega) = \tan\frac{\vartheta}{2}e^{-i\omega}\,, \end{aligned}\right\} \tag{1}$$

where i denotes the complex unit $\sqrt{-1}$. Then

$$z\overline{z} = \tan^2\frac{\vartheta}{2} - x^2 + y^2\ ,\ 1 + z\overline{z} = 1 + x^2 + y^2 = \frac{1}{\cos^2\frac{\vartheta}{2}}$$

and

$$\cos\vartheta = \frac{1 - z\overline{z}}{1 + z\overline{z}}\,,\quad \sin\vartheta = \frac{2\sqrt{z\overline{z}}}{1 + z\overline{z}}\,. \tag{2}$$

Diese Coordinaten $z, \overline{z}$ sind besonders deshalb von Bedeutung, weil sich in ihnen alle Drehungen der Kugel als lineare Transformationen darstellen. Für den Sehnenabstand zweier Kugelpunkte z_λ und z_μ erhält man:

$$r_{\lambda\mu}^2 = 4\frac{(z_\lambda - z_\mu)(\overline{z_\lambda} - \overline{z_\mu})}{(1 + z_\lambda\overline{z_\lambda})(1 + z_\mu\overline{z_\mu})}\,. \tag{3}$$

Denn der Ausdruck rechts ist invariant gegen alle Kugeldrehungen, d. h. gegen die Transformationen

$$z' = \frac{\alpha z - \overline{\beta}}{\beta z + \overline{\alpha}} \quad \text{(cf. Klein, a. a. O. p. 34)},$$

wo auch $\alpha, \overline{\alpha}, \beta, \overline{\beta}$ conjugiert complexe Größen sind, und er nimmt für $z_\mu = 0$, $\overline{z_\mu} = 0$ in der That den Wert an:

$$\frac{4z_\lambda\overline{z_\lambda}}{1 + z_\lambda\overline{z_\lambda}} = 4\tan^2\frac{\vartheta_\lambda}{2}\cos^2\frac{\vartheta_\lambda}{2} = 4\sin^2\frac{\vartheta_1}{2} = r_\lambda^2\,.$$

77 | Daher wird der Ausdruck für das *Selbstpotential* unseres Strudelsystemes (cf. (3) §3 auf S. 70):

$$\mathcal{H} = \sum_{\lambda,\mu}\epsilon_\lambda\epsilon_\mu \lg\frac{r_{\lambda\mu}}{2} = \frac{1}{2}\sum_{\lambda,\mu}\epsilon_\lambda\epsilon_\mu \lg\frac{(z_\lambda - z_\mu)(\overline{z_\lambda} - \overline{z_\mu})}{(1 + z_\lambda\overline{z_\lambda})(1 + z_\mu\overline{z_\mu})} = \text{ const} \tag{4}$$

oder, wenn $\mathcal{H} = \lg Q = \frac{1}{2}\lg Q^2$ gesetzt wird:

$$Q^2 = \prod_{\lambda,\mu}\left(\frac{(z_\lambda - z_\mu)(\overline{z_\lambda} - \overline{z_\mu})}{(1 + z_\lambda\overline{z_\lambda})(1 + z_\mu\overline{z_\mu})}\right)^{\epsilon_\lambda\epsilon_\mu} = \frac{\prod_{\lambda,\mu}(z_\lambda - z_\mu)^{\epsilon_\lambda\epsilon_\mu}(\overline{z_\lambda} - \overline{z_\mu})^{\epsilon_\lambda\epsilon_\mu}}{\prod_\lambda(1 + z_\lambda\overline{z_\lambda})^{\epsilon_\lambda(\epsilon - \epsilon_\lambda)}} \tag{4'}$$

$$(\lambda, \mu = 1, 2, \ldots, n; \quad \epsilon = \epsilon_1 + \epsilon_2 + \cdots + \epsilon_n).$$

Durch Umrechnung der Gleichungen (4) in §3 auf die „conjugierten Coordinaten“ $z, \overline{z}$ erhält man nun die *Bewegungsgleichungen* in der Form:

$$\left.\begin{aligned} \epsilon_\lambda\frac{dz_\lambda}{dt} &= \frac{i}{2}(1 + z_\lambda\overline{z_\lambda})^2\frac{\partial\mathcal{H}}{\partial\overline{z_\lambda}} \\ \epsilon_\lambda\frac{d\overline{z_\lambda}}{dt} &= -\frac{i}{2}(1 + z_\lambda\overline{z_\lambda})^2\frac{\partial\mathcal{H}}{\partial z_\lambda} \end{aligned}\right\} \quad (\lambda = 1, 2, \ldots, n) \tag{5}$$

und aus ihnen leicht die für stereographische Coordinaten x, y:

$$\left.\begin{aligned} \epsilon_\lambda\frac{dx_\lambda}{dt} &= -\frac{1}{4}(1 + x_\lambda^2 + y_\lambda^2)^2\frac{\partial\mathcal{H}}{\partial y_\lambda} \\ \epsilon_\lambda\frac{dy_\lambda}{dt} &= +\frac{1}{4}(1 + x_\lambda^2 + y_\lambda^2)^2\frac{\partial\mathcal{H}}{\partial x_\lambda}\,, \end{aligned}\right\} \tag{6}$$

die sich auch unmittelbar aus (8b) p. 35 ergeben hätten.

These coordinates $z, \overline{z}$ are of significance in particular because all rotations of the sphere are represented as linear transformations in terms of them. For the chord distance of two points on the sphere z_λ and z_μ one obtains

$$r_{\lambda\mu}^2 = 4\frac{(z_\lambda - z_\mu)(\overline{z_\lambda} - \overline{z_\mu})}{(1 + z_\lambda\overline{z_\lambda})(1 + z_\mu\overline{z_\mu})}\,. \tag{3}$$

For the expression on the right side is invariant with respect to all rotations of the sphere, i. e., with respect to the transformations

$$z' = \frac{\alpha z - \overline{\beta}}{\beta z + \overline{\alpha}} \quad \text{(see } \textit{F. Klein 1884}\text{, p. 34)},$$

where α, $\overline{\alpha}$, β, $\overline{\beta}$, too, are conjugate complex magnitudes, and, for $z_\mu = 0$, $\overline{z_\mu} = 0$ it in fact assumes the value

$$\frac{4z_\lambda\overline{z_\lambda}}{1 + z_\lambda\overline{z_\lambda}} = 4\tan^2\frac{\vartheta_\lambda}{2}\cos^2\frac{\vartheta_\lambda}{2} = 4\sin^2\frac{\vartheta_1}{2} = r_\lambda^2\,.$$

Thus the expression for the *self-potential* of our whirl system (see (3) §3 on p. 70) becomes

$$\mathcal{H} = \sum_{\lambda,\mu}\epsilon_\lambda\epsilon_\mu \lg\frac{r_{\lambda\mu}}{2} = \frac{1}{2}\sum_{\lambda,\mu}\epsilon_\lambda\epsilon_\mu \lg\frac{(z_\lambda - z_\mu)(\overline{z_\lambda} - \overline{z_\mu})}{(1 + z_\lambda\overline{z_\lambda})(1 + z_\mu\overline{z_\mu})} = \text{ const} \tag{4}$$

or, if one sets $\mathcal{H} = \lg Q = \frac{1}{2}\lg Q^2$,

$$Q^2 = \prod_{\lambda,\mu}\left(\frac{(z_\lambda - z_\mu)(\overline{z_\lambda} - \overline{z_\mu})}{(1 + z_\lambda\overline{z_\lambda})(1 + z_\mu\overline{z_\mu})}\right)^{\epsilon_\lambda\epsilon_\mu} = \frac{\prod_{\lambda,\mu}(z_\lambda - z_\mu)^{\epsilon_\lambda\epsilon_\mu}(\overline{z_\lambda} - \overline{z_\mu})^{\epsilon_\lambda\epsilon_\mu}}{\prod_\lambda(1 + z_\lambda\overline{z_\lambda})^{\epsilon_\lambda(\epsilon - \epsilon_\lambda)}} \tag{4'}$$

$$(\lambda, \mu = 1, 2, \ldots, n; \quad \epsilon = \epsilon_1 + \epsilon_2 + \cdots + \epsilon_n).$$

Through conversion of equations (4) in §3 to the "conjugate coordinates" $z, \overline{z}$ one now obtains the *equations of the motion* in the form

$$\left.\begin{aligned} \epsilon_\lambda\frac{dz_\lambda}{dt} &= \frac{i}{2}(1 + z_\lambda\overline{z_\lambda})^2\frac{\partial\mathcal{H}}{\partial\overline{z_\lambda}} \\ \epsilon_\lambda\frac{d\overline{z_\lambda}}{dt} &= -\frac{i}{2}(1 + z_\lambda\overline{z_\lambda})^2\frac{\partial\mathcal{H}}{\partial z_\lambda} \end{aligned}\right\} \quad (\lambda = 1, 2, \ldots, n) \tag{5}$$

and from them one easily obtains for stereographic coordinates x, y

$$\left.\begin{aligned} \epsilon_\lambda\frac{dx_\lambda}{dt} &= -\frac{1}{4}(1 + x_\lambda^2 + y_\lambda^2)^2\frac{\partial\mathcal{H}}{\partial y_\lambda} \\ \epsilon_\lambda\frac{dy_\lambda}{dt} &= +\frac{1}{4}(1 + x_\lambda^2 + y_\lambda^2)^2\frac{\partial\mathcal{H}}{\partial x_\lambda}\,, \end{aligned}\right\} \tag{6}$$

which could have been immediately obtained also from (8b) p. 35.

Endlich nehmen die *Schwerpunktsgleichungen* (6)″ p. 74 in conjugierten Coordinaten die Form an:

$$\left.\begin{aligned}
&\epsilon = (\xi_0 + i\eta_0) = \epsilon z_0 = \sum_{\lambda=1}^{n} \epsilon_\lambda \frac{2z_\lambda}{1 + z_\lambda \overline{z_\lambda}} = \text{ const} \\
&\epsilon = (\xi_0 - i\eta_0) = \epsilon \overline{z_0} = \sum_{\lambda=1}^{n} \epsilon_\lambda \frac{2\overline{z_\lambda}}{1 + z_\lambda \overline{z_\lambda}} = \text{ const} \\
&\epsilon\zeta_0 = \sum_{\lambda=1}^{n} \epsilon_\lambda \frac{1 - z_\lambda \overline{z_\lambda}}{1 + z_\lambda \overline{z_\lambda}} = \text{ const} \\
&\qquad\qquad\qquad\qquad (\epsilon = \epsilon_1 + \epsilon_2 + \cdots \epsilon_n)\,,
\end{aligned}\right\} \tag{7}$$

78 |

wo ξ_0, μ_0, ζ_0 die rechtwinkligen Coordinaten des Schwerpunktes S bezeichnen, von denen bei specieller Wahl des Coordinatensystemes (cf. p. 73) gleichzeitig $\xi_0 = 0$ und $\eta_0 = 0$ und $\zeta_0 = h$ gesetzt werden kann, wo h der Abstand des Schwerpunktes vom Kugelmittelpunkt bedeutet.

§6. Das allgemeine Gleichgewichtsproblem

Der einfachste Grenzfall der Bewegung ist der des *Gleichgewichtes* der Ruhe, der Fall, wo alle n Strudel ihre Anfangslage dauernd behalten und daher die ganze Strömung *stationär* ist. Nach unseren Bewegungsgleichungen (4) in §2 oder (5) und (6) in §5 ist dies nur möglich, wenn die partiellen Ableitungen des Selbstpotentiales $\mathcal{H}$ nach allen $2n$ sphärischen Coordinaten des Strudelsystemes gleichzeitig verschwinden, oder m. a. W. wenn das Selbstpotential den ersten Bedingungen eines Maximums oder Minimums genügt.

79 | Die Bedingungen des Gleichgewichtes wären also in Polarcoordinaten, „stereographischen“ und „conjugierten Coordinaten“:

$$\left.\begin{aligned}
&\text{a)} \quad \frac{\partial \mathcal{H}}{\partial \vartheta_\lambda} = 0, \quad \frac{\partial \mathcal{H}}{\partial w_\lambda} = 0 \\
&\text{b)} \quad \frac{\partial \mathcal{H}}{\partial x_\lambda} = 0, \quad \frac{\partial \mathcal{H}}{\partial y_\lambda} = 0 \qquad (\lambda = 1, 2, \ldots, n)\,. \\
&\text{c)} \quad \frac{\partial \mathcal{H}}{\partial z_\lambda} = 0, \quad \frac{\partial \mathcal{H}}{\partial \overline{z}_\lambda} = 0
\end{aligned}\right\} \tag{1}$$

Berücksichtigt man, daß nach (3) §3

$$\mathcal{H} = \sum_{\lambda,\mu} \epsilon_\lambda \epsilon_\mu \lg \frac{r_{\lambda\mu}}{2} = \lg Q = \lg \prod_{\lambda,\mu} \left(\frac{r_{\lambda\mu}}{2}\right)^{\epsilon_\lambda \epsilon_\mu}, \tag{2}$$

Finally, in conjugate coordinates, the *center-of-gravity equations* (6)″ p. 74 assume the form

$$\left.\begin{aligned} \epsilon = (\xi_0 + i\eta_0) = \epsilon z_0 &= \sum_{\lambda=1}^{n} \epsilon_\lambda \frac{2z_\lambda}{1 + z_\lambda \overline{z_\lambda}} = \text{ const} \\ \epsilon = (\xi_0 - i\eta_0) = \epsilon \overline{z_0} &= \sum_{\lambda=1}^{n} \epsilon_\lambda \frac{2\overline{z_\lambda}}{1 + z_\lambda \overline{z_\lambda}} = \text{ const} \\ \epsilon\zeta_0 &= \sum_{\lambda=1}^{n} \epsilon_\lambda \frac{1 - z_\lambda \overline{z_\lambda}}{1 + z_\lambda \overline{z_\lambda}} = \text{ const} \\ & \qquad (\epsilon = \epsilon_1 + \epsilon_2 + \cdots \epsilon_n)\,, \end{aligned}\right\} \tag{7}$$

where ξ_0, μ_0, ζ_0 denote the orthogonal coordinates of the center of gravity S of which it is possible to set simultaneously $\xi_0 = 0$, $\eta_0 = 0$, and $\zeta_0 = h$, given a special choice of the coordinate system (see p. 73), where h denotes the distance of the center of gravity from the center of the sphere.

§6. The general equilibrium problem

The simplest borderline case of the motion is that of the *equilibrium* of rest, the case where all n whirls permanently retain their initial position, and hence where the entire current is *stationary*. According to our equations of motion (4) in §2, or (5) and (6) in §5, this is possible only if the partial derivatives of the self-potential $\mathcal{H}$ with respect to all $2n$ spherical coordinates of the whirl system vanish simultaneously, or in other words, if the self-potential satisfies the first conditions of a maximum or minimum.

Thus, the conditions of the equilibrium in polar coordinates, "stereographic" and "conjugate coordinates" would be

$$\left.\begin{aligned} &\text{a)} \quad \frac{\partial \mathcal{H}}{\partial \vartheta_\lambda} = 0, \quad \frac{\partial \mathcal{H}}{\partial w_\lambda} = 0 \\ &\text{b)} \quad \frac{\partial \mathcal{H}}{\partial x_\lambda} = 0, \quad \frac{\partial \mathcal{H}}{\partial y_\lambda} = 0 \qquad (\lambda = 1, 2, \ldots, n)\,. \\ &\text{c)} \quad \frac{\partial \mathcal{H}}{\partial z_\lambda} = 0, \quad \frac{\partial \mathcal{H}}{\partial \overline{z}_\lambda} = 0 \end{aligned}\right\} \tag{1}$$

If one takes into account that, by (3) §3,

$$\mathcal{H} = \sum_{\lambda,\mu} \epsilon_\lambda \epsilon_\mu \lg \frac{r_{\lambda\mu}}{2} = \lg Q = \lg \prod_{\lambda,\mu} \left(\frac{r_{\lambda\mu}}{2}\right)^{\epsilon_\lambda \epsilon_\mu}, \tag{2}$$

und ferner, daß $\mathcal{H}$ auch die Kräftefunktion von n Massen $\epsilon_1, \ldots, \epsilon_n$ ist, wenn diese Massen Anziehungs- oder Abstoßungskräfte auf einander ausüben, welche den Massen direkt und den Entfernungen umgekehrt proportional sind, so erhält man den Satz:

Satz I. *Die Frage nach dem Gleichgewichte von gegebenen Strudeln* $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ *auf einer Kugel ist äquivalent jedem der beiden folgenden Probleme*

1) *dem* geometrischen *Problem, den gegebenen Coefficienten* $\epsilon_1, \ldots, \epsilon_n$ *entsprechend* n Punkte *so auf einer Kugel zu verteilen, dass das Product ihrer* $\frac{1}{2}n(n-1)$ *Sehnenabstände* $r_{\lambda\mu}$*, jeder potenziert mit dem Product* $\epsilon_\lambda\epsilon_\mu$ *der zugehörigen Coeffizienten, ein Maximum oder Minimum wird,*

80 | 2) *dem* mechanischen *Problem, n Massen* $\epsilon_1, \ldots, \epsilon_n$ *so auf der Kugel zu verteilen, dass sie sich im* Gleichgewichte *halten unter wechselseitiger Ausübung von Anziehungs- oder Abstossungskräften, welche den Massen direct und den Entfernungen umgekehrt proportional sind.*

Die Bedingungsgleichungen (1) schreiben sich explicit am einfachsten in der Form c) für conjugierte Coordinaten $z, \overline{z}$. Mit Hilfe des Ausdruckes (4) §5 für $\mathcal{H}$ erhält man nämlich:

$$\frac{\partial \mathcal{H}}{\partial z_\lambda} = \frac{1}{2}\epsilon_\lambda \sum_{\mu\neq\lambda}\left(\frac{\epsilon_\mu}{z_\lambda - z_\mu} - \frac{\epsilon_\mu \overline{z}_\lambda}{1 + z_\lambda \overline{z}_\lambda}\right)$$
$$= \frac{\epsilon_\lambda}{2}\sum_{\mu\neq\lambda}\frac{\epsilon_\mu}{z_\lambda - z_\mu} - \frac{\epsilon_\lambda}{2}(\epsilon - \epsilon_\lambda)\frac{\overline{z}_\lambda}{1 + z_\lambda\overline{z}_\lambda},$$

und die Gleichgewichtsbedingungen werden:

$$\sum_{\mu\neq\lambda}\frac{\epsilon_\mu}{z_\lambda - z_\mu} = (\epsilon - \epsilon_\lambda)\frac{\overline{z}_\lambda}{1 + z_\lambda\overline{z}_\lambda} \;;\; \sum_{\mu\neq\lambda}\frac{\epsilon_\mu}{\overline{z}_\lambda - \overline{z}_\mu} = (\epsilon - \epsilon_\lambda)\frac{z_\lambda}{1 + z_\lambda\overline{z}_\lambda} \qquad (3)$$
$$(\lambda = 1, 2, \ldots, n;\ \epsilon = \epsilon_1 + \epsilon_2 + \cdots + \epsilon_n)$$

und sind einander paarweise conjugiert. Doch lassen sie sich nach einer einfachen Umformung auch schreiben:

$$\sum_{\mu\neq\lambda}\epsilon_\mu\frac{1 + \overline{z}_\lambda z_\mu}{z_\lambda - z_\mu} = 0 \qquad \sum_{\mu\neq\lambda}\epsilon_\mu\frac{1 + z_\lambda\overline{z}_\mu}{z_\lambda - z_\mu} = 0 \qquad (4)$$
$$(\lambda = 1, \ldots, n)$$

und gehen, wenn man durch Drehung der Kugel den Punkt $\mathcal{P}_\lambda$ in das Projectionscentrum $z = \infty$ transformiert, über in:

$$\sum_{\mu\neq\lambda}\epsilon_\mu z_\mu = 0 \qquad \sum_{\mu\neq\lambda}\epsilon_\mu\overline{z}_\mu = 0 \qquad (z_\lambda = \infty). \qquad (5)$$

81 | Das bedeutet aber: der Schwerpunkt der $n-1$ Punkte $z_1, z_2, \ldots, z_n$ (außer z_λ), welche die stereographischen Bilder der Cugelpunkte $P_1, \ldots, P_n$ sind,

and furthermore that $\mathcal{H}$ is also the force function of n masses $\epsilon_1, \ldots, \epsilon_n$, assuming these masses exert forces of attraction or repulsion on one another that are directly proportional to the masses and inversely proportional to the distances, one obtains the theorem:

Theorem I. *The question of the equilibrium of given whirls $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ on a sphere is equivalent to each of the following two problems:*

1) *to the* geometric *problem of distributing n* points *on the sphere in correspondence with the given coefficients $\epsilon_1, \ldots, \epsilon_n$ so that the product of their $\frac{1}{2}n(n-1)$ chord distances $r_{\lambda\mu}$, each exponentiated by the product $\epsilon_\lambda \epsilon_\mu$ of the associated coefficients, becomes a maximum or minimum,*

2) *to the* mechanical *problem of distributing n masses $\epsilon_1, \ldots, \epsilon_n$ on the sphere so that they keep each other in* equilibrium *while exercising attractive or repulsive forces on one another that are directly proportional to the masses and inversely proportional to the distances.*

It is easiest to explicitly write the determining equations (1) in the form c) for conjugate coordinates $z, \overline{z}$. For by means of expression (4) §5 for $\mathcal{H}$ one obtains

$$\frac{\partial \mathcal{H}}{\partial z_\lambda} = \frac{1}{2}\epsilon_\lambda \sum_{\mu \neq \lambda} \left(\frac{\epsilon_\mu}{z_\lambda - z_\mu} - \frac{\epsilon_\mu \overline{z}_\lambda}{1 + z_\lambda \overline{z}_\lambda} \right)$$
$$= \frac{\epsilon_\lambda}{2} \sum_{\mu \neq \lambda} \frac{\epsilon_\mu}{z_\lambda - z_\mu} - \frac{\epsilon_\lambda}{2}(\epsilon - \epsilon_\lambda) \frac{\overline{z}_\lambda}{1 + z_\lambda \overline{z}_\lambda},$$

and the equilibrium equations become

$$\sum_{\mu \neq \lambda} \frac{\epsilon_\mu}{z_\lambda - z_\mu} = (\epsilon - \epsilon_\lambda) \frac{\overline{z}_\lambda}{1 + z_\lambda \overline{z}_\lambda} \;\; ; \;\; \sum_{\mu \neq \lambda} \frac{\epsilon_\mu}{\overline{z}_\lambda - \overline{z}_\mu} = (\epsilon - \epsilon_\lambda) \frac{z_\lambda}{1 + z_\lambda \overline{z}_\lambda} \qquad (3)$$
$$(\lambda = 1, 2, \ldots, n; \; \epsilon = \epsilon_1 + \epsilon_2 + \cdots + \epsilon_n)$$

and are pairwise conjugate to one another. But, upon a simple transformation, it is possible to write them also as follows:

$$\sum_{\mu \neq \lambda} \epsilon_\mu \frac{1 + \overline{z}_\lambda z_\mu}{z_\lambda - z_\mu} = 0 \qquad \sum_{\mu \neq \lambda} \epsilon_\mu \frac{1 + z_\lambda \overline{z}_\mu}{z_\lambda - z_\mu} = 0 \qquad (4)$$
$$(\lambda = 1, \ldots, n)$$

and, if the point $\mathcal{P}_\lambda$ is transformed into the center of projection $z = \infty$ by a rotation of the sphere, they pass over into:

$$\sum_{\mu \neq \lambda} \epsilon_\mu z_\mu = 0 \qquad \sum_{\mu \neq \lambda} \epsilon_\mu \overline{z}_\mu = 0 \qquad (z_\lambda = \infty). \qquad (5)$$

But this means that the center of gravity of the $n-1$ points $z_1, z_2, \ldots, z_n$ (except for z_λ) that are the stereographic images of the spherical points

genommen von P_λ als Centrum, fällt in den Mittelpunkt $z = 0$, und wir haben den Satz:

Satz II. *Die notwendige und hinreichende Bedingung für das Gleichgewicht eines Systemes von n Strudeln $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ besteht darin, dass wenn man von einem beliebigen ϵ_λ dieser Strudel als Centrum aus die $n-1$ übrigen Strudelpunkte ϵ_μ auf die Äquatorebene stereographisch projiciert und jedem Bildpunkte p_μ die Masse ϵ_μ beilegt, der Schwerpunkt dieses ebenen Systemes immer in den Kugelmittelpunkt fällt.*

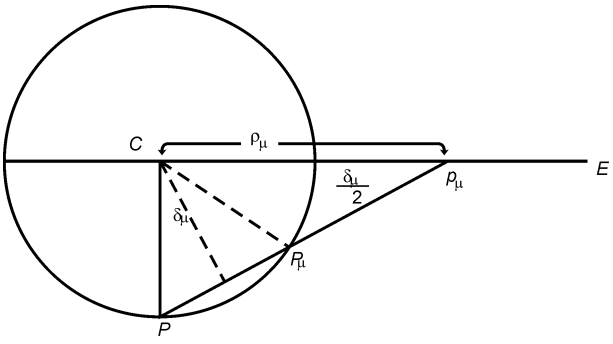


Fig. 2.

Dieser Satz läßt sich aber auch unmittelbar geometrisch einsehen. Es seien nämlich P und P_μ zwei Punkte auf der Kugel um C mit dem sphärischen Abstande $PCP_\mu = \delta_\mu$ von einander, E sei die zu CP gehörende Äquatorebene und p_μ sei das von P aus auf die Ebene E projicierte stereographische Bild von P_μ. Dann ist $\angle Cp_\mu P = \dfrac{\delta_\mu}{2}$ und daher $Cp_\mu = \varrho_\mu = \cot\dfrac{\delta_\mu}{2}$. Somit ist $\epsilon_\mu \varrho_\mu = \epsilon_\mu \cot\dfrac{\delta_\mu}{2}$ die doppelte Geschwindigkeit, welche ein Strudel P_μ mit
82 dem Momente $\pi\epsilon_\mu$ dem Punkte P er- | teilen würde in der Richtung senkrecht zur Ebene der Zeichnung PCP_μ (cf. Satz I des §3). Hält man nun den Punkt P fest und läßt P_μ eine Anzahl von Strudelpunkten $P_1, P_2, \ldots$ durchlaufen, so bekommt man in der Tangentialebene von P, welche der Ebene E paralel ist, ebenso viele Geschwindigkeitsvectoren q_μ, welche den entsprechenden Vectoren $\epsilon_\mu \varrho_\mu$ in den Richtungen C_{p_μ} proportional sind und immer nur um denselben Winkel 90° gegen sie gedreht. Diese Geschwindigkeitsvectoren werden daher dann und nur dann die Resultante 0 haben und P wird ein Gleichgewichtspunkt sein, wenn die sämtlichen durch C gehenden Vectoren $\epsilon_\mu \varrho_\mu$ in der Ebene E die Resultante 0 besitzen. In diesem Falle ist aber bekanntlich C der Schwerpunkt der in den Punkten p_μ localisierten Massen ϵ_μ. Diese Eigenschaft von P muß nun jedem Strudelpunkte P_λ in Bezug auf alle $n-1$ übrigen Strudel zukommen, wenn das ganze System im Gleichgewichte sein soll, und damit ist der behauptete Satz bewiesen.

$P_1, \ldots, P_n$, reckoned from P_λ as the center, falls into the center point $z = 0$, and we have the theorem:

Theorem II. *The necessary and sufficient condition for the equilibrium of a system of n whirls $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ consists in the fact that if from any ϵ_λ of these whirls as center one stereographically projects the remaining $n-1$ whirl points ϵ_μ onto the equatorial plane and attributes to each image point p_μ the mass ϵ_μ, then the center of gravity of this planar system always falls into the center point of the sphere.*

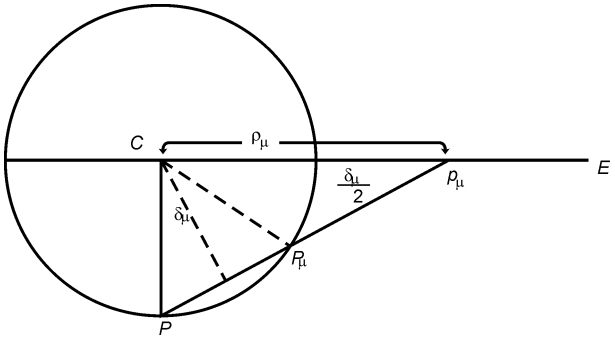


Fig. 2.

But this theorem also becomes immediately evident from the following geometric consideration. For let P and P_μ be two points on the sphere around C with the spherical distance $PCP_\mu = \delta_\mu$ from one another. Let E be the equatorial plane belonging to CP, and p_μ the stereographic image of P_μ projected from P onto the plane E. Then we have $\angle Cp_\mu P = \dfrac{\delta_\mu}{2}$, and hence $Cp_\mu = \varrho_\mu = \cot\dfrac{\delta_\mu}{2}$. Thus, $\epsilon_\mu \varrho_\mu = \epsilon_\mu \cot\dfrac{\delta_\mu}{2}$ is twice the velocity a whirl P_μ of momentum $\pi\epsilon_\mu$ would impart on the point P in the direction perpendicular to the plane of the figure PCP_μ (see Theorem I of §3). If one now keeps the point P fixed and lets P_μ run through a number of whirl points $P_1, P_2, \ldots$, then one obtains in the tangential plane of P parallel to the plane E the same number of velocity vectors q_μ that are proportional to the corresponding vectors $\epsilon_\mu \varrho_\mu$ in the directions C_{p_μ} and are always rotated against them only by the same angle of 90°. Therefore, these velocity vectors will have the resultant 0 and P will be an equilibrium point if and only if all the vectors $\epsilon_\mu \varrho_\mu$ passing through C possess the resultant 0 in the plane E. But in this case, as is well-known, C is the center of gravity of the masses ϵ_μ localized at the points p_μ. Now, every whirl point P_λ must have this property of P with respect to all remaining $n-1$ whirls if the entire system is to be in equilibrium, whereby the stated theorem is proved.

83 | Wir betrachten jetzt nur die einfachsten Fälle des Gleichgewichtes für $n = 1, 2, 3$.

1) Ein einziger Strudel ist mit sich selbst immer im Gleichgewicht. (Satz II §1 dieses Kap.)

2) (Satz III) Zwei Strudel sind dann und nur dann mit einander im Gleichgewicht, wenn sie sich an den beiden Enden eines Durchmessers befinden.

Dies ergibt sich unmittelbar aus jedem der oben gegebenen Kriterien. So ist z. B. in dem genannten Falle $\mathcal{H} = \epsilon_1 \epsilon_2 \lg r_{12}$ ein Maximum oder Minimum, weil der Durchmesser die größte Sehne ist. Die Gleichungen (3) aber reduzieren sich auf

$$\frac{\epsilon_2}{z_1 - z_2} = \frac{\epsilon_1 \overline{z_1}}{1 + z_1 \overline{z_1}}\,,$$
$$\frac{\epsilon_1}{z_2 - z_1} = \frac{\epsilon_1 \overline{z_2}}{1 + z_2 \overline{z_2}}\,,$$

oder

$$\overline{z_1} z_2 = \overline{z_2} z_1 = -1\,,$$

was nur für zwei entgegengesetzte Kugelpunkte der Fall ist.

3) Etwas complizierter ist der Fall *dreier* Strudel. Zunächst sieht man leicht, daß sie nur auf einem grössten Kugelkreise im Gleichgewicht stehen können. Denn da z. b. die beiden dem Punkte P_1 von den beiden anderen P_2 und P_3 erteilten Geschwindigkeiten auf den Ebenen CP_1P_2 und CP_1P_3
84 senkrecht stehen, so müssen diese beiden Ebenen zu- | sammenfallen, wenn die Geschwindigkeitsvectoren sich aufheben sollen.

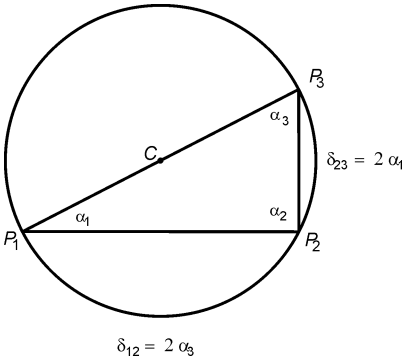


Fig. 3.

Es seien nun P_1, P_2, P_3 drei Strudel auf einem größten Kreise um C und $\alpha_1, \alpha_2, \alpha_3$ die Winkel des Dreieckes $P_1P_2P_3$. Dann stehen die erteilten Geschwindigkeiten alle senkrecht auf der Ebene der Zeichnung und sind proportional der Größen

$$\cot \frac{\delta_{23}}{2} = \cot \alpha_1\,, \quad \cot \frac{\delta_{31}}{2} = \cot \alpha_2\,, \quad \cot \frac{\delta_{12}}{2} = \cot \alpha_3\,,$$

We now consider only the simplest cases of the equilibrium for $n = 1, 2, 3$.

1) A single whirl is always in equilibrium with itself. (Theorem II §1 of this chap.)

2) (Theorem III) Two whirls are in equilibrium with one another if and only if they lie at the two ends of a diameter.

This immediately arises from each of the criteria specified above. Thus, e. g., in the case mentioned, $\mathcal{H} = \epsilon_1 \epsilon_2 \lg r_{12}$ is a maximum or minimum since the diameter is the greatest chord. But equations (3) reduce to

$$\frac{\epsilon_2}{z_1 - z_2} = \frac{\epsilon_1 \overline{z_1}}{1 + z_1 \overline{z_1}}\,,$$
$$\frac{\epsilon_1}{z_2 - z_1} = \frac{\epsilon_1 \overline{z_2}}{1 + z_2 \overline{z_2}}\,,$$

or

$$\overline{z_1} z_2 = \overline{z_2} z_1 = -1\,,$$

which is only the case for two opposite points on the sphere.

3) The case of *three* whirls is somewhat more intricate. First, it is readily seen that they can only be in equilibrium on a great spherical circle. For since, e. g., the two velocities given to the point P_1 by the two other points P_2 and P_3 are perpendicular to the planes CP_1P_2 and CP_1P_3, these two planes must coincide, if the velocity vectors are to cancel out one another.

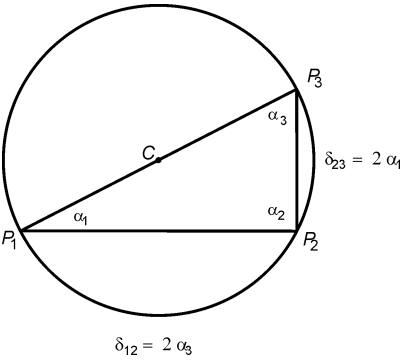


Fig. 3.

Now let P_1, P_2, P_3 be three whirls on a great circle around C, and $\alpha_1, \alpha_2, \alpha_3$ the angles of the triangle $P_1P_2P_3$. Then the attributed velocities are all perpendicular to the plane of the figure and proportional to the magnitudes

$$\cot \frac{\delta_{23}}{2} = \cot \alpha_1\,, \quad \cot \frac{\delta_{31}}{2} = \cot \alpha_2\,, \quad \cot \frac{\delta_{12}}{2} = \cot \alpha_3\,,$$

abgesehen von hinzukommenden Factoren $\epsilon_1, \epsilon_2, \epsilon_3$. Berücksichtigt man nun zugleich auch immer den Richtungssinn der Geschwindigkeitscomponenten, so ergeben sich die Gleichgewichtsbedingungen:

$$\epsilon_2 \cot\alpha_3 - \epsilon_3 \cot\alpha_2 = 0\,, \quad \epsilon_3 \cot\alpha_1 - \epsilon_1 \cot\alpha_3 = 0\,, \quad \epsilon_1 \cot\alpha_2 - \epsilon_2 \cot\alpha_1 = 0$$

oder, anders geschrieben:

$$\epsilon_1 : \epsilon_2 : \epsilon_3 = \cot\alpha_1 : \cot\alpha_2 : \cot\alpha_3\,. \tag{6}$$

Also haben wir den Satz

Satz IV. *Drei Strudel sind dann und nur dann im Gleichgewicht auf der Kugel, wenn sie auf einem grössten Kreise liegen und ihre Momente sich verhalten wie die Cotangenten der anliegenden Dreieckswinkel.*

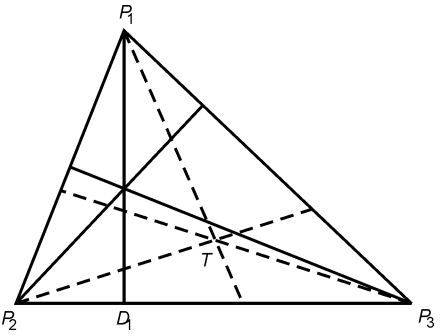


Fig. 4.

Fällt man im $\Delta P_1P_2P_3$ die Höhe P_1D_1, so verhält sich

$$P_2D_1 : P_3D_1 = \cot\alpha_2 : \cot\alpha_3$$

und im Falle des Gleichgewichtes müßten sich ja zwei Momente $\epsilon_2 : \epsilon_3$ verhalten wie die *anliegenden* Höhenabschnitte. Macht man aber $P_3E_1 = P_2D_1$, so ist

85 |
$$P_2E_1 : P_3E_1 = \cot\alpha_3 : \cot\alpha_2 \tag{7}$$

und im Falle des Gleichgewichtes wäre E_1 der *Schwerpunkt* von E_2 und E_3 in P_2 und P_3, in jedem Falle ist es der Schwerpunkt der Massen $\cot\alpha_2$ und $\cot\alpha_3$. Construiert man ebenso die Punkte E_2 und E_3, indem man auch die beiden anderen Seiten im umgekehrten Verhältnis der Höhenabschnitte teilt, so schneiden sich die drei Transversalen P_1E_1, P_2E_2 und P_3E_3 in dem Schwerpunkte T der drei in den Punkten P_1, P_2, P_3 localisierten Massen $\cot\alpha_1$, $\cot\alpha_2$, $\cot\alpha_3$, den wir (der späteren Anwendung wegen) als den „*Hauptpunkt*“ des Dreieckes bezeichnen wollen. Im Falle des Gleichgewichtes sind dann diese hypostasierten Massen den Strudelmomenten proportional,

except for additional factors $\epsilon_1, \epsilon_2, \epsilon_3$. If one also always takes into account the direction of the velocity components, then the following equilibrium conditions arise:

$$\epsilon_2 \cot\alpha_3 - \epsilon_3 \cot\alpha_2 = 0\,, \quad \epsilon_3 \cot\alpha_1 - \epsilon_1 \cot\alpha_3 = 0\,, \quad \epsilon_1 \cot\alpha_2 - \epsilon_2 \cot\alpha_1 = 0$$

or, written differently:

$$\epsilon_1 : \epsilon_2 : \epsilon_3 = \cot\alpha_1 : \cot\alpha_2 : \cot\alpha_3\,. \tag{6}$$

Thus we have the theorem

Theorem IV. *Three whirls are in equilibrium on a sphere if and only if they lie on a great circle and their momenta are related to one another as the cotangents of the adjacent angles of the triangle.*

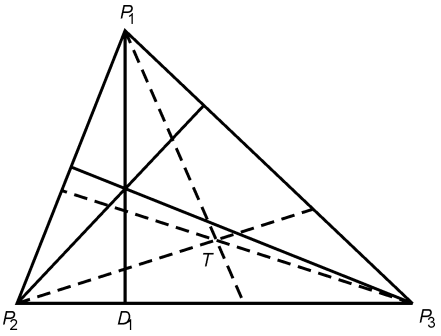


Fig. 4.

If one drops the height P_1D_1 in $\Delta\, P_1P_2P_3$, the proportions become

$$P_2D_1 : P_3D_1 = \cot\alpha_2 : \cot\alpha_3\,,$$

and, in the case of the equilibrium, the two momenta $\epsilon_2 : \epsilon_3$ would have to be related to one another as the *adjacent* height segments. But if one sets $P_3E_1 = P_2D_1$, then

$$P_2E_1 : P_3E_1 = \cot\alpha_3 : \cot\alpha_2$$

and, in the case of the equilibrium, E_1 would be the *center of gravity* of E_2 and E_3 in P_2 and P_3; in any case, it is the center of gravity of the masses $\cot\alpha_2$ and $\cot\alpha_3$. If one constructs the points E_2 and E_3 in the same way by splitting also the two other sides in inverse proportion of the height segments, then the three transversals P_1E_1, P_2E_2 and P_3E_3 intersect in the center of gravity T of the three masses $\cot\alpha_1$, $\cot\alpha_2$, $\cot\alpha_3$ localized at the points P_1, P_2, P_3. We shall call T the "*principal point*" of the triangle (on account of the later application). Then, in the case of the equilibrium, these postulated

und der „Hauptpunkt“ fällt mit dem Schwerpunkte der drei Strudel zusammen. Wir haben also den Satz:

Satz IVa. *Bezeichnet man als „Hauptpunkt“ eines Dreieckes den Schnittpunkt der drei Transversalen, welche die drei Seiten im umgekehrten Verhältnisse teilen wie die Höhen, oder den Schwerpunkt der Massen* $\cot\alpha_1, \cot\alpha_2, \cot\alpha_3$ *in den entsprechenden Eckpunkten, so muss, wenn drei Strudel auf*
86 *einem grössten | Kugelkreise im Gleichgewicht stehen, der Schwerpunkt ihrer Momente in den Hauptpunkt des von ihnen gebildeten Dreieckes fallen.*

§7. Das Gleichgewicht gleicher Strudel

Das Gleichgewichtsproblem vereinfacht sich in dem Falle, wo die Momente aller n Strudel einander gleich sind. Wir können dann annehmen:

$$\epsilon_1 = \epsilon_2 = \cdots = \epsilon_n = +1 = \frac{\epsilon}{n}$$

und erhalten so

$$Q = \prod_{\lambda,\mu} \frac{r_{\lambda,\mu}}{2}\,, \quad Q' = \prod_{\lambda,\mu} r_{\lambda,\mu}\,,$$

unser Problem ist daher nach Satz I des vorigen § so auszusprechen:

Problem. *Es sollen n Punkte auf einer gegeben Kugelfläche so verteilt werden, dass das Product Q' ihrer sämtlichen $\binom{n}{2}$ Sehnenabstände ein Maximum oder ein Minimum wird.*

In conjugierten Coordinaten $z_\lambda, \overline{z_\lambda}$ ist dann

$$Q^2 = \frac{\prod_{\lambda,\mu}(z_\lambda - z_\mu)(\overline{z_\lambda} - \overline{z_\mu})}{\prod_\lambda (1 + z_\lambda \overline{z_\lambda})^{n-1}} = \mathrm{Mx} \tag{1}$$

und die Gleichgewichtsbedingungen (3) nehmen die Form an:

87 | $$\sum_{\mu\neq\lambda} \frac{1}{z_\lambda - z_\mu} = (n-1)\frac{\overline{z}_\lambda}{1 + z_\lambda \overline{z}_\lambda}\,, \quad \sum_{\mu\neq\lambda} \frac{1}{\overline{z}_\lambda - \overline{z}_\mu} = (n-1)\frac{z_\lambda}{1 + z_\lambda \overline{z}_\lambda}\,. \tag{2}$$
$$(\lambda = 1, 2, \ldots, n)$$

Setzen wir nun:

$$\left.\begin{aligned} F(z) &= (z - z_1)(z - z_2)\cdots(z - z_n) \\ \overline{F}(z) &= (\overline{z} - \overline{z_1})(\overline{z} - \overline{z_2})\cdots(\overline{z} - \overline{z}_n)\,, \end{aligned}\right\} \tag{3}$$

masses are proportional to the whirl momenta, and the "principal point" coincides with the center of gravity of the three whirls. We thus have the theorem:

Theorem IVa. *If we denote by "principal point" of a triangle the point of intersection of the three transversals that cut the three sides in inverse proportion to the heights, or the center of gravity of the masses* $\cot\alpha_1$, $\cot\alpha_2$, $\cot\alpha_3$ *in the corresponding vertices, then the center of gravity of the momenta of three whirls must be the principal point of the triangle formed by them, assuming the three whirls are in equilibrium on a great circle.*

§7. The equilibrium of equal whirls

The equilibrium problem is simplified in the case where the momenta of all n whirls are equal to one another. We can then assume

$$\epsilon_1 = \epsilon_2 = \cdots = \epsilon_n = +1 = \frac{\epsilon}{n}$$

and thus obtain

$$Q = \prod_{\lambda,\mu} \frac{r_{\lambda,\mu}}{2}\,, \quad Q' = \prod_{\lambda,\mu} r_{\lambda,\mu}\,,$$

and hence, according to Theorem I of the previous §, our problem is to be expressed as follows:

Problem. *Given a spherical surface, n points are to be distributed on it so that the product* Q' *of all their* $\binom{n}{2}$ *chord distances become a maximum or a minimum.*

Then in conjugate coordinates $z_\lambda, \overline{z_\lambda}$

$$Q^2 = \frac{\prod_{\lambda,\mu}(z_\lambda - z_\mu)(\overline{z_\lambda} - \overline{z_\mu})}{\prod_\lambda (1 + z_\lambda \overline{z_\lambda})^{n-1}} = \mathrm{Mx}\,, \tag{1}$$

and equilibrium conditions (3) assume the form

$$\sum_{\mu\neq\lambda} \frac{1}{z_\lambda - z_\mu} = (n-1)\frac{\overline{z}_\lambda}{1 + z_\lambda \overline{z}_\lambda}\,, \quad \sum_{\mu\neq\lambda} \frac{1}{\overline{z}_\lambda - \overline{z}_\mu} = (n-1)\frac{z_\lambda}{1 + z_\lambda \overline{z}_\lambda}\,. \tag{2}$$
$$(\lambda = 1, 2, \ldots, n)$$

If we now set

$$\left.\begin{aligned} F(z) &= (z - z_1)(z - z_2)\cdots(z - z_n) \\ \overline{F}(z) &= (\overline{z} - \overline{z_1})(\overline{z} - \overline{z_2})\cdots(\overline{z} - \overline{z}_n)\,, \end{aligned}\right\} \tag{3}$$

wo z eine complexe Variable sein soll, d. h. denken wir uns die n Unbekannten $z_1, z_2, \ldots, z_n$ als Wurzeln einer algebraischen Gleichung n-ten Grades $F(z) = 0$, so ergiebt sich (wenn wir jetzt immer nur die eine der beiden conjugierten Gleichungen schreiben)

$$\sum_{\mu\neq\lambda} \frac{1}{z_\lambda - z_\mu} = \lim_{z=z_\lambda} \left[\sum_{\mu=1}^{n} \frac{1}{z - z_\mu} - \frac{1}{z - z_\lambda}\right] = \lim_{z=z_\lambda} \left(\frac{F'(z)}{F(z)} - \frac{1}{z - z_\lambda}\right)$$
$$= \frac{1}{2}\frac{F''(z_\lambda)}{F'(z_\lambda)},$$

so daß (3) übergeht in:

$$\frac{F''(z_\lambda)}{F'(z_\lambda)} = (2n-2)\frac{\overline{z_\lambda}}{1 + z_\lambda\overline{z_\lambda}} \quad (\lambda = 1, 2, \ldots, n), \tag{4}$$

also abgekürzt:

$$\frac{F''(z)}{F'(z)} = (2n-2)\frac{\overline{z}}{1 + z\overline{z}} \tag{4'}$$

immer, wenn $F(z) = 0$.

Dieser Bedingung (4) gemäß soll eine ganze Function n-ten Grades mit ungleichen Wurzeln $z_1, \ldots, z_n$ bestimmt werden. Die vollständige Lösung dieses Problemes, d. h. die Auffindung aller Lösungen mit dem Nachweise, daß
88 sie | die einzigen sind, ist mir freilich bisher noch nicht gelungen, doch läßt sich Folgendes beweisen.

Satz I. *Der wahre Schwerpunkt von n gleichen Strudeln muss für den Fall des Gleichgewichtes stets in den Kugelmittelpunkt fallen.*

Summiert man nämlich die Gleichungen (2) über $\lambda = 1, 2, \ldots, n$, so erhält man wegen $\dfrac{1}{z_\lambda - z_\mu} + \dfrac{1}{z_\mu - z_\lambda} = 0$ stets immer Null, also

$$\sum_{\lambda=1}^{n} \frac{\overline{z}_\lambda}{1 + z_\lambda\overline{z}_\lambda} = 0, \quad \sum_{\lambda=1}^{n} \frac{z_\lambda}{1 + z_\lambda\overline{z}_\lambda} = 0, \tag{5}$$

das heißt aber (vergl. (7) §5): Der Schwerpunkt S muß auf dem Durchmesser $O'CO$ des Projectionszentrums O' liegen, und zwar, da O' willkürlich ist, auf jedem Durchmesser, d. h. im Kugelmittelpunkt. Doch ist dies natürlich nur eine *notwendige* Bedingung.

Satz II. *Befinden sich n Strudelpunkte auf einem einzigen* Kugelkreise *im Gleichgewicht, so muss dies ein* grösster *Kreis sein, und die n Punkte bilden in ihm ein* reguläres Polygon.

Der folgende Nachweis dürfte von Interesse sein und vielleicht einer Verall-
89 gemeinerung fähig. Die Ebene des betrachteten Kreises | können wir parallel

where z is supposed to be a complex variable, i. e., if we conceive of the n unknowns $z_1, z_2, \ldots, z_n$ as roots of an algebraic equation of nth degree $F(z) = 0$, then (if we now always write only one of the two conjugate equations)

$$\sum_{\mu\neq\lambda} \frac{1}{z_\lambda - z_\mu} = \lim_{z=z_\lambda} \left[\sum_{\mu=1}^{n} \frac{1}{z - z_\mu} - \frac{1}{z - z_\lambda} \right] = \lim_{z=z_\lambda} \left(\frac{F'(z)}{F(z)} - \frac{1}{z - z_\lambda} \right)$$
$$= \frac{1}{2} \frac{F''(z_\lambda)}{F'(z_\lambda)},$$

so that (3) is transformed into

$$\frac{F''(z_\lambda)}{F'(z_\lambda)} = (2n-2) \frac{\overline{z_\lambda}}{1 + z_\lambda \overline{z_\lambda}} \quad (\lambda = 1, 2, \ldots, n), \tag{4}$$

and hence, for short,

$$\frac{F''(z)}{F'(z)} = (2n-2) \frac{\overline{z}}{1 + z\overline{z}} \tag{4'}$$

always when $F(z) = 0$.

According to this condition (4), an entire function of nth degree with unequal roots $z_1, \ldots, z_n$ is supposed to be determined. I have not yet managed to solve this problem completely, i. e., to find all solutions and show that they are the only ones. But it is possible to prove the following.

Theorem I. *In the case of the equilibrium, the true center of gravity of n equal whirls must always be the center point of the sphere.*

For if one sums equations (2) over $\lambda = 1, 2, \ldots, n$, then, on account of $\dfrac{1}{z_\lambda - z_\mu} + \dfrac{1}{z_\mu - z_\lambda} = 0$, one always obtains zero, and hence

$$\sum_{\lambda=1}^{n} \frac{\overline{z}_\lambda}{1 + z_\lambda \overline{z}_\lambda} = 0, \quad \sum_{\lambda=1}^{n} \frac{z_\lambda}{1 + z_\lambda \overline{z}_\lambda} = 0, \tag{5}$$

which, however, is tantamount to (see (7) §5): The center of gravity S must lie on the diameter $O'CO$ of the center of projection O', and in particular, on every diameter, i. e., in the center point of the sphere, since O' is arbitrary. But this, of course, is only a *necessary* condition.

Theorem II. *If n whirl points lie on a single* spherical circle *in equilibrium, then this must be a* great *circle, and the n points form a* regular polygon *in it.*

The following proof should be of some interest and may be capable of generalization. We can assume that the plane of the circle is parallel to the

der Äquatorebene annehmen, dann ist die Gleichung dieses Kreises

$$z\overline{z} = \alpha^2 = \text{ const }\,,$$

so daß die Bedingung $(4)'$ übergeht in:

$$\frac{F''(z)}{F'(z)} = (2n-2)\frac{\alpha^2}{1+\alpha^2}\frac{1}{z} \equiv \frac{m}{z} \qquad \left(m = \frac{\alpha^2}{1+\alpha^2}(2n-2)\right).$$

Diese Gleichung soll gelten für alle n (von einander verschiedenen) Wurzeln z_λ der Gleichung $F(z) = 0$, die betrachtete Gleichung ist aber selbst nur vom Grade $n-1$ und muß daher *identisch* bestehen für alle Werte z. Somit können wir auch beiderseits integrieren und erhalten

$$F(z) = c(z^{m+1} - a^{m+1})\,,$$

wo c und a Constanten sind. Nun soll aber $F(z)$ nach (3) vom Grade n sein und 1 zum höchsten Coefficienten haben, es muß also sein:

$$c = 1\,, \quad m+1 = n$$

und aus der letzten Beziehung folgt

$$m = (2n-2)\frac{\alpha^2}{1+\alpha^2} = n-1 \text{ oder } \alpha^2 = 1\,,$$

d. h. $z_\lambda\overline{z_\lambda} = 1$, der Kreis muß ein grösster sein, und es ist nunmehr

$$F(z) = z^n - a^n\,,$$

worin, wegen $z_\lambda\overline{z_\lambda} = 1$ auch $a\overline{a} = 1$, d. h. $a = e^{i\varphi}$ sein muß, also

$$z_\lambda = e^{i(\varphi + \frac{2\pi\lambda}{n})} = \cos\left(\varphi + \frac{2\pi\lambda}{n}\right) + i\sin\left(\varphi + \frac{2\pi\lambda}{n}\right)$$

und hiermit folgt die Behauptung.

90 | Hiermit ist zugleich der Satz bewiesen, dass n Punkte auf einem gegebenen *Kreise* immer ein reguläres Polygon bilden müssen, wenn das Product ihrer
Fig. 5 sämtlichen $\binom{n}{2}$ Sehnenabstände ein Maximum sein soll.

Daß n Strudel in gleichen Abständen auf einem größten Kugelkreis überhaupt im Gleichgewichte sind, folgt schon unmittelbar aus der *Symmetrie* der ganzen Figur um jeden Durchmesser des Polygones, denn hier muß bei der stereographischen Projection von jedem Polygonpunkte P_1 aus der Schwerpunkt des abgebildeten Systemes $p_2, p_3, \ldots, p_n$ in der That in den Mittelpunkt C fallen, wie im Satze II S. 81 gefordert. Dieselbe Symmetrie findet aber auch statt, wenn man noch 2 gleiche Strudel in den beiden Polen des größten Kreises anbringt, welchem das reguläre Polygon eingeschrieben ist, sodaß die sämtlichen Strudel eine reguläre Doppelpyramide bilden. Außerdem ist natürlich die Symmetrie gewahrt, wenn die sämtlichen Strudel sich
91 in den Eckpunkten eines regulären Polyeders befinden. | Somit können wir sagen:

equatorial plane. Then the equation of this circle is

$$z\overline{z} = \alpha^2 = \text{ const} ,$$

so that condition $(4)'$ is transformed into

$$\frac{F''(z)}{F'(z)} = (2n-2)\frac{\alpha^2}{1+\alpha^2}\frac{1}{z} \equiv \frac{m}{z} \qquad \left(m = \frac{\alpha^2}{1+\alpha^2}(2n-2)\right) .$$

This equation is supposed to hold for all n (mutually different) roots z_λ of the equation $F(z) = 0$. But the equation under consideration itself is only of degree $n-1$, and hence must hold *identically* for all values z. Thus, we can also integrate on both sides and obtain

$$F(z) = c(z^{m+1} - a^{m+1}) ,$$

where c and a are constants. But now, according to (3), $F(z)$ is supposed to be of degree n and is supposed to have 1 as its highest coefficient, and hence necessarily

$$c = 1 , \quad m+1 = n ,$$

and from the last relation it follows that

$$m = (2n-2)\frac{\alpha^2}{1+\alpha^2} = n-1 \text{ or } \alpha^2 = 1 ,$$

i. e., $z_\lambda\overline{z_\lambda} = 1$, the circle must be a great circle, and now

$$F(z) = z^n - a^n ,$$

wherein, on account of $z_\lambda\overline{z_\lambda} = 1$, also necessarily $a\overline{a} = 1$, i. e., $a = e^{i\varphi}$, and hence

$$z_\lambda = e^{i(\varphi + \frac{2\pi\lambda}{n})} = \cos\left(\varphi + \frac{2\pi\lambda}{n}\right) + i\sin\left(\varphi + \frac{2\pi\lambda}{n}\right) ,$$

whence the assertion follows.

Thus, at the same time, the theorem is proved that n points on a given *circle* must always form a regular polygon if the product of all their $\binom{n}{2}$ chord
Fig. 5 distances is supposed to be a maximum.

That n whirls lying at equal distances on a great spherical circle are at all in equilibrium already follows immediately from the *symmetry* of the whole figure around every diameter of the polygon since here, for the stereographic projection from each point P_1 of the polygon, the center of gravity of the image system $p_2, p_3, \ldots, p_n$ must be the center point C, as required in Theorem II p. 81. But the same symmetry occurs when 2 additional equal whirls are attached in the two poles of the great circle in which the regular polygon is inscribed so that all whirls together form a regular bipyramid. Moreover, the symmetry is of course preserved when all whirls lie at the vertices of a regular polyhedron. Thus we can say:

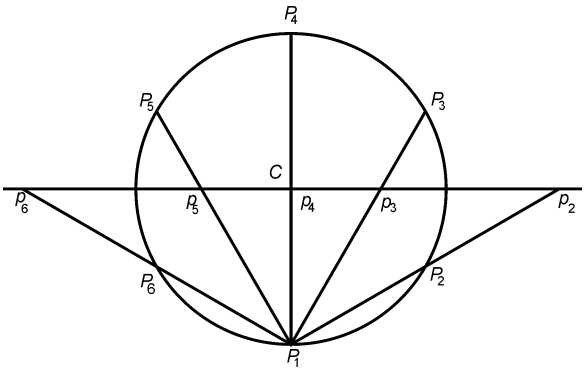


Fig. 5.

Satz III. *n gleiche Strudel sind auf einer Kugel sicher dann im Gleichgewichte, wenn jeder nach einem der Strudelpunkte gezogene Durchmesser eine Symmetrieaxe des ganzen Systemes darstellt, also in den folgenden 3 Fällen:*

1) *Wenn die n Strudelpunkte ein* reguläres Polygon *in einem grössten Kreise bilden,*

2) *wenn sie eine* reguläre $n-2$ seitige Doppelpyramide *bilden, d. h. ein reguläres $n-2$ eck im grössten Kreis mit den beiden Polen des auf einer Ebene senkrechten Durchmessers,*

3) *wenn sie ein* reguläres Polyeder *bilden, falls $n = 4, 6, 8, 12, 20$ ist.*

Formuliert man unser Problem als Maximumsaufgabe, so läßt sich zeigen, daß der Wert der Function $Q'_n = \prod r_{\lambda\mu}$ im Falle der $n-2$ seitigen Doppelpyramide $Q'_{n-2,2}$ immer grösser ist als der $Q'_{n,0}$ für das reg. Polygon

$$Q'_{n-2,2} = (n-2)^{\frac{n-2}{2}} 2^{n-1} > Q'_{n,0} = n^{\frac{n}{2}},$$

weil

$$k_n = \frac{Q'_{n-2,1}}{Q'_{n,0}} = \left(\frac{n-2}{n}\right)^{\frac{n}{2}} \frac{2^{n-1}}{n-2} \geq 1 \text{ für } n \geq 4\,.$$

(Die Berechnung von $Q'_{n,0}, Q'_{n-2,2}$ geschieht leicht mit Hilfe das Cotesschen Satzes). Ferner überzeugt man sich schon durch numerische Ausrechnung, daß
92 das reguläre n eckige Polyeder | falls es existiert zu einem noch größeren Werte $P'_n = P'_{(0)}$ führt, also wahrscheinlich das absolute Maximum der Function darstellt. Für $n = 8$ ist z. B.

$$P'_8 = P'_{8,0} = 2^{12} \qquad \text{(reguläres 8 Eck im größten Kreise 1)}$$
$$P'_8 = P'_{6,2} = 2^{10} \cdot 3^3 = 6,75 \cdot P'_{8,0} \qquad \text{(reguläre 6 seitige Pyramide)}$$
$$P'_8 = P'_{(8)} = \tfrac{2^{34}}{3^{12}} = 7,89 \cdot P'_{8,0} \qquad \text{(Würfel in der Kugel 1)}$$

Daß diese „symmetrischen“ Gleichgewichtsfälle die einzigen sind, läßt sich abgesehen von den im vorigen §6 behandelten Fällen $n = 1, 2, 3$ auch noch für $n = 4$ beweisen.

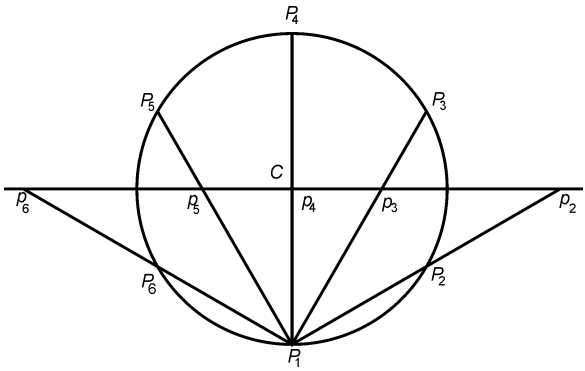


Fig. 5.

Theorem III. *n equal whirls on a sphere are certainly in equilibrium if every diameter drawn toward one of the whirl points is an axis of symmetry of the entire system, and hence in the following 3 cases:*

1) *When the n whirl points form a* regular polygon *in a great circle,*

2) *when they form a* regular, $(n-2)$-sided bipyramid *, i. e., a regular* $(n-2)$*-gon in the great circle with the two poles of the diameter orthogonal to a plane,*

3) *when they form a* regular polyhedron, *provided* $n = 4, 6, 8, 12, 20$.

Framing our problem as a maximum problem, one can show that, in the case of the $(n-2)$-sided bipyramid $Q'_{n-2,2}$, the value of the function $Q'_n = \prod r_{\lambda\mu}$ is always greater than the $Q'_{n,0}$ for the reg⟦ular⟧ polygon

$$Q'_{n-2,2} = (n-2)^{\frac{n-2}{2}} 2^{n-1} > Q'_{n,0} = n^{\frac{n}{2}},$$

since

$$k_n = \frac{Q'_{n-2,1}}{Q'_{n,0}} = \left(\frac{n-2}{n}\right)^{\frac{n}{2}} \frac{2^{n-1}}{n-2} \geq 1 \text{ for } n \geq 4\,.$$

(It is easy to calculate $Q'_{n,0}, Q'_{n-2,2}$ by means of Cotes's theorem.) Furthermore, numeric computation already shows that the regular n-gonal polyhedron, if it exists, leads to an even greater value $P'_n = P'_{(0)}$, and hence is very likely the absolute maximum of the function. E. g., for $n = 8$,

$$\begin{aligned} P'_8 &= P'_{8,0} = 2^{12} && \text{(regular 8-gon in great circle 1)} \\ P'_8 &= P'_{6,2} = 2^{10} \cdot 3^3 = 6,75 \cdot P'_{8,0} && \text{(regular 6-edged pyramid)} \\ P'_8 &= P'_{(8)} = \tfrac{2^{34}}{3^{12}} = 7,89 \cdot P'_{8,0} && \text{(cube in sphere 1)} \end{aligned}$$

That these "symmetric" equilibrium cases are the only ones can also be proved for $n = 4$ in addition to the cases $n = 1, 2, 3$ considered in the previous §6.

Satz IV. *Vier gleiche Strudel befinden sich dann und nur dann im Gleichgewicht, wenn sie entweder ein Quadrat im grössten Kreise oder ein reguläres Tetraeder bilden.*

Liegen die 4 Strudelpunkte in einer einzigen Ebene, so müssen sie schon nach Satz II ein Quadrat im größten Kreise bilden. Es bleibt also nur noch zu zeigen, daß wenn dies nicht der Fall ist, das von ihnen gebildete Tetraeder ein reguläres sein muß. Die Ebene der 3 ersten Strudelpunkte A, B, C werde mit E, ihr Schwerpunkt mit S bezeichnet. Dann muß die Verbindungslinie SD mit dem vierten Punkte D durch den Mittelpunkt C der Kugel gehen, der ja nach Satz I mit den Schwerpunkt aller vier Punkte zusammenfällt. Legen wir nun durch C eine Ebene E' senkrecht zu SCD, die von den Strahlen DA, DB, DC in den Punkten A', B', C' geschnitten wird, so muß nach Satz II von §6 der Schwerpunkt S' dieser neuen Punkte ebenfalls in den Mittelpunkt
92a C fallen. Wenn aber bei einer Centralprojektion | dem Schwerpunkte eines Dreieckes der Schwerpunkt des projicierten entspricht, so sind, wie man sich leicht überzeugt, die beiden Ebenen parallel.[1] Es steht also auch die Ebene E senkrecht auf der Geraden SCD und der Schwerpunkt S des Dreieckes ABC ist gleichzeitig Mittelpunkt des umschriebenen Kreises, den die Ebene E aus der Kugel ausschneidet. Daraus folgt aber, daß das Dreieck ABC gleichseitig ist, ebenso auch die Dreiecke ABD, ADC und DBC, da der Punkt D nicht bevorzugt ist, und das ganze Tetraeder $ABCD$ muß daher regulär sein, w.z.b.w.

93 |

Capitel IV
Das Problem der drei Strudel

Vorbemerkung. Sind nur zwei Strudel P_1, P_2 auf der Kugel gegeben, so ist ihr Selbstpotential (III §3 Satz II):

$$\mathcal{H} = \epsilon_1 \epsilon_2 \lg \frac{r_{1,2}}{2} = \text{ const},$$

ihr Abstand bleibt also constant, und da gleichzeitig ihr Schwerpunkt S im Raume fest bleiben muß und die Geschwindigkeit immer nur von der Configuration abhängt, so muß das ganze System um den durch den Schwerpunkt gehenden Durchmesser CS gleichförmig rotieren.

Während also das Bewegungsproblem für den Fall $n = 2$ ohne weiteres als vollständig gelöst betrachtet werden kann, bietet schon der nächst einfache Fall dreier Strudel sehr viel größere Schwierigkeiten. Hier ist $2n = 6$ die Anzahl der unbekannten Functionen der Zeit sowie auch der Differentialgleichungen. Davon kennen wir aber bereits 4 von einander unabhängige

[1] Diese Bemerkung verdanke ich der freundlichen Mitteilung von Hrn. Dr. Hessenberg.

Theorem IV. *Four equal whirls are in equilibrium if and only if they either form a square in the great circle or a regular tetrahedron.*

If the 4 whirl points lie in a single plane, then, already by theorem II, they must form a square in the great circle. Thus it only remains to show that if this is not the case, then the tetrahedron formed by them must be a regular one. Let us denote the plane of the first 3 whirl points A, B, C by E and its center of gravity by S. Then the connecting line SD with the fourth point D must pass through the center point C of the sphere, which, by theorem I, coincides with the center of gravity of all four points. If we now place a plane E' through C perpendicular to SCD, which is intersected by the rays DA, DB, DC in the points A', B', C', then, by theorem II of §6, the center of gravity S' of these new points must also be the center point C. If, however, to the center of gravity of some triangle there corresponds in a central projection that of the projected triangle, then, as is readily seen, the two planes are parallel.[1] Hence, the plane E is also perpendicular to the straight line SCD, and the center of gravity S of the triangle ABC is also the center point of the circumscribed circle cut out off the sphere by the plane E. But from this it follows that the triangle ABC is equilateral, and so are the triangles ABD, ADC and DBC, since the point D is not distinguished, and hence the entire tetrahedron $ABCD$ must be regular, q.e.d.

Chapter IV
The problem of three whirls

Preliminary remark. If only two whirls P_1, P_2 are given on the sphere, then their self-potential is (III §3 theorem II):

$$\mathcal{H} = \epsilon_1 \epsilon_2 \lg \frac{r_{1,2}}{2} = \text{ const},$$

and hence their distance remains constant, and, since, at the same time, their center of gravity S must remain fixed in space and the velocity always depends only on the configuration, the entire system has to rotate uniformly around the diameter CS passing through the center of gravity.

Hence, while the motion problem for the case $n = 2$ may actually be considered completely solved, already the next simplest case of three whirls poses much greater difficulties. In this case, $2n = 6$ is the number of the unknown functions of both time and the differential equations. But of those

[1] For this remark I am indebted to a friendly communication of Dr. Hessenberg.

94 Integrale (C. III §4), es | bleiben also nur noch zwei auszuführende Integrationen übrig, von denen nur die eine sich auf die Bahncurven, die andere aber allein auf den zeitlichen Verlauf der Erscheinung bezieht.

Zerlegung des Problems. Anstatt aber direct die Bewegungsgleichungen in der Form (4) p. 70 oder (5) od. (6) p. 77 zu benutzen, wollen wir zunächst das Problem in zwei Teile zerlegen, in das Problem der *relativen* und das der *absoluten* Bewegung, indem wir erstens fragen: Wie bewegen sich die drei Punkte relativ zu einander, d. h. wie ändern sich die Seiten und Winkel des aus ihnen gebildeten Dreieckes? und zweitens: wie bewegt sich das so bestimmte unveränderliche Dreieck auf der als fest vorausgesetzten Kugel? Die erste Frage soll uns in den drei nächsten §§ ausschließlich beschäftigen.

95 |

§1. Die Gleichungen der Relativbewegung

Zur Bestimmung der relativen Bewegung führen wir als neue Variable ein die drei Seitenlängen r_1, r_2, r_3 des (ebenen) Strudeldreieckes $P_1P_2P_3$ und haben nun zunächst die Differentialgleichungen aufzustellen, denen diese drei Größen als Functionen der Zeit genügen müssen. Zu diesem Zwecke könnten wir allerdings die allgemeinen Bewegungsgleichungen von §3 oder §5 des vorigen Cap. durch einfache Substitution auf die neuen Coordinaten transformieren, was einige Rechnung erfordern würde; wir ziehen es aber vor, vermittelst geometrischer Betrachtungen auf die ersten Principien (III §1 u. §3) unserer Theorie selbst zurückzugehen.

Da die Geschwindigkeit, welche ein Strudel einem anderen erteilt, immer senkrecht steht auf dem beide verbindenden größten Kreise, also auch senkrecht auf ihrer Verbindungsgeraden, so liefert diese gegenseitige Einwirkung keinen Beitrag zur Vergrößerung oder Verkleinerung ihres Abstandes. (In der
96 | That ist auch dieser Abstand constant in dem Falle, wo nur 2 Strudel vorhanden sind.) Geben wir also nur 3 Strudel P_1, P_2, P_3, so ist die Änderung des Abstandes $\overline{P_2P_3} = r_1$ lediglich dem gegenüberliegenden Strudel P_1 mit dem Moment $\epsilon_1\pi$ zuzuschreiben.

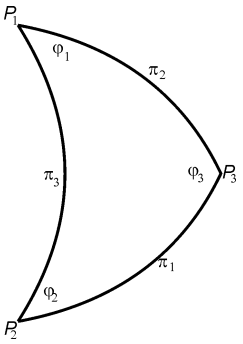


Fig. 6.

we already know 4 mutually independent integrals (C. III §4). Thus only two integrations remain to be evaluated, of which only one concerns the curves, while the other only concerns the temporal evolution of the phenomenon.

Analysis of the problem. Instead of making direct use of the equations of motion in the form (4) p. 70, or (5) or (6) p. 77, we will begin by splitting the problem into two parts, the problem of the *relative* and that of the *absolute* motion, by asking first: How do the three points move in relation to one another, i. e., how do the sides and angles of the triangle formed by them change? And, secondly: How does the unchangeable triangle thus determined move on the sphere, which is assumed to be fixed? In the following three §§, we will address only the first question.

§1. The equations of relative motion

In order to determine the relative motion we introduce as new variables the three side lengths r_1, r_2, r_3 of the (planar) whirl triangle $P_1P_2P_3$. Now, we first have to determine the differential equations that these three magnitudes, being functions of time, must satisfy. To this end, we could convert the general equations of motion of §3 or §5 of the previous chap⟦ter⟧ to the new coordinates by means of a simple substitution, which would require a good amount of calculation; but we prefer to return to the first principles (III §1 a⟦nd⟧ §3) of our theory themselves by way of geometric considerations.

Since the velocity one whirl imparts upon another is always perpendicular to the great circle joining both, and hence also perpendicular to their connecting line, this mutual interaction does not contribute to the increase or decrease of their distance. (In fact, this distance is also constant in the case where there are only 2 whirls.) Thus, if we consider only 3 whirls P_1, P_2, P_3, then the change in distance $\overline{P_2P_3} = r_1$ can only be attributed to the whirl P_1 on the opposite side with the momentum $\epsilon_1\pi$.

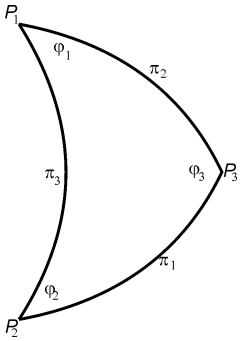


Fig. 6.

Bezeichnet man nun die Seiten des sphärischen Dreieckes $P_1P_2P_3$ mit π_1, π_2, π_3 und seine Winkel mit $\varphi_1, \varphi_2, \varphi_3$, so wird:

$$1 - \frac{r_1^2}{2} = \cos\pi_1 = \cos\pi_2 \ \cos\pi_3 \ r \ \sin\pi_2 \ \sin\pi_3 \ \cos\varphi_1 \tag{1}$$

und daher durch Differentiation:

$$r_1 \frac{dr_1}{dt} = \sin\pi_1 \ \frac{d\pi_1}{dt} = \sin\pi_2 \ \sin\pi_3 \ \sin\varphi_1 \ \frac{d\varphi_1}{dt}\,, \tag{1}'$$

denn hier können die Glieder mit $\frac{d\pi_2}{dt}$ und $\frac{d\pi_3}{dt}$ unterdrückt werden, weil π_2 und π_3 durch P_1 nicht verändert werden. Vielmehr erhalten P_2 und P_3 lediglich die Geschwindigkeitscomponenten

$$\frac{\epsilon_1}{2}\cot\frac{\pi_1}{2} \quad \text{und} \quad \frac{\epsilon_1}{2}\cot\frac{\pi_2}{2}$$

senkrecht zu den Bogen $\widehat{P_1P_2}$ und $\widehat{P_1P_3}$, und ihnen entsprechen die Winkelgeschwindigkeiten um den Radius CP_1 als Axe:

$$\frac{\epsilon_1}{2}\frac{\cot\frac{\pi_3}{2}}{\sin\pi_3} \quad \text{und} \quad \frac{\epsilon_1}{2}\frac{\cot\frac{\pi_2}{2}}{\sin\pi_2}\,,$$

welche die Werte annehmen:

97 |
$$\frac{\epsilon_1}{4\sin^2\frac{\pi_3}{2}} = \frac{\epsilon_1}{r_3^2} \quad \text{und} \quad \frac{\epsilon_1}{4\sin^2\frac{\pi_2}{2}} = \frac{\epsilon_1}{r_2^2}\,.$$

Hiermit wird ihre Differenz:

$$\frac{d\varphi_1}{dt} = \frac{\epsilon_1}{r_2^2} - \frac{\epsilon_1}{r_3^2}$$

und daher nach (1)′

$$\begin{aligned} r_1\frac{dr_1}{dt} &= \epsilon_1 \ \sin\pi_2 \ \sin\pi_3 \ \sin\varphi_1 \left(\frac{1}{r_2^2} - \frac{1}{r_3^2}\right) \\ &= \epsilon_1 R \left(\frac{1}{r_2^2} - \frac{1}{r_3^2}\right)\,. \end{aligned} \tag{2}$$

Nun ist aber im sphärischen Dreieck $P_1P_2P_3$: (cf. Baltzer Elemente d. M. VI §5, 11)

$$\begin{aligned} R &= \sin\pi_2\sin\pi_3\sin\varphi_1 = \sin\pi_3\sin\pi_1\sin\varphi_2 = \sin\pi_1\sin\pi_2\sin\varphi_3 \\ &= \pm\sqrt{1 - \cos^2\pi_1 - \cos^2\pi_2 - \cos^2\pi_3 + 2\cos\pi_1\cos\pi_2\cos\pi_3} \\ R &= \pm\sqrt{-p_1^2 - p_2^2 - p_3^2 + 2p_2p_3 + 2p_3p_1 + 2p_1p_2 - 2p_1p_2p_3}\,, \end{aligned} \tag{3}$$

If one denotes the sides of the spherical triangle $P_1P_2P_3$ by π_1, π_2, π_3, and its angles by $\varphi_1, \varphi_2, \varphi_3$, then

$$1 - \frac{r_1^2}{2} = \cos\pi_1 = \cos\pi_2 \ \cos\pi_3 \ r \ \sin\pi_2 \ \sin\pi_3 \ \cos\varphi_1 \tag{1}$$

and hence, by differentiation,

$$r_1 \frac{dr_1}{dt} = \sin\pi_1 \ \frac{d\pi_1}{dt} = \sin\pi_2 \ \sin\pi_3 \ \sin\varphi_1 \ \frac{d\varphi_1}{dt}\,, \tag{1}'$$

for here the terms with $\frac{d\pi_2}{dt}$ and $\frac{d\pi_3}{dt}$ can be suppressed since π_2 and π_3 are not altered by P_1. Rather, P_2 and P_3 only obtain the velocity components

$$\frac{\epsilon_1}{2}\cot\frac{\pi_1}{2} \quad \text{and} \quad \frac{\epsilon_1}{2}\cot\frac{\pi_2}{2}$$

perpendicular to the arcs $\widehat{P_1P_2}$ and $\widehat{P_1P_3}$, and to them there correspond the angular velocities around the radius CP_1 as axis:

$$\frac{\epsilon_1}{2}\frac{\cot\frac{\pi_3}{2}}{\sin\pi_3} \quad \text{and} \quad \frac{\epsilon_1}{2}\frac{\cot\frac{\pi_2}{2}}{\sin\pi_2}\,,$$

which assume the values

$$\frac{\epsilon_1}{4\sin^2\frac{\pi_3}{2}} = \frac{\epsilon_1}{r_3^2} \quad \text{and} \quad \frac{\epsilon_1}{4\sin^2\frac{\pi_2}{2}} = \frac{\epsilon_1}{r_2^2}\,.$$

Thus, their difference becomes

$$\frac{d\varphi_1}{dt} = \frac{\epsilon_1}{r_2^2} - \frac{\epsilon_1}{r_3^2}\,,$$

and hence, by (1)′,

$$\begin{aligned} r_1\frac{dr_1}{dt} &= \epsilon_1 \ \sin\pi_2 \ \sin\pi_3 \ \sin\varphi_1 \left(\frac{1}{r_2^2} - \frac{1}{r_3^2}\right) \\ &= \epsilon_1 R \left(\frac{1}{r_2^2} - \frac{1}{r_3^2}\right)\,. \end{aligned} \tag{2}$$

But now, in the spherical triangle $P_1P_2P_3$ (see *Baltzer 1883*, VI §5, 11), we have

$$\begin{aligned} R &= \sin\pi_2\sin\pi_3\sin\varphi_1 = \sin\pi_3\sin\pi_1\sin\varphi_2 = \sin\pi_1\sin\pi_2\sin\varphi_3 \\ &= \pm\sqrt{1 - \cos^2\pi_1 - \cos^2\pi_2 - \cos^2\pi_3 + 2\cos\pi_1\cos\pi_2\cos\pi_3} \\ R &= \pm\sqrt{-p_1^2 - p_2^2 - p_3^2 + 2p_2p_3 + 2p_3p_1 + 2p_1p_2 - 2p_1p_2p_3}\,, \end{aligned} \tag{3}$$

wenn gesetzt wird:

$$\left.\begin{aligned} p_1 &= \frac{1}{2}r_1^2 = 1-\cos\pi_1 \\ p_2 &= \frac{1}{2}r_2^2 = 1-\cos\pi_2 \\ p_3 &= \frac{1}{2}r_3^2 = 1-\cos\pi_3\,, \end{aligned}\right\} \tag{4}$$

und die Wurzel ist positiv oder negativ zunehmen, je nachdem man von außen das sphärische $\Delta\ P_1P_2P_3$ betrachtend und es im Sinne $P_1 \to P_2 \to P_3$ umfahrend sein Inneres (d. h. das kleinere Stück der Kugelfläche) oder einfach das Innere des ebenen $\Delta\ P_1P_2P_3$ zur linken oder zur rechten hat. R verschwindet nur, wenn die 3 Punkte P_1, P_2, P_3 in einem grössten Kreise
98 liegen, und | wechselt sein Vorzeichen, wenn die Ebene des Dreieckes den Kugelmittelpunkt durchschreitet.

Somit gewinnt man aus (2) die Differentialgleichungen des Problems in der Form:

$$\left.\begin{aligned} \frac{dp_1}{dt} &= \frac{\epsilon_1}{2}R\left(\frac{1}{p_2}-\frac{1}{p_3}\right) \\ \frac{dp_2}{dt} &= \frac{\epsilon_2}{2}R\left(\frac{1}{p_3}-\frac{1}{p_1}\right) \\ \frac{dp_3}{dt} &= \frac{\epsilon_3}{3}R\left(\frac{1}{p_1}-\frac{1}{p_2}\right) \end{aligned}\right\} \tag{5}$$

Hier ergeben sich unmittelbar zwei Integrale. Es ist nämlich ersichtlich:

$$\frac{1}{\epsilon_1}\frac{dp_1}{dt}+\frac{1}{\epsilon_2}\frac{dp_2}{dt}+\frac{1}{\epsilon_3}\frac{dp_3}{dt}=0 \quad \text{und}$$
$$\frac{1}{\epsilon_1\,p_1}\frac{dp_1}{dt}+\frac{1}{\epsilon_2\,p_2}\frac{dp_2}{dt}+\frac{1}{\epsilon_3\,p_3}\frac{dp_3}{dt}=0\,,$$

also durch Integration nach t:

$$\left.\begin{aligned} \frac{p_1}{\epsilon_1}+\frac{p_2}{\epsilon_2}+\frac{p_3}{\epsilon_3} &= c = \text{ const} \\ \frac{\lg\ p_1}{\epsilon_1}+\frac{\lg\ p_2}{\epsilon_2}+\frac{\lg\ p_3}{\epsilon_3} &= c = \text{ const} \end{aligned}\right\} \tag{6}$$

Das zweite Integral ist identisch mit der „Constanz des Selbstpotentiales“ $\mathcal{H}$ oder dem „Integral der lebendigen Kraft“ (cf. III §4 Satz I), da hier:

$$\begin{aligned} \mathcal{H} &= \epsilon_2\ \epsilon_3\ \lg\frac{r_1}{2}+\epsilon_3\ \epsilon_1\ \lg\frac{r_2}{2}+\epsilon_1\ \epsilon_2\ \lg\frac{r_3}{2} \\ &= \frac{1}{2}\ \epsilon_1\ \epsilon_2\ \epsilon_3\ \left(\frac{\lg\frac{p_1}{4}}{\epsilon_1}+\frac{\lg\frac{p_2}{4}}{\epsilon_2}+\frac{\lg\frac{p_3}{4}}{\epsilon_3}\right)= \text{ const}\,, \end{aligned}$$

when setting

$$\left.\begin{aligned} p_1 &= \frac{1}{2} r_1^2 = 1 - \cos \pi_1 \\ p_2 &= \frac{1}{2} r_2^2 = 1 - \cos \pi_2 \\ p_3 &= \frac{1}{2} r_3^2 = 1 - \cos \pi_3 \,, \end{aligned}\right\} \tag{4}$$

and the root must be assumed to be positive or negative, depending on whether, when considering the spherical $\Delta\ P_1 P_2 P_3$ from outside and circumnavigating it in the direction $P_1 \to P_2 \to P_3$, we have its interior (i. e., the smaller piece of the spherical surface) or simply the interior of the planar $\Delta\ P_1 P_2 P_3$ to the left or to the right. R vanishes only when the 3 points P_1, P_2, P_3 lie in a great circle, and changes its sign when the plane of the triangle crosses the center of the sphere.

We thus obtain from (2) the differential equations of the problem in the form

$$\left.\begin{aligned} \frac{dp_1}{dt} &= \frac{\epsilon_1}{2} R \left(\frac{1}{p_2} - \frac{1}{p_3} \right) \\ \frac{dp_2}{dt} &= \frac{\epsilon_2}{2} R \left(\frac{1}{p_3} - \frac{1}{p_1} \right) \\ \frac{dp_3}{dt} &= \frac{\epsilon_3}{3} R \left(\frac{1}{p_1} - \frac{1}{p_2} \right) \end{aligned}\right\} \tag{5}$$

Here, two integrals immediately arise. For it is evident that

$$\frac{1}{\epsilon_1} \frac{dp_1}{dt} + \frac{1}{\epsilon_2} \frac{dp_2}{dt} + \frac{1}{\epsilon_3} \frac{dp_3}{dt} = 0 \quad \text{and}$$
$$\frac{1}{\epsilon_1\, p_1} \frac{dp_1}{dt} + \frac{1}{\epsilon_2\, p_2} \frac{dp_2}{dt} + \frac{1}{\epsilon_3\, p_3} \frac{dp_3}{dt} = 0 \,,$$

and hence, by integration with respect to t,

$$\left.\begin{aligned} \frac{p_1}{\epsilon_1} + \frac{p_2}{\epsilon_2} + \frac{p_3}{\epsilon_3} &= c = \text{ const} \\ \frac{\lg\ p_1}{\epsilon_1} + \frac{\lg\ p_2}{\epsilon_2} + \frac{\lg\ p_3}{\epsilon_3} &= c = \text{ const} \end{aligned}\right\} \tag{6}$$

The second integral is identical with the "constancy of the self-potential" $\mathcal{H}$ or with the "integral of the living force" (see III §4 Theorem I), since here

$$\begin{aligned} \mathcal{H} &= \epsilon_2\ \epsilon_3\ \lg \frac{r_1}{2} + \epsilon_3\ \epsilon_1\ \lg \frac{r_2}{2} + \epsilon_1\ \epsilon_2\ \lg \frac{r_3}{2} \\ &= \frac{1}{2}\ \epsilon_1\ \epsilon_2\ \epsilon_3\ \left(\frac{\lg \frac{p_1}{4}}{\epsilon_1} + \frac{\lg \frac{p_2}{4}}{\epsilon_2} + \frac{\lg \frac{p_3}{4}}{\epsilon_3} \right) = \text{ const} \,, \end{aligned}$$

99 | das erste aber lediglich eine Folge der Schwerpunktsätze. Nach einem Satze von Lagrange nämlich (Baltzer, a. a. O. V §11, 7) ist hier, wenn C den Kugelmittelpunkt und S den Schwerpunkt bezeichnet,

$$\epsilon_1\overline{CP_1}^2 + \epsilon_2\overline{CP_2}^2 + \epsilon_3\overline{CP_3}^2 - (\epsilon_1+\epsilon_2+\epsilon_3)\overline{CS}^2 = \\ = \frac{\epsilon_2\epsilon_3\overline{P_2P_3}^2 + \epsilon_3\epsilon_1\overline{P_3P_1}^2 + \epsilon_1\epsilon_2\overline{P_1P_2}^2}{\epsilon_1+\epsilon_2+\epsilon_3}$$

oder

$$\epsilon_1+\epsilon_2+\epsilon_3 = (1-h^2) = \frac{2\ \epsilon_2\ \epsilon_3\ p_1 + 2\ \epsilon_3\ \epsilon_1\ p_2 + 2\ \epsilon_1\ \epsilon_2\ p_3}{\epsilon_1+\epsilon_2+\epsilon_3} = \text{const}\,, \quad (7)$$

wo $h = \overline{CS} = \text{const}$ den (constanten) Abstand des Schwerpunktes vom Kugelmittelpunkt bedeutet. Die rechte Seite aber ist

$$= \frac{2\ \epsilon_1\ \epsilon_2\ \epsilon_3}{\epsilon_1+\epsilon_2+\epsilon_3}\left(\frac{p_1}{\epsilon_1}+\frac{p_2}{\epsilon_2}+\frac{p_3}{\epsilon_3}\right).$$

Der Gleichung (7) läßt sich auch die Form geben

$$h^2 - 1 = -\frac{\epsilon_2\ \epsilon_3\ r_1^2 + \epsilon_3\ \epsilon_1 r_2^2 + \epsilon_1\ \epsilon_2\ r_3^2}{(\epsilon_1+\epsilon_2+\epsilon_3)^2} = \text{const}\,, \quad (7)'$$

wo die linke Seite die Potenz des Schwerpunktes auf die Kugel, die rechte Seite aber seine Potenz auf den dem $\Delta\ P_1P_2P_3$ umschriebenen Kreis ausdrückt.

Die beiden Integrale (6) leisten zusammen die vollständige Integration der beiden Differentialgleichungen

$$\frac{dp_1}{p_1(p_2-p_3)} = \frac{dp_2}{p_2(p_3-p_1)} = \frac{dp_3}{p_3(p_1-p_2)}\,, \quad (5)'$$

welche die Bahncurven der Relativbewegung bestimmen. Es bedarf also nur
100 | noch einer einzigen Integration, und zwar einer Quadratur, um die drei Größen p_1, p_2, p_3 und damit alle Bestimmungsstücke des Dreieckes als Functionen der Zeit zu bestimmen.

Die Differentialgleichungen (5) haben eine gewisse formale Analogie mit den Eulerschen Gleichungen für die Drehung eines starren Körpers um einen festen Punkt:

$$\frac{dp_1}{dt} = c_1\ p_2\ p_3\,, \quad \frac{dp_2}{dt} = c_2\ p_3\ p_1\,, \quad \frac{dp_3}{dt} = c_3\ p_1\ p_2\,,$$

welche sich bekanntlich durch elliptische Functionen integrieren lassen. Eine analoge einfache Integration ist aber für unser Problem *nur* in dem Specialfalle möglich, wo alle drei Strudelmomente einander gleich sind: $\epsilon_1 = \epsilon_2 = \epsilon_3$, weshalb wir auch diesen Fall zuerst ausführlicher behandeln wollen.

but the first one is only a corollary of the theorem on centers of gravity. For, by *Lagrange*'s theorem (*Baltzer 1883*, V §11, 7), we have, assuming C denotes the center point of the sphere, and S the center of gravity,

$$\epsilon_1 \overline{CP_1}^2 + \epsilon_2 \overline{CP_2}^2 + \epsilon_3 \overline{CP_3}^2 - (\epsilon_1 + \epsilon_2 + \epsilon_3)\overline{CS}^2 = \\ = \frac{\epsilon_2\epsilon_3 \overline{P_2P_3}^2 + \epsilon_3\epsilon_1 \overline{P_3P_1}^2 + \epsilon_1\epsilon_2 \overline{P_1P_2}^2}{\epsilon_1 + \epsilon_2 + \epsilon_3}$$

or

$$\epsilon_1 + \epsilon_2 + \epsilon_3 = (1 - h^2) = \frac{2\ \epsilon_2\ \epsilon_3\ p_1 + 2\ \epsilon_3\ \epsilon_1\ p_2 + 2\ \epsilon_1\ \epsilon_2\ p_3}{\epsilon_1 + \epsilon_2 + \epsilon_3} = \text{const}\,, \quad (7)$$

where $h = \overline{CS} = \text{const}$ denotes the (constant) distance of the center of gravity from the center point of the sphere. But the right side is

$$= \frac{2\ \epsilon_1\ \epsilon_2\ \epsilon_3}{\epsilon_1 + \epsilon_2 + \epsilon_3} \left(\frac{p_1}{\epsilon_1} + \frac{p_2}{\epsilon_2} + \frac{p_3}{\epsilon_3} \right) .$$

It is possible to render equation (7) also in the form

$$h^2 - 1 = -\frac{\epsilon_2\ \epsilon_3\ r_1^2 + \epsilon_3\ \epsilon_1 r_2^2 + \epsilon_1\ \epsilon_2\ r_3^2}{(\epsilon_1 + \epsilon_2 + \epsilon_3)^2} = \text{const}\,, \quad (7)'$$

where the left side expresses the potential of the center of gravity onto the sphere, whereas the right side expresses its potential onto the circle circumscribing the $\Delta\ P_1P_2P_3$.

Together, the two integrals (6) achieve the complete integration of the two differential equations

$$\frac{dp_1}{p_1(p_2 - p_3)} = \frac{dp_2}{p_2(p_3 - p_1)} = \frac{dp_3}{p_3(p_1 - p_2)}\,, \quad (5)'$$

which determine the paths of the relative motion. Thus, it only takes one more integration, namely a quadrature, in order to determine as functions of time the three magnitudes p_1, p_2, p_3, and hence all that is needed to determine the triangle.

Differential equations (5) exhibit a certain formal analogy with the Euler equations for the rotation of a rigid body around a fixed point:

$$\frac{dp_1}{dt} = c_1\ p_2\ p_3\,, \quad \frac{dp_2}{dt} = c_2\ p_3\ p_1\,, \quad \frac{dp_3}{dt} = c_3\ p_1\ p_2\,,$$

which, as is well-known, can be integrated by means of elliptic functions. But, for our problem, an analogous, simple integration is possible *only* in the special case where all three whirl momenta are equal to one another: $\epsilon_1 = \epsilon_2 = \epsilon_3$, which is why we will consider this case in greater detail first.

§2. Der Fall gleicher Strudelmomente

Ohne Beschränkung der Allgemeinheit können wir annehmen:

$$\epsilon_1 = \epsilon_2 = \epsilon_3 = 1\,,$$

indem wir über die Einheit der Zeit zweckmäßig verfügen. Unsere Differentialgleichungen (5) nehmen jetzt die Form an:

101 |
$$\left.\begin{aligned} \frac{dp_1}{dt} &= \frac{R}{2}\left(\frac{1}{p_2}-\frac{1}{p_3}\right) = \frac{R}{2}\,\frac{p_3-p_2}{p_2p_3} \\ \frac{dp_2}{dt} &= \frac{R}{2}\left(\frac{1}{p_3}-\frac{1}{p_1}\right) = \frac{R}{2}\,\frac{p_1-p_3}{p_3p_1} \\ \frac{dp_3}{dt} &= \frac{R}{2}\left(\frac{1}{p_1}-\frac{1}{p_2}\right) = \frac{R}{2}\,\frac{p_2-p_1}{p_1p_2} \end{aligned}\right\} \tag{1}$$

und die beiden Integrale (6) p. 98 lassen sich schreiben:

$$\left.\begin{aligned} p_1+p_2+p_3 &= m = \text{ const} \\ p_1\,p_2\,p_3 &= n = \text{ const} \end{aligned}\right\} \tag{2}$$

Mittelst dieser beiden Integrale kann man nun z. B. p_2 und p_3 eliminieren: $p_2+p_3 = m-p_1,\ \ p_1\,p_3 = \dfrac{n}{p_1},\quad p_2-p_3 = \sqrt{(m-p_1)^2-\dfrac{4n}{p_1}}$ (vergl. (3) des vorigen §) und

$$\left.\begin{aligned} R &= \sqrt{-(2p_1-m)^2+\frac{4n}{p_1}-2n}\,, \text{ also} \\ \frac{dp_1}{dt} &= \frac{p_1}{2n}\sqrt{\left[-(2p_1-m)^2+\frac{4n}{p_1}-2n\right]\left[(m-p_1)^2-\frac{4n}{p_1}\right]} \\ &\frac{1}{2n}\sqrt{[-p_1(2p_1-m)^2-2np+4n]\,[p_1(m-p_1)^2-4n]}\,, \end{aligned}\right\} \tag{3}$$

wo unter der $\sqrt{\ }$ eine ganze Function 6 ten Grades steht. Die Integration dieser Differentialgleichung (welche für jede der drei Größen p_1, p_2, p_3 gelten muß) führt also auf ein hyperelliptisches Integral, dessen Umkehrung nicht eindeutig wird. Wir können aber in unserem Falle die Einführung hyperelliptischer Functionen vermeiden und die Integration rein durch elliptische Functionen vollziehen, wenn wir als neue Variable einführen:

102 |
$$w = p_2p_3+p_3p_1+p_1p_2\,. \tag{2a}$$

§2. The case of equal whirl momenta

We can assume without loss of generality that

$$\epsilon_1 = \epsilon_2 = \epsilon_3 = 1\,,$$

by purposeful deployment of the unit of time. Our differential equations (5) now assume the form

$$\left.\begin{aligned} \frac{dp_1}{dt} &= \frac{R}{2}\left(\frac{1}{p_2} - \frac{1}{p_3}\right) = \frac{R}{2}\,\frac{p_3 - p_2}{p_2 p_3} \\ \frac{dp_2}{dt} &= \frac{R}{2}\left(\frac{1}{p_3} - \frac{1}{p_1}\right) = \frac{R}{2}\,\frac{p_1 - p_3}{p_3 p_1} \\ \frac{dp_3}{dt} &= \frac{R}{2}\left(\frac{1}{p_1} - \frac{1}{p_2}\right) = \frac{R}{2}\,\frac{p_2 - p_1}{p_1 p_2} \end{aligned}\right\} \tag{1}$$

and the two integrals (6) p. 98 can be written as follows:

$$\left.\begin{aligned} p_1 + p_2 + p_3 = m = \ \text{const} \\ p_1\, p_2\, p_3 = \ n \ = \ \text{const} \end{aligned}\right\} \tag{2}$$

By means of these two integrals it is now possible to eliminate, e. g., p_2 and p_3: $p_2 + p_3 = m - p_1$, $p_1\, p_3 = \dfrac{n}{p_1}$, $\quad p_2 - p_3 = \sqrt{(m-p_1)^2 - \dfrac{4n}{p_1}}$ (comp. (3) of the previous §) and

$$\left.\begin{aligned} R &= \sqrt{-(2p_1 - m)^2 + \frac{4n}{p_1} - 2n}\,, \ \text{and hence} \\ \frac{dp_1}{dt} &= \frac{p_1}{2n}\sqrt{\left[-(2p_1 - m)^2 + \frac{4n}{p_1} - 2n\right]\left[(m-p_1)^2 - \frac{4n}{p_1}\right]} \\ &\frac{1}{2n}\sqrt{[-p_1(2p_1 - m)^2 - 2np + 4n]\,[p_1(m-p_1)^2 - 4n]}\,, \end{aligned}\right\} \tag{3}$$

where an entire function of the 6th degree is below the $\sqrt{\ }$. Hence, the integration of this differential equation (which must hold for each of the three magnitudes p_1, p_2, p_3) leads to a hyperelliptic integral whose reverse is not single-valued. In our case, however, we can avoid the introduction of hyperelliptic functions and we can perform the integration by elliptic functions only, if we introduce as a new variable

$$w = p_2 p_3 + p_3 p_1 + p_1 p_2\,. \tag{2a}$$

Jetzt sind nämlich alle 3 elementaren symmetrischen Functionen der Größen p_1, p_2, p_3 als bekannt anzusehen, und wir können die Gleichung aufstellen:

$$z^3 - mz^2 + wz - n = 0\,, \tag{4}$$

welche die drei unbekannten Functionen p_1, p_2, p_3 zu Wurzeln hat und für jeden Wert von w ihre Werte zu bestimmen gestattet.

Wir brauchen also nur noch die Function $w = w(t)$ zu bestimmen und bilden aus (2a) zunächst:

$$\begin{aligned}\frac{dw}{dt} &= (p_2+p_3)\frac{dp_1}{dt} + (p_3+p_1)\frac{dp_2}{dt} + (p_1+p_2)\ \frac{dp_3}{dt}\\ &= \frac{1}{2}\frac{R}{p_1p_2p_3}\left[p_1(p_3^2-p_2^2) + p_2(p_1^2-p_3^2) + p_3(p_2^2-p_1^2)\right]\\ &= \frac{R}{2n}(p_3-p_2)\ (p_2-p_1)\ (p_1-p_3)\\ &= \pm\frac{R}{2n}\sqrt{D}\,,\end{aligned}$$

wo $D = D(w)$ die Diskriminante der Gleichung (4) bezeichnet und den Wert hat (cf. Weber, Algebra I §159)

$$\begin{aligned}D = D(w) &= -4w^3 + m^2w^2 + 18mnw - 4m^3n - 27n^2\\ &= -4(w-w_1)(w-w_2)(w-w_3)\,,\end{aligned} \tag{5}$$

also eine ganze Function dritten Grads von w. Zugleich ist aber nach (3) p. 87

$$\begin{aligned}R^2 &= -p_1^2 - p_2^2p_3^2 + 2p_2p_3 + 2p_3p_1 + 2p_1p_2 - 2p_1p_2p_3\\ &= -(p_1+p_2+p_3)^2 + 4p_2p_3 + 4p_3p_1 + 4p_1p_2 - 2p_1p_2p_3\\ R^2 &= -m^2 + 4w - 2n\,, \quad R = \pm\sqrt{4w - m^2 - 2n}\,.\end{aligned} \tag{6}$$

103 | Wir erhalten also zur Bestimmung von $w = w(t)$ die Differentialgleichung

$$\frac{dw}{dt} = \pm\frac{1}{2n}\sqrt{(4w - m^2 - 2n)D(w)}\,, \tag{7}$$

wo jetzt der Radicandus in der That nur vom 4ten Grade ist. So ergiebt sich die Zeit t als ein elliptisches Integral erster Gattung von w:

$$2nt = \int \frac{dw}{\pm\sqrt{(4w - m^2 - 2n)D(w)}} \tag{7a}$$

und umgekehrt w als eindeutige elliptische Function von t. Die drei halben Quadrate der Dreiecksseiten p_1, p_2, p_3 sind dann als Wurzeln von (4) dreideutige elliptische Functionen der Zeit.

For, now, all three 3 elementary symmetric functions of the magnitudes p_1, p_2, p_3 are to be considered known, and we can set up the equation

$$z^3 - mz^2 + wz - n = 0 \,, \tag{4}$$

which has the three unknown functions p_1, p_2, p_3 as roots and allows for the determination of its values for every value of w.

Hence, we only need to determine the function $w = w(t)$, and from (2a) we first form

$$\begin{aligned}
\frac{dw}{dt} &= (p_2 + p_3)\frac{dp_1}{dt} + (p_3 + p_1)\frac{dp_2}{dt} + (p_1 + p_2)\ \frac{dp_3}{dt} \\
&= \frac{1}{2}\frac{R}{p_1 p_2 p_3}\left[p_1(p_3^2 - p_2^2) + p_2(p_1^2 - p_3^2) + p_3(p_2^2 - p_1^2)\right] \\
&= \frac{R}{2n}(p_3 - p_2)\ (p_2 - p_1)\ (p_1 - p_3) \\
&= \pm\frac{R}{2n}\sqrt{D}\,,
\end{aligned}$$

where $D = D(w)$ denotes the discriminant of equation (4) and has the value (see *Weber 1895*, §159)

$$\begin{aligned}
D = D(w) &= -4w^3 + m^2w^2 + 18mnw - 4m^3n - 27n^2 \\
&= -4(w - w_1)(w - w_2)(w - w_3)\,,
\end{aligned} \tag{5}$$

and hence ⟦is⟧ an entire function of the third degree of w. But, at the same time, by (3) p. 87,

$$\begin{aligned}
R^2 &= -p_1^2 - p_2^2p_3^2 + 2p_2p_3 + 2p_3p_1 + 2p_1p_2 - 2p_1p_2p_3 \\
&= -(p_1 + p_2 + p_3)^2 + 4p_2p_3 + 4p_3p_1 + 4p_1p_2 - 2p_1p_2p_3 \\
R^2 &= -m^2 + 4w - 2n\,, \quad R = \pm\sqrt{4w - m^2 - 2n}\,.
\end{aligned} \tag{6}$$

Thus, we obtain for the determination of $w = w(t)$ the differential equation

$$\frac{dw}{dt} = \pm\frac{1}{2n}\sqrt{(4w - m^2 - 2n)D(w)}\,, \tag{7}$$

where, now, the radicand is actually only of the 4th degree. Time t thus emerges as an elliptic integral of the first kind of w:

$$2nt = \int \frac{dw}{\pm\sqrt{(4w - m^2 - 2n)D(w)}} \tag{7a}$$

and, conversely, w as a single-valued elliptic function of t. Then the three half-squares of the triangle sides p_1, p_2, p_3, being roots of (4), are three-valued elliptic functions of time.

Anstatt nun das elliptische Integral (7a) auf seine Normalform zu reducieren, wie es zum Zwecke der numerischen Berechnung erforderlich wäre, ziehen wir es vor, unsere Differentialgleichung (7) in der ursprünglichen Form zu diskutieren.

Zunächst haben wir die Bedingungen aufzusuchen, denen ein Wert von w genügen muß, damit ihm ein reelles Dreieck in der Kugel und damit eine mögliche Configuration der drei Strudel entspricht. Vor allem müssen die drei
104 Größen p_1, p_2, p_3 | als Quadrate der Dreiecksseiten reell und positiv sein, und nach (2) und (2a) ebenso auch die Größen m, n, w. Deswegen müssen auch die beiden Constanten m, n der Bedingung genügen

$$m^3 \geq 27n \ , \tag{8}$$

weil das arithmetische Mittel $\frac{m}{3}$ der drei positiven Werte p_1, p_2, p_3 nicht kleiner sein kann als ihr geometrisches Mittel $\sqrt[3]{n}$. Damit ferner die Gleichung (4) drei reelle Wurzeln hat, muß ihre Discriminante $D(w) \geq 0$ sein, $D(w) = 0$ würde ausdrücken, daß das Dreieck gleichschenklig ist. Da also der eine Factor $D(w)$ unter der Quadratwurzel in (7) positiv sein soll, so muß es auch der andere, wenn $\frac{dw}{dt}$ reell ist, also

$$w \geq \frac{m^2 + 2n}{4} = w_0 > 0 \ . \tag{9}$$

Diese Bedingung (9) aber drückt aus, daß das Dreieck p_1, p_2, p_3 der Kugel eingeschrieben werden kann, d. h. daß der Radius seines umschriebenen Kreises ≤ 1 ist. Dieser Radius ϱ wird nämlich allgemein

$$\begin{aligned} \varrho &= \sqrt{\frac{2\ p_1p_2p_3}{2p_2p_3 + 2p_3p_1 + 2p_1p_2 - p_1^2 - p_2^2 - p_3^2}} \\ &= \frac{\sqrt{2p_1p_2p_3}}{\Delta} = \sqrt{\frac{2p_1p_2p_3}{R^2 + 2p_1p_2p_3}} \end{aligned} \tag{10}$$

wo

$$\Delta = \sqrt{2p_2p_3 + 2p_3p_1 + 2p_1p_2 - p_1^2 - p_2^2 - p_3^2} = \sqrt{R^2 + 2p_1p_2p_3}$$

der doppelte Flächeninhalt des Dreieckes ist. In unserem Falle aber wird

$$\Delta = \sqrt{4w - m^2} \ , \quad \varrho = \sqrt{\frac{2n}{4w - m^2}} \leq 1 \tag{10$'$}$$

wenn $R^2 = 4w - m^2 - 2n = 4(w - w_0) \geq 0$ ist.
105 | Wir haben also:

Sind m und n positive Grössen, welche der Bedingung (8) genügen $m^3 \geq 27n$, so entspricht einem Werte von w mittelst (4) dann und nur dann ein

Now, instead of reducing the elliptic integral (7a) to its normal form, as would be necessary for the numeric computation, we prefer to discuss our differential equation (7) in the original form.

First, we have to explore the conditions a value of w must satisfy so that there is a real triangle in the sphere, and hence a possible configuration of the three whirls, corresponding to it. In particular, the three magnitudes p_1, p_2, p_3, being squares of the triangle sides, must be real and positive, and so, by (2) and (2a), must also be the magnitudes m, n, w. For this reason, the two constants m, n, too, must satisfy the condition

$$m^3 \geq 27n \ , \tag{8}$$

since the arithmetic mean $\dfrac{m}{3}$ of the three positive values p_1, p_2, p_3 cannot be smaller than their geometric mean $\sqrt[3]{n}$. Furthermore, for equation (4) to have three real roots, its discriminant $D(w)$ must be ≥ 0, $D(w) = 0$ expressing that the triangle is isosceles. Hence, since the factor $D(w)$ beneath the square root in (7) is supposed to be positive, the other factor must also be positive, if $\dfrac{dw}{dt}$ is real, and hence

$$w \geq \frac{m^2 + 2n}{4} = w_0 > 0 \ . \tag{9}$$

But this condition (9) expresses that the triangle p_1, p_2, p_3 can be inscribed into the sphere, i. e., that the radius of its circumscribed circle is ≤ 1. For this radius ϱ generally becomes

$$\begin{aligned} \varrho &= \sqrt{\frac{2\ p_1 p_2 p_3}{2p_2p_3 + 2p_3p_1 + 2p_1p_2 - p_1^2 - p_2^2 - p_3^2}} \\ &= \frac{\sqrt{2p_1p_2p_3}}{\Delta} = \sqrt{\frac{2p_1p_2p_3}{R^2 + 2p_1p_2p_3}} \end{aligned} \tag{10}$$

where

$$\Delta = \sqrt{2p_2p_3 + 2p_3p_1 + 2p_1p_2 - p_1^2 - p_2^2 - p_3^2} = \sqrt{R^2 + 2p_1p_2p_3}$$

is double the area of the triangle. But in our case

$$\Delta = \sqrt{4w - m^2} \ , \quad \varrho = \sqrt{\frac{2n}{4w - m^2}} \leq 1 \tag{10}'$$

if $R^2 = 4w - m^2 - 2n = 4(w - w_0)$ is ≥ 0.

We therefore have:

If m and n are positive magnitudes satisfying condition (8) $m^3 \geq 27n$, then to a value of w there corresponds a possible triangle in the sphere by

mögliches Dreieck in der Kugel, wenn (9) $w \geq w_0 = \dfrac{m^3+2n}{4} > 0$ und gleichzeitig die Discriminante (5) $D(w) = D(w; m, n) \geq 0$ ist. Der Grenzfall $D(w) = 0$ bedeutet, dass das Dreieck gleichschenklig ist, und der andere $w = w_0$ oder $R = 0$, dass es in einen grössten Kugelkreis fällt.

Betrachten wir nun die kubische Gleichung

$$D(w) = -4w^3 + m^2w^2 + 18mnw - 4m^3n - 27n^2 = 0 \ , \tag{5a}$$

so hat ihre Discriminante den Wert

$$D_1 = D_1(m, n) = 16n(m^3 - 27n)^3 \geq 0 \ , \tag{11}$$

welcher nach (8) immer positiv sein muß und nur verschwindet für ein gleichseitiges Dreieck $p_1 = p_2 = p_3 = \dfrac{m}{3} = \sqrt[3]{n}$. Also: Unsere Gleichung (5a) $D(w) = 0$, welche die gleichschenkligen Dreiecke für m und n bestimmt, hat immer drei reelle Wurzeln w_1, w_2, w_3, von denen die eine w_1 stets negativ und die anderen beiden w_2, w_3 positiv sind, weil die Anzahl der Zeichenwechsel in den Coefficienten von (5a) = 2 ist. Da nun gleichzeitig immer $D(\pm\infty) = \mp\infty$
106 ist, so hat die Curve der Function $D = D(w)$ den hier ge-|zeichneten Verlauf, und zwar ist der schraffierte Teil der Curve zwischen den beiden positiven

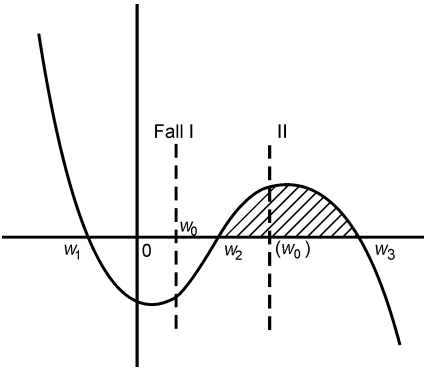


Fig. 7.

Wurzeln w_2, w_3 der einzige, welchem reelle Dreiecke entsprechen. Da aber in unserem Problem diese Dreiecke noch in der Kugel 1 liegen müssen und daher $w > w_0$, so sind die beiden Hauptfälle zu unterscheiden, in welchen 1) $0 < w_0 < w_2$, 2) $w_2 < w_0 < w_3$ ist, während der Fall $w_0 > w_3$ bei reellen Bewegungen überhaupt nicht vorkommen kann.

I) *Erster Hauptfall:* $0 < w_0 < w_2 \leq w \leq w_3$.

$$\frac{dw}{dt} = \pm\frac{2}{n}\sqrt{(w-w_0)(w-w_1)(w-w_2)(w_3-w)}$$

(4) if and only if (9) $w \geq w_0 = \dfrac{m^2 + 2n}{4} > 0$,[2] and, at the same time, the discriminant (5) $D(w) = D(w; m, n)$ is ≥ 0. The borderline case $D(w) = 0$ means that the triangle is isosceles, and the other, $w = w_0$ or $R = 0$, that it falls into a great spherical circle.

If we now consider the cubic equation

$$D(w) = -4w^3 + m^2w^2 + 18mnw - 4m^3n - 27n^2 = 0 \ , \tag{5a}$$

then its discriminant has the value

$$D_1 = D_1(m, n) = 16n(m^3 - 27n)^3 \geq 0 \ , \tag{11}$$

which, by (8), must always be positive and only vanishes for an equilateral triangle $p_1 = p_2 = p_3 = \dfrac{m}{3} = \sqrt[3]{n}$. And hence: Our equation (5a) $D(w) = 0$, which determines the isosceles triangles for m and n, always has three real roots w_1, w_2, w_3, of which w_1 is always negative and the other two w_2, w_3 are positive, since the number of changes in sign in the coefficients of (5a) is $= 2$. Since, at the same time, always $D(\pm\infty) = \mp\infty$, the graph of the function $D = D(w)$ takes the course drawn here, and in particular the shaded part of the graph between the two positive roots w_2, w_3 is the only one to which

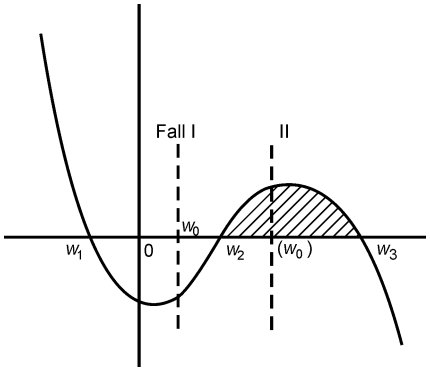


Fig. 7.

real triangles correspond. But since in our problem these triangles must still lie in sphere 1, and hence $w > w_0$, we need to distinguish the two main cases 1) $0 < w_0 < w_2$, 2) $w_2 < w_0 < w_3$, whereas the case $w_0 > w_3$ can never occur for real motions.

I) *First main case:* $0 < w_0 < w_2 \leq w \leq w_3$.

$$\frac{dw}{dt} = \pm\frac{2}{n}\sqrt{(w - w_0)(w - w_1)(w - w_2)(w_3 - w)}$$

[2] ⟦Zermelo erroneously writes "m^3" instead of "m^2".⟧

und w muß zwischen den Werten $w = w_2$ und $w = w_3$, d. h. zwischen zwei gleichschenkligen Dreiecksformen periodisch hin- und herschwanken, wobei das Dreieck niemals in einen grössten Kreis rückt ($w = w_0$). Die halbe Periode τ hat den Wert

$$\tau = \frac{n}{2} \int_{w_2}^{w_3} \frac{dw}{\sqrt{(w - w_0)(w - w_1)(w - w_2)(w_3 - w)}} ,$$

welcher nur im Grenzfalle (Ia) $w_0 = w_2$ unendlich gross wird, im anderen Grenzfall (Ib) $w_2 = w_3$ aber nach dem endlichen Werte

$$\tau_0 = \frac{n}{2} \lim_{w_2 = w_3} \int_{w_2}^{w_3} \frac{dw}{\sqrt{}} = \frac{n\pi}{4} \frac{1}{\sqrt{(w_2 - w_0)(w_2 - w_1)}}$$

convergiert.

107 | Während also mit w das Dreieck, als Ganzes betrachtet, in der Periode 2τ seine anfängliche Gestalt wieder annimmt, unterliegen seine einzelnen Seiten, $\overline{P_2P_3}$ z. B., und mit ihnen die Größen p_1, p_2, p_3 einer dreifach längeren Periode 6τ. Wenn nämlich das Dreieck bei $w = w_2$ oder $w = w_3$ gleichschenklig wird, z. B. $p_2 = p_3$, so hat hier zwar mit w auch die Basisseite p_1 ein Maximum oder Minimum (cf. (1) p. 101), die beiden anderen p_2 und p_3 aber nicht, sondern sie fahren fort zu wachsen, bzw. abzunehmen, wobei zunächst $\frac{dp_2}{dt} + \frac{dp_3}{dt} = 0$ ist. Das bedeutet: das Dreieck kehrt zwar mit w in seine früheren Gestalten zurück, aber nicht seine einzelnen Seiten p_2, p_3, sondern diese vertauschen sich gegen einander indem die (zunehmende) kleinere an die Stelle der (abnehmenden) größeren tritt, und umgekehrt. Diese Vertauschung findet jedesmal statt, wenn zwei Seiten gleich werden, und dabei muß, wie man sich ohne Schwierigkeit überzeugt, der ganze Cyclus aller 6 möglichen Permutationen wirklich durchlaufen werden, und zwar nach dem folgenden Schema:

108 |

$$\begin{array}{lccccl}
 & p_1 & p_2 & p_3 & & \\
 & & & \times & &) \; w = w_2 \\
 & p_1 & p_3 & p_2 & & \\
w = w_3 \; (& & \times & & & \\
 & p_3 & p_1 & p_2 & & \\
 & & & \times & &) \; w = w_2 \\
 & p_3 & p_2 & p_1 & & \\
w = w_3 \; (& & \times & & & \\
 & p_2 & p_3 & p_1 & & \\
 & & & \times & &) \; w = w_2 \\
 & p_2 & p_1 & p_3 & & \\
w = w_3 \; (& & \times & & & \\
 & p_1 & p_2 & p_3 & & \text{usw.}
\end{array}$$

Fig. 8

and w must periodically oscillate between the values $w = w_2$ and $w = w_3$, i. e., between two isosceles triangle forms, where the triangle never moves into a great circle ($w = w_0$). The half-period τ has the value

$$\tau = \frac{n}{2} \int_{w_2}^{w_3} \frac{dw}{\sqrt{(w - w_0)(w - w_1)(w - w_2)(w_3 - w)}} ,$$

which becomes infinitely large only in the borderline case (Ia) $w_0 = w_2$, but which, in the other borderline case (Ib) $w_2 = w_3$, converges toward the finite value

$$\tau_0 = \frac{n}{2} \lim_{w_2 = w_3} \int_{w_2}^{w_3} \frac{dw}{\sqrt{}} = \frac{n\pi}{4} \frac{1}{\sqrt{(w_2 - w_0)(w_2 - w_1)}} .$$

Hence, while along with w, the triangle, considered as a whole, regains its initial shape during the period 2τ, its individual sides, e. g. $\overline{P_2P_3}$, and along with them the magnitudes p_1, p_2, p_3, are subject to a period three times that length. For if the triangle becomes isosceles for $w = w_2$ or $w = w_3$, e. g. $p_2 = p_3$, then, along with w, the base p_1 also has a maximum or minimum (see (1) p. 101), but the other two sides p_2 and p_3 do not, but instead continue to grow or shrink, where at first $\frac{dp_2}{dt} + \frac{dp_3}{dt} = 0$. This means: while the triangle, along with w, returns to its previous shapes, its individual sides p_2, p_3, do not so return but are interchanged with one another so that the (increasing) smaller one takes the place of the (decreasing) greater one, and vice versa. This interchange always occurs when two sides become equal to one another. In this case, as is easily seen, the entire cycle of all 6 possible permutations must be actually traversed, namely according to the following schema:

$$\begin{array}{lcccl}
 & p_1 & p_2 & p_3 & \\
 & & \times & &) \; w = w_2 \\
 & p_1 & p_3 & p_2 & \\
w = w_3 \; (& \times & & & \\
 & p_3 & p_1 & p_2 & \\
 & & \times & &) \; w = w_2 \\
 & p_3 & p_2 & p_1 & \\
w = w_3 \; (& \times & & & \\
 & p_2 & p_3 & p_1 & \\
 & & \times & &) \; w = w_2 \\
 & p_2 & p_1 & p_3 & \\
w = w_3 \; (& \times & & & \\
 & p_1 & p_2 & p_3 & \text{etc.}
\end{array}$$

Fig. 8

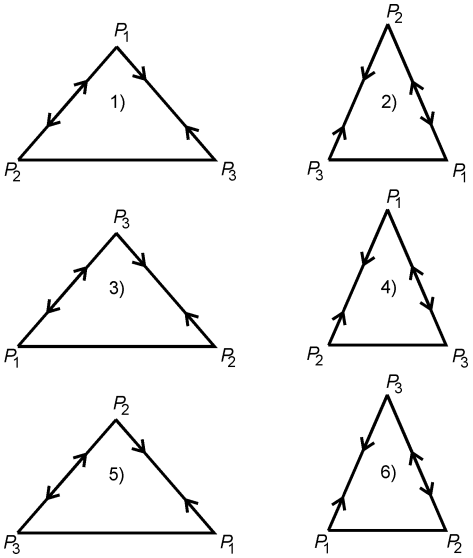


Fig. 8.

Die dabei auftretenden 6 gleichschenkligen Dreiecke, welche, je 3 und 3, zu zwei Typen ($w = w_2$ und $w = w_3$) gehören, soll die beistehende Figur veranschaulichen, in welcher das Wachsen und Abnehmen der Dreiecksseiten durch Pfeile ($\leftrightarrow$ und $\rightarrow\leftarrow$) angedeutet ist:

109 | Im *Grenzfall* (*Ia*): $0 < w_0 = w_2 \leq w \leq w_3$ wird die Periode τ unendlich, also die Bewegung aperiodisch, d. h. das Dreieck nähert sich asymptotisch einem gleichschenkligen im grössten Kreise ($w = w_0 = w_2$), welches eine Gleichgewichtsfigur, aber eine instabile darstellt. Wenn nämlich einmal w noch so wenig von $w_0 = w_2$ verschieden ist, so kann es zwar direct asymptotisch in w_0 übergehen, es kann aber auch zuerst in endlicher Zeit eine einmalige Schwankung nach w_3 vornehmen, um dann allerdings sicher sich w_0 asymptotisch zu nähern (die Unterscheidung der beiden Fälle hängt von dem Vorzeichen von $\sqrt{D(w)}$, also wesentlich von dem Sinn des Dreickes ab).

Man erkennt diesen Fall daran, daß von zwei möglichen gleichschenkligen Dreiecken (w_2 und w_3) das eine im grössten Kreise liegt.

Im *Grenzfall Ib:* $w_0 < w_2 = w = w_3$ ist das Dreieck im Anfang gleichseitig und bleibt starr für alle Zeiten. Dieses (relative) Gleichgewicht ist auch stabil. Ist hier zugleich auch $w_0 = w_2 = w_3$, so liegt das gleichseitige Dreieck noch im grössten Kreise, und nach Cap. III §6 fin. sind die drei Strudel dann sogar im absoluten Gleichgewichte.

110 | Im *zweiten Hauptfalle II:* $w_2 < w_0 \leq w \leq w_3$ giebt es nur *ein* brauchbares gleichschenkliges Dreieck und *ein* Dreieck im größten Kreise, und zwischen beiden Grenzfiguren schwankt das veränderliche Dreieck periodisch hin und

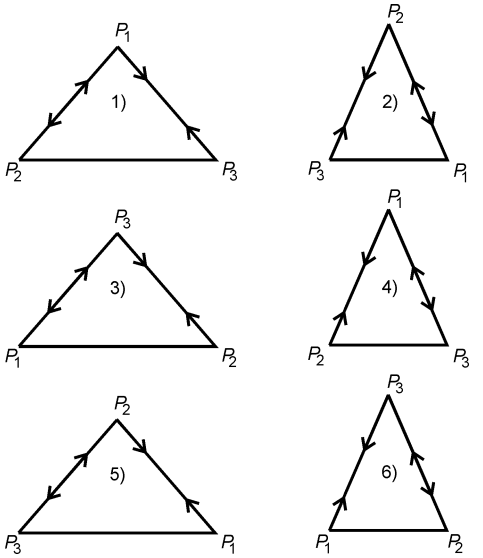


Fig. 8.

The adjacent figure is meant to illustrate the 6 isosceles triangles that occur in this case and that form two threesomes, each of which belongs to one of 2 types ($w = w_2$ or $w = w_3$). In the figure, the increase and decrease of the triangle sides is indicated by arrows ($\leftrightarrow$ and $\rightarrow\leftarrow$):

In the *borderline case* (*Ia*): $0 < w_0 = w_2 \leq w \leq w_3$ the period τ becomes infinite, and hence the motion aperiodic, i. e., the triangle asymptotically approach an isosceles triangle in the great circle ($w = w_0 = w_2$), which represents an equilibrium figure, albeit an unstable one. For as soon as w differs only slightly from $w_0 = w_2$, it can directly approach w_0 asymptotically, but also, first take a one-time turn toward w_3 during a finite period of time and then to steadily approach w_0 asymptotically (the distinction between the two cases depends on the sign of $\sqrt{D(w)}$, and hence essentially on the sense of direction of the triangle).

One recognizes this case by the fact that one of two possible isosceles triangles (w_2 and w_3) lies in the great circle.

In the *borderline case Ib:* $w_0 < w_2 = w = w_3$ the triangle is initially equilateral and remains rigid for all times. This (relative) equilibrium is also stable. If here also $w_0 = w_2 = w_3$, then the equilateral triangle still lies in the great circle, and, according to ch. III §6 fin., the three whirls are then even in absolute equilibrium.

In the *second main case II:* $w_2 < w_0 \leq w \leq w_3$ there is only *one* viable isosceles triangle and *one* triangle in the great circle, and the variable triangle

her mit der halben Periode:

$$\tau = \frac{n}{2}\int_{w_0}^{w_3} \frac{dw}{\sqrt{(w-w_0)(w-w_1)(w-w_2)(w_3-w)}} ,$$

welche für $w_0 = w_2$ unendlich groß wird, aber für $w_0 = w_3$ nach einem endlichen Werte convergiert:

$$\tau_0 = \frac{n\pi}{4} \frac{1}{\sqrt{(w_0-w_1)(w_0-w_2)}} .$$

Wird das Dreieck gleichschenklig bei $w = w_3$, so vertauschen sich wieder die beiden gleichen Seiten, bei $w = w_0$ aber, wo das Dreieck in den grössten Kreis rückt, wechselt $R = \sqrt{4w - m^2 - 2n} = 2\sqrt{w - w_0}$ sein Vorzeichen, und alle Seiten kehren ihre Bewegung um. Dabei kann immer nur dieselbe Seite z. B. p_1 Basis des gleichschenkligen Dreieckes werden und hat ihr Maximum als Basis und ihr Minimum im größten Kreis, oder umgekehrt. Die beiden anderen Seiten aber haben sowohl Maximum als Minimum im größten Kreis, und zwar abwechselnd. Die wahre Periode ist 4τ.

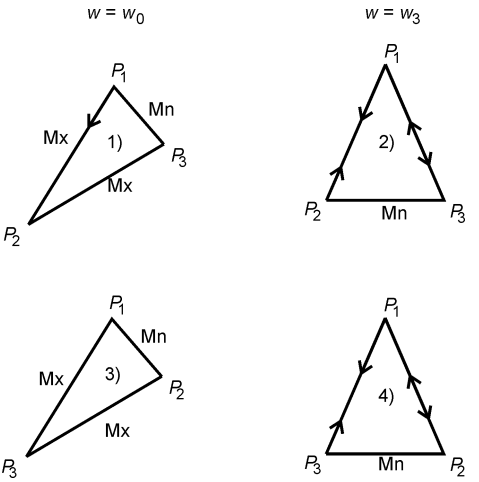


Fig. 9.

111 | Im *Grenzfall IIa:* $w_2 < w_0 = w_3$ ist auch $w = w_0 = w_3$, und das Dreieck ist gleichschenklig im grössten Kreise und starr für alle Zeiten, wobei das relative Gleichgewicht auch stabil ist.

Um also die verschiedenen Fälle zu unterscheiden, genügt es, die mit den Bedingungen (2) S. 101:

$$p_1 + p_2 + p_3 = m\,, \quad p_1 p_2 p_3 = n$$

verträglichen Wertsysteme p_1, p_2, p_3, aufzusuchen, welche gleichschenkligen Dreiecken entsprechen. Zu diesem Zwecke bedient man sich anstatt der Gleichung (5a) $D(w) = 0$ einfacher der directen Substitution $p_2 = p_3$ in die

periodically oscillates between the two limit figures with the half-period

$$\tau = \frac{n}{2} \int_{w_0}^{w_3} \frac{dw}{\sqrt{(w-w_0)(w-w_1)(w-w_2)(w_3-w)}} ,$$

which becomes infinitely large for $w_0 = w_2$, but for $w_0 = w_3$ converges toward a finite value:

$$\tau_0 = \frac{n\pi}{4} \frac{1}{\sqrt{(w_0-w_1)(w_0-w_2)}} .$$

If the triangle becomes isosceles when $w = w_3$, then the two equal sides are in turn interchanged, but when $w = w_0$, where the triangle moves into the great circle, $R = \sqrt{4w - m^2 - 2n} = 2\sqrt{w - w_0}$ changes its sign, and all sides reverse their motion. Here only the same side, e. g. p_1, can become the base of the isosceles triangle and has its maximum as base and its minimum in the great circle, or vice versa. The two other sides, however, have both maximum and minimum in the great circle, namely alternately. The true period is 4τ.

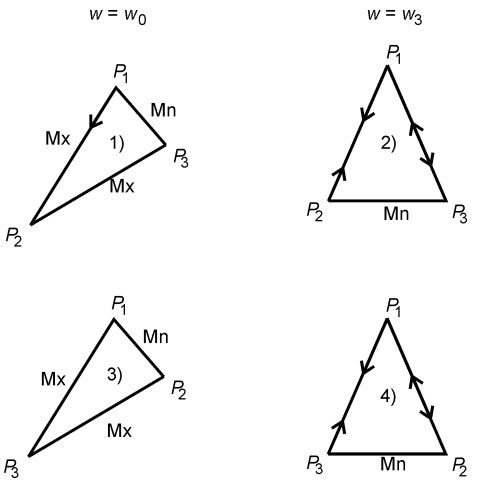


Fig. 9.

In the *borderline case IIa:* $w_2 < w_0 = w_3$ also $w = w_0 = w_3$, and the triangle is isosceles in the great circle and rigid at all times, where the relative equilibrium is also stable.

Hence, in order to distinguish between the different cases it suffices to find the value systems p_1, p_2, p_3 compatible with conditions (2) p. 101:

$$p_1 + p_2 + p_3 = m , \quad p_1 p_2 p_3 = n ,$$

which correspond to isosceles triangles. To this end one uses more simply the direct substitution $p_2 = p_3$ into the equations (2) instead of using the

Gleichungen (2), und erhält so

$$p_1 + 2p_2 = m\ , \quad p_1 p_2^2 = n\,,$$

also die kubische Gleichung für $p_2 = p_3$

$$p_2^2(m - 2p_2) = n$$

oder für p_1:

$$p_1(m - p_1)^2 = 4n\,.$$

Die so gefundenen beiden gleichschenkligen Dreiecke sind aber dann noch darauf zu untersuchen, ob ihr Radius ϱ nicht etwa > 1 wird, und in diesem Falle zu streichen, und dann erkennt man

1) Wenn *beide* gleichsch. Dr. brauchbar sind, so haben wir *Fall* I (S. 106) und event. *Grenzfall* Ia (S. 109) wenn für das eine $\varrho = 1$ wird.
112 | 2) Wenn beide *zusammenfallen* und das Dreieck gleichseitig ist: *Grenzfall* Ib.
3) Wenn nur das *eine* brauchbar ist: *Fall* II und bei $\varrho = 1$ event. IIa.

§3. Discussion des allgemeinen Falles

Wir haben den Specialfall $\epsilon_1 = \epsilon_2 = \epsilon_3$ schon deswegen zuerst ausführlicher behandelt, weil er in einfachster Form schon alles Wesentliche der möglichen Bewegungsformen des allgemeinen Falles aufweist, sodaß wir uns bei diesem jetzt sehr viel kürzer fassen können, als sonst erforderlich gewesen wäre.

Nach (5) S. 98 sind die Diffgln. der Relativbewegung

$$\left.\begin{aligned} \frac{dp_1}{dt} &= \frac{\epsilon_1}{2} R \left(\frac{1}{p_2} - \frac{1}{p_3}\right) \\ \frac{dp_2}{dt} &= \frac{\epsilon_2}{2} R \left(\frac{1}{p_3} - \frac{1}{p_1}\right) \\ \frac{dp_3}{dt} &= \frac{\epsilon_3}{2} R \left(\frac{1}{p_1} - \frac{1}{p_2}\right) \end{aligned}\right\} \tag{1}$$

und ihre beiden Integrale

$$\left.\begin{aligned} \frac{p_1}{\epsilon_1} + \frac{p_2}{\epsilon_2} + \frac{p_3}{\epsilon_3} &= \kappa_1 p_1 + \kappa_2 p_2 + \kappa_3 p_3 = c \\ \frac{\lg p_1}{\epsilon_1} + \frac{\lg p_2}{\epsilon_2} + \frac{\lg p_3}{\epsilon_3} &= \lg(p_1^{\kappa_1} p_2^{\kappa_2} p_3^{\kappa_3}) = \lg c'\,. \end{aligned}\right\} \tag{2}$$

equation (5a) $D(w) = 0$, and thus obtains

$$p_1 + 2p_2 = m\ , \quad p_1 p_2^2 = n\,,$$

and hence the cubic equation for $p_2 = p_3$

$$p_2^2(m - 2p_2) = n$$

or for p_1:

$$p_1(m - p_1)^2 = 4n\,.$$

But one still needs to establish whether the radius ϱ of the two isosceles triangles thus determined does not become > 1, and, if so, to eliminate them. One then finds that

1) If the isosceles triangles are *both* viable, then we have *case* I (p. 106) and possibly *borderline case* Ia (p. 109) if for one of them $\varrho = 1$.
2) If both of them *coincide* and the triangle is equilateral: *borderline case* Ib.
3) If only *one* of them is viable: *case* II and for $\varrho = 1$ possibly IIa.

§3. Discussion of the general case

One reason why we considered the special case $\epsilon_1 = \epsilon_2 = \epsilon_3$ more closely first was that it exhibits in the simplest form all that is essential to the possible forms of motion in the general case, so that we are now in a position to give an account of the latter that is much shorter than would have been necessary otherwise.

By (5) p. 98, the differential equations of the relative motions are

$$\left.\begin{aligned} \frac{dp_1}{dt} &= \frac{\epsilon_1}{2} R \left(\frac{1}{p_2} - \frac{1}{p_3}\right) \\ \frac{dp_2}{dt} &= \frac{\epsilon_2}{2} R \left(\frac{1}{p_3} - \frac{1}{p_1}\right) \\ \frac{dp_3}{dt} &= \frac{\epsilon_3}{2} R \left(\frac{1}{p_1} - \frac{1}{p_2}\right) \end{aligned}\right\} \tag{1}$$

and their two integrals are

$$\left.\begin{aligned} &\frac{p_1}{\epsilon_1} + \frac{p_2}{\epsilon_2} + \frac{p_3}{\epsilon_3} = \kappa_1 p_1 + \kappa_2 p_2 + \kappa_3 p_3 = c \\ &\frac{\lg p_1}{\epsilon_1} + \frac{\lg p_2}{\epsilon_2} + \frac{\lg p_3}{\epsilon_3} = \lg(p_1^{\kappa_1} p_2^{\kappa_2} p_3^{\kappa_3}) = \lg c'\,. \end{aligned}\right\} \tag{2}$$

112a | Auch hier ist wie im Specialfall für $\frac{dp_1}{dt} = 0$ entweder $p_2 = p_3$, d. h. das Dreieck wird gleichschenklig mit P_1 als Spitze, oder $R = 0$, d. h. es rückt in den grössten Kreise $\varrho = 1$.

Im ersteren Falle $p_2 = p_3$ hat auch wieder nur die eine Seite r_1 ein Maximum oder Minimum, die beiden anderen nehmen weiter zu oder ab, aber nicht mehr wie dort (S. 107) mit $\frac{dp_2}{dt} + \frac{dp_3}{dt} = 0$ sondern mit $\kappa_2 \frac{dp_2}{dt} + \kappa_3 \frac{dp_3}{dt} = 0$, das Dreieck kehrt also nicht wieder unmittelbar in seine früheren Formen zurück. Nur in dem Falle, wo das Dreieck gleichseitig ist, verschwinden alle drei Ableitungen $\frac{dp_1}{dt}$, $\frac{dp_2}{dt}$, $\frac{dp_3}{dt}$ gleichzeitig und zu gleich auch alle höheren Differentialquotienten, die sich anfänglich sämtlich durch die ersten und die 0-ten (p_1, p_2, p_3 selbst) und durch R rational ausdrücken lassen, und es wird dauernd $p_1 = p_2 = p_3 = \text{const}$.

Ist das Dreieck einmal gleichseitig, so bleibt es starr für alle Zeiten.

Im übrigen kann es für gegebene Werte c, c' niemals mehr als zwei gleichschenklige Dreiecke mit P_1 als Spitze geben. Denn für $p_2 = p_3$ hätte man nach (2)

$$\kappa_1 p_1 = (\kappa_2 + \kappa_3) p_2 = c\ , \quad p_1^{\kappa_1} p_2^{\kappa_2 + \kappa_3} = c',$$

also

$$f(p_1) = p_1^{\kappa_1} (c - \kappa_1 p_1)^{(\kappa_2 + \kappa_3)} = c' (\kappa_2 + \kappa_3)^{(\kappa_2 + \kappa_3)}, \tag{3}$$

wo wir der Kürze wegen immer $\kappa_2 + \kappa_3 > 0$ annehmen wollen, was sich event.
113 durch Umkeh- | rung der Zeitrichtung immer erreichen ließe. Hier kommen dann nur solche Wurzeln $p_1 > 0$ in Betracht, für welche $c - \kappa_1 p_1 = (\kappa_2 + \kappa_3) p_2 > 0$ ist. Nun ist aber

$$\frac{f'(p_1)}{f(p_1)} = \frac{\kappa_1}{p_1} - \frac{\kappa_1 (\kappa_2 + \kappa_3)}{c - \kappa_1 p_1}$$

und verschwindet für $0 < p_1 < \frac{c}{k}$ nur einmal nämlich für

$$p_1 = \frac{c}{\kappa_1 + \kappa_2 + \kappa_3}\,.$$

Daher kann auch die Gleichung (3) zwischen denselben Grenzen nicht mehr als zwei Wurzeln haben, q. e. d.

Zu den verschiedenen Spitzen P_1, P_2, P_3 gehören aber jedesmal *andere* Paare von gleichschenkligen Dreiecken, im Ganzen könnte es also im Maximum schon 6 geben.

Im zweiten Falle verschwindet $\frac{dp_1}{dt}$ und zugleich auch $\frac{dp_2}{dt}$, $\frac{dp_3}{dt}$ für $R = 0$ oder $\varrho = 1$, d. h. im grössten Kreise. Hierbei muß $R = \sqrt{-p_1^2 - p_2^2 \cdots}$

Here, as in the special case for $\frac{dp_1}{dt} = 0$, either $p_2 = p_3$, i. e., the triangle becomes isosceles with P_1 as apex, or $R = 0$, i. e., it moves into the great circle $\varrho = 1$.

In the first case $p_2 = p_3$, once again, only the side r_1 has a maximum or minimum, whereas the other two continue to increase or decrease, but not as in the previous case (p. 107) with $\frac{dp_2}{dt} + \frac{dp_3}{dt} = 0$ but with $\kappa_2 \frac{dp_2}{dt} + \kappa_3 \frac{dp_3}{dt} = 0$, and hence the triangle does not immediately return to its previous forms. Only in the case where the triangle is equilateral do all three derivatives $\frac{dp_1}{dt}, \frac{dp_2}{dt}, \frac{dp_3}{dt}$ vanish simultaneously, and, at the same time, also all higher derivatives, all of which can initially be expressed rationally by the former and the 0-th (p_1, p_2, p_3 themselves) and by R, and we have always $p_1 = p_2 = p_3 = \text{const}$.

Once the triangle is equilateral, it remains rigid for all times.

Incidentally, for given values c, c', there can never be more than two isosceles triangles with P_1 as apex. For when $p_2 = p_3$ one would have, by (2),

$$\kappa_1 p_1 = (\kappa_2 + \kappa_3) p_2 = c \,, \quad p_1^{\kappa_1} p_2^{\kappa_2 + \kappa_3} = c',$$

and hence

$$f(p_1) = p_1^{\kappa_1} (c - \kappa_1 p_1)^{(\kappa_2 + \kappa_3)} = c' (\kappa_2 + \kappa_3)^{(\kappa_2 + \kappa_3)}, \tag{3}$$

where we shall always assume $\kappa_2 + \kappa_3 > 0$ for short, which could possibly always be achieved by reversal of the direction of time. Here, only those roots $p_1 > 0$ are relevant for which $c - \kappa_1 p_1 = (\kappa_2 + \kappa_3) p_2 > 0$. But now

$$\frac{f'(p_1)}{f(p_1)} = \frac{\kappa_1}{p_1} - \frac{\kappa_1 (\kappa_2 + \kappa_3)}{c - \kappa_1 p_1}$$

which vanishes only once when $0 < p_1 < \frac{c}{k}$, namely when

$$p_1 = \frac{c}{\kappa_1 + \kappa_2 + \kappa_3} \,.$$

Therefore, equation (3) cannot have more than two roots between the same limits, q. e. d.

But to the different apices P_1, P_2, P_3 there always correspond *other* pairs of isosceles triangles, so that, on the whole, there could already be 6 of them in the maximum.

In the second case $\frac{dp_1}{dt}$ vanishes, and, at the same time, also $\frac{dp_2}{dt}, \frac{dp_3}{dt}$ when $R = 0$ or $\varrho = 1$, i. e., in the great circle. Here, $R = \sqrt{-p_1^2 - p_2^2 \cdots}$

(cf. (3) p. 97) sein Vorzeichen wechseln, falls nicht etwa gleichzeitig

$$\frac{dR}{dt} = \frac{\epsilon_1}{2}\left(\frac{1}{p_2} - \frac{1}{p_3}\right)(p_2 + p_3 - p_1) + \frac{\epsilon_2}{2}\left(\frac{1}{p_3} - \frac{1}{p_1}\right)(p_3 + p_1 - p_2) + \quad (4)$$
$$+ \frac{\epsilon_3}{2}\left(\frac{1}{p_1} - \frac{1}{p_2}\right)(p_1 + p_2 - p_3) = 0$$

ist, in welchem Falle auch alle höheren Ableitungen nach t verschwinden und das Dreieck wieder starr bleibt. Von diesem Falle abgesehen wird also das
114 Dreieck wieder genau in seine früheren Ge- | stalten zurückkehren, nur mit entgegengesetztem R und in umgekehrter Reihenfolge. Geschieht dies aber noch einmal in derselben Bewegung, so muß das Dreieck *periodisch* zwischen diesen beiden Grenzformen im größten Kreise hin und her schwanken.

Jetzt läßt sich bereits allgemein zeigen:

Satz. *Die Relativbewegung von 3 beliebigen Strudeln ist entweder selbst eine periodische oder der asymptotische Grenzfall einer periodischen Bewegung, und im letzteren Falle nähert sich das Dreieck asymptotisch einer relativen Gleichgewichtsfigur.*

Aus den Gl. (2) folgt nämlich, daß die positiven Größen p_1, p_2, p_3 gewisse obere Grenzen nicht überschreiten, sie können aber nur eine endliche Anzahl von Maximis oder Minimis haben, nämlich höchstens in 6 gleichschenkligen Dreiecken und in einem größten Kreise, wenn die Bewegung aperiodisch sein soll. Dann müssen sie sich aber ihren oberen Grenzen asymptotisch nähern und diese entsprechen einer Gleichgewichtsfigur, d. h. einem Dreieck mit $R = 0$ (im größten Kreis) und gleichzeitig $\frac{dR}{dt} = 0$ (cf. (4) S. 113).

115 | Von diesen aperiodischen Grenzfällen abgesehen sind also nur noch folgende 3 Hauptfälle möglich.

Fall I Das Dreieck geht niemals in den größten Kreis, dann müssen alle 3 Seiten ihre Maxima sowohl als ihre Minima in gleichschenkligen Dreiecken haben, d. h. *Das Dreieck durchläuft cyclisch alle 6 möglichen gleichschenkligen Formen*, und zwar gerade so wie nach S. 108 für den Fall gleicher Strudel, in welchem nur immer je 3 dieser Typen congruent waren.

Fall II Das Dreieck wird niemals gleichschenklig, muß also seine Maxima und Minima sämtlich in 2 größten Kreisen haben, also: *Das Dreieck schwankt periodisch zwischen 2 Grenzformen im grössten Kreise.*

Fall III Das Dreieck kann sowohl gleichschenklig werden als auch in den größten Kreis eintreten, und die Seiten haben ihre Maxima und Minima während einer Periode teilweise in den Grenzformen der einen und teilweise in denen der anderen Art. Auf die sich hierbei ergebende compliziertere Fallunterscheidung wollen wir jetzt nicht näher eingehen.

(see (3) p. 97) must change its sign, provided that not also, say,

$$\frac{dR}{dt} = \frac{\epsilon_1}{2}\left(\frac{1}{p_2} - \frac{1}{p_3}\right)(p_2 + p_3 - p_1) + \frac{\epsilon_2}{2}\left(\frac{1}{p_3} - \frac{1}{p_1}\right)(p_3 + p_1 - p_2) + \quad (4)$$
$$+ \frac{\epsilon_3}{2}\left(\frac{1}{p_1} - \frac{1}{p_2}\right)(p_1 + p_2 - p_3) = 0 \ ,$$

in which case also all higher derivatives with respect to t vanish, and the triangle again remains rigid. Hence, apart from this case, the triangle will regain exactly its previous shapes, only with opposite R and in reverse order. But if this happens one more time in the same motion, then the triangle must *periodically* oscillate between these two limit forms in the great circle.

It is now already possible to show generally:

Theorem. *The relative motion of any 3 whirls is either itself a periodic motion or the asymptotic limit case of one, and, in the latter case, the triangle asymptotically approaches a relative equilibrium figure.*

For from equation (2) it follows that the positive magnitudes p_1, p_2, p_3 do not exceed certain upper limits. But they can only have a finite number of maxima or minima, namely in at most 6 isosceles triangles and in one great circle, if the motion is to be aperiodic. But then they must asymptotically approach their upper limits, which correspond to an equilibrium figure, i. e., a triangle with $R = 0$ (in the great circle) and, at the same time, $\dfrac{dR}{dt} = 0$ (see (4) S. 113).

Therefore, apart from these aperiodic limit cases only the 3 following main cases are still possible.

Case I The triangle never enters into the great circle. Then all 3 sides must have both their maxima and their minima in isosceles triangles, i. e., *the triangle passes cyclically through all 6 possible isosceles forms*, namely in the same way as for the case of equal whirls in accordance with p. 108, in which always only 3 of these types each were congruent.

Case II The triangle never becomes isosceles, and hence must have all of its maxima and minima in 2 great circles, thus: *The triangle periodically oscillates between 2 borderline forms in the great circle.*

Case III The triangle can both become isosceles and enter into the great circle, and the sides have their maxima and minima during one period partly in the borderline forms of the one kind and partly in those of the other kind. We shall not address here the more involved distinctions between cases that arise in this connection.

116 | **Schlußbemerkungen über die Bewegung des Dreieckes**

Nachdem auf Grund der bisherigen Betrachtungen die Relativbewegung, d. h. die Veränderung der Dreiecksgestalt jetzt als bekannt angesehen werden kann, handelt es sich nur noch um die absolute Bewegung eben dieses bekannten veränderlichen Dreiecks auf der Kugel. Bezüglich dieser absoluten Bewegung ist aber meine Untersuchung bisher noch nicht zum Abschlusse gekommen; ich beschränke mich daher darauf, einige der gefundenen Resultate hier ohne Beweis kurz anzugeben.

1) Nach Ablauf einer Periode der Relativbewegung ist die ganze Configuration wieder dieselbe, nur hat sich dabei die ganze Figur um den durch den Schwerpunkt S gehenden Durchmesser (die „Centralaxe“) um einen bestimmten Winkel β gedreht, und dasselbe Spiel wiederholt sich von neuem.

117 | 2) Die Ebene des Strudeldreieckes dreht sich in jedem Augenblick um seine jeweilige „Hauptaxe“, d. h. um die Verbindungsgerade ST seines Schwerpunktes S und seines „Hauptpunktes“ T (cf. C. III §6, S. 85) mit einer Winkelgeschwindigkeit λ, welche direct proportional ist der Summe der drei Stru-

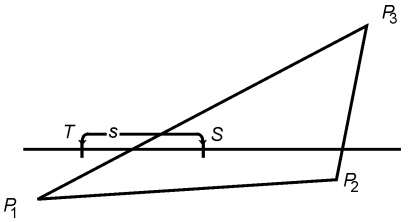


Fig. 10.

delmomente, der Summe der Quadrate ihrer gegenseitigen Abstände und der „Hauptlänge“ s (d. h. der Länge ST) und umgekehrt proportional dem Quadrate des Flächeninhaltes $\frac{1}{2}\Delta$ des Strudeldreieckes $P_1P_2P_3$:

$$\lambda = \frac{1}{2}(\epsilon_1 + \epsilon_2 + \epsilon_3)(p_1 + p_2 + p_3)\frac{s}{\Delta^2}\,.$$

Die Dreiecksebene umhüllt also einen Kegel mit dem (festen) Schwerpunkt S als Spitze und der jeweiligen „Hauptaxe“ ST als erzeugender Geraden.

3) Ein System von drei Strudeln mit beliebigen Momenten bleibt starr und rotiert gleichförmig um seine „Centralaxe“ CS,

1. wenn das Strudeldreieck $P_1P_2P_3$ gleichseitig ist,
2. wenn sein „Hauptpunkt“ T auf der „Centralaxe“ CS liegt (seine Ebene also im grössten Kreise),

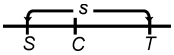


Fig. 11.

3. wenn der Schwerpunkt S in den Kugelmittelpunkt C fällt (Specialfall von 2).

Concluding remarks on the motion of the triangle

Since the previous considerations have made us familiar with the relative motion, i. e., the change in the form of the triangle, it only remains to explore the absolute motion of this now familiar variable triangle on the sphere. As for this absolute motion, however, my investigation has not yet been finalized; therefore, I confine myself to indicating here briefly without proof some of the results found.

1) Having completed one period of the relative motion, the entire configuration is the same again, except that the whole figure is rotated around the diameter (the "central axis") passing through the center of gravity S by a certain angle β, and the same game starts afresh.

2) At every instant the plane of the whirl triangle turns around its respective "main axis", i. e., around the straight line ST joining its center of gravity S and its "main point" T (see ch. III §6, p. 85) with an angular velocity λ that is directly proportional to the sum of the three whirl momenta, the sum

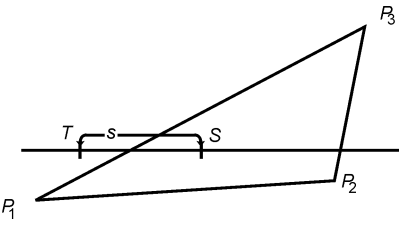


Fig. 10.

of the squares of their mutual distances and of the "main length" s (i. e., the length ST) and inversely proportional to the square of the area $\frac{1}{2}\Delta$ of the whirl triangle $P_1P_2P_3$:

$$\lambda = \frac{1}{2}(\epsilon_1 + \epsilon_2 + \epsilon_3)(p_1 + p_2 + p_3)\frac{s}{\Delta^2}\,.$$

Hence, the triangle plane envelops a cone with the (fixed) center of gravity S as apex and the respective "main axis" ST as generating straight line.

3) A system of three whirls with arbitrary momenta remains rigid and uniformly rotates around its "central axis" CS,

1. when the whirl triangle $P_1P_2P_3$ is equilateral,
2. when its "main point" T lies on the "central axis" CS (and hence its plane in the great circle),

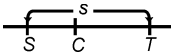


Fig. 11.

3. when the center of gravity S falls into the center of the sphere C (special case of 2).

§5. Die absolute Bewegung

s1902c

Wir nehmen jetzt die „Relativbewegung“ des Strudelsystemes als bekannt an, d. h. wir betrachten die drei Seiten des an den drei Strudelpunkten ebenen Dreiecks als gegebene Funktionen der Zeit und fragen nunmehr nach der absoluten Bewegung, nach der Verschiebung des veränderlichen Dreiecks in Bezug auf ein mit der Kugel fest verbundenes Coordinatensystem.

Zur Lösung dieses Problemes bedienen wir uns vor allem des Satzes von der Erhaltung des Schwerpunktes (III §4).

Satz I. *Der wahre Schwerpunkt eines Strudelsystems ist während der ganzen Bewegung ein fester Punkt im Raume.*

Dabei sind natürlich die Momente der drei Strudel oder ihnen proportionale Größen als Massen zu betrachten.

Da der Schwerpunkt S jedenfalls in der Ebene der drei Strudelpunkte liegen muß, so geht auch diese Ebene bei der ganzen Bewegung beständig durch den festen Punkt S.

120 | Um also die Bewegung der Dreiecksebene zu bestimmen, genügt es, die ihrer Normalen zu kennen. Wir wollen diejenige Seite der Ebene und die entsprechende Normale als *positiv* bezeichnen, welche *oben* liegen muß, wenn wir beim Durchlaufen des Dreiecksumfanges in ein für alle Mal festgesetzter Reihenfolge $A_1A_2A_3$ das *Innere* des ebenen Dreiecks zur *Linken* haben wollen. In der Richtung dieser positiven Normalen ziehen wir einen Kugelradius und nennen seinen Endpunkt N auf der Kugel den „Mittelpunkt“, genauer den „sphärischen Mittelpunkt“ unseres Dreieckes (denn es ist einer der Mittelpunkte des umschriebenen Kugelkreises), während sein Mittelpunkt in der Dreiecksebene mit M bezeichnet wird, und durch die Bewegung dieses Mittelpunktes, welche eine vollständig continuierliche sein wird, ist auch die der Ebene *eindeutig* bestimmt. Hätten wir dagegen einfach den Endpunkt des vom Mittelpunkte auf die Ebene gefällten Lotes gewählt, so würde dieser von dem einen Ende des Durchmessers zum anderen plötzlich umspringen müssen, sobald die Dreiecksebene den Kugelmittelpunkt passiert. Durch die Bewegung der Ebene ist dann aber auch die des Strudeldreieckes selbst bestimmt; denn dieses Dreieck, dessen ursprüngliche Gestalt wir vermöge der Relativbewegung als bekannt ansehen, muß in seinem umschriebenen Kreise immer so liegen, daß der Schwerpunkt in den Punkt S fällt. Eine Unbestimmt-

§5. The absolute motion*

s1902c

⟦The introductory note just before *1902a* also addresses *s1902c*.⟧

We now assume as known the "relative motion" of the whirl system, i. e., we consider the three edges of the triangle planar at the three whirl points as given functions of time, and address the question about the absolute motion, about the displacement of the variable triangle with respect to a coordinate system that is rigidly associated with the sphere.

In order to solve the problem we primarily use the theorem about the conservation of the center of gravity (III §4).

Theorem I. *During the entire motion the true center of gravity of a whirl system is a fixed point in space.*

Here, it is of course necessary to view the momenta of the three whirls, or magnitudes proportional to them, as masses.

Since the center of gravity S must lie in the plane of the three whirl points anyway, this plane, too, always passes through the fixed point S during the entire motion.

Thus, in order to determine the motion of the triangle plane, it is sufficient to know that of its normal. We shall denote by *positive* that side of the plane and the corresponding normal that must lie *on top* if we are to have the *interior* of the planar triangle to our *left* as we trace the circumference of the triangle in a sequence $A_1A_2A_3$ established once and for all. We draw a spherical radius in the direction of this positive normal and denote its end point N on the sphere the "center point", or, to be more precise, the "spherical center point", of our triangle (for it is one of the center points of the circumscribed spherical circle), while its center point in the triangle plane is denoted by M. The motion of this center point, which will be completely continuous, also *uniquely* determines that of the plane. If, by contrast, we had simply picked the end point of the perpendicular by the center point to the plane, then it would have to shift suddenly from one end of the diameter to the other once the triangle plane passes through the center point of the sphere. But the motion of the plane also determines that of the whirl triangle itself; for this triangle, whose original shape we assume as known by virtue of the relative motion, must always be situated in its circumscribed circle so that the center of gravity lies in the point S. Only in the exceptional case

* ⟦*s1902c* is an earlier draft than *s1902b*. When revising the earlier draft of *s1902b*, Zermelo obviously changed the division into sections. In particular, Chapter IV of *s1902b* now ends with "§3", whereas the earlier draft may have ended with "§4". Hence, the "§5" in the title of *s1902c*.⟧

heit entsteht nur in dem Ausnahmefalle, wo der Schwerpunkt beständig in den Mittelpunkt des Kreises fällt.

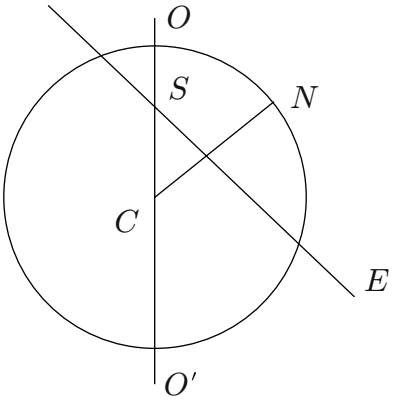


Fig. 1.

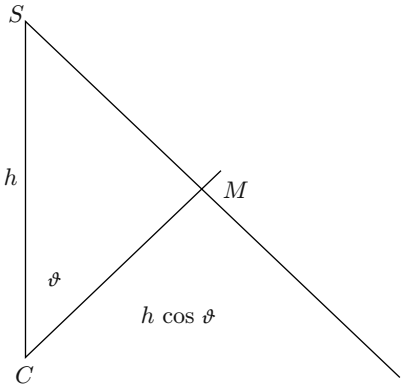


Fig. 2.

121 | Wir wählen nun ein festes Polarcoordinatensystem ϑ, ω, dessen Hauptachse $O'O$ durch den (festen) Schwerpunkt S geht, (wie schon ⟦gap⟧) und als Centralachse bezeichnet werden möge und beziehen den Punkt M auf dieses Coordinatensystem. Dann wird $\lambda = h\cos\vartheta$ dem Betrage nach gleich dem Abstande der Ebene E vom Kugelmittelpunkt und positiv oder negativ je nachdem N auf der „oberen“ oder „unteren“ d. h. der zu $\vartheta < \frac{\pi}{2}$ oder $\vartheta > \frac{\pi}{2}$ gehörenden Halbkugel liegt. Der Radius ρ des umschriebenen Kreises wird dann

$$\rho = \sqrt{1-\lambda^2} = \sqrt{1-h^2\cos^2\vartheta}\ ,$$

where the center of gravity always lies in the center point of the circle does an indeterminacy arise.

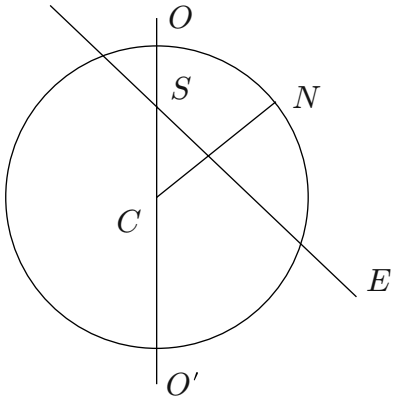


Fig. 1.

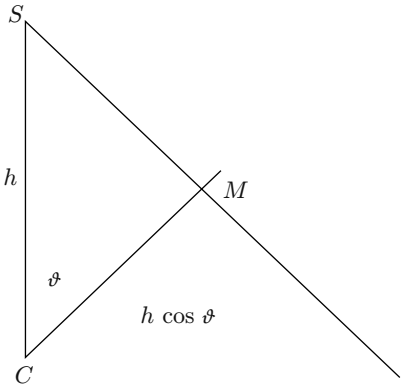


Fig. 2.

We now choose a fixed polar coordinate system ϑ, ω, whose principal axis $O'O$ passes through the (fixed) center of gravity S, (as already [[gap]]), and call it the central axis, and relate the point M to this coordinate system. Then the absolute value of $\lambda = h\cos\vartheta$ is equal to the distance of the plane E from the center point of the sphere, and it is positive or negative depending on whether N lies on the "upper" or "lower" hemisphere, i. e., in the hemisphere belonging to $\vartheta < \frac{\pi}{2}$ or to $\vartheta > \frac{\pi}{2}$. The radius ρ of the circumscribed circle then becomes

$$\rho = \sqrt{1-\lambda^2} = \sqrt{1-h^2\cos^2\vartheta}\,,$$

ist also ebenfalls eine Funktion von ϑ allein. Umgekehrt ist auch

$$\begin{aligned}\lambda = h\cos\vartheta &= \pm\sqrt{1-\rho^2} \\ &= \pm\sqrt{1-\frac{2pqr}{\Delta^2}} = \frac{R}{\Delta}\end{aligned} \qquad (1)^{1}$$

auch dem Vorzeichen nach, wenn Δ ([[unreadable word]] doppelte Inhalt des Dreieckes) immer *positiv* gerechnet wird. Denn unsere Zeichenregel für λ stimmt mit der in §2 für R gegebenen vollständig überein.

Nun sind aber R und Δ, wie wir in dem vorigen § gesehen haben, im allgemeinen periodische Funktionen der Zeit, also sind auch λ und ϑ periodische Funktionen. Dagegen braucht das Azimuth ω *nicht* periodisch zu sein, daher wird dem Punkte $N(\vartheta,\omega)$ nach Ablauf einer Periode ein anderer Punkt $N_1(\vartheta,\omega_1)$ auf demselben Parallelkreise entsprechen, wo die Differenz $\iota = \omega_1 - \omega$ im Allgemeinen von Null verschieden ist.

122 | Doch ist jetzt, abgesehen von dieser Drehung ι um die Centralaxe die ganze Configuration jetzt wieder genau dieselbe wie im Anfang, und das ganze Spiel wird sich periodisch wiederholen. Wir haben also den Satz:

Satz II. *Nach jeder Periode der Relativbewegung hat sich die ganze Figur lediglich um einen constanten Winkel um die Centralaxe gedreht.*

Daher besteht auch die Bahn des Mittelpunktes N aus lauter congruenten Stücken, welche im allgemeinen eine ganze Kugelzone überall dicht umspinnen (ganz wie die Curven der Kreiselbewegung oder die geodätischen Linien eines Rotationsellipsoides). Tritt das Dreieck einmal in den größten Kreis, so

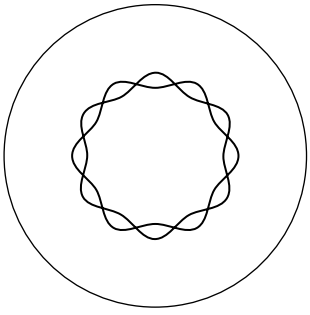


Fig. 3.

Fig. 4
Fig. 5 kann dieser, weil seine Ebene jedenfalls durch S geht, nur ein Meridian O, O' sein, und der sphärische Mittelpunkt N liegt auf dem Aequator. Die Bahn-

[1] [[Remark on the margin: "(cf. (9) §3). $= \pm\sqrt{\frac{4\omega-m^2-2n}{4\omega-m^2}} = \sqrt{1-\frac{2n}{4\omega-m^2}}$ für den Fall gleicher Strudel".]]

and hence is also a function of ϑ only. Conversely, also

$$\begin{aligned}\lambda &= h\cos\vartheta = \pm\sqrt{1-\rho^2}\\ &= \pm\sqrt{1-\frac{2pqr}{\Delta^2}} = \frac{R}{\Delta}\end{aligned} \qquad (1)^1$$

also with respect to the sign, if Δ ([[unreadable word]] twice the area of the triangle) is always reckoned *positively*. For our rule of signs for λ completely corresponds to the one specified in §2 for R.

But now R and Δ are, as we have seen in the previous §, in general periodic functions of time, and hence λ and ϑ, too, are periodic functions. By contrast, the azimuth ω does *not* need to be periodic, and hence to the point $N(\vartheta,\omega)$ there will correspond another point $N_1(\vartheta,\omega_1)$ on the same parallel circle after one period, where the difference $\iota = \omega_1 - \omega$ is in general different from zero.

But now, except for this rotation ι around the central axis, the entire configuration is precisely as it was at the beginning, and the entire game will repeat itself periodically. We thus have the theorem:

Theorem II. *After each period of the relative motion the entire figure only has turned by a constant angle around the central axis.*

It is for this reason that the path of the center point N only consists of congruent pieces that, in general, are spun around an entire spherical zone everywhere densely (just like the curves of the gyroscopic motion or the geodesic lines of a rotational ellipsoid). Once the triangle moves into the

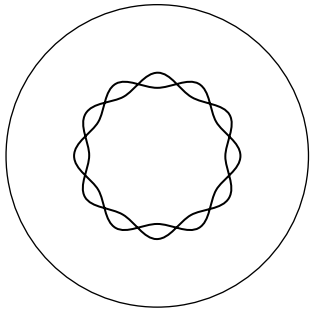


Fig. 3.

Fig. 4
Fig. 5

great circle, then the latter can only be a meridian O, O' since its plane passes through S anyhow, and the spherical center point N lies on the equator. The

[1] [[Remark on the margin: "(see (9) §3). $= \pm\sqrt{\frac{4\omega-m^2-2n}{4\omega-m^2}} = \sqrt{1-\frac{2n}{4\omega-m^2}}$ for the case of equal whirls".]]

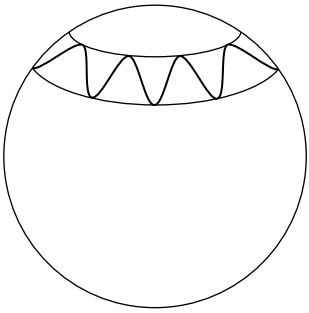


Fig. 4.

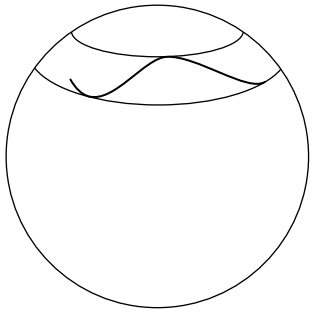


Fig. 5.

curve von N bleibt also ganz auf der einen (der „oberen“ oder der „unteren“) Halbkugel, wenn das Dreieck niemals in einen größten Kreis tritt. Hier werden die beiden Grenzkreise der Zone, d. h. die Maxima und Minima ϑ_2 und ϑ_3 bestimmt durch

$$h\cos\vartheta_2 = \sqrt{1 - \frac{2n}{4w_2 - m^2}}\,,\quad h\cos\vartheta_3 = \sqrt{1 - \frac{2n}{4w_3 - m^2}}\,,$$

und im andern Falle, wo das Dreieck zwischen $w = w_0$ und $w = w_3$ schwankt und bei $w = w_0$ mit Zeichenwechsel von $R = [\![\text{gap}]\!]$ in den größten Kreis tritt, also zwischen $+\sqrt{4(w_2 - w_0)}$ u. $-\sqrt{4(w_2 - w_0)}$ oszilliert, reicht die Zone auf
123 den *beiden* Seiten vom Äquator bis $h\cos\vartheta = \pm\sqrt{1 - \frac{2n}{4w_3 - m^2}}$. | Da aber auch im allgemeinen Falle am größten Kreise ($R = 0$) die [[unreadable word; perhaps “ganze”]] Bewegung nur mit entgegengesetztem Vorzeichen von R wieder umkehrt, so folgt:

Satz III. *Jede Bahnkurve des sphärischen Mittelpunktes, welche den Aequator überhaupt überschreitet, besteht aus lauter symmetrischen Stücken abwechselnd oberhalb und unterhalb des Aequators.*

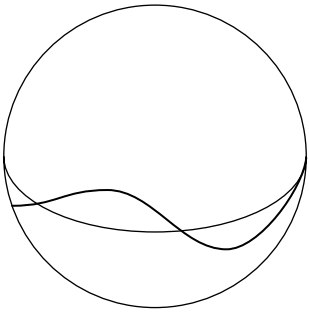


Fig. 6.

Fig. 7

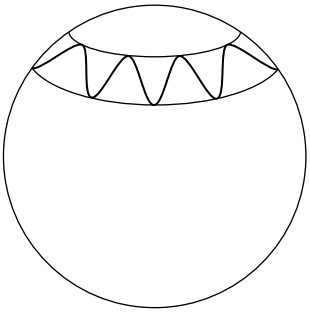


Fig. 4.

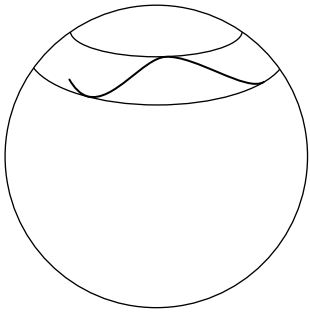


Fig. 5.

path of N thus remains entirely on the ("upper" or "lower") hemisphere if the triangle never moves into a great circle. Here, the two limit circles of the zone, i. e., the maxima and minima ϑ_2 and ϑ_3, are determined by

$$h\cos\vartheta_2 = \sqrt{1 - \frac{2n}{4w_2 - m^2}}\,,\ h\cos\vartheta_3 = \sqrt{1 - \frac{2n}{4w_3 - m^2}}\,,$$

and in the other case where the triangle oscillates between $w = w_0$ and $w = w_3$ and moves into the great circle with a change in sign of $R = [\![\text{gap}]\!]$ when $w = w_0$, and hence oscillates between $+\sqrt{4(w_2 - w_0)}$ a[[nd]] $-\sqrt{4(w_2 - w_0)}$, the zone reaches $h\cos\vartheta = \pm\sqrt{1 - \dfrac{2n}{4w_3 - m^2}}$ on *both* sides of the equator. But since also in the general case the [[unreadable word; perhaps "entire"]] motion at the great circle ($R = 0$) reverses direction again only with the opposite sign of R, it follows that:

Theorem III. *Any path of the spherical center point, if it crosses the equator at all, only consists of symmetric pieces lying alternately above and below the equator.*

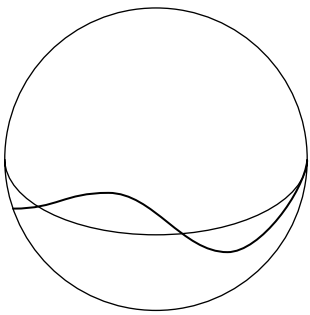


Fig. 6.

Fig. 7

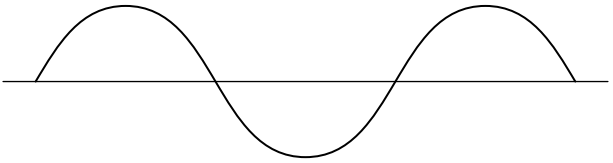


Fig. 7.

Um nun die Bewegung des Poles N, oder, was dasselbe ist, die Drehung der Ebene $A_1A_2A_3$ oder E um den Schwerpunkt S zu bestimmen, genügt es, die Geschwindigkeits*componenten* $\zeta_1, \zeta_2, \zeta_3$ der drei Strudel A_1, A_2, A_3 in der Richtung *senkrecht zur Ebene* E zu kennen. Es sei wieder M der Mittelpunkt des dem Dreieck A_1, A_2, A_3 umschriebenen Kreises mit dem Radius ρ und C der Kugelmittelpunkt. Dann steht $CM \perp E$, und es ist $\overline{CA_1} = \overline{CA_2} = \overline{CA_3} = 1$. Wie in §2 mögen die Seiten des *sphärischen* Dreieckes A_1, A_2, A_3, also die Winkel A_2CA_3, A_3CA_1, A_1CA_2 mit v_1, v_2, v_3, dagegen die Seiten des ebenen Dreieckes mit r_1, r_2, r_3 und ihre halben Quadrate mit p_1, p_2, p_3 bezeichnet
124 werden. Endlich seien κ_1, κ_2, κ_3 die Neigungs- | winkel dieser drei Ebenen gegen die Ebene E.[2]

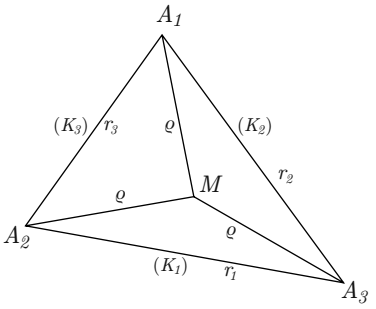


Fig. 8.

Da nun das $\Delta\, MA_2A_3 = \frac{1}{2}\Delta_1$ die senkrechte Projection des Dreieckes $CA_2A_3 = [\![\text{unreadable}]\!] = \frac{1}{2}\sin v_1$ ist und analog für die $[\![\text{unreadable letter}]\!]^n$, so wird

$$\cos\kappa_1 = \frac{\Delta_1}{\sin v_1}\,,\ \cos\kappa_2 = \frac{\Delta_2}{\sin v_2}\,,\ \cos\kappa_3 = \frac{\Delta_3}{\sin v_3}\,.$$

Gleichzeitig ist aber nach III §3 die vom Strudel A_2 dem Strudel A_3 erteilte Geschwindigkeit, falls beide das Moment π besäßen,

$$= \frac{1}{2}\cot\frac{v_1}{2}$$

[2] [[Above "Neigungswinkel dieser drei Ebenen" stands "an den Kanten A_2A_3, A_3A_1, A_1A_2".]]

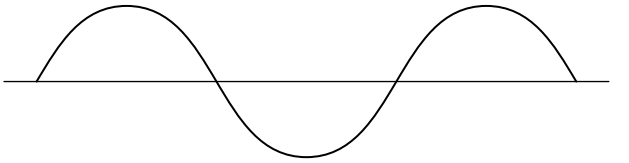


Fig. 7.

Now in order to determine the motion of the pole N, or, what amounts to the same, the rotation of the plane $A_1A_2A_3$, or E, around the center of gravity S, it is sufficient to know the velocity *components* $\zeta_1, \zeta_2, \zeta_3$ of the three whirls A_1, A_2, A_3 in the direction *perpendicular to the plane* E. Suppose that, once again, M is the center point of the circle circumscribing the triangle A_1, A_2, A_3 with the radius ρ, and that C is the center point of the sphere. Then $CM \perp E$, and $\overline{CA_1} = \overline{CA_2} = \overline{CA_3} = 1$. As in §2, let us denote the edges of the *spherical* triangle A_1, A_2, A_3, and hence the angles A_2CA_3, A_3CA_1, A_1CA_2 by v_1, v_2, v_3, but the edges of the planar triangle by r_1, r_2, r_3, and their half-squares by p_1, p_2, p_3. Finally, let κ_1, κ_2, κ_3 be the angles of inclination of these three planes with respect to the plane E.[2]

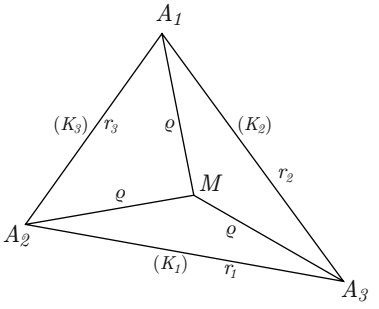


Fig. 8.

Now since the $\Delta\ MA_2A_3 = \frac{1}{2}\Delta_1$ is the perpendicular projection of the triangle $CA_2A_3 = [\![\text{unreadable}]\!] = \frac{1}{2}\sin v_1$, and analogously for the $[\![\text{unreadable letter}]\!]^n$, we obtain

$$\cos\kappa_1 = \frac{\Delta_1}{\sin v_1}\,,\ \cos\kappa_2 = \frac{\Delta_2}{\sin v_2}\,,\ \cos\kappa_3 = \frac{\Delta_3}{\sin v_3}\,.$$

But, at the same time, by III §3, the velocity given by the whirl A_2 to the whirl A_3 is, provided that both have the momentum π,

$$= \frac{1}{2}\cot\frac{v_1}{2}\,,$$

[2] [[Above "angles of inclination of these three planes" stands "at the edges A_2A_3, A_3A_1, A_1A_2".]]

und ihre Richtung steht senkrecht auf der Ebene CA_2A_3. Daher wird die Komponente k_1 dieser Geschwindigkeit senkrecht zur Ebene E

$$\begin{aligned} k_1 = \frac{1}{2}\cot\frac{v_1}{2}\,\cos\kappa_1 &= \frac{\Delta_1}{4\sin^2\frac{v_1}{2}} = \frac{\Delta_1}{r_1^2} \\ &= \frac{1}{2}\cot A_1 \;=\; \frac{p_2+p_3-p_1}{-p^2-r^2+2qr+2rp+2pq} \end{aligned}$$

und ebenso cyklisch

$$\begin{aligned} 2k_1 = \cot A_1 &= \frac{p_2+p_3-p_1}{\Delta} \\ 2k_2 = \cot A_2 &= \frac{p_3+p_1-p_2}{\Delta} \\ 2k_3 = \cot A_3 &= \frac{p_1+p_2-p_3}{\Delta}\,, \end{aligned} \tag{4}$$

wenn wie früher der doppelte Inhalt des Dreieckes $A_1A_2A_3$ mit Δ bezeichnet wird, so daß

$$\Delta \;=\; \left|\sqrt{2p_2p_3+2p_3p_1+2p_1p_2-p_1^2-p_2^2-p_3^2}\right|\,.$$

Somit werden die drei Geschwindigkeitskomponenten in der Richtung der „positiven“ Normalen des Dreieckes $A_1A_2A_3$:

125 |
$$\begin{aligned} \zeta_1 &= \epsilon_3k_2-\epsilon_2k_3 \\ \zeta_2 &= \epsilon_1k_3-\epsilon_3k_1 \\ \zeta_3 &= \epsilon_2k_1-\epsilon_21k_2\,, \end{aligned} \tag{5}$$

auch dem Vorzeichen nach, wie man sich leicht überzeugt. Aus diesen Gleichungen folgt aber

$$\begin{aligned} \epsilon_1\zeta_1+\epsilon_2\zeta_2+\epsilon_3\zeta_3 &= 0 \\ k_1\zeta_2+k_2\zeta_2+k_3\zeta_3 &= 0\,, \end{aligned} \tag{6}$$

und die erste dieser Gleichungen bedeutet, daß der Schwerpunkt der Massen $\epsilon_1,\epsilon_2,\epsilon_3$, und die zweite, daß der Schwerpunkt der Massen k_1,k_2,k_3, die man sich in den Punkten A_1,A_2,A_3 vorhanden denkt, *keine* Geschwindigkeitskomponente $\perp E$ besitzt. Der erste Schwerpunkt ist der eigentliche Schwerpunkt S, der andere aber, der Schwerpunkt der Massen k_1,k_2,k_3 oder $\cot A_1,\cot A_2,\cot A_3$ in den gleichbenannten Eckpunkten ist nichts anderes als der früher III §6 fin. definierte „Hauptpunkt“ T des Dreieckes $A_1A_2A_3$, welcher von den Strudelmomenten unabhängig und nur durch die *Gestalt* des Dreieckes bestimmt ist. Schneidet man daher die Ebene E mit ihrer neuen Lage, die sie nach einem kleinen Zeitteilchen annimmt, so muß die Grenzlage der Schnittlinie durch die beiden Punkte S und T gehen, die Ebene E dreht
126 sich also um die Gerade ST | mit einer Winkelgeschwindigkeit, welche direkt

and their direction is perpendicular to the plane CA_2A_3. Therefore, the component k_1 of this velocity becomes perpendicular to the plane E

$$\begin{aligned} k_1 = \frac{1}{2}\cot\frac{v_1}{2}\cos\kappa_1 &= \frac{\Delta_1}{4\sin^2\frac{v_1}{2}} = \frac{\Delta_1}{r_1^2} \\ = \frac{1}{2}\cot A_1 &= \frac{p_2+p_3-p_1}{-p^2-r^2+2qr+2rp+2pq} \end{aligned}$$

and also cyclical

$$\begin{aligned} 2k_1 = \cot A_1 &= \frac{p_2+p_3-p_1}{\Delta} \\ 2k_2 = \cot A_2 &= \frac{p_3+p_1-p_2}{\Delta} \\ 2k_3 = \cot A_3 &= \frac{p_1+p_2-p_3}{\Delta}\,, \end{aligned} \tag{4}$$

if, as before, Δ denotes twice the area of the triangle $A_1A_2A_3$, so that

$$\Delta = \left|\sqrt{2p_2p_3+2p_3p_1+2p_1p_2-p_1^2-p_2^2-p_3^2}\right|.$$

Thus, the three velocity components in the direction of the "positive" normal of the triangle $A_1A_2A_3$ become

$$\begin{aligned} \zeta_1 &= \epsilon_3k_2-\epsilon_2k_3 \\ \zeta_2 &= \epsilon_1k_3-\epsilon_3k_1 \\ \zeta_3 &= \epsilon_2k_1-\epsilon_21k_2\,, \end{aligned} \tag{5}$$

also with respect to the sign, as is readily seen. But from these equations it follows that

$$\begin{aligned} \epsilon_1\zeta_1+\epsilon_2\zeta_2+\epsilon_3\zeta_3 &= 0 \\ k_1\zeta_2+k_2\zeta_2+k_3\zeta_3 &= 0\,, \end{aligned} \tag{6}$$

and the first of these equations means that the center of gravity of the masses $\epsilon_1, \epsilon_2, \epsilon_3$ possesses *no* velocity component $\perp E$, as does the second equation for the center of gravity of the masses k_1, k_2, k_3, which are assumed to occur at the points A_1, A_2, A_3. The first center of gravity is the actual center of gravity S, while the other one, the center of the gravity of the masses k_1, k_2, k_3, or $\cot A_1, \cot A_2, \cot A_3$ in the homonymous vertices, is but the previously III §6 fin. defined "principal point" T of the triangle $A_1A_2A_3$, which is determined independently of the whirl momenta only by the *shape* of the triangle. Hence, if one intersects the plane E with its new position, which it assumes after a small fraction of time, then the limit position of the line of intersection must pass through the two points S and T, and hence the plane E turns around the straight line ST at an angular velocity directly proportional to

proportional ist der Summe der drei Strudelmomente, der Summe der Quadrate ihrer Abstände und der „Länge der Hauptaxe“, d. h. der Entfernung von Schwerpunkt und Hauptpunkt, und umgekehrt proportional dem Quadrate des Flächeninhaltes des von ihnen gebildeten ebenen Dreieckes.

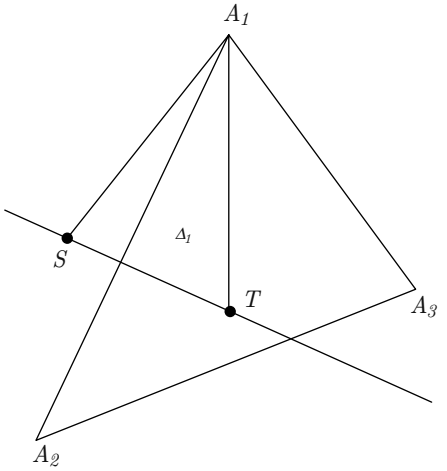


Fig. 9.

Um nun die Winkelgeschwindigkeit λ dieser Drehung zu bestimmen, berechnen wir zunächst den Inhalt $\frac{1}{2}\Delta$ des Dreieckes A_1ST und bezeichnen zu diesem Zwecke die rechtwinkligen Coordinaten der Punkte A_1, A_2, A_3 in der Ebene E der Reihe nach mit

$$x_1, y_1\,;\; x_2, y_2\,;\; x_3, y_3$$

und die der Punkte S und T mit $x_0, y_0\,,\; x_0', y_0'$, so daß

$$\epsilon x_0 = (\epsilon_1 + \epsilon_2 + \epsilon_3)x_0 = \epsilon_1 x_1 + \epsilon_2 x_2 + \epsilon_3 x_3$$
$$\epsilon y_0 = (\epsilon_1 + \epsilon_2 + \epsilon_3)y_0 = \epsilon_1 y_1 + \epsilon_2 y_2 + \epsilon_3 y_3$$

und

$$kx_0' = (k_1 + k_2 + k_3)x_0' = k_1 x_1 + k_2 x_2 + k_3 x_3$$
$$ky_0' = (k_1 + k_2 + k_3)y_0' = k_1 y_1 + k_2 y_2 + k_3 y_3$$

wo nach (4) $2k = p_1 + p_2 + p_3$. Demnach wird

$$\Delta_1 = \pm \begin{vmatrix} x_1 & y_1 & 1 \\ x_0 & y_0 & 1 \\ x_0' & y_0' & 1 \end{vmatrix} = \frac{\pm 1}{\epsilon k} \begin{vmatrix} x_1 & y_1 & 1 \\ \epsilon_1 x_1 + \epsilon_2 x_2 + \epsilon_3 x_3 & \epsilon_1 y_1 + \epsilon_2 y_2 + \epsilon_3 y_3 & 1 \\ k_1 x_1 + k_2 x_2 + k_3 x_3 & k_1 y_1 + k_2 y_2 + k_3 y_3 & 1 \end{vmatrix}$$
$$= \pm \frac{\epsilon_2 k_3 - \epsilon_3 k_2}{\epsilon k} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = \pm w_1 \frac{\Delta}{\epsilon k}$$

the sum of the three whirl momenta, the sum of the squares of their distances and the "length of the principal axis", i. e., the distance of center of gravity and principal point, and inversely proportional to the square of the area of the planar triangle formed by them.

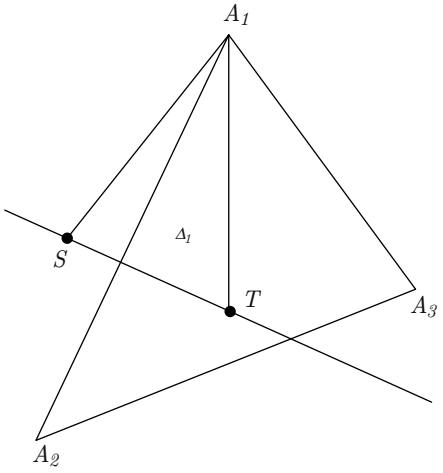


Fig. 9.

Now in order to determine the angular velocity λ of this rotation, we first calculate the area $\frac{1}{2}\Delta$ of the triangle A_1ST. To this end we denote the orthogonal coordinates of the points A_1, A_2, A_3 in the plane E successively by

$$x_1, y_1\,;\; x_2, y_2\,;\; x_3, y_3$$

and those of the points S and T by $x_0, y_0\,,\; x_0', y_0'$, so that

$$\epsilon x_0 = (\epsilon_1 + \epsilon_2 + \epsilon_3)x_0 = \epsilon_1 x_1 + \epsilon_2 x_2 + \epsilon_3 x_3$$
$$\epsilon y_0 = (\epsilon_1 + \epsilon_2 + \epsilon_3)y_0 = \epsilon_1 y_1 + \epsilon_2 y_2 + \epsilon_3 y_3$$

and

$$kx_0' = (k_1 + k_2 + k_3)x_0' = k_1 x_1 + k_2 x_2 + k_3 x_3$$
$$ky_0' = (k_1 + k_2 + k_3)y_0' = k_1 y_1 + k_2 y_2 + k_3 y_3$$

where, by (4), $2k = p_1 + p_2 + p_3$. Consequently,

$$\Delta_1 = \pm\begin{vmatrix} x_1 & y_1 & 1 \\ x_0 & y_0 & 1 \\ x_0' & y_0' & 1 \end{vmatrix} = \frac{\pm 1}{\epsilon k}\begin{vmatrix} x_1 & y_1 & 1 \\ \epsilon_1 x_1 + \epsilon_2 x_2 + \epsilon_3 x_3 & \epsilon_1 y_1 + \epsilon_2 y_2 + \epsilon_3 y_3 & 1 \\ k_1 x_1 + k_2 x_2 + k_3 x_3 & k_1 y_1 + k_2 y_2 + k_3 y_3 & 1 \end{vmatrix}$$
$$= \pm\frac{\epsilon_2 k_3 - \epsilon_3 k_2}{\epsilon k}\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = \pm w_1 \frac{\Delta}{\epsilon k}$$

wenn Δ der doppelte Inhalt von $A_1A_2A_3$ ist. Setzt man nun $\overline{ST} = s$, so wird die Länge des von A_1 auf ST gefällten Lotes $l_1 = \dfrac{\Delta_1}{s} = \pm\dfrac{w_1\Delta}{\epsilon k s}$, und daher die Drehungsgeschwindigkeit

$$\lambda = \frac{w_1}{l_1} = \frac{\epsilon k s}{\Delta} = \frac{\epsilon}{2}(p_1 + p_2 + p_3)\frac{s}{\Delta^2}. \tag{7}$$

127 | Der *Sinn* der Drehung kann nur abhängen von der Richtung ST und von dem Vorzeichen der Summe der Strudelmomente. Nimmt man aber das Dreieck im größten Kreise *gleichschenklig* an mit A_1 als Spitze und $\epsilon_1 = 1$, $\epsilon_2 = \epsilon_3 = 0$, so fällt der Schwerpunkt nach A_1, der Hauptpunkt aber in die Diagonale A_1D_1, und die Figur dreht sich in positivem Sinne um die Achse TA_1 oder TS, da nur der eine (positive Strudel) A_1 vorhanden ist.

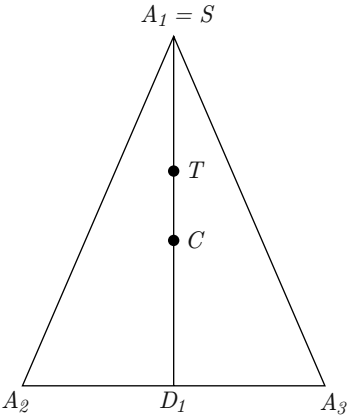


Fig. 10.

Es muss also die Richtung von T nach S als die Drehungsaxe angesehen werden, um welche die Drehung im positiven oder negativen Sinne erfolgt, je nachdem $\epsilon = \epsilon_1 + \epsilon_2 + \epsilon_3$ positiv oder negativ ist.

Bezeichnet man mit s_1, s_2, s_3 die Projectionen der Linien TA_1, TA_2, TA_3 auf die Gerade TS, so wird immer

$$\epsilon s = \epsilon_1 s_1 + \epsilon_2 s_2 + \epsilon_3 s_3,$$

also

$$\lambda = \frac{1}{2}\frac{p_1 + p_2 + p_3}{\Delta^2}(\epsilon_1 s_1 + \epsilon_2 s_2 + \epsilon_3 s_3)\,, \tag{7a}$$

und diese Formel ist auch dann anwendbar, wenn $\epsilon = \epsilon_1 + \epsilon_2 + \epsilon_3 = 0$ ist und der Schwerpunkt S ins Unendliche fällt.

Da der eine Punkt S der momentanen Drehungsachse TS ein fester Punkt ist, so können wir sagen:

if Δ is twice the area of $A_1A_2A_3$. If we now set $\overline{ST} = s$, then the length of the perpendicular dropped from A_1 to ST becomes $l_1 = \dfrac{\Delta_1}{s} = \pm\dfrac{w_1\Delta}{\epsilon k s}$, and hence the rotational velocity,

$$\lambda = \frac{w_1}{l_1} = \frac{\epsilon k s}{\Delta} = \frac{\epsilon}{2}(p_1 + p_2 + p_3)\frac{s}{\Delta^2}. \tag{7}$$

The *sense* of the rotation can only depend on the direction ST and on the sign of the sum of the whirl momenta. But if we assume the triangle in the great circle to be *isosceles* with A_1 as apex and $\epsilon_1 = 1$, $\epsilon_2 = \epsilon_3 = 0$, then the center of gravity lies in A_1, while the principal point lies on the diagonal A_1D_1, and the figure turns in positive sense around the axis TA_1 or TS, since only the one (positive whirl) A_1 exists.

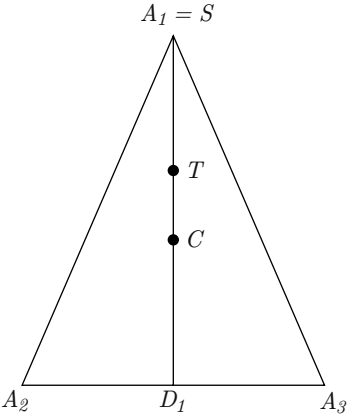


Fig. 10.

Thus, we have to assume that the direction from T to S is the axis of rotation around which the rotation occurs in positive or negative sense, depending on whether $\epsilon = \epsilon_1 + \epsilon_2 + \epsilon_3$ is positive or negative.

If s_1, s_2, s_3 denote the projections of the lines TA_1, TA_2, TA_3 onto the straight line TS, then always

$$\epsilon s = \epsilon_1 s_1 + \epsilon_2 s_2 + \epsilon_3 s_3,$$

and hence

$$\lambda = \frac{1}{2}\frac{p_1 + p_2 + p_3}{\Delta^2}(\epsilon_1 s_1 + \epsilon_2 s_2 + \epsilon_3 s_3)\,, \tag{7a}$$

and this formula applies also when $\epsilon = \epsilon_1 + \epsilon_2 + \epsilon_3 = 0$ and the center of gravity S lies at infinity.

Since the one point S of the respective axis of rotation TS is a fixed point, we may say:

128 | **Satz IV.** *Die Ebene E des Strudeldreieckes umhüllt bei ihrer Bewegung einen Kegel mit dem Schwerpunkte S als Spitze und der jeweiligen „Hauptaxe“ ST als erzeugender Geraden.*

Die Bahnkurve des ⟦unreadable word⟧ ist dann die sphärische Abbildung dieses Kegels durch parallele Normalen.

Durch unsere Drehungsgeschwindigkeit λ läßt sich auch die Drehungscomponente um die Centralaxe CS, d. h. der Wert von $\frac{d\omega}{dt}$ für den „Mittelpunkt“ N ohne Schwierigkeit berechnen. Doch ist es mir bisher noch nicht gelungen, dieser Differentialgleichung zur Bestimmung von ω eine analytisch einfache Form zu geben, welche eine genauere Discussion der entstehenden Kegelformen gestattete. Ich beschränke mich daher auf einige allgemeine Bemerkungen.

Die Drehungsgeschwindigkeit *verschwindet* nur, wenn $\epsilon s = 0$ ist, d. h.
129 wenn der Schwerpunkt und der Hauptpunkt zusammenf., | d. h. wenn die Cotangenten der Dreieckswinkel sich verhalten wie die Strudelmomente der entsprechenden Eckpunkte. Soll dies immer geschehen und somit die Dreiecksebene sich immer selbst parallel sein, so muß das Dreieck sich selbst ähnlich und da nach §2 gleichzeitig auch $\frac{p_1}{\epsilon_1} + \frac{p_2}{\epsilon_2} + \frac{p_3}{\epsilon_3}$ constant ist, sich selbst congruent bleiben. Dies ist nur in zwei Fällen möglich.

a) Die drei Strudel liegen im grössten Kreise und ihre Momente verhalten sich wie die Cotangenten der Dreieckswinkel; dann besteht absolutes Gleichgewicht (cf. III, §6).

b) Die Strudel sind gleich und bilden ein gleichseitiges Dreieck. Dann rotiert die ganze Figur starr in dem umschriebenen Parallelkreis mit constanter leicht zu berechnender Geschwindigkeit.

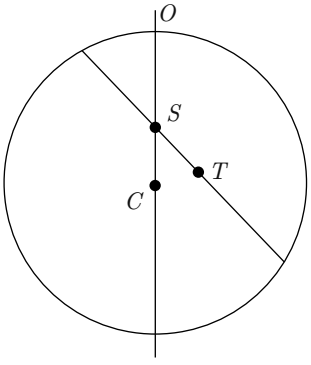


Fig. 11.

Theorem IV. *As it moves, the plane E of the whirl triangle envelops a cone with the center of gravity S as apex and the respective "principal axis" ST as generating straight line.*

The path of the ⟦unreadable word⟧ is then the spherical mapping of this cone by parallel normals.

By means of our rotational velocity λ it is also possible to compute easily the rotation component around the central axis CS, i. e., the value of $\frac{d\omega}{dt}$ for the "center point" N. But I have not yet managed to give this differential equation for the determination of ω an analytically simple form, which would allow for a more precise discussion of the emerging cone shapes. I therefore restrict myself to a few general remarks.

The rotational velocity *vanishes* only when $\epsilon s = 0$, i. e., when the center of gravity and the principal point coinc⟦ide⟧, i. e., when the cotangents of the triangle angles are proportional to the whirl momenta of the corresponding vertices. For this to be always the case, and hence for the triangle plane to be always parallel to itself, the triangle must always remain similar to itself, and since, by §2 [3] at the same time also $\frac{p_1}{\epsilon_1} + \frac{p_2}{\epsilon_2} + \frac{p_3}{\epsilon_3}$ is constant, congruent to itself. This is only possible in two cases.

a) The three whirls lie in the great circle, and their momenta are proportional to the cotangents of the triangle angles; in this case, there obtains absolute equilibrium (see III, §6).[4]

b) The whirls are equal and form an equilateral triangle. Then the entire figure rigidly rotates in the circumscribed parallel circle at a constant, easily computed velocity.

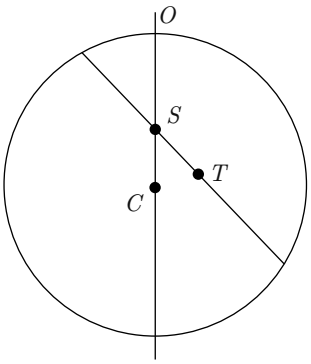


Fig. 11.

[3] ⟦Zermelo seems to refer to equations (6) in §1 of Chapter IV; these equations stand on page 98 of *s1902b*.⟧

[4] ⟦Zermelo refers to Satz (Theorem) IV on page 84 of *s1902b*.⟧

130 | 2) Liegt das Strudeldreieck $A_1A_2A_3$ einmal im größten Kreise ($R = 0$), so dreht sich abgesehen von dem eben erwähnten Gleichgewichtsfalle $S = T$ seine Ebene um die Achse ST und dabei wird der Pol N notwendig von der „oberen“ in die „untere“ Halbkugel übertreten oder umgekehrt und somit R sein Vorzeichen wechseln (cf. oben p. ⟦gap⟧), es sei denn, daß die Hauptachse *ST mit der „Zentralachse“ CO zusammenfällt.* Ist aber das letztere der Fall, liegen irgend einmal C, S, T auf einer geraden Linie, durch die dann natürlich auch die Dreiecksebene gehen muß, so dreht sich die Ebene nur um eben diesen *Durchmesser* und es muß auch $\frac{dR}{dt} = 0$ sein. Dann muß aber nach §4 p. ⟦gap⟧ das Dreieck *starr* sein und *beständig* im größten Kreise liegen, und zwar wird dann die ganze Figur einfach mit constanter Geschwindigkeit um die Centralachse $CTSO$ rotieren. Der a.a.O. angeführte Fall, wo der Schwerpunkt S in den Kugelmittelpunkt fällt, ist hiervon nur ein Unterfall, in welchem der durch den Hauptpunkt gehende Durchmesser die Rotationsachse darstellt.

131 | Auch in dem Fall des starren gleichseitigen Dreieckes bleibt die Funk. R und damit nach (1) auch ϑ für den Pol *constant*, der von der Ebene umhüllte Kegelmantel ist ein gerader Kreiskegel und das ganze System rotiert wieder gleichförmig um die Centralachse.

Wir haben also den

Satz V. *Das System von drei Strudeln mit beliebigen Momenten bleibt starr und rotiert gleichförmig um seine Centralaxe, d. h. um den durch den Schwerpunkt gehenden Durchmesser CS:*

1) *wenn das Dreieck gleichseitig ist*

2) *wenn der „Hauptpunkt“ des Dreieckes auf der Centralachse liegt* (*seine Ebene also im grössten Kreise*) *und als* ⟦*unreadable word*⟧ *fall von* 2)

3) *wenn sein Schwerpunkt in den Kugelmittelpunkt fällt.*

Die Winkelgeschwindigkeit bestimmt sich in allen diesen Fällen leicht mit Hilfe des Satzes IV. Diese Winkelgeschwindigkeit verschwindet aber und es ist *absolutes Gleichgewicht* vorhanden, wenn das Dreieck im grössten Kreis liegt, und sein Schwerpunkt mit seinem Hauptpunkt zusammenfällt.

2) [5] Once the whirl triangle $A_1A_2A_3$ lies in the great circle ($R = 0$), then, except for the equilibrium case $S = T$ just mentioned, its plane turns around the axis ST, and, in this case, the pole N will necessarily cross over from the "upper" into the "lower" hemisphere, or vice versa, and hence R will change its sign (see above p. ⟦gap⟧) [6], unless the principal axis ST *coincides with the "central axis"* CO. But if the latter is the case, if C, S, T lie on a straight line at some point in time, in which case, of course, also the triangle plane must pass through that line, then the plane only turns around precisely this *diameter*, and necessarily also $\frac{dR}{dt} = 0$. But then, by §4 p. ⟦gap⟧,[7] the triangle must be *rigid* and it must *always* lie in the great circle. In particular, the entire figure will then simply rotate around the central axis $CTSO$ at constant velocity. The case cited above, where the center of gravity S lies in the center point of the sphere, is but a subcase in which the diameter passing through the principal point is the axis of rotation.

The func⟦tion⟧ R, and hence, by (1) also ϑ, remains *constant* for the pole also in the case of the rigid, equilateral triangle. The lateral surface area of the cone enveloped by the plane is ⟦that of⟧ a straight circular cone, and the entire system again rotates uniformly around the central axis.

We thus have the

Theorem V. *The system of three whirls with arbitrary momenta remains rigid and rotates uniformly around its central axis, i. e., around the diameter* CS *passing through the center of gravity:*

1) *when the triangle is equilateral*

2) *when the "principal point" of the triangle lies on the central axis* (*and hence its plane in the great circle*), *and, as* ⟦*unreadable word*⟧ *case of* 2)

3) *when the center of gravity lies in the center point of the sphere.*

In all these cases, the angular velocity is easily determined by means of theorem IV. But this angular velocity vanishes, and there is *absolute equilibrium* when the triangle lies in the greatest circle and its center of gravity coincides with its principal point.

[5] ⟦Point "1)" is missing.⟧

[6] ⟦Zermelo refers to pages 114 and 115 of *s1902b*.⟧

[7] ⟦Zermelo refers to pages 114 and 115 of *s1902b*. These pages belong to §3 of Chapter IV of *s1902b*; see fn. ∗ at the beginning of the paper.⟧

Introductory note to 1902d

Rüdiger Thiele

In 1899, Zermelo completed his *Habilitation* in Göttingen, becoming there a *Privatdozent* with the right to give lectures. Zermelo had also lectured on the calculus of variations a number of times at the Göttingen Mathematical Society, beginning in 1897, and also, in particular, on Kneser's article "Zur Variationsrechnung" ("On the calculus of variations") (*Kneser 1898*), an important related work. Zermelo's first lectures in Göttingen on the calculus of variations in the summer semester of 1902 and his paper "Zur Theorie der kürzesten Linien" ("On the theory of shortest paths¨") (*Zermelo 1902d*) should be viewed in relationship to the Society lectures; Zermelo's next lecture on the calculus of variations took place in the winter semester of 1907.

The problem of shortest paths or, as one has said since Joseph Liouville (1844), the problem of geodesics on a surface, was identified by Zermelo as one of the classical problems in the calculus of variations, and indeed, questions relating to intuitively geometric geodesics belong among its standard problems. The variational problem leading to geodesics for a parametrically given shortest curve C_0: $(x^0(t), y^0(t))$, $t \in [t_1, t_2]$, on a surface F between the points P_1 and P_2 is as follows:

$$\begin{aligned} J(C) &= \int_C \sqrt{dx^2 + dy^2}\, dt \\ &= \int_{t_1}^{t_2} \sqrt{E\dot{x}^2 + 2F\dot{x}\dot{y} + G\dot{y}^2}\, dt\,, \end{aligned} \tag{1}$$

where E, F, and G in the second formulation of the problem are the Gaussian fundamental coefficients of the surface F.[1]

Let us begin by briefly considering how the problem of shortest paths was treated in Göttingen. In his Paris address "Mathematische Probleme" (1900), David Hilbert, in his Problem 4 ("Problem of the line as the shortest distance between two points"),[2] called "attention to a theorem that is given by many authors as the definition of a straight line, which asserts that the line is the shortest distance between two points" (*1900a*, 268). Hilbert's investigations of this minimality condition are closely related to his work on the foundations of

[1] Zermelo discusses the extension of the variational problem for shortest paths from geometric viewpoints, which is only natural. We therefore formulate the variational problem parametrically.

[2] Along with thirteen further problems, this problem was not presented in the Paris lecture, but appeared only in the printed version of the lecture (*Hilbert 1900a*).

H.-D. Ebbinghaus, A. Kanamori (Eds.), *Ernst Zermelo - Collected Works/ Gesammelte Werke II*, DOI 10.1007/978-3-540-70856-8_8, Schriften der Mathematisch-naturwissenschaftlichen Klasse der Heidelberger Akademie der Wissenschaften 23,

geometry (*1899*). As a result of this work, he developed, among other things, a geometry in which all the axioms of Hermann Minkowski's pseudogeometry are satisfied except for the parallel postulate (*1895*, 91). Hilbert remarked, "In the case of the plane and assuming the axiom of continuity, this problem leads to the question treated by Darboux (*1894*, 54) of finding all variational problems in the plane whose solutions are all the straight lines in the plane."

In his Göttingen dissertation *Über die Geometrien, in denen die Geraden die Kürzesten sind* (*On geometries in which straight lines are shortest paths*) (*Hamel 1901*), written under the supervision of Hilbert, which appeared two years later as *Hamel 1903*, Georg Hamel studied Hilbert's problem from the viewpoint of the calculus of variations. He thus used methods whose analytic assumptions are unsuited to geometry. To be sure, measuring lengths is a very elementary process, but the functional relationships among measurements pertain to arithmetic, not geometric, relationships; the length of a curve can be defined by an integral, as in (1), or even more generally. Using direct methods, Hilbert showed (*1900b*) under weakened assumptions that a given continuous curve on a surface $z = f(x, y)$ that joins two distinct points of the surface is rectifiable; f is taken to be continuous and to have continuous partial derivatives of first order.

The broadly posed fourth problem was made more precise in numerous papers, including those by Paul Funk (*1929*) and Herbert Busemann (*1943*). The desire to formulate assertions of the calculus of variations geometrically was realized, for example, in Finsler geometry, which originated in Paul Finsler's Göttingen dissertation *Über Kurven und Flächen* (*On curves and surfaces*) (*1918*) written under the supervision of Constantin Carathéodory.

Zermelo mentions three possible ways to extend the variational problem $J(C) \to$ minimum associated with (1), which can be characterized by the following key words:

(a) absolute minima,
(b) restrictions on surfaces,
(c) differential inequalities as constraints.

The last two cases appear naturally in practical questions, and in this case, Zermelo introduced the problem of road and rail construction. The construction of roads and rail lines can be limited by geographical conditions to a surface patch $F_0 \subseteq F$; moreover, the steepness or curvature of a road or track section can be constrained. Since Zermelo assumes simply connected surface patches F_0, obstacles on F_0 are excluded from consideration. The constraints (b) and (c), Zermelo observes, can also lead to the result that on the surface F itself, no solutions are possible (and from an engineering perspective, this leads to tunnels and viaducts).[3] However, in the case of constraints, it would be necessary to deal with the question whether under such conditions, a set

[3] About 40 years later, Zermelo would choose this topic as one of the chapters of a planned book *Mathematische Miniaturen* (*Mathematical miniatures*) under

of admissible comparison functions (comparison curves) is available to make possible the variation and its infinitesimal techniques. The standard methods of the calculus of variations require, in contrast to the theory of optimization, that open sets be used. In the case of domain restriction, where also parts of the boundary ∂F_0 of the surface patch F_0 may need to be considered, some special ideas are necessary, which Zermelo sketches for simply connected surface patches F_0 (see also *Bolza 1909*, 392–407, 527).

Zermelo calls curves of constant steepness *Kletterkurven* (*climbing curves*). At every point of the surface there are two simply infinite families of climbing curves, where by alternate use of these curves from both families, a zigzag line can be created with whose help one can approximate curves with larger steepness (see also *Bolza 1909*, 126f.).

A functional $J(x)$ defined on a set A of functions (or curves) can be associated with a variational problem

$$J(x) \to \text{ extremum} \qquad \text{on } A_0 \subseteq A.$$

Here the subset A_0, with which also a notion of neighborhood is associated, determines the type of solution, that is, whether one has a strong or weak, relative or absolute, and so forth, extremum (see *du Bois-Reymond 1879a*, 283; *Kneser 1900*, §17; *Osgood 1900/01*, 105). In his dissertation, Zermelo worked with very general notions of distance and in this way dealt with the question of the type of an extremum in variational problems (*Zermelo 1894*, 24). Of the notions of distance discussed by Zermelo, today only the distances in the space of continuous or continuously differentiable functions are used, which lead to strong and weak extrema.

In the present work, Zermelo looks at relative and absolute minima of geodesics, which he calls *shortest* and *by far shortest* paths, respectively. The infinitesimal variational technique first takes into account only sufficiently close comparison functions (respectively comparison curves) that lie in a given neighborhood of the extremal E (geometrically in a strip around the extremal E), leading thereby first to necessary conditions for relative extrema, in this case minima. Here it is necessary in addition to include non-analytic

the title "Straßenbau im Gebirge" ("Road construction in the mountains"); cf. *Ebbinghaus 2007*, 14.

functions. The extremals E obtained in relation to the strip, however, even if they exhibit minimality, need not lead to (relative) minima for the entire function (or along the entire curve); rather, the presence of minimality can definitely be limited to parts of the function (respectively the curve). In the latter case, the minimality on E is limited by a point P_1^*, called the conjugate to the initial point P_1 of the extremal E. (One would have respectively $P_1^* = (x(t_1^*), y(t_1^*))$ and $P_1^* = (x_1^*, y(x_1^*))$ for parametrically and non-parametrically formulated variational problems.) Conjugate points can be computed from nontrivial solutions $u(t)$ of a Jacobian accessory differential equation if one determines t_1^* with $u(t_1^*) = 0$. The so-called Jacobi differential equation follows from consideration of the second variation (see *Bolza 1909*, §29, for parametric problems, and §12 for non-parametric problems; *Bliss 1946*, 27).

The geometric locus of the conjugate points P_1^* forms a curve that is the enveloping curve of the family of extremals E (solutions of the Euler-Lagrange equation). As soon as a curve of the family touches the enveloping curve belonging to the family, it is possible to use as well pieces of the envelope to solve the corresponding boundary value problem of the Euler-Lagrange equation. In other words, the extremity in these cases is lost. A vivid and concise discussion can be found in *Bliss 1946*, §§10–13. The necessary condition that the endpoint P_2 of an extremal E (here a geodesic) must come before the conjugate point P_1^* to the initial point of the extremal P_1 (with certain exceptions $P_2 = P_1^*$; see *Kneser 1898*, 50; *Osgood 1901a*, 166) is called the Jacobi condition.

Conjugate points limit (relative) extrema and thereby a fortiori also absolute extrema. Gaston Darboux had earlier emphasized, in his *Leçons sur la théorie générale des surfaces* (*Lectures on the general theory of surfaces*) (*1894*, 89), that the absolute extremum would always end before the relative extremum, or more precisely, that the points P_1^{**} that determine the end of the absolute extremum on the extremal E lie on E before those points P_1^* that determine the end of the relative extremum (therefore within the arc $P_1P_1^*$ of E). Analogously to the geometric locus of the conjugate point P_1^* to the initial point P_1 of the extremal, the enveloping curve of the extremal family, Zermelo introduces for the "by far shortest" paths (that is, for absolute extrema of the shortest paths) a curve, which he calls a *Doppelabstandskurve* (*double distance curve*); it presents the general situation for geodesics mentioned by Darboux in a more precise form.

Zur Theorie der kürzesten Linien

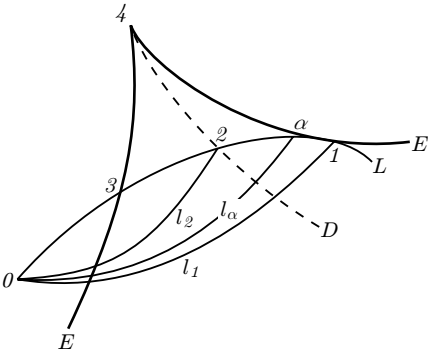
1902d

Das Problem der kürzesten Linien auf einer Fläche kann als das klassische Problem der Variationsrechnung bezeichnet werden, die zu seiner Lösung verwendeten Methoden sind vorbildlich und charakteristisch für die Methoden der Variationsrechnung überhaupt. Im Folgenden sei es mir daher gestattet, auf einige *Erweiterungen* dieses Problems hinzuweisen, die in ihrer anschaulichen Form zur Aufklärung allgemeinerer Variationsprobleme beitragen können.

1. Zunächst unterscheide ich zwischen *„kürzesten“* und *„allerkürzesten“* Linien, je nachdem sie nur in einem gewissen Flächenstücke ihrer Umgebung oder auf der ganzen Fläche die kürzeste Verbindung ihrer Endpunkte darstellen.

Jede von einem Punkte 0 ausgehende geodätische Linie L ist kürzeste bis zum nächsten „konjugierten Punkt“ 1, in dem sie die zu 0 gehörende Enveloppe E berührt. Aber schon zwischen 0 und 1 ist sie (nach *Kneser*) keine kürzeste mehr, und die allerkürzeste Verbindung beider Punkte ist eine von L verschiedene zweite geodätische Linie von der Länge $l_1 < L_1$. Dann werden auch die an l_1 sich stetig anschließenden geodätischen Linien l_α, deren Endpunkte α sehr nahe | vor 1 auf L liegen, immer noch kürzer sein als die auf L gemessenen geodätischen Abstände L_α dieser Punkte vom Anfangspunkt. Die geodätische Linie L hört also schon *vor* ihrem ersten konjugierten Punkte auf, eine „allerkürzeste“ zu sein, und es giebt eine Übergangsstelle 2, wo sie von einer gleichlangen geschnitten wird, sodaß $L_2 = l_2$ ist. Der geometrische Ort dieser Übergangspunkte 2 bildet die zu 0 gehörige „Doppelabstandskurve“ D, in der immer zwei gleich lange von 0 ausgehende kürzeste sich schneiden. Jede geodätische Linie ist also allerkürzeste nur bis zu einem gewissen Schnittpunkt mit der Doppelabstandskurve, der immer *vor* ihrem Berührungspunkt mit der Enveloppe liegen muß. (Für den Fall, daß sich in 0 keine kürzeste Linie selbst durchschneidet, können die geodätischen Linien l_α nicht kürzeste bleiben, wenn α an 0 heranrückt, sondern es giebt einen Grenzpunkt 3, der

185



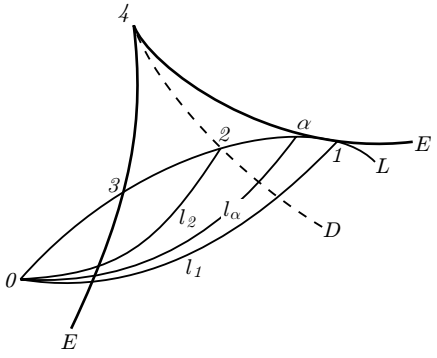
On the theory of shortest lines

1902d

The problem of the shortest lines on a surface may be considered the classic problem of the calculus of variations. The methods used for solving it are exemplary for and characteristic of the methods of the calculus of variations as a whole. Therefore, in what follows I may be permitted to suggest several *extensions* of this problem that, due to their distinct form, can help elucidate more general variational problems.

1. I first distinguish between "*shortest*" and "*by far shortest*" lines, depending on whether they represent the shortest connection between their endpoints only in a certain surface region of their neighborhood or on the entire surface.

Every geodetic line L starting from a point 0 is the shortest one up to the next "conjugate point" 1 at which it makes contact with the envelope E belonging to 0. However, already between 0 and 1 it is no longer the shortest one (according to *Kneser*), and by far the shortest connection between the two points is a second geodetic line different from L of length $l_1 < L_1$. Then the geodetic lines l_α which are continuously attached to l_1 and whose terminal points α lie very close before 1 on L will still be shorter than the geodetic distances L_α, measured along L, of these points from the starting point. Hence, the geodetic line L ceases to be a "by far shortest" line already *before* its first conjugate point, and there exists a transitional position 2, where it is intersected by an equally long line so that we have $L_2 = l_2$. The geometric location of this transitional point 2 forms the "double distance curve" D belonging to 0 in which two shortest lines of equal length starting from 0 always intersect. Hence, every geodetic line is a by far shortest line only up to a certain point where it intersects with the double distance curve and which must always lie *before* its point of contact with the envelope. (In the case where no shortest line intersects itself in 0 the geodetic lines l_α cannot remain shortest lines if α approaches 0. Instead, there then exists a boundary



auf l_3 die Rolle des konjugierten spielt und einen Schnittpunkt der Linie L mit der Enveloppe E darstellt.) Schließlich geht die Doppelabstandskurve D durch die Spitzen 4 der Enveloppe, in denen die entsprechenden Punkte 1, 2, 3 und die Linien L_4 und l_4 zusammenfallen.

2. Sucht man die kürzesten Linien innerhalb eines beliebig *begrenzten Flächenstückes*, so hat man ein Variationsproblem, in dem *Ungleichungen* als Nebenbedingungen auftreten. Ich beschränke mich hier auf ein einfach zusammenhängendes *ebenes* Flächenstück C, das also von einer geschlossenen sich selbst nicht schneidenden Kurve $\mathfrak{C}$ begrenzt ist, welche ihrerseits aus Stücken mit stetiger Tangente und endlicher Krümmung bestehen möge, im übrigen aber beliebig verlaufen kann. Im Innern von C liege ein Punkt 0, und gesucht werden die kürzesten ganz innerhalb C verlaufenden Verbindungslinien L_x dieses Punktes 0 mit allen Punkten x der Peripherie, also die verschiedenen Lagen, die ein in 0 befestigter gespannter Faden annimmt, wenn man ihn über die ganze Peripherie $\mathfrak{C}$ herumführt, die für den Faden als undurchdringlich vorausgesetzt wird.

Jede solche kürzeste Linie setzt sich, wie aus der Theorie der ersten Variation bekannt ist, abwechselnd aus geradlinigen Stücken und aus Teilen der Peripherie zusammen, die *nach innen konvex* sind, und an den Übergangs-
186 stellen muß Berührung stattfinden. Nun läßt sich | zeigen, daß das System dieser Linien *stetig* über die ganze Peripherie fortgesetzt werden kann, und daß zwei solcher Linien wohl im Anfang ein Stück gemein haben, aber, einmal getrennt, sich niemals wieder schneiden können. Wäre nämlich das letztere der Fall, so würden die beiden kürzesten Linien zwischen den Punkten 1 und 2 ein Flächenstück C' einschließen, das ganz im Innern von C läge und dessen Begrenzung mit Ausnahme der beiden Endpunkte *nirgends konvex* wäre. Das ist aber unmöglich, da unter den Parallelen zu der Geraden 12, die das Flächenstück durchsetzen, immer eine *letzte* existieren und diese dann notwendig an einer *konvexen* Stelle 3 die Peripherie berühren müßte. Die gefundenen kürzesten Linien L_x sind also die einzigen und damit auch die allerkürzesten innerhalb C und zerlegen *fächerförmig* das ganze Flächenstück C. Durch eine beliebig kleine Abänderung dieser Figur könnte man nun die zusammenfallenden Linienstücke durch getrennte ersetzen und mit Hilfe des entstehenden Strahlensystems das ganze Innere des Flächenstückes C ein-eindeutig und stetig auf das Innere eines Kreises abbilden. Doch bedarf dies noch eines strengen Beweises.

3. Anstatt einfacher Ungleichungen kann man auch Ungleichheiten zwischen Differentialausdrücken, „Differentialungleichungen“, als Nebenbedingungen einführen, man kann z. B. fordern, daß die *Steilheit* der Kurve eine gegebene Grenze nicht überschreitet: $\varkappa = \dfrac{dz}{ds} \leqq \varkappa_0$, wenn die Fläche gegeben ist durch eine Gleichung: $z = \varphi(x, y)$. Man sucht also für zwei Punkte der Fläche die kürzeste unter allen Verbindungslinien von „begrenzter Steilheit“, ein Problem, wie es praktisch angenähert verwirklicht ist beim Bau von Gebirgsstraßen. Die Lösungen dieses Problems setzen sich, wie leicht ersichtlich,

point 3 that plays the role of the conjugate point along l_3 and represents an intersection point of the line L with the envelope E.) The double distance curve D eventually runs through the cusps 4 of the envelope in which the corresponding points 1, 2, 3 and the lines L_4 and l_4 coincide.

2. If we seek the shortest lines in the interior of an arbitrarily *bounded surface region*, then we are confronted with a variational problem involving *inequalities* as auxiliary conditions. I shall restrict myself here to a simply connected *planar* surface region C, and hence to one that is bounded by a closed curve $\mathfrak{C}$ not intersecting itself which, in turn, may consist of regions with continuous tangent and finite curvature but whose course is otherwise arbitrary. Let 0 be a point in the interior of C. Now find those shortest lines L_x that lie entirely within C and connect this point 0 with all points x of the periphery, and hence the various positions taken by a taut thread fixed at 0, if we pass it round the entire periphery $\mathfrak{C}$, which we take to be impenetrable to the thread.

As we know from the theory of the first variation, each such shortest line consists of alternately rectilinear portions and of parts of the periphery that are *inwardly convex*, and contact must be made at the transitional positions. Now, it is possible to show that the system of these lines is capable of a continuous continuation across the entire periphery and that once separated any two such lines, while certainly capable of sharing an initial segment, can never intersect one another again. For if the latter were the case, then the two shortest lines between the points 1 and 2 would enclose a surface region C' which would lie entirely in the interior of C and whose boundary except for the two terminal point would be *nowhere convex*. This, however, is impossible since among the lines parallel to the straight line 12 that lace the surface region there would always have to be a *last one* which then, by necessity, would have to make contact with the periphery at a *convex* point 3. The shortest lines L_x found are therefore the only ones in existence, and hence also the by far shortest lines within C, and they divide the entire surface region C *like a fan*. By an arbitrarily small modification of this figure, we could now replace the coinciding line segments by separate ones and, by means of the emerging system of rays, map one-to-one and continuously the entire interior of the surface region C onto the interior of a circle. This, however, still stands in need of rigorous proof.

3. Instead of simple inequalities, we can also introduce inequalities between differential expressions, "differential inequalities", as auxiliary conditions. We can, e.g., require that the *steepness* of the curve not exceed a given limit: $\varkappa = \dfrac{dz}{ds} \leqq \varkappa_0$, if the surface is given in terms of an equation: $z = \varphi(x, y)$. Hence, what is sought is, given two points of the surface, the shortest among all connecting lines of "limited steepness", a problem that is approximately realized in the construction of mountain roads. As is readily evident, the solutions to this problem are composed of shortest *geodetic*

zusammen aus kürzesten *geodätischen* Linien, soweit sie nirgends zu steil sind, und aus Kurven von *konstanter Steilheit* $\varkappa = \varkappa_0$, und an den Übergangsstellen müssen auch die geodätischen Linien dieselbe maximale Steilheit besitzen. Jedes Stück einer solchen Linie konstanter Steilheit, die ich zur Abkürzung einfach als „Kletterkurve“ bezeichnen will, ist nun auch wirklich immer die kürzeste Verbindung seiner Endpunkte unter allen denen, die nirgends steiler sind, und zwei Kletterkurven von derselben Steilheit zwischen denselben Punkten sind auch immer gleich lang. Da es nun in jedem Flächenpunkte im allgemeinen zwei reelle oder imaginäre Richtungen giebt, für welche die Steilheit einen gegebenen Wert $\varkappa$ annimmt, so giebt es zwei einfach-unendliche Scharen von Kletterkurven $\varkappa_0$, die alle Teile der Fläche, in denen diese Steilheit $\varkappa_0$ überhaupt möglich ist, netzförmig überdecken und an den Grenz-
187 | kurven dieser Gebiete in Spitzen endigen. In einem solchen Gebiete kann man nun zwei Punkte, die nicht auf derselben Kletterkurve liegen, durch abwechselnde Verwendung der beiden Scharen gleichwohl sehr häufig durch eine Linie konstanter Steilheit $\varkappa_0$ verbinden, also durch eine Zickzacklinie oder Serpentine, wie wir sie auf den meisten Gebirgspässen beobachten können. Nun läßt sich zeigen: jede Kurve von überall *größerer* Steilheit $\varkappa > \varkappa_0$ kann immer durch eine zickzackförmige Kletterkurve $\varkappa_0$ ersetzt werden, die in beliebiger Nähe zwischen denselben Endpunkten verläuft und allen Bedingungen des Problems genügt. Solche Kurven $\varkappa > \varkappa_0$ giebt es aber von einem Anfangspunkte O aus nach allen Punkten eines Gebietes, das von den beiden von O ausgehenden einfachen Kletterkurven und außerdem von der Grenzkurve des betreffenden Steilheitsgebietes oder des Kletterkurven-Netzes $\varkappa_0$ begrenzt wird.

Wollte man dagegen, wie beim Eisenbahnbau, plötzliche Richtungsänderungen ausschließen und auch für die *Krümmung* des Weges eine obere Grenze k_0 festsetzen, so wäre das Problem in vielen Fällen überhaupt unlösbar, weil es solche Kurven zwischen den betrachteten Punkten auf der Fläche überhaupt nicht gäbe. In solchen Fällen müßte man notwendig die Fläche verlassen und in den Raum eindringen, also, technisch gesprochen, Tunnels und Viadukte anlegen. Dabei ergeben sich denn auch unter anderem die schraubenförmigen „Kehrtunnels“ verschiedener Gebirgsbahnen als mathematisch korrekte Lösungen des Variationsproblems.

lines, provided that they are nowhere too steep, and of curves of *constant steepness* $\varkappa = \varkappa_0$, and at the transitional positions the geodetic lines, too, must have the same maximal steepness. Now, every segment of such a line of constant steepness, or "climbing curve" for short, is really always the shortest connection of its endpoints among all those that are nowhere steeper, and two climbing curves of identical steepness between the identical points are also always of the same length. Since now in every surface point there are, in general, two real or imaginary directions for which the steepness takes a given value $\varkappa$, there are two simply-infinite families of climbing curves $\varkappa_0$ that cover all parts of the surface in which this steepness $\varkappa_0$ is possible at all like a net and that terminate in cusps at the boundary curves of these regions. Now, in such a region, it is often possible to connect two points that do not lie on the same climbing curve by a line of constant steepness $\varkappa_0$ by alternately using the two families, and hence by a zig-zag line or serpentine, as we can find them at most mountain passes. We can now show: every curve of everywhere *greater* steepness $\varkappa > \varkappa_0$ can always be replaced by a zig-zag-like climbing curve $\varkappa_0$ that runs at arbitrary proximity between the same endpoints and meets all conditions of the problem. Such curves $\varkappa > \varkappa_0$ exist, however, running from a starting point O to all points of a region that is bounded by the two simple climbing curves starting from O and, in addition, also by the boundary curve of the respective steepness region or of the climbing curve net $\varkappa_0$.

If, in contrast, sudden changes in direction were to be excluded, as in the case of railway construction, and an upper limit k_0 were to be fixed also for the *curvature* of the path, then, in many cases, the problem would be simply unsolvable, for such curves would simply not exist between the considered points on the surface. In such cases, we would, by necessity, have to leave the surface and move into space, and hence, technically speaking, build tunnels and overpasses. Here, we find as mathematically correct solutions to the variational problem also the coil-like "helical tunnels" of various mountain railways.

Introductory note to 1904a

Rüdiger Thiele

The question of feasible necessary conditions for solutions of a variational problem is as old as the calculus of variations itself, as for example in the variational problem

$$J(y) = \int_a^b f(x, y(x), y'(x))\, dx \;\rightarrow\; \text{extremum}\,. \tag{1}$$

Joseph Louis de Lagrange developed the elegant method of transforming the necessary condition $\delta J = 0$ (vanishing of the first variation) by partial integration so that under the integral there stands a product of the variation of the comparison function $\delta y = y(x) - y^0(x)$ (with solution $y^0(x)$) and a differential expression $E(y)$:

$$\int_a^b E(y)\, \delta y\, dx = 0\,.$$

From this one derives, with the help of a so-called fundamental lemma that relies on continuity arguments, the Euler-Lagrange equation of second order as a necessary condition for a twice continuously differentiable solution $y^0(x)$ of (1):

$$E(y) = \frac{d}{dx} f_{y'} - f_y = 0\,. \tag{2}$$

In the course of the development of mathematics, the question has arisen whether the assumptions for the derivation of this necessary condition can be weakened, for a critical glance at (1) shows that in the problem itself, no second derivatives are required, rather that it involves assumptions that were made for technical reasons in carrying out the proof. Compared to a twice differentiable solution $y^0(x)$ of (2), in the case of a minimum problem, comparison functions with less differentiability could return possibly smaller values than the extremal $y^0(x)$. One can show that such is not the case (see *Courant and Hilbert 1924*, 167f.). But it is natural to ask whether for extremals of (1), one could do without the hypothesis of twice continuous differentiability. Zermelo certainly concerned himself with such questions when he was lecturing on the calculus of variations in Göttingen in the summer semester of 1902.

Paul du Bois-Reymond (*1879a,b*) derived, using a differently executed partial integration in $\delta J = 0$, necessary conditions for only continuously differentiable extremals of (1) and was able to show that with this weakening of the conditions, there are no additional solutions. To do so, he used a

H.-D. Ebbinghaus, A. Kanamori (Eds.), *Ernst Zermelo - Collected Works/ Gesammelte Werke II*, DOI 10.1007/978-3-540-70856-8_9, Schriften der Mathematisch-naturwissenschaftlichen Klasse der Heidelberger Akademie der Wissenschaften 23,

lemma (now known as the du Bois-Reymond lemma) that in Oskar Bolza's formulation reads as follows:[1]

If $g(x)$ in $[a,b]$ is continuous and if for every continuously differentiable function $\eta(x)$ in $[a,b]$ that vanishes at the endpoints a and b, we have

$$\int_a^b g(y)\eta'(x)\,dx = 0\,, \tag{3}$$

then $g(x) =$ constant in $[a,b]$.

In the paper under consideration, Zermelo generalizes du Bois-Reymond's result in *1879a* and also simplifies du Bois-Reymond's proof. The generalization aims for higher derivatives (see below). In 1909, Bolza authored the epochal enlarged German edition *Vorlesungen über Variationsrechnung* (*1909*) of his *Lectures on the calculus of variations* (*1904*), in which he presents the du Bois-Reymond lemma with a different proof. Constantin Carathéodory remarks in his review *1910* of the *Vorlesungen* that in this meritorious work, the proofs are given with the necessary exactness and conciseness, but that "the greatest possible simplicity is not always achieved." He continues that regarding the du Bois-Reymond lemma, Zermelo's proof, as "unsurpassable with respect to simplicity, conciseness, and classical elegance," should have been presented.[2]

Zermelo gives two proofs. The first uses a (variable) point of continuity of $g(x)$ to show with a special variation $\eta(x)$ that $g(x)$ is a polynomial of degree at most $(n-1)$. The second proof is close to that of du Bois-Reymond (*1879a*, 307ff.), which considers an isoperimetric variational problem.[3] In addition to (3), one has, on account of the constraints,

$$\int_a^b \eta'(x)\,dx = \eta(b) - \eta(a) = 0\,.$$

The second proof is again presented with an altered conclusion. For this, Zermelo uses an observation by Erhard Schmidt, who had come to Göttingen in 1901 and would soon finish his doctorate under David Hilbert in 1905.

Zermelo's generalized result is as follows:

Let n be an arbitrary natural number. Furthermore, let $g(x)$ in $[a,b]$ be continuous, and suppose that for every n-times continuously differentiable

[1] The American school of the calculus of variations (Bolza) reduced such general analytic assumptions as those that assume the integrand in (1) to be an analytic function, while Zermelo still assumed the analyticity of the integrand (p. 560). Bolza's representation can be found in *Bolza 1904*, 22ff., *Bolza 1909*, 25ff.

[2] *Carathéodory 1910*, 222, *Carathéodory 1954/57*, vol. V, 305; *Ebbinghaus 2007*, 14.

[3] Through the substitutions $y^{(i)}(x) = u_i$ $(i = 1,\dots,n-1)$, $y^{(n)}(x) = u'_{n-1}$, one can reduce the variational problem to an equivalent variational problem with first derivative and $n-1$ constraints $G_i \equiv u'_i - u_{i+1} = 0$ $(i = 0,\dots,n-2)$ (that is, to an isoperimetric variational problem).

function $\eta(x)$ in this interval that vanishes together with its derivatives up to order $n-1$ at the endpoints of the interval $[a,b]$, we have

$$\int_a^b g(y)\eta^{(n)}(x)\,dx = 0\,. \tag{4}$$

Then $g(x)$ is a polynomial of degree at most $n-1$,

$$g(x) = c_1 + c_2 x + \cdots + c_n x^{n-1}\,. \tag{5}$$

Zermelo works with analytic functions $g(x)$ (p. 560). Incidentally, at the suggestion of his teacher Karl Weierstrass, he gave a justification for the abovementioned fundamental lemma (*Zermelo 1894*, 41); the fundamental lemma arises when one formally replaces $\eta^{(n)}$ in (4) by η and then in (5) takes $c_1 = 0$.

The argument against analytic assumptions is of even greater importance in the case of multiple integrals. A generalization to n-dimensional integrals can be found in *Giaquinta and Hildebrandt 1996*, vol. I, 32:

Über die Herleitung der Differentialgleichung bei Variationsproblemen

1904a

Nach der Methode von *Lagrange* pflegt man die erste Variation δS eines bestimmten Integrales S, dessen Maximum oder Minimum gesucht wird, durch partielle Integration so umzuformen, daß unter dem Integralzeichen nur noch die Variation δy der unbekannten Funktion selbst, aber keine ihrer Ableitungen übrig bleibt. Durch Einführung einer speziellen Variation schließt man dann auf das Verschwinden des mit δy multiplizierten Ausdruckes und damit auf das Bestehen der Lagrangeschen Differentialgleichung als einer notwendigen Bedingung für das Eintreten eines Maximums oder Minimums. Bei dieser Herleitung wird aber die Existenz und die Stetigkeit der ersten $2n$ Ableitungen der unbekannten Funktion vorausgesetzt, wenn der Integrandus von S die Ableitungen bis zur n^{ten} Ordnung enthält. Es wäre also denkbar, daß bei diesem Verfahren alle diejenigen Lösungen des Variationsproblemes verloren gingen, welche nicht so oft differentiierbar sind. Dieses Bedenken veranlaßte bereits im Jahre 1879 *P. du Bois-Reymond* (Math. Ann. 15) zu einer von der gebräuchlichen abweichenden Umformung der ersten Variation, welche nur die Integrierbarkeit des n^{ten} Differentialquotienten der gesuchten Funktion voraussetzt, an allen Stetigkeitsstellen dieser Ableitung aber gleichfalls auf die Notwendigkeit der Lagrangeschen Differentialgleichung führt und nun

Let G be a domain of $\mathbb{R}^n$, f a Lebesgue integrable function on G, and $D = (D_1, \ldots, D_n)$ the differential operator for the first partial derivatives. If for every analytic comparison function $\eta \in C_c^\infty(G)$, one has

$$\int_G f(x) D_i \eta \, dx = 0 \quad (i = 1, \ldots, n),$$

then $f =$ constant *almost everywhere on* G.

It remains to mention that Hans Hahn also dealt with questions regarding the degree to which the assumption of a continuous second derivative of the extremals could be weakened. In *1903b*, Hahn showed that a reduction of this assumption is also possible for Lagrange problems (problems with differential equations as constraints), for which he used the Mayer form of this variational problem. Hahn was able to weaken the strong assumption of an analytic integrand to the derivatives that actually appear. Finally, in *1906*, he extended this result to rectifiable curves with unique tangents.

On the derivation of the differential equation in variational problems

1904a

When seeking the maximum or minimum of a definite integral S, it is, according to *Lagrange*'s method, customary to transform the first variation δS by partial integration so that only the variation δy of the unknown function itself but none of its derivatives still remains under the integral sign. By a special variation one then infers the disappearance of the expression multiplied by δy, and hence the existence of the *Lagrange* differential equation as a necessary condition for the occurrence of a maximum or minimum. But for this deduction one presupposes the existence and the continuity of the first $2n$ derivatives of the unknown function if the integrand S contains the derivatives up to the nth order. Thus, it would be conceivable that all those solutions of the variational problem that cannot be differentiated that many times are lost in this procedure. This concern already prompted *P. du Bois-Reymond* in 1879 (*1879a,b*) to adopt a transformation of the first variation which, unlike the customary transformation, only assumes that the nth derivative of the sought function is integrable, but which also leads to the necessity of the *Lagrange* differential equation at all continuity points of this derivative and now licences the reverse inference to the unlimited differentia-

rückwärts auf die unbegrenzte Differentiierbarkeit der gesuchten Funktion zu schließen gestattet. Indessen ist diese Methode trotz ihrer Leistungsfähigkeit und prinzipiellen Bedeutung bisher fast unbeachtet geblieben.*) Aus diesem Grunde schien es mir wünschenswert, der etwas komplizierten und nicht ganz elementaren *du Bois-Reymond*schen Beweisführung eine neue möglichst vereinfachte Gestalt zu geben, welche auch in jeder elementaren Darstellung der Variationsrechnung verwendbar sein soll.

559 | Es sei

$$S = \int_a^b V\left(x, y, y', \cdots, y^{(n)}\right) dx$$

das vorgelegte Integral, in welchem die Argumente $y, y', \cdots, y^{(n)}$ die gesuchte Funktion mit ihren ersten n nach x genommenen Ableitungen ausdrücken sollen. Dann muß für den Fall eines Extremums die erste Variation verschwinden, d. h. es ist:

$$\delta S = \int_a^b \left(Y\delta y + Y_1\delta y' + \cdots + Y_n\delta y^{(n)}\right) dx = 0\,, \tag{1}$$

wo $Y, Y_1, \cdots, Y_n$ die partiellen nach $y, y', \cdots, y^{(n)}$ genommenen Ableitungen von V bedeuten und $\delta y = \eta(x)$ eine beliebige nach n facher Differentiation noch integrierbare Funktion von x ist, welche zugleich mit ihren $n-1$ ersten Ableitungen $\delta y' = \eta'(x), \cdots, \delta y^{(n-1)} = \eta^{(n-1)}(x)$ an beiden Grenzen $x = a$ und $x = b$ verschwindet und mit diesen Ableitungen im ganzen Integrationsintervall ausnahmslos stetig sein soll.

Durch partielle Integration versucht man nun nach *du Bois-Reymond*, alle Glieder bis auf das höchste, mit $\delta y^{(n)}$ multiplizierte aus dem Integranden der ersten Variation zu entfernen, indem man ansetzt:

$$Y\delta y + Y_1\delta y' + \cdots + Y_n\delta y^{(n)} = \frac{d}{dx}\left(\Lambda_1\delta y + \Lambda_2\delta y' + \cdots + \Lambda_n\delta y^{(n-1)}\right) + \Lambda\delta y^{(n)}$$

und erhält successive durch Koeffizientenvergleichung:

$$\Lambda_1' = Y, \quad \Lambda_2' = Y_1 - \Lambda_1, \cdots, \Lambda_n' = Y_{n-1} - \Lambda_{n-1} \tag{2}$$

und schließlich:

$$\Lambda = \Lambda(x) = Y_n - \Lambda_n = Y_n - \int Y_{n-1}dx + \int dx \int Y_{n-2}dx - + \cdots \pm {}^{(n)}\int Y dx^n\,. \tag{3}$$

*) Doch vergl. neuerdings *H. Hahn*, Monatsh. Math. Phys. 14 (1903) p. 325, wo die angedeutete Schlußweise auf die Multiplikatorenmethode des „allgemeinen“ Problemes ausgedehnt wird, und für den allereinfachsten Fall *J. E. Whittemore*, Ann. of math. (2) 2 (1901) p. 125.

bility of the sought after function. Its efficiency and fundamental significance notwithstanding, this method has been almost completely ignored heretofore.[1] For this reason, I thought it desirable to recast the somewhat intricate and not entirely elementary proof by *du Bois-Reymond* in a new form that is as simple as possible and that can also be used in any elementary account of the calculus of variations.

Let

$$S = \int_a^b V\left(x, y, y', \cdots, y^{(n)}\right) dx$$

be the given integral in which the arguments $y, y', \cdots, y^{(n)}$ are supposed to express the sought after function with its first n derivatives with respect to x. Then for an extremum the first variation must vanish, i. e.,

$$\delta S = \int_a^b \left(Y\delta y + Y_1\delta y' + \cdots + Y_n\delta y^{(n)}\right) dx = 0\,, \tag{1}$$

where $Y, Y_1, \cdots, Y_n$ denote the partial derivatives of V with respect to $y, y', \cdots, y^{(n)}$, and $\delta y = \eta(x)$ any function of x that is still integrable after n-fold differentiation and vanishes, along with its $n-1$ first derivatives $\delta y' = \eta'(x), \cdots, \delta y^{(n-1)} = \eta^{(n-1)}(x)$, at both limits $x = a$ and $x = b$ and that, along with these derivatives, is to be continuous in the entire integration interval without exception.

Following *du Bois-Reymond*, one now uses partial integration in order to eliminate from the integrand of the first variation all terms except for the highest term, which is multiplied by $\delta y^{(n)}$, by setting

$$Y\delta y + Y_1\delta y' + \cdots + Y_n\delta y^{(n)} = \frac{d}{dx}\left(\Lambda_1\delta y + \Lambda_2\delta y' + \cdots + \Lambda_n\delta y^{(n-1)}\right) + \Lambda\delta y^{(n)}$$

and successively obtain by comparison of coefficients

$$\Lambda_1' = Y, \quad \Lambda_2' = Y_1 - \Lambda_1, \cdots, \Lambda_n' = Y_{n-1} - \Lambda_{n-1} \tag{2}$$

and eventually

$$\Lambda = \Lambda(x) = Y_n - \Lambda_n = Y_n - \int Y_{n-1}dx + \int dx \int Y_{n-2}dx - + \cdots \pm {}^{(n)}\int Y dx^n\,. \tag{3}$$

[1] But see the recent *Hahn 1903b*, where the indicated inference is extended to the multiplicator method of the "general" problem, and, for the simplest case, *Whittemore 1901a*.

Wir haben also mit den so bestimmten Funktionen $\Lambda, \Lambda_1, \cdots, \Lambda_n$:

$$\delta S = \left[\Lambda_1 \delta y + \Lambda_2 \delta y' + \cdots + \Lambda_n \delta y^{(n-1)}\right]_a^b + \int_a^b \Lambda \delta y^{(n)} dx = 0\,, \qquad (4)$$

und für alle Variationen $\delta y = \eta(x)$, die auch den angegebenen Grenzbedingungen genügen:

$$\int_a^b \Lambda(x)\eta^{(n)}(x)dx = 0\,. \qquad (5)$$

Aus dieser Beziehung kann nun, was nachher ausgeführt ⟦werden⟧ soll, geschlossen werden, daß Λ ein Polynom $n - 1^{\text{ten}}$ Grades von x sein muß. Dann folgt aus der unbegrenzten Differentiierbarkeit von Λ auch die von y wenigstens an jeder Stetigkeitsstelle von $y^{(n)}$. Da nämlich

$$Y_n = Y_n\left(x, y, y', \cdots, y^{(n)}\right)$$

560 | ebenso wie V als eine analytische Funktion ihrer sämtlichen $n+2$ Argumente vorausgesetzt werden soll, so ist auch $y^{(n)}$ eine analytische Funktion von $x, y, y', \cdots, y^{(n-1)}$, Y_n und als solche regulär an jeder regulären Stelle von Y_n, an welcher $Y_{nn} = \dfrac{\partial Y_n}{\partial y^{(n)}} \neq 0$ ist. Als Funktion von x betrachtet ist daher $y^{(n)}$ an jeder solchen Stelle $x = x_0$ ebenso oft differentiierbar wie Y_n, da sich jede weitere Ableitung $y^{(n+r)}$ durch die niederen Ableitungen von y und durch die nach x genommenen Ableitungen von Y_n bis zur r^{ten} ausdrücken läßt. Mit $y^{(n)}$ sind aber auch alle Funktionen $Y_1, Y_2, \cdots, Y_{n-1}$ und ferner gemäß (2) alle Funktionen $\Lambda_1', \Lambda_2', \cdots, \Lambda_n'$ r mal differentiierbar, und dasselbe gilt, wenn Λ ein Polynom ist, auch von der Funktion $Y_n' = \Lambda' + \Lambda_n'$. Y_n und mit ihm $y^{(n)}$ ist demnach sogar $r+1$ mal differentiierbar, also einmal mehr, als vorausgesetzt, d. h. beliebig oft.

Durch n-malige Differentiation von Λ erhält man nun für jede Stetigkeitsstelle von $y^{(n)}$, an welcher V regulär und $Y_{nn} \neq 0$ ist,

$$\Lambda^{(n)} = Y_n^{(n)} - Y_n^{(n-1)} + \cdots + Y = 0\,, \qquad (6)$$

also die bekannte Lagrangesche Differentialgleichung für die gesuchte Funktion $y = f(x)$.

Für den Satz, *daß $\Lambda(x)$ unter der Voraussetzung* (5) *ein Polynom $n-1^{\text{ten}}$ Grades sein muß*, gebe ich jetzt zwei verschiedene Beweise, deren erster, wie in solchen Fällen üblich, eine spezielle diskontinuierliche Variation zu Grunde legt, während der zweite, kürzer und kunstvoller, der „isoperimetrischen" Beweismethode *du Bois-Reymond*s näher kommt.

Thus, for the functions $\Lambda, \Lambda_1, \cdots, \Lambda_n$ so determined, we have

$$\delta S = \left[\Lambda_1 \delta y + \Lambda_2 \delta y' + \cdots + \Lambda_n \delta y^{(n-1)}\right]_a^b + \int_a^b \Lambda \delta y^{(n)} dx = 0\,, \tag{4}$$

and for all variations $\delta y = \eta(x)$ also satisfying the specified limit conditions

$$\int_a^b \Lambda(x)\eta^{(n)}(x)dx = 0\,. \tag{5}$$

As shall ⟦be⟧ shown below, it is now possible to infer from this relation that Λ must be a polynomial of $(n-1)$ th degree of x. From the unlimited differentiability of Λ then follows the unlimited differentiability of y at least at every continuity point of $y^{(n)}$. For since it is assumed that

$$Y_n = Y_n\left(x, y, y', \cdots, y^{(n)}\right)\,,$$

like V, is an analytic function of all of its $n+2$ arguments, $y^{(n)}$, too, is an analytic function of $x, y, y', \cdots, y^{(n-1)}$, Y_n and hence regular at every regular point of Y_n at which $Y_{nn} = \dfrac{\partial Y_n}{\partial y^{(n)}} \neq 0$. Thus, when considered as a function of x, $y^{(n)}$ is differentiable at each such point $x = x_0$ as many times as Y_n, since any further derivative $y^{(n+r)}$ can be expressed by the lower derivatives of y and by the derivatives of Y_n with respect to x up to the rth one. But along with $y^{(n)}$ also all functions $Y_1, Y_2, \cdots, Y_{n-1}$ and, furthermore, by (2), all functions $\Lambda_1', \Lambda_2', \cdots, \Lambda_n'$ are differentiable r times, and the same also holds true of the function $Y_n' = \Lambda' + \Lambda_n'$, assuming Λ is a polynomial. Hence, Y_n and, along with it, $y^{(n)}$ is differentiable even $r+1$ times, and thus one more time than was assumed, i. e., arbitrarily many times.

By an n-fold differentiation of Λ, one now obtains for each continuity point of $y^{(n)}$ at which V is regular and $Y_n \neq 0$ [2]

$$\Lambda^{(n)} = Y_n^{(n)} - Y_n^{(n-1)} + \cdots + Y = 0\,, \tag{6}$$

and hence the well-known Lagrange differential equation for the sought function $y = f(x)$.

For the theorem stating *that, given assumption* (5), $\Lambda(x)$ *must be a polynomial of* $(n-1)$*th degree*, I now provide two different proofs, the first of which, as is customary in such cases, is based on a special discontinuous variation, while the second one, which is briefer and more artful, comes closer to *du Bois-Reymond*'s "isoperimetric" proof method.

[2] ⟦Zermelo erroneously writes "Y_{nn}" instead of "Y_n".⟧

Erster Beweis.

Ist $g(x) = u_1 + u_2 x + \cdots + u_n x^{n-1}$ ein beliebiges Polynom $n-1^{\text{ten}}$ Grades, so ist gemäß der Lagrangeschen Interpolationsformel für beliebige $n+1$ Argumente $x_0, x_1, \cdots, x_n$ und für $\varphi(x) = (x - x_0)(x - x_1) \cdots (x - x_n)$:

$$\frac{g(x_0)}{\varphi'(x_0)} + \frac{g(x_1)}{\varphi'(x_1)} + \cdots + \frac{g(x_n)}{\varphi'(x_n)} = 0 \tag{7}$$

und umgekehrt, wenn für $n+1$ Argumente $x_0, x_1, \cdots, x_n$ diese Bedingung erfüllt ist, so werden die entsprechenden Funktionswerte $g(x_0), g(x_1), \cdots, g(x_n)$ durch ein und dasselbe Polynom $g(x)$ dargestellt. Es ist aber auch, wenn $g(\xi, x)$ ein Polynom $n - 1^{\text{ten}}$ Grades in ξ mit dem willkürlichen Parameter x ist,

$$\frac{g(x_0, x)}{\varphi'(x_0)} + \frac{g(x_1, x)}{\varphi'(x_1)} + \cdots + \frac{g(x_n, x)}{\varphi'(x_n)} = 0\,. \tag{8}$$

Es seien nun $x_0, x_1, \cdots, x_n$ beliebige $n+1$ Stetigkeitsstellen der Funktion $\Lambda(x)$ im Intervalle $a \cdots b$ und h eine beliebige positive Größe, welche kleiner ist als die kleinste Differenz dieser Größen x_i. Dann betrachten | wir die 561
folgende im ganzen Intervalle $(a \cdots b)$ mit ihren ersten $n-1$ Ableitungen stetige Funktion $\eta = \eta(x)$, deren n^{te} Ableitung in allen Teilintervallen zwischen $x_i - h$ und $x_i + h$ konstant, im ganzen Reste des Integrationsgebietes aber $= 0$ ist. Es sei nämlich für

$$\begin{aligned}
&a \leqq x \leqq x_0 - h && \eta = 0 && \eta^{(n)} = 0\\
&x_0 - h \leqq x \leqq x_0 + h && \eta = \overline{\eta_0}(x) = \frac{(x - x_0 + h)^n}{\varphi'(x_0)}\,, && \eta^{(n)} = \frac{n!}{\varphi'(x_0)}\\
&x_0 + h \leqq x \leqq x_1 - h && \eta = \eta_0(x) = && \\
& && = \frac{(x - x_0 + h)^n - (x - x_0 - h)^n}{\varphi'(x_0)} && \eta^{(n)} = 0\\
&x_1 - h \leqq x \leqq x_1 + h && \eta = \overline{\eta}_1(x) && \\
& && = \eta_0(x) + \frac{(x - x_1 + h)^n}{\varphi'(x_1)} && \eta^{(n)} = \frac{n!}{\varphi'(x_1)}\\
&x_1 + h \leqq x \leqq x_2 - h && \eta = \eta_1(x) && \\
& && = \eta_0(x) + \frac{(x - x_1 + h)^n - (x - x_1 - h)^n}{\varphi'(x_1)} && \eta^{(n)} = 0
\end{aligned}$$

usw. Dann ist im letzten Intervalle:

$$x_n + h \leqq x \leqq b \quad \eta = \eta_n(x) = \eta_{n-1}(x) + \frac{(x - x_n + h)^n - (x - x_n - h)^n}{\varphi'(x_n)} = 0\,,$$

denn es ist

$$g\,(\xi, x) = (x - \xi + h)^n - (x - \xi - h)^n$$

First proof.

If $g(x) = u_1 + u_2 x + \cdots + u_n x^{n-1}$ is any polynomial of $(n-1)$ th degree, then, according to the Lagrange interpolation formula, for any $n+1$ arguments $x_0, x_1, \cdots, x_n$ and for $\varphi(x) = (x-x_0)(x-x_1)\cdots(x-x_n)$,

$$\frac{g(x_0)}{\varphi'(x_0)} + \frac{g(x_1)}{\varphi'(x_1)} + \cdots + \frac{g(x_n)}{\varphi'(x_n)} = 0 \tag{7}$$

and, conversely, if this condition is satisfied for $n+1$ arguments $x_0, x_1, \cdots, x_n$, then the corresponding function values $g(x_0), g(x_1), \cdots, g(x_n)$ are represented by one and the same polynomial $g(x)$. But if $g(\xi, x)$ is a polynomial of $(n-1)$ th degree in ξ in some parameter x, then also

$$\frac{g(x_0, x)}{\varphi'(x_0)} + \frac{g(x_1, x)}{\varphi'(x_1)} + \cdots + \frac{g(x_n, x)}{\varphi'(x_n)} = 0\,. \tag{8}$$

Suppose now that $x_0, x_1, \cdots, x_n$ are any $n+1$ continuity points of the function $\Lambda(x)$ in the interval $a \cdots b$, and h any positive magnitude smaller than the least difference of these magnitudes x_i. We then consider the following function $\eta = \eta(x)$ which, along with its first $n-1$ derivatives, is continuous in the entire interval $(a \cdots b)$ and whose nth derivative is constant in all partial intervals between $x_i - h$ and $x_i + h$ but $= 0$ in the entire remainder of the integration domain. For suppose that for

$$\begin{aligned}
&a \leqq x \leqq x_0 - h && \eta = 0 && \eta^{(n)} = 0\\
&x_0 - h \leqq x \leqq x_0 + h && \eta = \overline{\eta_0}(x) = \frac{(x - x_0 + h)^n}{\varphi'(x_0)}\,, && \eta^{(n)} = \frac{n!}{\varphi'(x_0)}\\
&x_0 + h \leqq x \leqq x_1 - h && \eta = \eta_0(x) = && \\
& && = \frac{(x - x_0 + h)^n - (x - x_0 - h)^n}{\varphi'(x_0)} && \eta^{(n)} = 0\\
&x_1 - h \leqq x \leqq x_1 + h && \eta = \overline{\eta}_1(x) && \\
& && = \eta_0(x) + \frac{(x - x_1 + h)^n}{\varphi'(x_1)} && \eta^{(n)} = \frac{n!}{\varphi'(x_1)}\\
&x_1 + h \leqq x \leqq x_2 - h && \eta = \eta_1(x) && \\
& && = \eta_0(x) + \frac{(x - x_1 + h)^n - (x - x_1 - h)^n}{\varphi'(x_1)} && \eta^{(n)} = 0
\end{aligned}$$

etc. Then, in the last interval, we have

$$x_n + h \leqq x \leqq b \quad \eta = \eta_n(x) = \eta_{n-1}(x) + \frac{(x - x_n + h)^n - (x - x_n - h)^n}{\varphi'(x_n)} = 0\,,$$

since

$$g\,(\xi, x) = (x - \xi + h)^n - (x - \xi - h)^n$$

ein Polynom $n-1^{\text{ten}}$ Grades in ξ, also nach (8)

$$\eta_n(x) = \frac{(x-x_0+h)^n-(x-x_0-h)^n}{\varphi'(x_0)} + \cdots + \frac{(x-x_n+h)^n-(x-x_n-h)^n}{\varphi'(x_n)} = 0\,.$$

Unsere Funktion η genügt also allen Grenz- und Stetigkeitsbedingungen der „erlaubten“ Variation δy, und es ist nach Voraussetzung (5):

$$\delta S = \int_a^b \Lambda\eta^{(n)}dx = \sum_{i=0}^{n}\int_{x_i-h}^{x_i+h}\Lambda\overline{\eta}_i^{(n)}dx = \sum_{i=0}^{n}\frac{n!}{\varphi'(x_i)}\int_{x_i-h}^{x_i+h}\Lambda(x)dx = 0\,, \qquad (9)$$

also, wenn man durch $n!2h$ dividiert und für $h=0$ zur Grenze übergeht:

$$\frac{\Lambda(x_0)}{\varphi'(x_0)} + \frac{\Lambda(x_1)}{\varphi'(x_1)} + \cdots + \frac{\Lambda(x_n)}{\varphi'(x_n)} = 0\,, \qquad (10)$$

die Funktion $\Lambda(x)$ wird also gemäß der zu (7) gemachten Bemerkung an einer *variabelen* Stetigkeitsstelle z. B. x_0 durch dasselbe Polynom

$$\Lambda(x) = u_1 + u_2x + \cdots + u_nx^{n-1}$$

dargestellt, dessen Koeffizienten schon durch die Funktionswerte an den n *festen* Stellen $x_1, \cdots, x_n$ bestimmt sind.

Zweiter Beweis.

Für jedes Polynom $g(x) = u_1 + u_2x + \cdots + u_nx^{n-1}$ und für die n^{te} Ableitung $\eta^{(n)}$ einer Funktion $\eta(x)$, deren $n-1^{\text{te}}$ Ableitung zwischen a und b stetig ist, ergibt sich durch partielle Integration:

562 |
$$\int_a^b g(x)\eta^{(n)}(x)dx = \Big[g\eta^{(n-1)} - g'\eta^{(n-2)} + \cdots \pm g^{(n-1)}\eta\Big]_a^b \mp \int_a^b g^{(n)}\eta\,dx\,, \qquad (11)$$

wo wegen $g^{(n)}(x)=0$ das letzte Glied verschwindet, und der ganze Ausdruck ist $=0$, wenn die Funktion $\eta = \eta(x)$ allen Bedingungen der „erlaubten Variation“ genügt, also mit ihren ersten $n-1$ Ableitungen an beiden Grenzen verschwindet.

Ist umgekehrt z eine integrierbare Funktion von x, welche für ein beliebiges Polynom $g(x)$ vom $n-1^{\text{ten}}$ Grade der Gleichung genügt

$$\int_a^b g(x)z\,dx = 0\,, \qquad (12)$$

is a polynomial of $(n-1)$ th degree in ξ, and hence, by (8),

$$\eta_n(x) = \frac{(x-x_0+h)^n - (x-x_0-h)^n}{\varphi'(x_0)} + \cdots + \frac{(x-x_n+h)^n - (x-x_n-h)^n}{\varphi'(x_n)} = 0\,.$$

Our function η thus satisfies all boundary and continuity conditions of the "admissible" variation δy, and, by assumption (5), we have

$$\delta S = \int_a^b \Lambda\eta^{(n)}dx = \sum_{i=0}^{n}\int_{x_i-h}^{x_i+h}\Lambda\overline{\eta}_i^{(n)}dx = \sum_{i=0}^{n}\frac{n!}{\varphi'(x_i)}\int_{x_i-h}^{x_i+h}\Lambda(x)dx = 0\,, \qquad (9)$$

and hence, by dividing by $n!2h$ and by taking the limit for $h=0$,

$$\frac{\Lambda(x_0)}{\varphi'(x_0)} + \frac{\Lambda(x_1)}{\varphi'(x_1)} + \cdots + \frac{\Lambda(x_n)}{\varphi'(x_n)} = 0\,. \qquad (10)$$

Thus, according to the remark made concerning (7), the function $\Lambda(x)$ is represented at a *variable* continuity point, e. g. x_0, by the same polynomial

$$\Lambda(x) = u_1 + u_2x + \cdots + u_nx^{n-1}$$

whose coefficients are already determined by the function values at the n *fixed* points $x_1, \cdots, x_n$.

Second proof.

For every polynomial $g(x) = u_1+u_2x+\cdots+u_nx^{n-1}$ and for the nth derivative $\eta^{(n)}$ of a function $\eta(x)$ whose $(n-1)$ th derivative is continuous between a and b, one obtains by partial integration

$$\int_a^b g(x)\eta^{(n)}(x)dx = \Big[g\eta^{(n-1)} - g'\eta^{(n-2)} + \cdots \pm g^{(n-1)}\eta\Big]_a^b \mp \int_a^b g^{(n)}\eta\,dx\,, \qquad (11)$$

where, on account of $g^{(n)}(x) = 0$, the last term vanishes and the entire expression is $= 0$, if the function $\eta = \eta(x)$ satisfies all conditions for the "admissible variation", and hence vanishes along with its first $n-1$ derivatives at both limits.

If, conversely, z is an integrable function of x that, for any polynomial $g(x)$ of the $(n-1)$ th degree, satisfies the equation

$$\int_a^b g(x)z\,dx = 0\,, \qquad (12)$$

so ist z immer die n^{te} Ableitung einer erlaubten Variation η. Denn man kann durch n-malige Integration von z zwischen den Grenzen a und x eine Funktion η bestimmen, die mit ihren ersten $n-1$ Ableitungen im Intervalle stetig ist und den *unteren* Grenzbedingungen genügt:

$$\eta(a) = 0, \cdots, \eta^{(n-1)}(a) = 0\,.$$

Es ist daher wegen (11) und (12):

$$\begin{aligned}\int_a^b gz\,dx &= \int_a^b g\eta^{(n)}dx \\ &= g(b)\eta^{(n-1)}(b) - g'(b)\eta^{(n-2)}(b) + \cdots \pm g^{(n-1)}(b)\eta(b) = 0\end{aligned}$$

für ein beliebiges Polynom g, d. h. für willkürliche Werte

$$g(b), g'(b), \cdots, g^{(n-1)}(b)\,.$$

Daraus folgt aber, daß auch

$$\eta(b) = 0, \eta'(b) = 0, \cdots, \eta^{(n-1)}(b) = 0$$

sein muß, d. h. $\eta(x)$ ist eine erlaubte Variation.

Ist nun $\Lambda = \Lambda(x)$ eine im Intervall $(a \cdots b)$ integrierbare Funktion, und wieder $g(x) = u_1 + u_2 x + \cdots + u_n x^{n-1}$, so ist das Integral

$$U = \int_a^b (\Lambda - g)^2\,dx = U(u_1, u_2, \cdots, u_n) \geqq 0 \tag{13}$$

eine wesentlich positive ganze Funktion zweiten Grades der sämtlichen Größen $u_1, u_2, \cdots, u_n$ und besitzt daher ein Minimum $\overline{U}$ für irgend ein im Endlichen liegendes Wertesystem $u_1 = \overline{u}_1,\ u_2 = \overline{u}_2, \cdots,\ u_n = \overline{u}_n$, oder für das Polynom:

$$g = \overline{g}(x) = \overline{u}_1 + \overline{u}_2 x + \cdots + \overline{u}_n x^{n-1}\,.$$

Es ist daher auch, wenn g wieder ein beliebiges Polynom $n-1^{\text{ten}}$ Grades und t einen von x unabhängigen Parameter bezeichnet,

$$U(t) = \int_a^b (\Lambda - \overline{g} - tg)^2\,dx \geqq \overline{U} = U(0)\,,$$

563 | d. h. die quadratische Funktion $U(t)$ hat ein Minimum an der Stelle $t = 0$, und es muß dort ihre Ableitung verschwinden:

$$-\frac{1}{2}U'(0) = \int_a^b (\Lambda - \overline{g})\,g\,dx = 0 \tag{14}$$

then z is always the nth derivation of an admissible variation η. For, by n-fold integration of z between the limits a and x, one can determine a function η that, along with its first $n-1$ derivatives, is continuous in the interval and satisfies the *lower* boundary conditions

$$\eta(a) = 0, \cdots, \eta^{(n-1)}(a) = 0\,.$$

Hence, on account of (11) and (12),

$$\int_a^b gz\,dx = \int_a^b g\eta^{(n)}dx$$
$$= g(b)\eta^{(n-1)}(b) - g'(b)\eta^{(n-2)}(b) + \cdots \pm g^{(n-1)}(b)\eta(b) = 0$$

for any polynomial g, i. e., for any values

$$g(b), g'(b), \cdots, g^{(n-1)}(b)\,.$$

But from this it follows that also necessarily

$$\eta(b) = 0, \eta'(b) = 0, \cdots, \eta^{(n-1)}(b) = 0\,,$$

i. e., $\eta(x)$ is an admissible variation.

If now $\Lambda = \Lambda(x)$ is a function integrable in the interval $(a \cdots b)$, and again $g(x) = u_1 + u_2x + \cdots + u_nx^{n-1}$, then the integral

$$U = \int_a^b (\Lambda - g)^2\,dx = U(u_1, u_2, \cdots, u_n) \geqq 0 \tag{13}$$

is an essentially positive entire function of the second degree of all magnitudes $u_1, u_2, \cdots, u_n$, and hence possesses a minimum $\overline{U}$ for some finite value system $u_1 = \overline{u}_1$, $u_2 = \overline{u}_2, \cdots$, $u_n = \overline{u}_n$, or for the polynomial

$$g = \overline{g}(x) = \overline{u}_1 + \overline{u}_2x + \cdots + \overline{u}_nx^{n-1}\,.$$

Hence, if again g is any polynomial of the $(n-1)$ th degree and if t denotes a parameter independent of x, also

$$U(t) = \int_a^b (\Lambda - \overline{g} - tg)^2\,dx \geqq \overline{U} = U(0)\,,$$

i. e., the quadratic function $U(t)$ has a minimum at the point $t = 0$, where its derivation must vanish:

$$-\frac{1}{2}U'(0) = \int_a^b (\Lambda - \overline{g})\,g\,dx = 0 \tag{14}$$

für willkürliche Koeffizienten des Polynoms $g(x)$.
Also ist nach dem oben Bewiesenen

$$\lambda(x) = \Lambda(x) - \overline{g}(x) = \eta^{(n)}(x)$$

die n^{te} Ableitung einer erlaubten Variation, und aus der allgemeinen Voraussetzung (5) unseres Satzes:

$$\int_a^b \Lambda \eta^{(n)} dx = 0$$

folgt für diese spezielle Variation:

$$\int_a^b \Lambda(x)\lambda(x)\, dx = \int_a^b (\overline{g} + \lambda)\, \lambda\, dx = 0\,,$$

oder, da man in (14) für g auch $\overline{g}$ einsetzen kann:

$$\int_a^b \lambda \overline{g}\, dx = 0$$

weiter:

$$\int_a^b \lambda^2 dx = \int_a^b \left(\Lambda(x) - \overline{g}(x)\right)^2 dx = 0\,, \tag{15}$$

und hieraus in bekannter Weise jedenfalls für alle Teilintervalle, in denen $y^{(n)}$ und damit auch Λ stetig ist:

$$\lambda(x) = \Lambda(x) - \overline{g}(x) = 0$$

oder, wie behauptet:

$$\Lambda(x) = \overline{u}_1 + \overline{u}_2 x + \cdots + \overline{u}_n x^{n-1}\,.$$

Da das Polynom $\overline{g}(x)$, das den Ausdruck (13) zu einem Minimum macht, nichts anderes ist als die beim n^{ten} Gliede abgebrochene Entwicklung der Funktion $\Lambda(x)$ nach den *Kugelfunktionen* des Intervalles, so kann man dem zweiten Beweise auch die folgende Form geben, die ich der freundlichen Mitteilung des Herrn *Erhard Schmidt* verdanke.

Für ein Polynom $n - 1^{\text{ten}}$ Grades kann man immer ansetzen:

$$\begin{aligned}\overline{g}(x) &= c_0 + c_1 \frac{d}{dx}(x-a)(b-x) + c_2 \frac{d^2}{dx^2}(x-a)^2(b-x)^2 + \cdots \\ &\qquad\qquad \cdots + c_{n-1}\frac{d^{n-1}}{dx^{n-1}}(x-a)^{n-1}(b-x)^{n-1} \\ &= c_0 + c_1 P_1(x) + c_2 P_2(x) + \cdots + c_{n-1} P_{n-1}(x)\end{aligned}$$

for any coefficients of the polynomial $g(x)$.
Thus, according to what was proved above,

$$\lambda(x) = \Lambda(x) - \overline{g}(x) = \eta^{(n)}(x)$$

is the nth derivative of an admissible variation, and from the general assumption (5) of our theorem:

$$\int_a^b \Lambda \eta^{(n)} dx = 0$$

it follows for this particular variation that

$$\int_a^b \Lambda(x)\lambda(x)\, dx = \int_a^b (\overline{g} + \lambda)\, \lambda\, dx = 0\,,$$

or, since we can also substitute $\overline{g}$ for g in (14):

$$\int_a^b \lambda \overline{g}\, dx = 0$$

furthermore:

$$\int_a^b \lambda^2 dx = \int_a^b (\Lambda(x) - \overline{g}(x))^2\, dx = 0\,, \tag{15}$$

and from this in familiar fashion, at least for all partial intervals in which $y^{(n)}$, and hence also Λ, is continuous, that

$$\lambda(x) = \Lambda(x) - \overline{g}(x) = 0$$

or, as asserted,

$$\Lambda(x) = \overline{u}_1 + \overline{u}_2 x + \cdots + \overline{u}_n x^{n-1}\,.$$

Since the polynomial $\overline{g}(x)$, which turns the expression (13) into a minimum, is but the expansion of the function $\Lambda(x)$, aborted at the nth term, with respect to the *spherical functions* of the interval, it is possible to render the second proof also in the following form, which Mr. *Erhard Schmidt* was kind enough to communicate to me.

For a polynomial of $(n-1)$ th degree it is always possible to set

$$\begin{aligned}\overline{g}(x) &= c_0 + c_1 \frac{d}{dx}(x-a)(b-x) + c_2 \frac{d^2}{dx^2}(x-a)^2(b-x)^2 + \cdots \\ &\qquad \cdots + c_{n-1}\frac{d^{n-1}}{dx^{n-1}}(x-a)^{n-1}(b-x)^{n-1} \\ &= c_0 + c_1 P_1(x) + c_2 P_2(x) + \cdots + c_{n-1} P_{n-1}(x)\end{aligned}$$

564 | und die Koeffizienten $c_0, c_1, \cdots, c_{n-1}$ so bestimmen, daß die Funktion $\Lambda - \overline{g}(x)$ die n^{te} Ableitung einer erlaubten Variation $\delta y = \eta(x)$ darstellt. Hierzu ist nämlich nur erforderlich, daß das einfache, zweifache bis r-fache zwischen den Grenzen a und b genommene Integral der Funktion verschwindet. Man erhält also durch $r+1$-malige Integration:

$$\int_a^b dx \int_a^x (\Lambda - c_0 - c_1 P_1 - \cdots - c_{r-1} P_{r-1})\, dx^r - c_r \int_a^b (x-a)^r (b-x)^r dx = 0\,,$$

wo alle Glieder mit $c_{r+1}, c_{r+2}, \cdots, c_n$ fortfallen, weil die $r+1$-fachen Integrale über die Funktionen $P_{r+1}, P_{r+2}, \cdots, P_{n-1}$ auf Ableitungen von $(x-a)^{r+s}$ $(b-x)^{r+s}$ führen und daher zwischen den Grenzen a und b sämtlich verschwinden. Unsere Gleichung dient also, wenn die vorhergehenden Koeffizienten $c_0, c_1, \cdots, c_{r-1}$ bekannt sind, zur Bestimmung des Koeffizienten c_r, der hier mit einem von 0 verschiedenen Integral multipliziert erscheint. Somit lassen sich durch successive Berechnung der Koeffizienten alle n Integralgleichungen erfüllen, und aus der Annahme

$$\int_a^b \Lambda \eta^{(n)} dx = 0 \quad \text{für} \quad \eta^{(n)} = \Lambda - \overline{g}(x)$$

folgt genau wie oben

$$\int_a^b (\Lambda - \overline{g})^2 dx = 0$$

und schließlich:

$$\Lambda = \overline{g} = c_0 + c_1 P_1(x) + \cdots + c_{n-1} P_{n-1}(x)\,,$$

dieser Beweis führt also zu einer direkten Darstellung der Funktion $\Lambda(x)$ als Polynom $n-1^{\text{ten}}$ Grades.

Göttingen, den 1. Dezember 1903.

and to determine the coefficients $c_0, c_1, \cdots, c_{n-1}$ so that the function $\Lambda - \overline{g}(x)$ represents the nth derivative of an admissible variation $\delta y = \eta(x)$. For this it is only necessary that the simple, double, up to the r-fold integral of the function taken between the limits a and b vanish. Hence, by $r+1$-fold integration, one obtains

$$\int_a^b dx \int_a^x (\Lambda - c_0 - c_1 P_1 - \cdots - c_{r-1} P_{r-1})\, dx^r - c_r \int_a^b (x-a)^r (b-x)^r dx = 0\,.$$

Since the $r+1$-fold integrals over the functions $P_{r+1}, P_{r+2}, \cdots, P_{n-1}$ lead to derivatives of $(x-a)^{r+s}$ $(b-x)^{r+s}$, all terms with c_{r+1}, $c_{r+2}, \cdots, c_n$ drop out here, and hence all vanish between the limits a and b. So, if the previous coefficients $c_0, c_1, \cdots, c_{r-1}$ are known, our equation serves the determination of the coefficient c_r, which occurs here multiplied by an integral different from 0. Thus, by successive calculation of the coefficients, all n integral equations can be satisfied, and from the assumption

$$\int_a^b \Lambda \eta^{(n)} dx = 0 \quad \text{for} \quad \eta^{(n)} = \Lambda - \overline{g}(x)$$

it follows, just as above, that

$$\int_a^b (\Lambda - \overline{g})^2 dx = 0$$

and eventually

$$\Lambda = \overline{g} = c_0 + c_1 P_1(x) + \cdots + c_{n-1} P_{n-1}(x)\,.$$

Hence, this proof leads to an immediate representation of the function $\Lambda(x)$ as a polynomial of $(n-1)$ th degree.

Göttingen, December 1, 1903.

Introductory note to Hahn and Zermelo 1904

Rüdiger Thiele

By *extremal thinking* we mean a way of thinking that in considering certain scientific observations, selects from the possible cases of an observed phenomenon the exceptional cases that exhibit a particular property to the greatest or least degree. Such extremal thinking has long been a well-established property of human thought, and one might even say with David Hilbert that it "has been a lodestar to the inquiring mind". But it was only with the development of infinitesimal mathematics beginning in the eighteenth century that extremal problems in mathematics and the sciences could be subjected to a systematic treatment using the classical calculus of variations.

The calculus of variations is generally considered to have been born with Johann Bernoulli's posing of the brachistochrone problem in 1696. It took almost a century until Leonhard Euler, in his paper "Methodus nova et facilis calculum variationum tractandi" (1771), showed that beyond differentiation, no new transcendent operation of "variation" had been brought into being by the calculus of variations, but that instead, variation could be reduced to differentiation. Another century passed before finally, in 1887, in the early development of functional analysis, Vito Volterra realized the typical functional of the calculus of variations as a function of functions (or curves) and thereby cleared the way for a new approach to the calculus of variations.

In the nineteenth century, applications of the calculus of variations in physics, in particular to mechanics (Joseph Louis de Lagrange, Carl Gustav Jacobi, William Rowan Hamilton), led to notable discoveries. But aside from these, there were many formal and discursive investigations carried out in the calculus of variations (Martin Ohm, Georg Wilhelm Strauch), with the result that this discipline appeared to be moving in two different directions.

Euler's reduction of variation to differentiation had the unintended effect of encouraging the treatment of questions in the calculus of variations through analogous methods of the differential calculus, such as the idea of studying extremal properties of functionals by studying the second variation, in analogy with the study of the second derivatives of functions. From the functional-analytic point of view, it is clear today that one will never achieve a satisfactory result by extending the notion of function from the differential calculus to that of functionals. How, for example, does one extend a neighborhood of a function (curve) with respect to the notions of distance in the n-dimensional space $\mathbb{R}^n$, all of which are equivalent?

The calculus of variations was about two hundred years old during Zermelo's student years; during the nineteenth century, its significance and applicability were judged in different ways at different times and in different places. The approach to the calculus of variations in Berlin was more stable,

H.-D. Ebbinghaus, A. Kanamori (Eds.), *Ernst Zermelo - Collected Works/ Gesammelte Werke II*, DOI 10.1007/978-3-540-70856-8_10, Schriften der Mathematisch-naturwissenschaftlichen Klasse der Heidelberger Akademie der Wissenschaften 23, © Springer-Verlag Berlin Heidelberg 2013

being associated with the names Carl Gustav Jacobi, Karl Weierstrass, and Hermann Amandus Schwarz. Weierstrass remarked in a celebrated lecture on the calculus of variations given in the summer semester 1879 that the character of an extremum in a variational problem depended on the notion of neighborhood that was used for the admissible comparison function. Zermelo of course had not heard Weierstrass's lecture, but on the one hand, there were quite a few transcripts and elaborations of Weierstrass's lectures, and on the other hand, Zermelo attended lectures on the calculus of variations by Weierstrass's student Schwarz (summer semester 1892).

Zermelo devoted the greater part of his dissertation *1894* to possible definitions of neighborhood that depart from the usually employed metrics. Two notions of neighborhood have survived to this day, that of the C^0 norm and that of the C^1 norm (for strong and weak extrema, following Adolf Kneser). At the end of his dissertation (pp. 77–97), Zermelo considered sufficient conditions for (strong) extrema (C^0 norm) using the notion of field introduced by Weierstrass (the latter's "surface bands"). When in 1897 Zermelo moved to Göttingen, where in 1900 David Hilbert began to take an interest in the calculus of variations, he was therefore quite up-to-date.

1 Prehistory and setting

The working out of the notion of analytic function (functions of numbers) and the emergence of variational problems in an analytic setting (functions of functions = functionals) occurred simultaneously at the beginning of the eighteenth century, so it makes sense that at the time, the two subjects were not clearly differentiated and that at first, attempts were made to translate methods that had proved successful in the treatment of functions of differential calculus to the solution of variational problems. In this regard, Hilbert commented:

> [...] if we call the calculus of variations the differential calculus of functions.[1]

As an example of such a variational problem let us make a simple integral J that depends on one unknown function $y(x)$ into an extremum through a suitable choice of this function,

$$J(y) = \int_a^b f(x, y, y')dx \;\; \rightarrow \;\; \text{extremum},$$

where

$$f(x, y, p) \in C^2 \quad \text{(that is, twice continuously differentiable)}.$$

[1] "[...] wenn wir die Variationsrechnung die Differentialrechnung der Funktionen nennen." Transcript of the lecture "Gewöhnliche Differentialgleichungen" ("Ordinary differential equations"), summer semester 1912, p. 138, Mathematisches Institut der Universität Göttingen.

For the admissible comparison functions $y(x)$, certain boundary values are prescribed, and perhaps there are some side conditions for such functions as well. The admissible comparison functions $y(x)$ belong to the space $C^2[a,b]$ of twice continuously differentiable functions defined on the interval $[a,b]$. Here extremality for strong extremals is understood in the sense of the space $C^0[a,b]$, that is, with the metric

$$|y(x) - y_0(x)| < \varepsilon \quad \text{for all } x \in [a,b]\,.$$

For fixed boundary conditions, that is, for $y(a) = A$ and $y(b) = B$, as well as in the case of no side conditions, the variational problem as just described is frequently called a classical variational problem or standard problem.

No matter how a variational problem is defined, not only does the mapping itself play a role (in the classical case realized by a simple integral J, that is, by a functional), but in a complete formulation of the problem, the set of admissible comparison functions (comparison curves) and the metric assigned to them (as a rule, in the sense of C^0 or C^1) are also involved. Moreover, the set of admissible comparison functions is characterized by analytic conditions (differentiability), by boundary conditions, and possibly by side conditions, obstacle conditions, and so forth. In other words, the formulation of a variational problem depends on the character of the exercise or practical problem to be solved.

It is such an aspect of a variational problem that interested Zermelo in his dissertation, since he wished to understand a central point of variational problems, namely the suitable form of an extremum in a variational problem. His first assertion in the dissertation (*1894*, 98) reads as follows:

> It is necessary to place greater emphasis than previously on the precise definition of the maximum or minimum in the calculus of variations.

Although a variational problem formulated for functions requires some alterations when formulated for curves, we shall, following Hahn and Zermelo, speak for the sake of geometric understanding also of admissible curves, the solution curve, and so forth (in particular, the geometric point of view is helpful in the case of critical and conjugate points). Among the assumptions made, possible solutions $y = y^0(x)$ of the Euler-Lagrange equations in $[a,b]$ satisfy

$$\frac{d}{dx} f_p(x, y, y') - f_y(x, y, y') = 0\,.$$

A one-parameter family $y = y(x, \alpha)$ of solutions of this differential equation (the *extremals*) will be called, following Kneser, a field or extremal field if $\partial y/\partial\alpha \neq 0 \in (a,b]$ (one unknown function).

The attempt to formulate sufficient conditions for extrema in the calculus of variations using the second variation $\delta^2 J$ of the variational problem

$$J(x) \;\rightarrow\; \text{minimum (maximum)}$$

was inspired by the possibility of establishing the extreme value of a function F using assertions about its second derivative ($F'' > 0$). Such attempts, despite the clarifying 1879 lecture by Weierstrass, persisted well into the years in which Zermelo was a student. Hans Hahn's teacher Gustav von Escherich attempted all his life in vain to obtain sufficient conditions for strong extrema via the second variation. Weierstrass (1879) and Kneser (*1900*) introduced for this purpose the notion of field.

When Zermelo arrived in Göttingen in 1897, one may assume that he soon became acquainted with Hilbert's abstract definition of the calculus of variations, which he used in his lecture on the calculus of variations (in the summer semester 1899) and which even today seems strikingly modern:

> Given is a set of mathematical objects. With each of them is associated in a certain given way a real number. One would like to select the object or objects with which the largest or smallest number is associated [...] if such a number exists.[2]

From this general point of view, one may formulate a variational problem in a functional-analytic way as follows: A variational problem consists of determining, given a (real) functional $J(y)$ defined on a set of comparison elements y, an extremum (maximum or minimum):

$$J(y) \;\rightarrow\; \text{extremum on } V.$$

The variational problem is determined by boundary, side, or obstacle conditions as well as analytic requirements, so that the extremum is sought on an algebraic structure $V_0 \subseteq V$. A solution $y^0(x)$ yields

$$J(y) \leq J(y^0(x)) \quad \text{or} \quad J(y) \geq J(y^0(x)) \quad \text{for all } y \in V_0\,,$$

according to which a maximum or minimum is sought. In a larger sense, contained in a variational problem are the question of existence of a solution and a discussion of its uniqueness. The classical calculus of variations uses as a functional, as a rule, the variational integral

$$J(y) = \int_G F\big(x, y(x), Dy(x)\big)dx \quad \big(x = (x_1, \ldots, x_\alpha),\ y = (y_1, \ldots, y_n)\big)\,,$$

where G is an open domain in $\mathbb{R}^n$, $y(x)$ a function defined in G, and Dy the first (partial) derivatives of y.

Weierstrass in his 1879 lecture, edited and distributed (but not authorized by Weierstrass) by the Berlin Mathematical Society, had taken a new tack and

[2] "Gegeben sind irgendwelche mathe. Dinge. Jedem ist in bestimmter geg. Weise eine reelle Zahl zugeordnet. Man soll das Ding oder solche Dinge heraussuchen, denen die kleinste oder größte Zahl zugeordent ist [...] sofern diese Zahl existiert." Draft of the lecture "Variationsrechnung" ("The Calculus of Variations"), summer semester 1899, Universitätsbibliothek Göttingen, Cod. Ms. D. Hilbert 555, p. 1.

replaced study of the second variation $\delta^2 J$ with the introduction of a surface strip, since it seemed impossible to obtain sufficient conditions for strong extrema (C^0 norm). What determines strong extrema is that in the associated metric for the comparison curves, only the coordinates of the curves, and not the direction of the tangents, play a role. A fixed sign of the total variation ΔJ for all comparison functions y yields sufficient conditions, that is, criteria follow from

$$\Delta J = \Delta J(y^0, \xi) = J(y) - J(y^0) = \delta J(y^0)\xi + \delta^2 J(y^0, \xi)\xi^2 + \cdots \quad (y - y^0 = \xi),$$

or since for extrema y^0, the associated variation $\delta J(y^0)$ vanishes, the criteria follow from

$$\delta J = \delta^2 J(y^0, \xi)\xi^2 + \cdots .$$

To the extent that in a variational problem there are no restrictions placed on the derivatives of the functions (strong extremum), the definiteness of $\delta^2 J(y^0, \xi)$ (respectively, of ΔJ) can be determined only from the neighborhood of the coordinates of y and y^0, that is, not from the Taylor expansion, since its validity in the space C^1 follows (respectively, assumes neighboring derivatives). Neighboring derivatives (velocities in mechanics) are useful in many physical problems. However, as a rule, geometric problems will require the admissibility of all directions for comparison curves (that is, they must allow vertical tangents if required).

It is well known that Weierstrass published his results only reluctantly, and in the case under consideration here not at all, so that his findings were promulgated only through transcripts of his lectures. Volume 7 of his collected works (*1927*), in which can be found a melange of his lectures on the calculus of variations, came too late to have any influence on the development of the subject.[3] Textbooks that contained Weierstrass's new direction did not exist in 1900; they appeared later, first *Kneser 1900*, then *Bolza 1904* in English, considerably expanded in German (*Bolza 1909*), and *Hadamard 1910*. The encyclopedia articles on the calculus of variations (*Kneser 1904* and *Hahn and Zermelo 1904*) were of particular import, as were their French translation and expansion by Maurice Lecat (*1913*). Additional orienting survey articles were written by William Osgood (*1900/01*).[4]

On a hike in the Harz Mountains in 1895 Felix Klein, Franz Meyer, and Felix Müller devised a plan to write a comprehensive reference work on the mathematics of the nineteenth century, a plan that was realized as *Ency-*

[3] See the discussions by Carathéodory in *1928, 1928/29,* and *1929*.

[4] William Fogg Osgood, who learned at first hand from studies in Germany and resulting contacts of the new ideas of the calculus of variations (Kneser), worked for a while on the calculus of variations and wrote a very instructive survey article, "Sufficient conditions in the calculus of variations," *1900/01*. Later surveys were written by Gilbert Bliss (*1920*), Arnold Dresden (*1926*), Hans Hahn (*1927*), and Byron Cosby (*1937*).

klopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen (*Encyclopedia of the mathematical sciences with the inclusion of their applications*). The first volume of the planned seven volumes appeared in seven installments between 1898 and 1901, with the completion of the entire work coming in 1935.

From the beginning, an attempt was made to update papers that had already appeared. Kneser submitted the article *1904* on the calculus of variations in 1900; in 1903 it was in press, and the publishing academies wanted to bring Kneser's article up to date during their session at the end of September 1903 in Göttingen. The academies were represented by Felix Klein (Göttingen), Walter von Dyck (Munich), and Ludwig Boltzmann (Vienna), with Boltzmann substituting for Gustav von Escherich. In the relevant report of the Göttingen academy, we read the following:

> Volume II/1. A supplement to article 8 (Variationsrechnung by A. Kneser) taking account of recent developments appears to be needed.[5]

The brief text gives no information as to who desired the updating, but it clearly shows that Kneser's article was seen as incomplete and not at the level of current knowledge. However, an appendix to the report anticipates the publication of Kneser's article with additions by 1904.

Although in 1900 Kneser was unusually busy completing his textbook and the encyclopedia article, it is noteworthy that he neither incorporated nor even mentioned the new supporting and trend-setting ideas of his textbook in the encyclopedia article, and it is therefore no wonder that the lack of new ideas was rightly seen by the publisher as a failing. A possible author was not yet mentioned by name. In the report of the publishing academies, which provided an overview of the planned contributions through August 1, 1904, there was therefore for the printing of volume II/1 also included, in addition to Kneser's contribution (which was already available separately), an additional article.

The next meeting of the academies, in April 1904, led to a memorandum in which now the publisher states that it will take responsibility for the continuation of Kneser's article:

> In book 5 of the first half of volume II, the end of the article by Sommerfeld (boundary value problems) and the article by Kneser (calculus of variations) were printed some time ago. The article Zermelo-Hahn, further developments of the calculus of variations in recent years, and the article by Burkhardt, trigonometric interpolation [...]

[5] "Zum Artikel 8 (Variationsrechnung: A. Kneser) erscheint eine Ergänzung, die dem neuesten Standpunkt Rechnung trägt, als notwendig." Göttinger Akademie der Wissenschaften, Archivnr. Scient 305, 1, Vol. 1, Nr. 37.

> are set in type. Presumably, they will come off the press in a few days.[6]

The date of submission of the article *Hahn and Zermelo 1904* in volume II/1 was given as January 1904 (p. 9), as also noted at the end of the article. It follows, then, that both authors worked expeditiously (October 1903 to January 1904). The incompleteness of *Kneser 1904* is noteworthy, for in 1900, Kneser not only completed this article (released in autumn 1904), but also saw the publication of his *Lehrbuch der Variationsrechnung* (*A course in the calculus of variations*) (*Kneser 1900*) in March 1900.

In 1935, Carathéodory remarked on Kneser's textbook, "The methods of Weierstrass were initially disseminated only by word of mouth; they became available to a wider public through Zermelo's dissertation [...] and then primarily through the work of Kneser."[7] Kneser's book was a milestone in the history of the calculus of variations, and in a review of the book, Paul Stäckel rightly wrote,

> Following such a unified plan, to produce such a presentation based on unified principles, was no mean task [...] and much was left to the author's own initiative. Therefore, what he has achieved is far more than a textbook in the usual sense of the word. It is [...] at once a contribution to the further development of the calculus of variations that will be of lasting value.[8]

Kneser has also introduced into the calculus of variations a number of well-chosen terms such as "extremal," "neighborhood in the wide and narrow senses" with the corresponding "weak and strong extrema," and "field." As a result, an attempt was made to convince Kneser himself to supply a report on the current state of the calculus of variations or a continuation of his article, but Kneser was unwilling, even though he would have been the most suitable candidate for the task.[9]

---

[6] "Von Heft 5 des ersten Halbbandes ist der Schluß des Artikels Sommerfeld (Randwertaufgaben) sowie der Artikel von Kneser (Variationsrechnung) seit längerer Zeit gedruckt. Der Artikel Zermelo-Hahn, Ergänzungen zur Variationsrechnung, sowie der Artikel Burkhardt, trigonometrische Interpolation [...] sind umgebrochen; voraussichtlich werden sie in den nächsten Tagen druckfrisch erklärt." Editors' report on the state of publication of *Encyklopädie der mathematischen Wissenschaften*, April 22/23, 1904, p. 1.

[7] *Carathéodory 1935*, 398.

[8] *Archiv der Mathematik und Physik*, 3rd series, **2** (1902), 185–189.

[9] One can only speculate on his reasons, but it seems that it was a dispute with Hilbert that tipped the scales. First of all, Hilbert's and Kneser's views on the foundations of geometry were not quite in agreement, which led, for example, to irrational responses in articles on the developing topic of integral equations. Secondly, there arose professional disagreements between Hilbert and Kneser regarding the calculus of variations, a visible manifestation of which can be seen in Kneser's insistence on his contribution to Hilbert's Paris lecture "Mathematische

Thus Zermelo and Hahn came under consideration. Zermelo was well acquainted with the most recent developments in the calculus of variations through his dissertation and had been charged in the Göttingen Mathematical Society, in 1900 at its 16th session, to give a talk on Kneser's recently published book, with which he was therefore familiar. Also, Zermelo had given a first lecture on the calculus of variations in Göttingen in 1902; in 1907, he again lectured on this topic. Hahn, who obtained his doctorate in 1903 in Vienna under the direction of von Escherich with the dissertation *Zur Theorie der zweiten Variation einfacher Integrale* (*On the theory of the second variation of simple integrals*) (*1903a*), remained in Göttingen on account of a stipend in the winter semester 1903/04. As he himself wrote in a biography for the Vienna Academy on June 21, 1921, he was particularly influenced by Hilbert's lectures and seminars. He must have been referring to the four-hour lecture on partial differential equations in addition to the associated tutorials. In December 1903, Hahn lectured in Göttingen on the work of von Escherich and on his own work in the calculus of variations; he had already published on the calculus of variations (3 papers) and analysis (2 papers), so that he was also seen as an additional author for the group in Göttingen (in addition to Hilbert, in particular for Klein).

It seems that both young mathematicians accepted the offer at once and rapidly completed *Hahn and Zermelo 1904*, a work of 15 pages with 65 footnotes.[10] As mentioned above, the article was submitted in January 1904. Both Zermelo (33 years old) and Hahn (25 years old) quickly became known in the mathematical world after their overview appeared in such a prominent publication.

2 The paper

The article by Zermelo and Hahn is divided into six sections:

1. The foundations of *Weierstrass*'s theory
2. Necessary and sufficient conditions in the simplest case
3. Isoperimetric problems
4. More general problems
5. Examples and applications
6. Existence questions

Probleme" (*1900a*), which ultimately led to the footnote inserted by Hilbert when it was already in press. See, for example, *Thiele 2007*, 387ff., 441. Also, a renewed offer from Hilbert to Kneser in a letter of February 19, 1909, to bring *Hahn and Zermelo 1904* up to date was again declined by Kneser (Universitätsbibliothek Göttingen, Cod. Ms. Hilbert 180).

[10] Compare with Kneser's article *1904*: 155 pages with 198 footnotes.

Remark on problem representation. The analytic formulation of a variational problem depends, as mentioned above, on the nature of the problem. In geometric variational problems, in which curves appear, parametric representations will as a rule come into play, since every direction must be admissible, which need not be the case for a problem in mechanics. If we call these two types of formulation *parameter* and *function* problems, the following historical remark can be made: Weierstrass had organized his lectures on the calculus of variations completely around parameter problems, while Hilbert, in his lectures, developed the calculus of variations for functionally formulated problems. Carathéodory, in his textbook *1935*, kept the two problem types strictly separated and gave them parallel treatments, where from Carathéodory's point of view, strong extrema for which all directions are allowed are possible only for parameter problems. For a discussion of the relationship between the two problem types, see *Giaquinta and Hildebrandt 1996*, vol. II, 166ff.

The methods of the calculus of variations are easier to explain in the functional formulation, since for the parametric formulation, one has to take the parametric representations into account in the development of the theory. The geometric point of view requires, for example, that a change in the parametric representation of a curve not change the curve and that additional technical requirements be satisfied, which has nothing directly to do with variational techniques and keeps from full view the main ideas of the methods.

For these reasons, Zermelo and Hahn chose the functional formulation, which in any case was how it was done in Göttingen. Kneser preferred the parametric formulation. Furthermore, at the end of his teaching career, even Weierstrass had formulated the plan of writing a textbook on the calculus of variations, which, for the sake of simplicity and understandability, would be based throughout on the functional representation. The details on this are contained in a letter to Schwarz of March 14, 1885.[11]

2.1 The foundations of *Weierstrass*'s theory

The further development of the calculus of variations proceeded mainly from Weierstrass's theory, whose central point was the total variation $\Delta J = J(y) - J(y^0)$, that is, the integral difference between an arbitrary comparison function and the extremals, to be expressed in a field by an integral over the so-called excess function E.

This section presents Kneser's and Hilbert's approaches to this result of Weierstrass. Kneser developed Weierstrass's theory in geometric form; more precisely, he passed from Weierstrass's surface bands to the extremal field and arrived thereby at general geodesic coordinates (a field of extremals and associated transversals, a generalization of Gauss's othogonality theorem for

[11] Archiv der Berlin-Brandenburgischen Akademie der Wissenschaften, Nachlass Hermann Amandus Schwarz 1175, 1224; see also *Thiele 2007*, 251ff.

geodesic lines). Hilbert placed his independence integral J^* at the center, which caused the integrability conditions for the path-independence of J^* to come into the foreground, yielding a partial differential equation for transversal directions (see below).

Later, these two methods were joined by a third direction, which Carathéodory (*1935*) brought into final form in only a few lines, namely the equivalent variational problem.[12] Footnote 25 to the encyclopedia article, which rests on a verbal communication from Carathéodory, is related to the beginning of this development, wherefrom for Carathéodory the full scope of his insight became clearer and clearer over time.

Let us summarize here these three approaches. To simplify the presentation, let us consider the standard problem, that is, we set, in particular, fixed boundary conditions for the comparison functions. With suitable modifications, for simple integrals, the results hold as well for a number of unknown functions; with the field concept, several alterations are necessary—in particular, multidimensional problems require new ideas.

(1) Field of extremals (Weierstrass-Kneser)

There exists a field of extremals $\mathbb{F} : y = y(x, \alpha)$ into which the extremal under consideration $C_0 : y = y^0(x)$ for $\alpha = \alpha_0$ is embedded. For comparison functions in the field $\mathbb{F}$, one has

$$\Delta J = \Delta J(y^0, \xi) = J(y) - J(y^0) = \int_a^b E(x, y, y')dx\,,$$

where E is the excess function. The extremals of the field $\mathbb{F}$ are intersected transversally by an additional family of curves; the two families of curves form a curvilinear coordinate system with coordinates α and u, where u is defined by $u = \int_{x_0}^{x} f(x, y)dx$; for information on the transversal family, see (2).

(2) Independence integral (Hilbert)

Hilbert formally introduced a path-independent integral J^*,

$$J^* = \int_{P_0}^{P} \big[f(x, y, p) - pf_p(x, y, p) + y' f_p(x, y, p)\big]dx\,.$$

[12] Carathéodory worked for about 30 years on this approach before he was able to write it down in three lines. He was inspired by a lecture in Göttingen given by Hahn, in which Hahn mentioned an "unsolvable" variational problem that Carathéodory made solvable by extending the notion of solvability (broken extremals) and built this idea into a dissertation (*1904*). In an appendix to his dissertation, there appears a method developed by Johann Bernoulli that is used for a sufficiency proof. This historical discovery of Carathéodory's then became the basis for his later equivalence ideas (see *Carathéodory 1935*, §§227ff., *Thiele 2007*, sections 2.3 and 2.4).

Here, using a suitable slope function $p = p(x, y)$, the independence integral J^* is constructed from J in such a way that

$J^* = 0$ on transversals,

$J^* = J$ on extremals (in particular on $C_0 : y = y(x, 0) = y^0(x)$).

Then in the independence domain of J^*, one has

$$\Delta J = J(C) - J(C_0) = J(C) - J^*(C) = \int_a^b E(x, y, y')dx\,.$$

In embryonic form, this idea appears first in the work of Johann Bernoulli (1716).

If one considers the independence integral J^*, where the path of integration is inside a field $\mathbb{F}$ (defined by the slope function $p = p(x, y)$) and can run from a fixed point P_0 to an arbitrary point P, then J^* is a well-defined function of the coordinates x, y of this point P:

$$J^* = J^*(P) = J^*(x, y)\,.$$

The curves $J^*(x, y) = \text{const}$ form the transversal family to the extremals of the field $\mathbb{F}$ (the transversals), and their differential equation is

$$\frac{\partial J^*}{\partial x}dx + \frac{\partial J^*}{\partial y}dy = 0,$$

and with slope $\tilde{y}'$,

$$f(x, y, p) + (\tilde{y}' - p)f_p(x, y, p) = 0\,.$$

The direction $\tilde{y}'$ is said to be transversal to the direction $p(x, y)$ of the extremals in the field.

Gauss's theorem on geodesics is generalized by Kneser's transversal theorem: two transversals cut off arcs in the field on the extremals with the same value of the variation integral.

(3) Equivalent variational problems (Carathéodory)

Carathéodory began with the simple idea that for a variational problem

$$I = \int_a^b f dx \to \text{minimum}$$

with a Lagrangian function f such that

$f \geq 0$ for all admissible comparison functions,

$f = 0$ for the function $y = y^0(x)$,

$y^0(x)$ is a solution (and thus in the case of twofold differentiability is also an extremal, or a solution of the Euler-Lagrange differential equation). The

properties just given cannot be expected in general, but one can try, using an altered Lagrangian function f^*, to transform the variational problem $I \to$ minimum in such a way that on the one hand, the extremals for both problems are the same (equivalence of problems) but on the other hand, the altered Lagrangian function f^* of the new variational problem exhibits the above properties of sufficiency. Such requirements lead, for a new Lagrangian function $f^* = f - ds(x, y(x))/dx$, to a partial differential equation for the function $s = s(x, y)$:

$$s_x + H(x, y, s_y) = 0$$

$$(H = -f + pf_p \text{ the Hamiltonian function associated with } f),$$

where the curves $s(x, y) = \text{const}$ are the transversals to the extremals.

Compared with (2) and (3), the point of view is inverted with Carathéodory: he begins with the transversals and then determines a family whose elements must be extremals. The differences between the various starting points are clearly seen in the field of optics: Carathéodory begins with the wave surfaces (transversals), while the other methods begin with the light rays (extremals). For Carathéodory, a field consisting of extremals and transversals (s-functions) forms a complete picture. For simple variational problems there is an n-parameter family of extremals that intersects a one-parameter family of s-surfaces transversally; for one unknown function, however, every extremal field locally leads to a complete picture, while for several unknown functions, integrability conditions (Lagrange bracket equal to zero) must be satisfied for the transversal family or a field to exist. In optics, the s-surfaces of the transversal family are also called the optic distance, since they cut off arcs on the light rays that are traversed by the light in the same amount of time.

2.2 Necessary and sufficient conditions in the simplest case

The Weierstrass representation of the total variation

$$\Delta J = J(C) - J(C_0) = \int_a^b E(x, y, y')dx$$

using the excess function E makes it possible, especially for strong extrema (in the C^0 norm), to formulate, using the sign of the excess function, additional necessary and sufficient conditions. Weierstrass's representation assumes the existence of a suitable field that embeds the extremals in question. On the assumption of suitable analytic conditions (as with the standard problem), the field construction underlying the criterion can always be carried out locally. How far it can be carried out along the extremals is governed by the Jacobi criterion. The criteria (that is, the field construction, to begin with) can also be extended to more general variational problems such as those with free boundary values.

This section of the article does not mention the discovery by Zermelo in 1894 that the Legendre criterion

$$f_{pp}(x, y^0(x), y^{0\prime}(x)) \geq 0$$

determining the expression f_{pp} is the quadratic term in the Taylor expansion of the excess function (*Zermelo 1894*, 60f.). A definite sign in the Legendre criterion allows one to conclude only a weak extremum. However, if one requires in the field, for arbitrary p,

$$f_{pp}(x, y, p) > 0\,,$$

then the positivity (negativity) of the excess function E is ensured as well. Moreover, this form of the criterion has technical advantages, since no knowledge of the slope function $p = p(x, y)$ is required. However, the criterion

$$E(x, y, p) > 0$$

is more far-reaching. The extension of the criterion to n unknown functions of a variable requires that one consider

$$\sum_{i,j=1}^{n} f_{p_i p_j} \zeta^i \zeta^j$$

and the associated altered excess function.

2.3 Isoperimetric problems

In the analytic formulation of a variational problem that we have called a standard problem, fixed boundary values of the comparison function and boundary conditions ensure a simple formulation of the problem. The requirements on a variational problem can be expanded, however. For example, one could additionally require of the admissible comparison function that its arc length always have some fixed value. Such a problem (isoperimetric problem) can be interpreted simply geometrically, and such a point of view goes back a long way in the history of the subject.[13] In isoperimetric variational problems one seeks the maximal surface area J of a plane figure whose boundary K has a prescribed length:

$$J = \max,\quad K = \text{const} = c\,.$$

The area is determined by the boundary curve of length c, and also by the line segment joining the fixed given boundary points P_1 and P_2.

[13] In honor of the legendary Phoenician queen Dido, who used this method to determine the borders of Carthage, the isoperimetric problem in the narrow sense is known as Dido's problem.

Suppose we now let the two fixed endpoints of the comparison curve coincide, so that now with the closed curve, one is dealing with a geometric problem whose solution is known to be the circle of circumference c. In view of the admissible comparison curves, the problem should be treated in parametric form.

The idea of a field can be extended as well to the isoperimetric problem, even when K is replaced by general functionals (isoperimetric problem in the wide sense). The extremals must now form a spatial field, where the value c of the isoperimetric integral enters the picture as the third, space coordinate joining the known two-dimensional extremal field $\mathbb{F}$. The Weierstrass criterion, which rests on the representation theorem using the excess function (strong extremals), as well as the Jacobi criterion (conjugate points) can be analogously carried over, to the extent that the given extremal of the problem $J \to \max$ is not simultaneously an extremal of the isoperimetric integral K, that is, does not satisfy the necessary condition for $K \to \text{extr}$. Von Escherich pointed out the exceptional case, and his student Hahn had reported on it during his stay in Göttingen.

Carathéodory, who had attended Hahn's lecture, immediately understood how Weierstrass's method should be changed so that it remained applicable. The modification consisted in the extension of the solution class (solution curves). Carathéodory extended the domain of allowed solution curves to those with finitely many discontinuities (corner points or discontinuous extremals) and constructed for the extended class of functions a field (*Carathéodory 1904*, §§6, 14). He further explicated his idea geometrically with an optical example, namely with the question how a curve of fixed length can be drawn on a spherical lamp shade between two given endpoints in a way that its shadow on the floor is as short or as long as possible.[14]

2.4 More general problems

Isoperimetric problems in the narrow and wide senses are simply special cases of a general variational problem, for one can extend the problem statements to several unknown functions of one or more independent variables and thereby impose further conditions (such as side conditions in differential or integral form) that the comparison functions must satisfy. At that time, one liked to speak in this connection of the most general problem of the calculus of variations; today, one prefers to consider certain problem types, such as Lagrange

[14] One speaks of discontinuous solutions, where "discontinuous" is used in the sense of the eighteenth century. At that time, functions were defined by an analytic expression (*expressio analytica*) and therefore a priori smooth, which is what one meant by continuity. Corner points were from this point of view already discontinuities, since it was here that one cobbled together functions (curves), or equivalently the analytic expressions. In the calculus of variations, this terminology harks back even today to how things were understood at that time.

and Mayer problems, both of which derive from the variational problem

$$J(x,u) = G(x,u(x)) + \int F(x,u,u')dx$$

of Bolza, where one sets for the comparison functions additional differential equations and general boundary conditions

$$\begin{aligned} \Phi_\alpha(x,u(x),u'(x)) = 0, \quad \alpha = 1,\ldots,s, \\ \Psi_\beta(x,u(x)) = 0, \quad \beta = 1,\ldots,r. \end{aligned}$$

For $F = 0$ one obtains the Mayer problem, while for $G = 0$, one has the Lagrange problem with variable endpoints.

The encyclopedia article limits itself at the end to a glimpse of the "general" problem of simple integrals. However, it does not discuss a difficult question that is bound up in it and arises from the side conditions: Can one still vary the function (curve) $y^0(x)$ under discussion or are the comparison functions so limited by the side condition(s) that the result is the extreme case of a rigid solution (that is, that there is no competing comparison function)?

In 1903, Hahn had worked on variational problems with differential and finite equations as side conditions. In such problems, which generalize the Lagrange problems (only differential side conditions), Hahn was unable to overcome, like his teacher von Escherich in the case of the actual Lagrange problems, the case of so-called anomalous curves (*Hahn 1904*; see footnote 32 of the encyclopedia article). The phenomenon of "anomalous curves", which appears in the extension of the standard problem, such as, for example, the Lagrange problem, was first observed by Adolph Mayer (*1885*) in relation to the validity of the Lagrange method of multipliers, and it was treated intensively by von Escherich after 1899 (see in this regard, for example, *Carathéodory 1935, Bliss 1946, Klötzler 1994*).

At the end of this section, the authors observe that even worse complications arose for n unknown functions ($n > 1$), but that things behaved fundamentally as in the case of a single unknown function. Here shines through the optimism voiced by Hilbert in his Paris address *1900a* (Problem 23), an optimism that he hoped to see realized in a dissertation by Nadeschda Gernet (*1902*), but which was not fully vindicated. In variational problems with n unknown functions of a single independent variable, the Euler-Lagrange differential equation, which is of second order, yields a $2n$-parameter family of extremals $y = y(x,\alpha)$. On the other hand, a field $\mathbb{F}$ comprises only an n-parameter family. Which extremals are then to be chosen? Mayer described the situation in two articles (*1903, 1905*) and specified the relevant requirements for what is now called a Mayer field using expressions now called Lagrange brackets: $[\alpha_i, \alpha_j] = 0$, $i,j = 1,\ldots,n$.

A comparison with light rays in optics is instructive: not every family of lines (light rays) can be interpreted as coming from a light source (point or more general source). Therefore, no Huygens wave surfaces can be specified,

nor can transversal families be constructed. Similar arguments form the basis for the restrictions noted by Jacques Hadamard (*1902*) for multiple integrals: the solution manifold of the corresponding Euler-Lagrange partial differential equation is of such high dimension that it is difficult to construct an extremal field corresponding to the one-dimensional case.

The application of Weierstrass's method requires the construction of a field; the difference of the integrals $J(C_0)$ and $J(C)$, where C_0 is the extremal in question and C is an admissible comparison function, can then be expressed simply by an integral over the excess function E, provided that C_0 and C are inside the field. A field F is determined by the surface elements $(x, y, p(x, y))$ in a domain B of $\mathbb{R}^{n+\nu+n\nu}$. Should the field contain a μ-parameter family of n-dimensional surfaces $y = y(x, \gamma)$, that is, if the elements of the family satisfy in B the embedding condition, or the differential equation

$$\frac{\partial y_i}{\partial x_\alpha} = p_{i\alpha}(x, y)\,, \qquad (*)$$

then the differential equation must be integrable on this family, since because of the independence of the partial derivatives, the mixed second partial derivatives of the surface $y = y(x)$ are equal by Schwarz's theorem:

$$\frac{\partial^2 y_i}{\partial x_\alpha \partial x_\beta} = \frac{\partial p_{i\alpha}}{\partial x_\beta} = \frac{\partial^2 y_i}{\partial x_\beta \partial x_\alpha} = \frac{\partial p_{i\beta}}{\partial x_\alpha}\,.$$

This integrability condition is a strong requirement on the field F. However, as Hermann Boerner (*1936*) noted, this assumption can be weakened. Aside from the extremal C_0 in question, no other surface need be embeddable in the field, so that the integrability condition $(*)$ is required only on C_0 and need not be satisfied for the field in general. Among these fields are some, called geodesic fields, for which one can specify by an E-function a Hilbert independence integral and the Weierstrass representation formula, which correspond to those described in 2.1. Thereby sufficient conditions can again be formulated. There are various approaches to geodesic fields, for which Théophile Lepage (*1941*) elegantly provided a unified formulation using the theory of differential forms.

As Carathéodory put it, in the geodesic field, the complete figure created by fields for simple integrals is thereby given up. Whether complete figures for multiple integrals with several unknown functions can be constructed is, on account of their high-dimensional structure, an open question.

2.5 Examples and applications

Examples are motivating forces for all mathematical disciplines, but this platitude is especially true for the calculus of variations, where initially, there existed a collection of methods for solving certain problems that gradually began to develop into a theory. Euler's classical work *Methodus inveniendi* offers exemplary documentary evidence of this. Kneser ended his article on the

calculus of variations in the *Encyklopädie* (*Kneser 1904*) with an overview of examples (geometric, mechanical, and theoretical; pp. 620–625); Zermelo and Hahn did not provide a list of their examples.[15] But on the one hand, the collection of individual examples given by Zermelo and Hahn shows how far at the time a general and unified theory had been achieved, and on the other hand, these problems also reflect the historical path of development of the calculus of variations.

The seven examples are taken from the Newtonian problem of surfaces of revolution axially exposed to a stream with minimal energy, and they refer to Kneser's treatment. The problem is considered in the framework of Newtonian fluid mechanics (that is, with Newton's law of resistance), for which reason a remarkable truncated vertex (planar front surface) appears.

Under Newton's assumptions on the resistance, a weak minimum can be proved. Significant here is the Newtonian requirement that the rate of ascent of the curve of revolution be bounded from below, for the requirement $|y'| > 1$ excludes the domain in which the requisite Legendre condition is invalid. If one ignores this requirement, then one could use the well-known zigzag curve of Adrien Marie Legendre to generate the surface of revolution and thereby arrive at the well-known paradox of the calculus of variations, namely that the resistance of the solid of revolution goes to zero the more distinct the zigzag form of the curve becomes (*Legendre 1786*). The fact of this paradox, which is determined by the character of the extremal requirement, is neither discussed nor even mentioned by the authors. This problem can, of course, be seen as the first problem of the classical calculus of variations, analytically formulated and solved by Newton in 1685 (*Principia,* 1687), but because the proof was suppressed, it had little influence on the development of the calculus of variations. As "draft horse" for the calculus of variations, the brachistochrone problem formulated a decade later (1696) together with the associated lively quarrel between the brothers Jakob and Johann Bernoulli—which was given a great deal of attention in the learned world, and with it the brachistochrone problem as well—turned out to be much more important.

The next example offered by Zermelo and Hahn goes beyond the legendary brachistochrone problem and attacks the isoperimetric problem, set as a challenge in the quarrel of 1697 between Jakob and his brother mentioned above. Again, *Kneser 1904* serves as a reference. Reference is made to Dido's problem in a generalized form; first, the endpoints of the so-called isoperimetric curve are themselves movable along curves, and second, the problem is shifted from the plane to a surface, and here in particular modified for tori.

[15] Carathéodory, for whom the interplay in the calculus of variations between the individual example and the general theory was important, always gave in his articles a list of the examples used, doing so even in his commentary to Euler's *Methodus inveniendi*. A similar protocol was followed by the authors Mariano Giaquinta and Stefan Hildebrandt in their standard reference *Calculus of variations* (*1996*).

The latter questions were studied primarily by the Chicago school of the calculus of variations (particularly Oskar Bolza's student Gilbert Bliss). Finally, we mention the dissertation by Hilbert's student Johann Oswald Müller (formerly called Richter) (*Müller 1903*), which deals with isoperimetric problems for double integrals. Müller was to develop sufficient criteria for multiple integrals in the framework of Hilbert's program in field theory, while Nadeschda Gernet was to deal with the complementary area of simple integrals with several unknown functions.[16] But Hilbert had underestimated the complexity of this question, so that he had to be satisfied with Müller's result that, using a field of spheres, proved only the known isoperimetric property of the sphere's surface, by means of methods of the calculus of variations. Mayer (*1903*) was the first to notice that Gernet had worked with special central families for which the Lagrange brackets vanish, so that Hilbert attributed to Mayer the first rigorous proof.

As a further problem, questions are discussed that relate to conjugate (that is, critical) points in variational problems. Such points arise from the Jacobi criterion and relate to the domain in which extremality is ensured. Here reference is made also to *Zermelo 1902d*. Among other things, Zermelo shows that in a simply connected plane region, shortest lines are uniquely determined, and he passes to surfaces of shortest lines of bounded growth, where discontinuous solutions appear.

Shortest lines are the subject of the dissertation *Über die Geometrien, in denen die Geraden die Kürzesten sind* (*On geometries in which the lines are shortest*) by Georg Hamel (*1901*). In it, Hamel considers, as the title succinctly expresses, metrics that lead to straight lines as shortest connecting lines. Hamel uses here the armamentarium of sufficient conditions of the calculus of variations. The problem of shortest lines on surfaces is also the subject of an article by Hilbert on direct methods (*Hilbert 1900b*), and it prompted Charles Noble to write his dissertation under Hilbert (*Noble 1901*).

In addition to the minimization of air resistance for a surface of rotation (a projectile), a second mechanical problem is introduced: the stability of the equilibrium of hanging (inextensible) threads. The solution is obtained ultimately via the principle of the minimum of the potential energy, but historically, this simple insight was problematic in that the inextensible thread brought a side condition into the analytic formulation of the variational problem (therefore an arbitary mathematical variability limited to a fixed thread length by a physical requirement). This condition required a different approach from that used in the standard problem (such as the brachistochrone problem or the problem of least resistance), which at first was not noticed.

[16] The dissertations of Müller and Gernet have the titles *Über die Minimaleigenschaft der Kugel* (*On the minimality of the sphere*) (*Müller 1903*) and *Untersuchungen zur Variationsrechnung* (*Studies in the calculus of variations*) (*Gernet 1902*). For information on this program and its results, see *Thiele 2007*, 609ff.

2.6 Existence questions

The main task in direct approaches consists in obtaining the solution to a variational problem from a minimal sequence via a suitable method. Hilbert formulated his 20th problem at his Paris lecture *1900a* thus:

> It is my conviction that it will be possible to prove these existence theorems by means of a general principle whose nature is indicated by Dirichlet's principle. This general principle will then perhaps enable us to approach the question: Has not every *regular* variation problem a solution?[17]

If one begins with Hilbert's general formulation of the calculus of variations, that by some rule, one associates a number with elements of a set that consists of functions defined on some domain, then for our set of issues, a connection with classical differential calculus as it relates to Weierstrass's theorem (every continuous function has an absolute extremum on a closed interval I) is of interest. It would be desirable to find an analogy for the concept of closed interval, namely an intrinsically compact system of functions in which one can find a uniformly convergent sequence u_n^* in every sequence of functions u_n. The set of continuous or continuously differentiable functions (C^0 and C^1) does not, unfortunately, form such a system, but according to Ascoli's theorem (1884), every set M of uniformly bounded or continuous sequences of functions has this property, that is, in every such sequence from M there is always a uniformly convergent sequence.

A further fundamental property, semicontinuity, was introduced by Leonida Tonelli (*Tonelli 1921/23*), an idea that was also used by André Roussel in 1927, whose roots, however, go back to Lebesgue (1907) and finally to Weierstrass. The connection to the classical calculus of variations is provided by the Legendre condition, from which the semicontinuity of the associated functional follows, that is, from $f_{pp} \geq 0$ follows the boundedness from below of the functional $J(y)$. With the Legendre condition and Ascoli's theorem one has at hand the means to prove the extremality of a function by direct methods.

The methods of the classical calculus of variations reduce a variational problem to a boundary value problem of an Euler-Lagrange differential equation or system of such differential equations so as then to establish that a solution of the boundary value problem of this differential equation is also

[17] "Ich bin überzeugt, daß es möglich sein wird, diese Existenzbeweise durch einen allgemeinen Grundgedanken zu führen, auf den das Dirichletsche Princip hinweist und der uns dann vielleicht in den Stand setzen wird, der Frage näherzutreten, ob nicht jedes *reguläre* Variationsproblem eine Lösung besitzt." (*1900a*, end of the 20th problem, 289; English translation from *Hilbert 1902a*, 470.)—Hilbert varied the assumption of regularity in his various articles. However, in essence, the Legendre condition or the ellipticity of the Euler-Lagrange equation is required; see especially *Hilbert 1904*.

a solution of the variational problem. Another strategy uses the direct approach. Here one begins with the boundedness (from below) of the variation integral for a certain sequence of admissible comparison functions (minimal sequence) in order then to obtain a solution (minimum).

As an example of such a direct approach, Hilbert's treatment of the Dirichlet problem

$$D(u) = \int_a^y \int_b^x (u_x^2 + u_y^2)\, dx\, dy \;\rightarrow \text{minimum}$$

under certain boundary conditions for a function $u = u(x, y)$ is introduced (solution of a boundary value problem for the Laplacian differential equation $\Delta u(x, y) = 0$). One chooses such a minimal sequence $u_1, u_2, \ldots$, for which we clearly have $D(u_i) \geq 0$. With this problem it can be shown that the limiting value

$$v(x, y) = \lim_{n\to\infty} \int_a^x \int_b^y u_n\, dx\, dy$$

not only exists, but also represents the unknown potential function of the given boundary value problem. However, one cannot assume a priori that one has convergence for a given minimal sequence $u_1, u_2, \ldots$. Passage to the limit for a given minimal sequence also does not necessarily lead to a solution of the variational problem. This was investigated by Richard Courant in *1926*.

The actual construction of a minimal sequence is important for numerical purposes, and for this, Walter Ritz (*1908*) has provided an important method.

The direct methods that already offer techniques for obtaining solutions to variational problems make it possible as well to consider the matter from the opposite viewpoint, namely to obtain with these techniques the solution of a boundary value problem of a differential equation if one can interpret this differential equation as the Euler-Lagrange equation of a variational problem.

Epilogue

Let us back up a bit to gain some perspective on *Hahn and Zermelo 1904*. The first thing that strikes us is that we are not dealing with a research article, but with an overview that extends Kneser's contribution *1904*, though to be sure, it was written from the vantage point of the most current research. This survey article was published in a prominent venue, namely the *Encyklopädie der mathematischen Wissenschaften*, a century work,[18] and it quickly obtained a wide readership and visibility in the mathematical world. Such an attractive publishing opportunity was an important reason why the two young authors jumped at the chance. They would each separately publish a joint paper only once again.

[18] A more general overview of the development and importance of the the encyclopedia is given in, for example, *Thiele 2011*, 227ff.

Plans to bring out the encyclopedia in English and French translations materialized only in France, where there was an expanded and updated translation of the German contributions to the encyclopedia, including the articles on the calculus of variations by Kneser and by Zermelo and Hahn. However, work on the French edition of the encyclopedia was abandoned with the outbreak of the First World War. The article on the calculus of variations expanded by Maurice Lecat under the editorial supervision of Jules Molk[19] (*Lecat 1913*) was one of the works appearing in the French torso of published

[19] Jules Molk (1857–1914) studied in Zurich, Paris, and Berlin (under Weierstrass, 1882); doctorate 1884, Paris; habilitation 1884, Rennes; professor in Besançon, 1888, and in Nancy, 1890; scientific editor of the French version of the encyclopedia.

Weiterentwicklung der Variationsrechnung in den letzten Jahren

Hahn and Zermelo 1904

Inhaltsübersicht.

1. Die Grundlagen der Weierstrass'schen Theorie (vgl. II A 8, Nr. **21**, *Kneser*). Die Weiterentwickelung der Variationsrechnung seit dem Jahre 1900 steht vorwiegend unter dem Einflusse der *Weierstrass*'schen Theorie, die in *A. Kneser*'s Lehrbuche[1] zuerst eine systematische Darstellung gefunden hat. Die formale Grundlage dieser Theorie, die Umformung der Integraldifferenz ΔJ vermittelst der E-Funktion, wird hier in ihrer geometrischen Bedeutung für sich untersucht und es wird ihr Zusammenhang mit der *Hamilton-Jacobi*'schen Theorie der Dynamik nachgewiesen. Ein „Feld" wird nach *Kneser*[2] definiert durch eine einparametrige Schar $y = f(x, a)$ von

[1] Lehrbuch der Variationsrechnung, Braunschweig 1900. (Im Folgenden stets als „Lehrbuch" zitiert.) Eine sehr übersichtliche und exakte Darstellung dieser Theorie giebt auch *W. F. Osgood*, Ann. of math. (2) 2 (1901), p. 105.

[2] Lehrbuch § 14. Doch bedient sich *Kneser* hier immer der „homogenen" Parameterdarstellung der Kurven, von der wir hier absehen, um die Formeln zu vereinfachen und besseren Anschluss an die Arbeiten anderer Forscher zu gewinnen.

texts. With such revisions, it furthered the ideal of the German publishers to bring the calculus of variations up to date as much as possible. If one looks at Lecat's article from this viewpoint, then an additional fact becomes clear: the subject of the article is essentially a presentation of the classical calculus of variations, in particular a discussion of necessary and sufficient conditions for a solution (Weierstrass's theory), but there also can be seen the roots of questions that in the subsequent interwar period would alter the calculus of variations considerably, in that now questions of existence of solutions and a beginning functional-analytic representation of the theory (as already appeared in *Hadamard 1910*) would come to the fore. Indeed, one should keep in mind that Hahn was intimately involved in the creation of linear functional analysis.

Further development of the calculus of variations in recent years

Hahn and Zermelo 1904

Table of contents.

1. The foundations of Weierstrass's theory (see *Kneser 1904*, no. **21**). The further development of the calculus of variations since 1900 has been most influenced by *Weierstrass*'s theory, of which *A. Kneser*'s textbook[1] gives the first systematic account. The formal foundation of this theory, the transformation of the integral difference ΔJ by means of the E-function, is here investigated with respect to its geometric significance, and its connection with the *Hamilton-Jacobi* theory of dynamics is demonstrated. According to *Kneser*[2] a "field" is defined by a one-parameter family $y = f(x, a)$ of

[1] *Kneser 1900* (henceforth referred to as "Lehrbuch"). A very clear and precise account of this theory is also given in *Osgood 1900/01*.

[2] Lehrbuch § 14. But *Kneser* always uses the "homogeneous" parameter representation of the curves, from which we shall refrain here in order to simplify the formulas and improve the interconnections with the work of other scientists.

Extremalen, d. h. von Lösungen der zu dem Integral:

$$J = \int f(x, y, y')\, dx$$

627 | gehörenden *Lagrange*'schen Differentialgleichung:

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} = 0 \,,$$

soweit auf ihnen der Ausdruck $\dfrac{\partial y}{\partial a}$ nicht verschwindet, sodass sie einen Bereich der Ebene einfach überdecken. Vergleicht man innerhalb eines solchen Feldes die Integralwerte J auf benachbarten Extremalen, so reduziert sich die erste Variation δJ vermöge der Differentialgleichung auf die vom Integralzeichen freien „Grenzglieder", welche von der Verschiebung der Endpunkte herrühren. In jedem Punkte des Feldes giebt es aber eine Verschiebungsrichtung, für die auch die entsprechenden Grenzglieder verschwinden, und die Kurven, welche in jedem ihrer Punkte diese Richtung besitzen, bedecken wieder das ganze Feld und werden als die „Transversalen" des Feldes bezeichnet[3]. Dabei gilt der Satz, dass alle Extremalenbögen zwischen zwei Transversalen gleiche Integralwerte J ergeben[4], eine Verallgemeinerung des *Gauss*'schen „Orthogonalitätssatzes" über geodätische Linien[5]. Extremalen und Transversalen bilden somit ein neues krummliniges Koordinatensystem innerhalb des Feldes, wobei der Parameter a der Extremalenschar und der Integralwert u längs einer Extremalen, gerechnet von einer festen Transversalen bis zu einem variabelen Endpunkte[6], als Koordinaten erscheinen. Mit Hülfe dieser Transformation lässt sich nun die Differenz ΔJ der Integrale J längs einer Feldextremalen und einer beliebigen innerhalb des Feldes verlaufenden „Vergleichskurve" $\mathfrak{c}$

[3] *Kneser* braucht diesen Ausdruck nicht, sondern spricht von „Kurven, die von den Extremalen transversal geschnitten werden".

[4] Lehrbuch § 15. Man vgl. auch *Darboux*, Théorie des surfaces 2, Nr. 521 ff. und Nr. 544 ff.

[5] *Gauss*, Disq. c. superf. curv. 15–16.

[6] Als Funktion von x und y betrachtet genügt u einer gewissen partiellen Differentialgleichung erster Ordnung, die eine Verallgemeinerung der *Jacobi-Hamilton*'schen Gleichung der Dynamik darstellt. Ist umgekehrt $u(x, y)$ eine beliebige Lösung dieser Differentialgleichung, so können die Kurven $u = \text{const.}$ als Transversalen einer Schar von Extremalen aufgefasst werden, die durch Integration einer gewöhnlichen Differentialgleichung erster Ordnung gefunden wird. Kennt man eine Lösung $u(x, y, c)$ unsrer partiellen Differentialgleichung, die eine willkürliche Konstante c enthält, so liefert die Gleichung $\dfrac{\partial u}{\partial c} = \text{const.}$ die Gesamtheit der Extremalen. Lehrbuch § 19. Vgl. auch *Darboux*, Théorie des surfaces 2, Nr. 538 f. und II A 8, Nr. **11** (*Kneser*).

extremals, i. e., of solutions of the *Lagrange* differential equation

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} = 0$$

belonging to the integral

$$J = \int f(x, y, y')\, dx \ ,$$

insofar as the expression $\dfrac{\partial y}{\partial a}$ does not vanish on them so that they simply cover a region of the plane. If we compare the integral values J on neighboring extremals within such a field, then, by virtue of the differential equation, the first variation δJ is reduced to the "boundary terms" free from the integral sign, which arise from the displacement of the endpoints. But at each point of the field there is one direction of displacement for which the corresponding boundary terms also vanish, and the curves having this direction at each of its points again cover the entire field and are called the "transversals" of the field.[3] For this case, that all extremal arcs between two transversals yield the same integral values J holds as a theorem,[4] which is a generalization of *Gauss*'s "orthogonality theorem" on geodesic lines.[5] Extremals and transversals thus form a new, curvilinear coordinate system within the field, where the parameter a of the extremal family and the integral value u along an extremal, taken from a fixed transversal up to a variable endpoint,[6] figure as coordinates. By means of this transformation we can now represent the difference ΔJ of the integrals J along a field extremal and any "comparison

[3] *Kneser* does not use this expression but speaks of "curves transversally intersected by the extremals".

[4] Lehrbuch § 15. See also *Darboux 1889*, no. 521 ff. and no. 544 ff.

[5] *Gauß 1828*, sections 15–16.

[6] Considered as a function of x and y, u satisfies a certain partial differential equation of first order representing a generalization of the *Jacobi-Hamilton* equation of dynamics. If, conversely, $u(x, y)$ is any solution of this differential equation, then we can conceive of the curves $u = \text{const.}$ as transversals of a family of extremals, which is found by integration of an ordinary differential equation of first order. If a solution $u(x, y, c)$ of our partial differential equation containing any constant c is known, then the equation $\dfrac{\partial u}{\partial c} = \text{const.}$ yields the totality of the extremals. Lehrbuch § 19. See also *Darboux 1889*, no. 538 f. and *Kneser 1904*, no. **11**).

darstellen durch das Integral der *Weierstrass*'schen E-Funktion, genommen über die Vergleichskurve[7].

628 | Dieselbe Umformung erreicht *D. Hilbert*[8] auf folgendem Wege. An Stelle des gegebenen Integrales:

$$J = \int f(x, y, y')\, dx$$

betrachtet er das folgende:

$$J^* = \int \big\{ f(x, y, p) + (y' - p) f_p(x, y, p) \big\}\, dx$$

und sucht p als Funktion von x und y so zu bestimmen, dass dieses Integral J^* vom Integrationswege unabhängig wird. So ergiebt sich für p die partielle Differentialgleichung:

$$\Big(\frac{\partial p}{\partial x} + p\frac{\partial p}{\partial y}\Big) f_{pp} + p f_{py} + f_{px} - f_y = 0\ ,$$

deren Charakteristiken die Extremalen des Problems, d. h. die Lösungen der entsprechenden *Lagrange*'schen Differentialgleichung sind, und man erhält eine Lösung p, indem man eine beliebige einparametrige Schar von Extremalen zu Grunde legt und in jedem Punkte x, y ihres Feldes die Ableitung $\frac{dy}{dx}$ längs der durch den Punkt gehenden Extremalen bildet. Für jede so bestimmte Funktion p wird also das Integral J^* vom Wege unabhängig; dieses Theorem bezeichnet *Hilbert* als den „Unabhängigkeitssatz". Das „Feldintegral" J^* verschwindet dann auf jeder Transversalen des Feldes; auf jeder Feldextremalen aber reduziert es sich auf das „Grundintegral" J[9] und ist daher, als Funktion des oberen Endpunktes betrachtet, identisch mit der *Kneser*'schen Funktion u. Vermöge des Unabhängigkeitssatzes lässt sich nun das Grundintegral über einen Extremalenbogen ausdrücken als das Feldintegral J^* über eine Vergleichskurve $\mathfrak{c}$ (siehe Fussnote 7)), und die Subtraktion der beiden Integrale J und J^* über dieselbe Kurve $\mathfrak{c}$ liefert schliesslich die *Weierstrass*'sche Transformation der Integraldifferenz ΔJ.

2. Notwendige und hinreichende Bedingungen im einfachsten Falle. Aus der *Weierstrass*'schen Umformung der Integraldifferenz schliesst
629 man, dass ein Extremalenstück einen grösseren oder kleineren | Integralwert

[7] Lehrbuch § 20. Als Vergleichskurven gelten dabei alle Kurven, deren Anfangs- und Endpunkt auf denselben Transversalen des Feldes liegen, wie Anfangs- und Endpunkt des zu vergleichenden Extremalenbogens.

[8] Mathematische Probleme, Vortrag gehalten auf dem internationalen Mathematikerkongress zu Paris 1900, Gött. Nachr. 1900, p. 291; Comptes rendus du deuxième congrès international des mathématiciens p. 106; Archiv Math. Phys. (3) 1 (1901), p. 231 (mit einigen Zusätzen).

[9] Beide Bezeichnungen finden sich bei *Hilbert* nicht.

curve" $\mathfrak{c}$ running within the field by the integral of *Weierstrass*'s E-function, taken over the comparison curve.[7]

D. Hilbert[8] achieves the same transformation as follows. Instead of the given integral

$$J = \int f(x, y, y')\, dx$$

he considers the integral

$$J^* = \int \left\{ f(x, y, p) + (y' - p) f_p(x, y, p) \right\} dx$$

and seeks to determine p as a function of x and y so that this integral J^* becomes independent of the path of integration. One thus obtains for p the partial differential equation

$$\Big(\frac{\partial p}{\partial x} + p \frac{\partial p}{\partial y} \Big) f_{pp} + p f_{py} + f_{px} - f_y = 0 \,,$$

whose characteristics are the extremals of the problem, i. e., the solutions of the corresponding *Lagrange* differential equation, and one obtains a solution p by using as a basis any one-parameter family of extremals and by forming at each point x, y of its field the derivative $\dfrac{dy}{dx}$ along the extremal passing through the point. Hence, for each function p thus determined the integral J^* becomes independent of the path; *Hilbert* refers to this theorem as the "independence theorem". The "field integral" J^* then vanishes on every transversal of the field; but, on every field extremal, it is reduced to the "basic integral" J,[9] and hence, when considered as function of the upper endpoint, it is identical with *Kneser*'s function u. Now, by virtue of the independence theorem, we can express the basic integral over an extremal arc as the field integral J^* over a comparison curve $\mathfrak{c}$ (see footnote 7), and subtraction of the two integrals J and J^* over the same curve $\mathfrak{c}$ eventually yields the *Weierstrass* transformation of the integral difference ΔJ.

2. Necessary and sufficient conditions in the simplest case. From the *Weierstrass* transformation of the integral difference it follows that an extremal segment yields a greater or smaller integral value than any comparison

[7] Lehrbuch § 20. Here, any curves whose initial point and endpoint lie on the same transversals of the field as the initial point and endpoint of the extremal arc which is to be compared are considered comparison curves.

[8] *Hilbert 1900a*, p. 291; *Hilbert 1902b*, p. 106; *Hilbert 1901a*, p. 231 (with several additions).

[9] Both designations are not to be found in *Hilbert*.

ergiebt als jede Vergleichskurve $\mathfrak{c}$, längs deren die E-Funktion ein beständig positives oder negatives Vorzeichen besitzt.

Unterscheidet man nun mit *Kneser*[10] ein „starkes“ und ein „schwaches“ Extremum, je nachdem nur die Punkte oder auch die Tangentenrichtungen der Vergleichskurven den entsprechenden der Extremalen benachbart sein müssen, so ergiebt sich innerhalb des Feldes für das *starke* Extremum das *Weierstrass*'sche Kriterium als notwendig und hinreichend, nach welchem das Vorzeichen der E-Funktion definit sein muss für eine willkürliche Richtung der Vergleichskurve[11]. Für das *schwache* Extremum dagegen braucht dieses Kriterium nur für hinreichend kleine Abweichungen der Tangentenrichtungen erfüllt zu sein, und ist dann äquivalent dem *Legendre*'schen Kriterium[11a], das ein definites Vorzeichen der Funktion $\dfrac{\partial^2 f}{\partial y'^2}$ längs der Extremalen erfordert.

Das Problem reduziert sich hiernach, im Falle vorgeschriebener Endpunkte, auf die Frage, ob das betrachtete Extremalenstück mit einem Felde umgeben werden kann. Bei einem hinreichend kleinen Extremalenbogen ist dies immer der Fall; und die Grenzen, an denen diese Eigenschaft aufhört, bestimmen sich durch das *Jacobi*'sche Kriterium der „konjugierten Punkte“ (II A 8, Nr. **16** u. **17**, *Kneser*). Dass dieses Kriterium auch notwendig ist, beweist
630 *Kneser*[12] durch Betrach-|tung der Enveloppe, die zu den vom Anfangspunkte der Integration ausgehenden Extremalen gehört: er zeigt, dass es in jeder Nähe eines von zwei konjugierten Punkten begrenzten Extremalenbogens Kurven giebt, die dem Integrale denselben Wert erteilen, wie die Extremale. Dieser Beweis versagt, wenn die Enveloppe auf der betrachteten Extremalen

[10] Lehrbuch § 16–17.

[11] Doch vgl. hierzu ein Beispiel von *O. Bolza* (Bull. Am. math. soc. (2) 9 (1901), p. 1), sowie die sich darauf beziehenden Bemerkungen von *E. R. Hedrick* (Bull. Am. math. soc. (2) 9 (1901), p. 245). *W. F. Osgood* (Amer. Trans. 2 (1901), p. 273) beweist für den Fall des starken Extremums weiter: Bezeichnet $\overline{J}$ den von der Extremale gelieferten Integralwert und S ein gewisses die Extremale umgebendes Gebiet, so muss jede Schar ganz in S gelegener Kurven, die mit der Extremale Anfangs- und Endpunkt gemein haben, und deren Integralwerte gegen $\overline{J}$ konvergieren, die Extremale zur Grenzkurve haben. Auf Grund dieses Satzes erweitert *Osgood* die Klasse der zulässigen Vergleichskurven mit Hülfe einer geeigneten Verallgemeinerung des Begriffes des Kurvenintegrals. Vereinfachungen des *Osgood*'schen Beweises geben *O. Bolza* (Amer. Trans. 2 (1901), p. 422) und *E. Goursat* (Amer. Trans. 5 (1904), p. 110).

[11a] Diese Bezeichnungsweise weicht von der in *Kneser*'s Lehrbuche ab, da dort auch von einem *Legendre*'schen Kriterium für das *starke* Extremum gesprochen wird.

[12] Lehrbuch § 25; vgl. auch II A 8, Nr. **17**, *Kneser*, sowie *E. Zermelo*, Diss. Berl. 1894, p. 96. Über das Verhältnis dieser Methode zu den älteren, die zweite Variation benützenden vgl. Lehrbuch, p. 105. — Die Untersuchung eines von zwei konjugierten Punkten begrenzten Extremalenbogens durch Betrachtung der dritten und vierten Variation hat *A. Korn* (Münch. Ber. 1901, p. 76) kürzlich von Neuem aufgenommen, wobei er wieder zu den *Erdmann*'schen Kriterien (II A 8, Nr. **22**, *Kneser*) gelangte.

curve $\mathfrak{c}$ along which the E-function possesses a sign that is always positive or always negative.

Now, if we, following *Kneser*,[10] make a distinction between a "strong" and a "weak" extremum, depending on whether only the points or also the tangent directions of the comparison curves must be adjacent to the corresponding ones of the extremals, then *Weierstrass*'s criterion turns out to be necessary and sufficient for the *strong* extremum within the field, according to which the sign of the E-function must be definite for any direction of the comparison curve.[11] In the case of a *weak* extremum, by contrast, this criterion only needs to be satisfied for sufficiently small deviations of the tangent directions, and it is then equivalent to *Legendre*'s criterion,[12] which demands a definite sign of the function $\dfrac{\partial^2 f}{\partial y'^2}$ along the extremal.

Accordingly, in the case of prescribed endpoints, the problem is reduced to the question whether it is possible to surround the extremal segment under consideration with a field. This is always the case for a sufficiently small extremal arc; and the boundaries at which this property ceases to hold are determined by *Jacobi*'s criterion for "conjugate points" (*Kneser 1904*, nos. **16** a. **17**). That this criterion is also necessary is proved by *Kneser*[13] by considering the envelope that belongs to the extremals starting at the initial point of the integration: he demonstrates that in each vicinity of an extremal arc bounded by two conjugate points there are curves assigning the same value to the integral as the extremal. This proof fails when the envelope possesses

[10] Lehrbuch § 16–17.

[11] But, on this, see an example by *O. Bolza* (*1901a*), and the relevant remarks by *E. R. Hedrick* (*1901a*). *W. F. Osgood* (*1901b*), furthermore, proves for the case of the strong extremum: If $\overline{J}$ denotes the integral value provided by the extremal and if S denotes a certain region surrounding the extremal, then every family of curves which are entirely situated in S and share initial point and endpoint with the extremal and whose integral values converge to $\overline{J}$, must have the extremal as limit curve. In light of this theorem, *Osgood* extends the class of admissible comparison curves by means of a suitable generalization of the concept of curve integral. *O. Bolza* (*1901b*) and *E. Goursat* (*1904*) give simplified versions of *Osgood*'s proof.

[12] This terminology departs from that used by *Kneser* in his Lehrbuch, since he also refers to a *Legendre* criterion for the *strong* extremum.

[13] Lehrbuch § 25; see also *Kneser 1904*, no. **17**, and *Zermelo 1894*, p. 96. On the relationship between this method and the older methods using the second variation, see Lehrbuch, p. 105.—*A. Korn* (*1902*) has recently resumed the investigation of an extremal arc bounded by two conjugate points by considering the third and fourth variation, which led him to *Erdmann*'s criteria again (*Kneser 1904*, no. **22**). ⟦Hahn and Zermelo erroneously write "*A. Korn* (Münch. Ber. 1901, p. 76)" instead of "*A. Korn* (Münch. Ber. 1902, p. 75)".⟧

einen Rückkehrpunkt besitzt, der seine Spitze dem Integrationsintervall zukehrt. *W. F. Osgood*[13], der unabhängig von *Kneser* diesen Umstand bemerkte, zeigt, dass in diesem Ausnahmefalle der von zwei konjugierten Punkten begrenzte Extremalenbogen noch ein Extremum liefert.

Eine ganz analoge Schlussweise gilt, wenn der Anfangspunkt der Integration auf einer vorgeschriebenen „Grenzkurve“ $\mathfrak{K}$ variieren kann; man hat dann ein Feld von Extremalen zu verwenden, welche die Kurve $\mathfrak{K}$ transversal schneiden[14]. Der erste Punkt, in dem die betrachtete Extremale die Enveloppe dieses Feldes berührt, heisst der zur Kurve $\mathfrak{K}$ gehörige „kritische Punkt“[15]; in ihm hört die Extremale auf, ein Extremum zu liefern[16]. Er liegt immer diesseits des zu dem Schnittpunkte unsrer Extremale mit der Kurve $\mathfrak{K}$ gehörigen konjugierten Punktes; seine Lage hängt nur ab von der Krümmung der Kurve $\mathfrak{K}$ in diesem Schnittpunkte[17].

Sind beide Endpunkte 1, 2 auf zwei Kurven $\mathfrak{K}_1$, $\mathfrak{K}_2$ variabel, so müssen
631 beide Grenzkurven die Extremale transversal schneiden, und | der kritische Punkt d_1, der zur Kurve $\mathfrak{K}_1$ gehört, darf, wenn ein Extremum stattfinden soll, weder in das Integrationsintervall 1 2 noch zwischen 2 und den zu $\mathfrak{K}_2$ gehörenden kritischen Punkt d_2 fallen. Die Notwendigkeit dieser letzten neu hinzukommenden Bedingung beweist *G. A. Bliss*[18], indem er einen Punkt 3 zwischen d_1 und d_2 auf der Extremalen annimmt, sodass diese bei der Anordnung $1\,2 d_1 3 d_2$ zwar zwischen $\mathfrak{K}_2$ und 3, aber nicht mehr zwischen $\mathfrak{K}_1$ und 3 ein Minimum des Integrales liefert. Dann giebt es also eine benachbarte Kurve $1'2'3$, welche $\mathfrak{K}_1$ und $\mathfrak{K}_2$ in $1'$ und $2'$ schneidet und für welche $J_{1'2'3'} < J_{123}$,

[13] Amer. Trans. 2 (1901), p. 166. Es wird dort auch gezeigt, dass beim Probleme der kürzesten Linien auf dem Rotationsellipsoide dieser Ausnahmefall thatsächlich eintritt. Übrigens findet sich ein Hinweis auf die Möglichkeit eines solchen Rückkehrpunktes schon bei *H. Poincaré*, „Les méthodes nouvelles de la mécanique céleste“ 3 (1899), Nr. 371, anlässlich einer Diskussion des Prinzips der kleinsten Wirkung. Ein regulärer Punkt der Enveloppe aller durch einen Punkt gehenden Extremalen (trajectoires) wird dort als „foyer ordinaire“ bezeichnet; ein Rückkehrpunkt, der seine Spitze dem Integrationsintervalle zukehrt, als „foyer en pointe“, ein nach der entgegengesetzten Richtung liegender Rückkehrpunkt als „foyer en talon“.

[14] Dass sich ein die Kurve $\mathfrak{K}$ transversal schneidender Extremalenbogen in der Umgebung dieser Kurve stets mit einem solchen Felde umgeben lässt, beweist *A. Kneser*, Lehrbuch § 30. Die transversale Lage der Kurve $\mathfrak{K}$ und der Extremale ergiebt sich bekanntlich als notwendig durch Betrachtung der ersten Variation (vgl. II A 8, Nr. **6**, *Kneser* und Lehrbuch § 12).

[15] *Kneser* bezeichnet diesen Punkt als „konjugierten Brennpunkt“; die Terminologie des Textes rührt von *G. A. Bliss* her.

[16] Lehrbuch §25. Der daselbst mitgeteilte Beweis unterliegt derselben Ausnahme, wie im Falle fester Endpunkte. Ein von *G. A. Bliss*, Amer. Trans. 3 (1902), p. 132 durch Betrachtung der zweiten Variation geführter Beweis gilt allgemein, sofern der Extremalenbogen über den kritischen Punkt hinausgeht.

[17] Lehrbuch § 30; *G. A. Bliss*, l. c. p. 139.

[18] Math. Ann. 58 (1904), p. 70.

a turning point on the extremal under consideration whose pike turns towards the integration interval. *W. F. Osgood*,[14] who noted this fact independently of *Kneser*, shows that, in this exceptional case, the extremal arc bounded by two conjugate points yields another extremum.

A quite analogous line of reasoning is valid when the initial point of the integration can vary on a prescribed "boundary curve" $\mathfrak{K}$; in this case, we must use a field of extremals transversally intersecting the curve $\mathfrak{K}$.[15] The first point in which the extremal under consideration makes contact with the envelope of this field is called the "critical point" belonging to the curve $\mathfrak{K}$;[16] in it, the extremal ceases to yield an extremum.[17] It always lies on this side of the conjugate point belonging to the intersection point of our extremal with the curve $\mathfrak{K}$; its position only depends on the curvature of the curve $\mathfrak{K}$ in this intersection point.[18]

If the two endpoints 1, 2 are variable on two curves $\mathfrak{K}_1$, $\mathfrak{K}_2$, then the two limit curves must intersect the extremal transversally, and the critical point d_1 belonging to the curve $\mathfrak{K}_1$ must fall neither in the integration interval 1 2 nor between 2 and the critical point d_2 belonging to $\mathfrak{K}_2$, if an extremum is to occur. That this last, newly added condition is necessary is proved by *G. A. Bliss*,[19] who posits a point 3 between d_1 and d_2 on the extremal so that the latter still yields a minimum of the interval for the arrangement $1\,2 d_1 3 d_2$ between $\mathfrak{K}_2$ and 3, but no longer between $\mathfrak{K}_1$ and 3. Thus, there is a neighboring curve $1'2'3$ which intersects $\mathfrak{K}_1$ and $\mathfrak{K}_2$ in $1'$ and $2'$ and for

[14] *Osgood 1901a*, where it is also shown that this exceptional case actually arises for the problem of the shortest lines on the ellipsoid of revolution. Incidentally, *H. Poincaré*, in *1899*, no. 371, already alludes to the possibility of such a turning point in a discussion regarding the principle of least action. There, a regular point of the envelope of all extremals (trajectoires) passing through a point is called a "foyer ordinaire"; a turning point whose pike turns towards the integration interval is called a "foyer en pointe", a turning point lying in the opposite direction a "foyer en talon".

[15] That it is always possible to surround an extremal arc transversally intersecting the curve $\mathfrak{K}$ in the vicinity of this curve with such a field is shown by *A. Kneser*, in Lehrbuch § 30. It is well-known that the transversal position of the curve $\mathfrak{K}$ and of the extremal turns out to be necessary when we consider the first variation (see *Kneser 1904*, no. **6**, and Lehrbuch § 12).

[16] *Kneser* calls this point the "conjugate focal point"; the terminology in the text is due to *G. A. Bliss*.

[17] Lehrbuch § 25. The proof conveyed there is subject to the same exception as in the case of fixed endpoints. A proof conducted by *G. A. Bliss*, *1902a*, p. 132, by considering the second variation is generally valid, provided the extremal arc exceeds the critical point.

[18] Lehrbuch § 30; *G. A. Bliss*, l. c. p. 139.

[19] *Bliss 1904*.

zugleich aber $J_{2'3} > J_{23}$, sodass sich in der That:

$$J_{1'2'} = J_{1'2'3} - J_{2'3} < J_{123} - J_{23} = J_{12}$$

ergiebt. Liegt dagegen d_2 zwischen 2 und d_1, während im ganzen betrachteten Intervalle das *Legendre*'sche, bezw. das *Weierstrass*'sche Kriterium erfüllt ist, und ist $1'2'$ eine willkürliche benachbarte Vergleichskurve zwischen $\mathfrak{K}_1$ und $\mathfrak{K}_2$, so kann man ihren Endpunkt $2'$ so mit dem zwischen d_1 und d_2 liegenden Punkte 3 verbinden, dass umgekehrt $J_{2'3} < J_{23}$, aber $J_{1'2'3} > J_{123}$ und somit $J_{1'2'} > J_{12}$ wird, d. h. die Extremale 1 2 liefert hier wirklich ein Minimum[18a]. Ein zugleich notwendiges und hinreichendes Kriterium, das auch für den Fall gilt, wo d_1 und d_2 zusammenfallen, erhält *Bliss* endlich durch Betrachtung der Kurve T, welche durch 2 geht und von allen Extremalen transversal geschnitten wird, die auf $\mathfrak{K}_1$ transversal stehen. Es wird nämlich ein Minimum dann und nur dann stattfinden, wenn diese Transversale T, welche $\mathfrak{K}_2$ in 2 berührt, in der Umgebung von 2 ganz auf der einen, $\mathfrak{K}_1$ zugekehrten Seite von $\mathfrak{K}_2$ liegt, vorausgesetzt, dass auf 1 2 überall das *Legendre*'sche bezw. das *Weierstrass*'sche Kriterium erfüllt ist und im Punkte 2 ausserdem $f > 0$ ist.

3. Isoperimetrische Probleme[19] (II A 8, Nr. **9**, *Kneser*). Die Anwendung der *Weierstrass*'schen Ideen auf isoperimetrische Probleme findet sich
632 in vollständiger Durchführung zuerst bei *Kneser*[20]. Das | einfachste Problem dieser Art verlangt, unter allen, zwei gegebene Punkte verbindenden Kurven, welche dem „isoperimetrischen Integrale“:

$$K = \int g(x, y, y')\, dx$$

einen vorgeschriebenen Wert erteilen, diejenige zu finden, für welche das Integral:

$$J = \int f(x, y, y')\, dx$$

[18a] *G. A. Bliss* selbst (l. c.) beweist sein *hinreichendes* Kriterium nicht wie angegeben, sondern durch Benutzung der auf $\mathfrak{K}_1$ transversal stehenden Feldextremalen $1'2'$ und durch zweimalige Differentiation des entsprechenden Integrales $J_{1'2'} = J(\gamma)$ nach dem Parameter γ des Feldes. Bezüglich analoger Untersuchungen von *G. Erdmann* und *A. Mayer* vgl. II A 8, Nr. **22** (*Kneser*).

[19] Die isoperimetrischen Probleme sind als Spezialfall in dem im nächsten Abschnitte behandelten allgemeinen Probleme enthalten, so dass alle dort angegebenen Sätze auch im Falle des isoperimetrischen Problemes ihre Anwendung finden. An dieser Stelle werden nur diejenigen Untersuchungen erwähnt, die sich speziell mit dem einfachsten isoperimetrischen Probleme befassen.

[20] Lehrbuch § 32–42.

which $J_{1'2'3'} < J_{123}$[20] but also $J_{2'3} > J_{23}$, so that, indeed,

$$J_{1'2'} = J_{1'2'3} - J_{2'3} < J_{123} - J_{23} = J_{12}\,.$$

If, by contrast, d_2 lies between 2 and d_1, while *Legendre*'s criterion, or *Weierstrass*'s criterion respectively, is satisfied in the entire interval under consideration, and if $1'2'$ is any neighboring comparison curve between $\mathfrak{K}_1$ and $\mathfrak{K}_2$, then it is possible to connect its endpoint $2'$ with the point 3 lying between d_1 and d_2 so that, conversely, $J_{2'3} < J_{23}$, but $J_{1'2'3} > J_{123}$, and hence $J_{1'2'} > J_{12}$, i. e., the extremal 1 2 really yields a minimum in this case.[21] *Bliss* eventually obtains a criterion that is both necessary and sufficient, and that is also valid where d_1 and d_2 coincide by examining the curve T which passes through 2 and which is transversally intersected by all extremals transversal to $\mathfrak{K}_1$. For a minimum occurs if and only if this transversal T, which makes contact with $\mathfrak{K}_2$ in 2, lies entirely on that side of $\mathfrak{K}_2$ in the vicinity of 2 which faces $\mathfrak{K}_1$, provided that *Legendre*'s criterion, or *Weierstrass*'s criterion respectively, is satisfied everywhere on 1 2, and that also $f > 0$ in the point 2.

3. Isoperimetric problems[22] (*Kneser 1904*, no. **9**). The first full account of the application of *Weierstrass*'s ideas to isoperimetric problems is given by *Kneser*.[23] The simplest problem of this kind consists in finding among all the curves joining two given points and yielding a prescribed value for the "isoperimetric integral"

$$K = \int g(x, y, y')\,dx$$

the curve for which the integral

$$J = \int f(x, y, y')\,dx$$

[20] ⟦Hahn and Zermelo erroneously write "$J_{1'2'3'}$" instead of "$J_{1'2'3}$".⟧

[21] *G. A. Bliss* himself (l. c.) does not prove his *sufficient* criterion as given but via use of the field extremal $1'2'$ transversal to $\mathfrak{K}_1$ and via twofold differentiation of the corresponding integral $J_{1'2'} = J(\gamma)$ with respect to the parameter γ of the field. For analogous investigations by *G. Erdmann* and *A. Mayer*, see *Kneser 1904*, no. **22**.

[22] The isoperimetric problems are a special case of the general problem considered in the next section so that all theorems specified there can also be applied in the case of the isoperimetric problem. We shall only note those investigations here that are particularly concerned with the simplest isoperimetric problems.

[23] Lehrbuch § 32–42.

ein Extremum wird. Da hier die zugehörige *Lagrange*'sche Differentialgleichung:

$$\frac{\partial(f+\lambda g)}{\partial y} - \frac{d}{dx}\frac{\partial(f+\lambda g)}{\partial y'} = 0$$

den Multiplikator λ (die „isoperimetrische Konstante") als Parameter enthält, so bilden die von einem festen Punkte ausgebenden Extremalen eine zweiparametrige Schar und in der Umgebung des Ausgangspunktes ein räumliches Feld[21], wenn man den Wert des isoperimetrischen Integrales als dritte Raumkoordinate einführt. In diesem Felde gelten die analogen Transformationen wie im einfachsten Falle. Doch ist diese „*Weierstrass*'sche Konstruktion" nur möglich, wenn die Vergleichskurve, als Raumkurve betrachtet, ganz innerhalb des räumlichen Feldes verläuft[22]. Wenn daher auch das *Weierstrass*'sche und das *Legendre*'sche Kriterium für das starke und das schwache Extremum auf den isoperimetrischen Fall übertragen werden können[23], so ist dabei das „starke" Extremum nur in dem modifizierten Sinne aufzufassen, der eine *räumliche* Nachbarschaft der verglichenen Kurven, also auch hinreichend kleine Abweichungen des isoperimetrischen Integrales in entsprechenden Punkten erfordert.

Auch das *Jacobi*'sche Kriterium lässt sich in entsprechender Form auf isoperimetrische Probleme übertragen. Dass die so erhaltene Bedingung notwendig ist, beweist *Kneser*, indem er aus der zweiparametrigen Extremalenschar, die das Feld bildet, eine einparametrige Schar so auswählt, dass sie eine Enveloppe besitzt, wodurch dann dieselben Schlüsse ermöglicht werden wie im einfachsten Falle[24]. In einer späteren Abhandlung[24a] giebt *Kneser* einen anderen Beweis durch Betrachtung der zweiten Variation.

633 | Es sei noch bemerkt, dass bei allen diesen Entwickelungen vorausgesetzt ist, dass die untersuchte Extremale nicht gleichzeitig Extremale des isoperimetrischen Integrales im Sinne des absoluten Extremums sei; eine Behandlung dieses Ausnahmefalles liegt bisher nicht vor[25].

[21] Lehrbuch § 41.

[22] Vgl. Lehrbuch § 38.

[23] Lehrbuch § 36. Das *Legendre*'sche Kriterium kann auch durch Betrachtung der zweiten Variation erhalten werden. Vgl. *O. Bolza*, Amer. Trans. 3 (1902), p. 305 = Decennial publications of the university of Chicago 9 (1902).

[24] Lehrbuch § 40. In gewissen Ausnahmefällen versagt dieser Beweis.

[24a] Math. Ann. 55 (1902), p. 86. Der Grundgedanke stammt aus Vorlesungen von *Weierstrass*. Auch dieser Beweis umfasst einen gewissen Ausnahmefall nicht, den aber *O. Bolza* mit Hülfe einer von *H. A. Schwarz* in seinen Vorlesungen vorgetragenen Methode erledigt (Math. Ann. 57 (1903), p. 44). Ein ganz allgemeiner Beweis für das *Jacobi*'sche Kriterium ergiebt sich aus den unten besprochenen Arbeiten von *G. v. Escherich*. Vgl. *H. Hahn*, Math. Ann. 58 (1904), p. 166.

[25] Nach mündlichen Mitteilungen von *C. Carathéodory* können in solchen Ausnahmefällen diskontinuierliche Lösungen des Variationsproblems auftreten, die den entsprechend modifizierten *Weierstrass*'schen Methoden noch zugänglich sind.

becomes an extremum. Since the associated *Lagrange* differential equation

$$\frac{\partial(f+\lambda g)}{\partial y} - \frac{d}{dx}\frac{\partial(f+\lambda g)}{\partial y'} = 0$$

contains the multiplier λ (the "isoperimetric constant") as parameter, the extremals issuing from a fixed point form a two-parameter family and, in the vicinity of the point of origin, a spatial field,[24] assuming we introduce the value of the isoperimetric integral as the third space coordinate. In this field, the transformations analogous to the simplest case hold. But "*Weierstrass*'s construction" is only possible if the comparison curve, considered as a space curve, lies entirely within the spatial field.[25] Hence, even though *Weierstrass*'s and *Legendre*'s criterion for the strong and the weak extremum can be extended to the isoperimetric case,[26] the "strong" extremum can only be understood in the modified sense, which requires a *spatial* neighborhood of the curves compared, and hence also sufficiently small deviations of the isoperimetric integral in appropriate points.

Correspondingly, *Jacobi*'s criterion, too, can be adapted to isoperimetric problems. *Kneser* proves that the condition thus obtained is necessary by picking from the two-parameter family of extremals forming the field a one-parameter family so that it possesses an envelope, which makes it possible to draw the same conclusions as in the simplest case.[27] In a later work,[28] *Kneser* provides another proof via consideration of the second variation.

It should also be noted that all these developments assume that the extremal under examination is not, at the same time, an extremal of the isoperimetric integral in the sense of the absolute extremum; a treatment of this exceptional case is as yet unavailable.[29]

[24] Lehrbuch § 41.

[25] See Lehrbuch § 38.

[26] Lehrbuch § 36. We can also obtain *Legendre*'s criterion also via consideration of the second variation. See *Bolza 1902a = Bolza 1902b*.

[27] Lehrbuch § 40. This proof fails in certain exceptional cases.

[28] *Kneser 1902a*. The basic idea originates in lectures by *Weierstrass*. This proof also fails to cover a certain exceptional case, which is settled, however, by *O. Bolza* by means of a method presented in lectures by *H. A. Schwarz* (*Bolza 1903a*). A fully general proof of *Jacobi*'s criterion emerges from the works by *G. v. Escherich* discussed below. See *Hahn 1904*, p. 166.

[29] *C. Carathéodory* suggested in oral communications that, in such exceptional cases, discontinuous solutions of the variational problem can arise that are still amenable to *Weierstrass*'s methods, when accordingly modified.

4. Allgemeinere Probleme. Eine vollständig durchgeführte Übertragung der *Weierstrass*'schen Methoden auf einfache Integrale mit beliebig vielen unbekannten Funktionen, zwischen denen Bedingungsdifferentialgleichungen bestehen, liegt bisher nicht vor. Doch liesse sich unter Verwendung aller von den verschiedenen Autoren bisher erhaltenen Resultate, auf deren Wiedergabe wir uns hier beschränken, eine solche Theorie wohl herstellen. Einen Ansatz nach dieser Richtung geben die Untersuchungen von *A. Kneser*[26]. Er beweist, dass innerhalb eines Feldes von Extremalen, dessen Existenz vorausgesetzt wird, die dem *Weierstrass*'schen und *Legendre*'schen Kriterium analogen Bedingungen für ein starkes, bezw. schwaches Extremum hinreichen. Unter derselben Voraussetzung beweist *A. Mayer* für dieses „allgemeine" Problem den *Hilbert*'schen Unabhängigkeitssatz[27] sowie seinen Zusammenhang mit der Integration der *Jacobi-Hamilton*'schen partiellen Differentialgleichung, die zu dem Variationsprobleme oder dem entsprechenden *Lagrange*'schen Differentialgleichungssysteme gehört.

Mit demselben allgemeinen Probleme beschäftigt sich *G. v. Escherich* in einer Reihe von Abhandlungen[28]. Er geht aus von der Betrachtung des durch Variation der *Lagrange*'schen Gleichungen entstehenden linearen Differentialgleichungssystemes, das er als „accessorisches System" bezeichnet[29]. Aus einer zwischen gewissen Lösungen desselben bestehenden Identität ergiebt
634 sich unmittelbar die Über-|führung der zweiten Variation in die von *Clebsch* angegebene „reduzierte" Form[30] (II A 8, Nr. **15**, *Kneser*), aus der nun weiter das *Legendre*'sche Kriterium als notwendige Bedingung folgt[31]. Für das Folgende wird nun diese Bedingung als erfüllt vorausgesetzt, und es wird ein „Hauptfall" und ein „Ausnahmefall" unterschieden[32]. Im Hauptfalle hat, wie gezeigt wird, die *Mayer*'sche Determinante $\Delta(x, x_0)$ (II A 8, Nr. **16**, p. 596, *Kneser*) (die „Determinante des dem Punkte x_0 konjugierten Systemes") im

[26] Lehrbuch §§ 55–61. Das dort behandelte Problem ist etwas allgemeiner als das im Text angegebene.

[27] Leipz. Ber. 1903, p. 131 = Math. Ann. 58 (1904), p. 235. Der Fall zweier unbekannter Funktionen auch bei *N. Gernet*, Diss. Gött. 1902.

[28] Zusammengefasst in Wien. Ber. 110 (1901), p. 1355, wo sich *v. Escherich* der homogenen Darstellung durch einen Parameter bedient. Vgl. Fussn. 2).

[29] Über seine Lösungen wird eine grosse Anzahl von Sätzen bewiesen: Wien. Ber. 107 (1898), p. 1234 ff., 1294 ff., ibid. 110 (1901), p. 1367 ff. Als besonders wichtig erweist sich dabei der für dieses System geltende *Green*'sche Satz (l. c. 107, p. 1244), insbesondere die in demselben auftretende bilineare Differentialform $\psi(z, r; u, \varrho)$.

[30] l. c. 110, p. 1390 ff. Vgl. auch l. c. 107, p. 1242 ff. und 108, p. 1278 ff.

[31] l. c. 107, p. 1383 ff.; 110, p. 1396 ff.

[32] Der Ausnahmefall tritt dann und nur dann ein, wenn die betrachtete Extremale zugleich Extremale für ein einfacheres Variationsproblem ist (*H. Hahn*, Math. Ann. 58 (1904), p. 148), oder anders gesprochen, wenn ein gewisses überbestimmtes System linearer Differentialgleichungen eine von Null verschiedene Lösung besitzt (*G. v. Escherich*, Wien. Ber. 108 (1899), p. 1287 ff.).

4. More general problems. A completely executed adaption of *Weierstrass*'s methods to simple integrals with arbitrarily many unknown functions between which differential equations of conditions hold is not yet available. But it should be possible to produce such a theory using all previous results obtained by the various authors, which are only reproduced here. A first step in this direction is taken in the investigations by *A. Kneser*.[30] He proves that the conditions for a strong and a weak extremum that are analogous to *Weierstrass*'s and *Legendre*'s criterion respectively are sufficient within a field of extremals, whose existence is assumed. Given the same assumptions, *A. Mayer* proves *Hilbert*'s independence theorem for this "general" problem[31] as well as its connection with the integration of the *Jacobi-Hamilton* partial differential equation, which belongs to the variational problem or the corresponding *Lagrange* system of differential equations.

G. v. Escherich is concerned with the same general problem in a series of publications.[32] He begins by considering the linear differential equation system arising from variation of the *Lagrange* equations, which he calls an "accessory system".[33] From an identity holding between certain solutions of it one immediately obtains the transformation of the second variation into the "reduced" form specified by *Clebsch*[34] (*Kneser 1904*, no. **15**), from which, furthermore, *Legendre*'s criterion as sufficient condition follows now.[35] For what follows, he assumes that this condition is satisfied and distinguishes between a "principal case" and an "exceptional case".[36] He shows that, in the principal case, the *Mayer* determinant $\Delta(x, x_0)$ (*Kneser 1904*, no. **16**, p. 596) (the "determinant of the system conjugate to the point x_0") has an isolated zero point in the point x_0, i. e., it is possible to surround a sufficiently

[30] Lehrbuch §§ 55–61. The problem with which he is concerned there is somewhat more general than the one presented in the text.

[31] *Mayer 1903, 1904*. The case of two unknown functions is also in *Gernet 1902*.

[32] Summarized in *von Escherich 1901*, where *v. Escherich* uses the homogeneous representation by one parameter. See footn. 2.

[33] A great number of theorems about its solutions are proved: *von Escherich 1898a*, p. 1234 ff., *von Escherich 1898b*, p. 1294 ff., *von Escherich 1901*, p. 1367 ff. *Green*'s theorem, which holds for this system (*von Escherich 1898a*, p. 1244), turns out to be of particular significance, especially the bilinear differential form $\psi(z, r; u, \varrho)$ occurring in it.

[34] *von Escherich 1901*, p. 1390 ff. See also *von Escherich 1898a*, p. 1242 ff. and *von Escherich 1898b*, p. 1278 ff.

[35] *von Escherich 1898c*, p. 1383 ff.; *von Escherich 1901*, p. 1396 ff.

[36] The exceptional case occurs if and only if the considered extremal is also extremal for a simpler variational problem (*Hahn 1904*), or, put differently, if a certain overdetermined system of linear differential equations possesses a solution different from zero (*von Escherich 1899*, p. 1287 ff.).

Punkte x_0 eine isolierte Nullstelle, d. h. ein genügend kleiner von x_0 ausgehender Extremalenbogen lässt sich mit einem Felde im *Weierstrass*'schen Sinne umgeben, während im Ausnahmefalle $\Delta(x, x_0)$ identisch verschwindet[33]. Im Hauptfalle lässt sich also (durch die Nullstellen von $\Delta(x, x_0)$) zu jedem Punkte x_0 ein konjugierter Punkt definieren, und die Notwendigkeit des *Jacobi*'schen Kriteriums wird nun auf zwei verschiedenen Wegen bewiesen[34]. Ferner giebt *v. Escherich* durch Betrachtung der zweiten Variation einen Beweis, dass auf einem keine zwei konjugierten Punkte enthaltenden Extremalenbogen ein definites Zeichen der quadratischen Form, in welche die zweite Variation transformiert wird, ein schwaches Extremum zur Folge hat, m. a. W., dass innerhalb des Feldes das *Legendre*'sche Kriterium für ein schwaches Extremum hinreicht[35].

Mit analogen Methoden behandelt *H. Hahn* den Fall, dass unter den Bedingungsgleichungen des Problems auch endliche Gleichungen vorkommen[36].
635 | Einen auf wesentlich anderen Grundlagen beruhenden, allerdings nicht ganz allgemeinen Beweis des *Jacobi*'schen Kriteriums liefert *A. Kneser*[37] durch eine Verallgemeinerung der von ihm im einfachsten Falle benutzten Schlussweise, welche auf der Existenz gewisser Enveloppen beruht.

Es sei endlich noch eine von *H. Hahn*[38] gegebene Methode zur Aufstellung des *Lagrange*'schen Differentialgleichungssystemes erwähnt, welche nicht von vornherein voraussetzt, dass die gesuchte Kurve zweimal differentiierbar sei, sondern aus der Existenz stetiger erster Ableitungen auch die der zweiten Derivierten folgert[39].

Aus allen diesen Untersuchungen geht hervor, dass, solange es sich um einfache Integrale handelt, das Auftreten mehrerer unbekannter Funktionen unter dem Integralzeichen zwar eine grössere Komplikation der Rechnungen

[33] Wien. Ber. 108 (1899), p. 1299; 110 (1901), p. 1405.

[34] Wien. Ber. 110 (1901), p. 1409 ff. Der eine dieser Beweise ist eine Übertragung des *Erdmann-Weierstrass*'schen Beweises (Lehrbuch § 28) auf unser allgemeines Problem; der zweite giebt eine exakte Durchführung der von *L. Scheeffer* (Math. Ann. 25 (1885), p. 522) verwendeten Methode. Die dabei festgehaltene Voraussetzung, dass die Bedingungsgleichungen, denen die Variationen unterliegen, durch die linearen Glieder ihrer *Taylor*'schen Entwickelungen ersetzt werden können, ist im „Hauptfalle" immer zulässig (*H. Hahn*, Math. Ann. 58 (1904), p. 158).

[35] Math. Ann. 55 (1902), p. 108.

[36] Monatsh. Math. Phys. 14 (1903), p. 3.

[37] Charkow Math. Mitt. (2) 7 (1902).

[38] Monatsh. Math. Phys. 14 (1903), p. 325

[39] Diese Problemstellung geht zurück auf *P. du Bois-Reymond*, der den Fall eines Integrales mit n Ableitungen ohne Nebenbedingung untersucht (Math. Ann. 15 (1879), p. 564). Eine dieselbe Frage behandelnde Abhandlung von *E. Zermelo* wird demnächst in den Math. Ann. erscheinen. Für den einfachsten Fall der Variationsrechnung gab *D. Hilbert* eine sehr anschauliche Methode an, die wiedergegeben ist bei *J. R. Whittemore*, Ann. of math. (2) 2 (1901), p. 130 und *N. Gernet*, Diss. Gött. 1902, p. 15.

small extremal arc issuing from x_0 with a field in *Weierstrass*'s sense, whereas $\Delta(x, x_0)$ vanishes identically in the exceptional case.[37] Thus, in the principal case, it is possible to define a conjugate point for every point x_0 (by means of the zero points of $\Delta(x, x_0)$), and that *Jacobi*'s criterion is necessary is now proved in two different ways.[38] Furthermore, by considering the second variation, *v. Escherich* proves that on an extremal arc containing no two conjugate points a definite sign of the quadratic form into which the second variation is transformed has as its consequence a weak extremum, in other words, that within the field *Legendre*'s criterion for a weak extremum is sufficient.[39]

H. Hahn uses analogous methods to treat the case where also finite equations occur under the constraint equations.[40]

A. Kneser[41] gives a proof of *Jacobi*'s criterion that rests on entirely different foundations but is not fully general by generalizing the line of reasoning that he uses in the simplest case and that rests on the existence of certain envelopes.

Finally, we should mention a method presented by *H. Hahn*[42] for determining the *Lagrange* system of differential equations, where it is not assumed from the outset that the sought curve can be differentiated twice but where from the existence of continuous first derivatives follows the existence of the second derivatives in turn.[43]

What emerges from all these investigations is that, while the occurrence of several unknown functions under the integral sign leads to a greater complexity in computations and proofs, the results remain quite analogous to

[37] *von Escherich 1899*, p. 1299; *von Escherich 1901*, p. 1405.

[38] *von Escherich 1901*, p. 1409 ff. One of these proofs is an adaptation of the *Erdmann-Weierstrass* proof (Lehrbuch § 28) to our general problem; the second one precisely implements the method used by *L. Scheeffer* (*1885*). The assumption specified there that the constraint equations governing the variations can be replaced by the linear terms of their *Taylor* expansions is always permissible in the "principal case" (*Hahn 1904*, p. 158).

[39] *von Escherich 1902*.

[40] *Hahn 1903a*.

[41] *Kneser 1902c*.

[42] *Hahn 1903b*.

[43] This problem is due to *P. du Bois-Reymond*, who investigates the case of an integral with n derivatives without ancillary condition (*du Bois-Reymond 1879b*). A paper by *E. Zermelo* addressing the same question is scheduled to appear soon in the Math. Ann. [[*Zermelo 1904a*]]. *D. Hilbert* presented a very clear method for the simplest case of the calculus of variations, which is reproduced in *Whittemore 1901a* and *Gernet 1902*, p. 15.

und Beweise zur Folge hat, dass aber die Resultate den für den Fall einer unbekannten Funktion erhaltenen durchaus analog bleiben. *J. Hadamard*[40] macht nun darauf aufmerksam, dass dies bei mehrfachen Integralen nicht mehr durchaus der Fall ist. So ist es bei mehrfachen Integralen mit mehreren unbekannten Funktionen für das Eintreten eines Extremums nicht notwendig, dass die „reduzierte" quadratische Form, in welche nach *Clebsch*[40a] die zweite Variation transformiert werden kann, definit sei[41].

5. Beispiele und Anwendungen. Die Fruchtbarkeit der von *Weierstrass* eingeführten neuen Methoden zeigt sich besonders darin, dass sie die vollständige Erledigung klassischer Variationsprobleme gestatten, die vorher nur unvollkommen behandelt werden konnten, und umgekehrt geben diese
636 speziellen Aufgaben wieder sehr häufig | Anlass zu neuen Fragestellungen und neuen Gesichtspunkten und damit auch zur Weiterentwickelung der allgemeinen Theorie.

So behandelt *A. Kneser*[42] das *Newton*'sche Problem von der Rotationsfläche kleinsten Widerstandes und zeigt, dass ein schwaches Extremum vorhanden ist, wenn der Rotationskörper an der Spitze unter dem Randwinkel 45° durch eine kreisförmige Scheibe abgestumpft wird[43]. Derselbe Autor erledigt[44] das „Problem der Dido", zwei nicht vorgeschriebene Punkte einer gegebenen Kurve durch eine zweite Kurve von gegebener Länge so zu verbinden, dass der von beiden eingeschlossene Flächeninhalt ein Maximum wird. Das sogenannte „isoperimetrische Problem auf einer Fläche" oder das „Problem der Kurven kürzesten Umrings" wird behandelt von *J. R. Whittemore*[45] und *O. Bolza*[45a]: es soll auf einer gegebenen Fläche die kürzeste Linie gefunden werden, die zwei vorgeschriebene Punkte einer gegebenen Kurve verbindet und mit dieser einen gegebenen Flächeninhalt einschliesst; als Lösung ergeben sich bekanntlich die Kurven konstanter geodätischer Krümmung. Ei-

[40] Bull. soc. math. 30 (1903), p. 253.

[40a] J. f. Math. 56 (1859), p. 122 (vgl. II A 8, Nr. **25**, *Kneser*).

[41] Mit Doppelintegralen beschäftigen sich ferner ausser dem einschlägigen Abschnitte in *Kneser*'s Lehrbuche (§§ 62–69): *D. Hilbert* (Mathem. Probleme (siehe Fussn. 8)) und *W. F. Osgood* (Ann. of math. (2) 2 (1901), p. 125), die auf dieselben den Unabhängigkeitssatz übertragen, und *A. Sommerfeld* (Deutsche Math.-Ver. 8 (1900), p. 188) der das Analogon des *Jacobi*'schen Kriteriums beweist.

[42] Arch. Math. Phys. (3) 2 (1902), p. 267.

[43] Dass im Falle eines Extremums der Winkel zwischen „Stirnfläche" und „Mantelfläche" notwendig 45° beträgt, wurde von *E. Armanini* bewiesen (Ann. di mat. (3) 4 (1900), p. 131). Dort findet sich auch die Litteratur über den Gegenstand zusammengestellt (l. c. p. 149).

[44] Math. Ann. 56 (1903), p. 169.

[45] Ann. of math. (2) 2 (1901), p. 176.

[45a] Decennial publications of the university of Chicago 9 (1902) und Math. Ann. 57 (1903), p. 48. Vgl. auch *Darboux*, Théorie des surfaces 3, Nr. 651 ff. und *Kneser*, Lehrbuch § 34.

those obtained in the case of one unknown function as long as we are dealing with simple integrals. Now, *J. Hadamard*[44] points out that this is no longer always the case for multiple integrals. So, in the case of multiple integrals with several unknown functions it is not necessary for the occurrence of an extremum that the "reduced" quadratic form, into which it is possible to transform the second variation according to *Clebsch*,[45] be definite.[46]

5. Examples and applications. The fruitfulness of the new methods *Weierstrass* introduced is particularly evident in the fact that they allow the complete resolution of classic variational problems that previously were amenable to only incomplete treatment. Conversely, these special problems very often give rise to new questions and new angles, and hence also to the further development of the general theory.

For instance, *A. Kneser*[47] addresses *Newton*'s problem of the surface of revolution with least resistance and shows that there exists a weak extremum if the body of revolution is truncated at the apex below the contact angle 45° by a circular disk.[48] The same author solves[49] the "problem of Dido" of joining two points on a given curve that are not prescribed by a second curve of given length so that the area enclosed by both becomes a maximum. The so-called "isoperimetric problem on a surface" or the "problem of the shortest curves" is treated by *J. K. Whittemore*[50] and by *O. Bolza*:[51] the shortest line on a given surface is sought which joins two prescribed points on a given curve and which, together with the latter, encloses a given area; it is well-known that the problem is solved by the curves of constant geodesic

[44] *Hadamard 1902*.

[45] *Clebsch 1859* (see *Kneser 1904*, no. **25**).

[46] Besides the relevant section in *Kneser*'s Lehrbuch (§§ 62–69), double integrals are also discussed by *D. Hilbert* (*1900a* (see footn. 8)) and *W. F. Osgood 1900/01*, p. 125), who adapt the independence theorem to double integrals, and by *A. Sommerfeld* (*1900a*), who proves the analogue to *Jacobi*'s criterion.

[47] *Kneser 1902b*.

[48] That in the case of an extremum the angle between the "frontal surface" and "lateral surface" is necessarily 45° was proved by *E. Armanini* (*1900*), who also assembled references to the relevant literature there (l. c. p. 149).

[49] *Kneser 1903a*.

[50] *Whittemore 1901b*, p. 176. ⟦Hahn and Zermelo erroneously write "J. R. Whittemore" instead of "J. K. Whittemore".⟧

[51] *Bolza 1902c* and *Bolza 1903b*. See also *Darboux 1894*, no. 651 ff., and *Kneser*, Lehrbuch § 34.

ne ausführliche Theorie der geodätischen Linien auf einer Ringfläche giebt *G. A. Bliss*[46].

Ferner beweist *A. Kneser*[47] die Stabilität des Gleichgewichtes schwerer hängender Fäden. Wird nämlich um die Kettenlinie, welche die potentielle Energie des Fadens zu einem Minimum macht, ein Gebiet S abgegrenzt, und wird der Faden so gebogen, dass er nicht mehr ganz in S liegt, so besitzt die Zunahme seiner potentiellen Energie eine von Null verschiedene untere Grenze[48]. Mit Hilfe dieses Satzes führt die von Systemen mit einer endlichen Anzahl von Freiheitsgraden her bekannte *Dirichlet*'sche Schlussweise (IV 1, Nr. **45**, Fussn. 275, *Voss*) auch hier zum Ziele.

637 Als eine wichtige Ergänzung des *Jacobi*'schen Kriteriums beweist | *G. Darboux*[49] und später *E. Zermelo*[50] einen Satz, nach welchem der Bogen einer geodätischen Linie im allgemeinen schon vor dem nächsten konjugierten Punkte aufhören muss, die „allerkürzeste" Verbindungslinie seiner Endpunkte zu sein, sondern zuletzt nur die kürzeste unter den benachbarten darstellt. *Zermelo* zeigt ferner, dass innerhalb eines begrenzten einfach zusammenhängenden ebenen Bereiches die kürzesten Verbindungslinien zweier Punkte (II A 8, Nr. **23**, *Kneser*)[50a] immer eindeutig bestimmt sind, und er betrachtet schliesslich auf einer vorgelegten Fläche $z = \varphi(x, y)$ die „kürzesten Linien von begrenzter Steilheit" $\left(\frac{dz}{ds} \leqq k\right)$ als Beispiel eines Problems mit Differentialungleichungen. Es treten hier diskontinuierliche Lösungen auf, wie dies bei Ungleichungen in der Regel der Fall zu sein scheint.

Um auch die isoperimetrischen Probleme mit Doppelintegralen der allgemeinen Methode zugänglich zu machen, behandelt *J. O. Müller*[51] als wichtigstes Beispiel die schon von *H. A. Schwarz*[52] bewiesene Eigenschaft der Kugel, eine kleinere Oberfläche zu besitzen als alle geschlossenen Flächen vom gleichen Volumen, indem er, entsprechend dem *Weierstrass*'schen Ideengange, ein Feld von Kugeln konstruiert und mittelst desselben den Wert der Kugeloberfläche darstellt durch ein Flächenintegral über eine beliebige geschlossene Vergleichsfläche gleichen Volumens, wobei der Integrand beständig < 1 ist. Dasselbe Theorem ergiebt sich übrigens bei *H. Minkowski*[53] als Spezialfall eines viel allgemeineren Satzes über Volumina und Oberflächen konvexer Körper.

---

[46] Ann. of math. (2) 4 (1902), p. 1.
[47] J. f. Math. 125 (1903), p. 189.
[48] Dieser Satz ist vollständig analog dem in Fussn. 11) angegebenen Satze von *Osgood*.
[49] Théorie des surfaces 3, Nr. 623.
[50] Deutsche Math.-Ver. 11 (1902), p. 184.
[50a] Vgl. *Darboux* l. c. Nr. 632 und *Kneser*, Lehrbuch § 44.
[51] Gött. Nachr. 1902, p. 176; Diss. Gött. 1903.
[52] Gött. Nachr. 1884, p. 1 = Ges. Abh. 2, p. 327.
[53] Deutsche Math.-Ver. 9 (1901), p. 115; Math. Ann. 57 (1903), p. 474. Dabei sind als Vergleichskörper nur konvexe Körper zugelassen.

curvature. An elaborate theory of geodesic lines on a ring surface is provided by *G. A. Bliss*.[52]

In addition, *A. Kneser*[53] proves the stability of the equilibrium of heavy strings held in suspension. For if we delineate a region S about the catenary curve that turns the potential energy of the string into a minimum and bend the string so that it does not entirely lie in S any longer, then the increase in its potential energy has a lower limit different from zero.[54] With the aid of this theorem, *Dirichlet*'s line of reasoning (*Voss 1901*, no. **45**, footn. 275), which is familiar from systems with a finite number of degrees of freedom, leads to the desired result also in this case.

Significantly supplementing *Jacobi*'s criterion, *G. Darboux*,[55] and later *E. Zermelo*,[56] prove a theorem according to which the arc of a geodesic line must in general cease to be the "by far shortest" connecting line of its endpoints already before the next conjugate point, but only is the shortest among the neighboring curves eventually. In addition, *Zermelo* shows that within a bounded simply connected planar domain the shortest connecting lines of two points (*Kneser 1904*, no. **23**)[57] are always uniquely determined, and he finally considers the "shortest lines of limited steepness" $\left(\frac{dz}{ds} \leqq k\right)$ on a prescribed surface $z = \varphi(x, y)$ as an example of a problem with differential inequalities. In this case, discontinuous solutions occur, as it seems to be the rule with inequalities.

In order to make the isoperimetric problems with double integrals amenable to treatment by the general method, *J. O. Müller*[58] takes up as the most important example the property of the sphere, which was already proved by *H. A. Schwarz*,[59] of having a smaller surface area than all closed surfaces of the same volume by constructing a field of spheres, in accordance with *Weierstrass*'s ideas, by means of which he represents the magnitude of the sphere area by a surface integral over any closed comparison surface of the same volume, where the integrand is always < 1. Incidentally, the same theorem occurs in *H. Minkowski*[60] as a special case of a much more general theorem on volumes and surface areas of convex bodies.

---

[52] *Bliss 1902b.*

[53] *Kneser 1903b.*

[54] This theorem is completely analogous to *Osgood*'s theorem stated in footn. 11.

[55] *Darboux 1894*, no. 623.

[56] *Zermelo 1902d.*

[57] See *Darboux 1894*, no. 632, and *Kneser*, Lehrbuch § 44.

[58] *Müller 1902; 1903.*

[59] *Schwarz 1884* = *Schwarz 1890b*, p. 327 ff.

[60] *Minkowski 1900; 1903*, where the only bodies admitted for purposes of comparison are convex bodies.

Eine Aufgabe die als Umkehrung eines Problemes der Variationsrechnung bezeichnet werden kann, löst *G. Hamel*[54], indem er alle Geometrien, d. h. alle Massbestimmungen in der Ebene aufstellt, für welche die geraden Linien die kürzesten sind. Es kommt dies darauf hinaus, das allgemeinste Integral $\int f(x, y, y')\, dx$ aufzufinden, welches durch die Geraden der Ebene zu einem Minimum gemacht wird. Hierdurch ist der Anschluss an diejenigen Untersuchungen gewonnen, die zu einer vorgelegten Differentialgleichung die zugehörigen Variationsprobleme suchen[55]. *Hamel* geht insofern über diese hinaus,
638 als er | verlangt, dass auch die hinreichenden Bedingungen der Variationsrechnung erfüllt seien.

Eine Anwendung der Variationsrechnung auf die Theorie der partiellen Differentialgleichungen geben *Yoshiye*[56] und *E. R. Hedrick*[57], indem sie, einen von *D. Hilbert* in seinen Vorlesungen gegebenen Gedanken ausführend, nachweisen, dass die Gleichungen der charakteristischen Streifen einer partiellen Differentialgleichung erster oder zweiter Ordnung immer übereinstimmen mit den *Lagrange*'schen Gleichungen eines gewissen Variationsproblemes.

6. Existenzfragen. Während die klassische Methode der Variationsrechnung von einer gegebenen Lösung der *Lagrange*'schen Differentialgleichung ausgeht, das Randwertproblem also als gelöst betrachtet und nun untersucht, ob diese Lösung wirklich ein Extremum liefert, stellt sich *D. Hilbert*[58] die Aufgabe, in gewissen Fällen von vornherein die Existenz einer

[54] Diss. Gött. 1901; Math. Ann. 57 (1903), p. 231.

[55] Für den Fall einer gewöhnlichen Differentialgleichung zweiter Ordnung ist diese Theorie durchgeführt bei *Darboux*, Théorie des surfaces 3, Nr. 604–606, wo u. a. der Satz bewiesen wird, dass zu jeder Gleichung $y'' = \varphi(x, y, y')$ sich unendlich viele Integrale $\int f(x, y, y')\, dx$ finden lassen, deren Extremalen die Lösungen der Gleichung sind (vgl. hierüber auch *Kürschák*, Math. Ann. 56 (1903), p. 163). — *A. Hirsch* (Math. Ann. 49 (1897), p. 49) beweist weiter: Wenn der aus der Gleichung $F\big(x, y, y', \ldots, y^{(2n)}\big) = 0$ abgeleitete lineare Differentialausdruck $\sum_{k=0}^{n} \frac{\partial F}{\partial y^{(k)}} u^{(k)}$ „sich selbst adjungiert" ist (II A 4 b, Nr. **26**, *Vessiot*), lässt sich durch Quadraturen ein Integral $\int f(x, y, y', \ldots, y^{(n)})\, dx$ finden, dessen Extremalen durch die Gleichung $F = 0$ gegeben sind. Es wird ferner die Gültigkeit eines analogen Satzes für partielle Differentialausdrücke der zweiten Ordnung mit zwei oder drei unabhängigen Veränderlichen bewiesen; die Vermutung, es gelte dies für eine beliebige Anzahl unabhängiger Veränderlicher, wurde von *W. Hertz* (Diss. Kiel 1903) für $n = 4, 5$ bestätigt. Weitere Litteratur dieses Gegenstandes giebt *Hamel*. — Hierher gehört auch das von *A. Guldberg* (Christiania Skrift. 1902, Nr. 7) behandelte Problem, das allgemeinste Integral $\int f(x, y, y')\, dx$ aufzufinden, das invariant bleibt bei einer kontinuierlichen Gruppe von Transformationen. Es zeigt sich, dass für solche Integrale die *Lagrange*'sche Gleichung sich auf Quadraturen oder auf eine Gleichung erster Ordnung zurückführen lässt.

[56] Math. Ann. 57 (1903), p. 185.

[57] Ann. of math. (2) 4 (1903), p. 141, 157.

[58] Deutsche Math.-Ver. 8 (1900), p. 184.

G. Hamel[61] solves a problem that can be considered the converse of a problem of the calculus of variations by specifying all geometries, i.e., all measure determinations in the plane for which the straight lines are the shortest ones. This is a matter of finding the most general integral $\int f(x, y, y')\,dx$ that is made a minimum by the straight lines of the plane. Thus the connection is made with the investigations that seek to determine the variational problems belonging to a given differential equation.[62] *Hamel* goes beyond them insofar as he demands that the sufficient conditions of the calculus of variations be also satisfied.

Using one of the ideas presented in *D. Hilbert*'s lectures *Yoshiye*[63] and *E. R. Hedrick*[64] apply the calculus of variations to the theory of partial differential equations by showing that the equations of the characteristic bands of a partial differential equation of first or second order are always in accord with the *Lagrange* equations of a certain variational problem.

6. Existence questions. While the classic method of the calculus of variations proceeds from a given solution of the *Lagrange* differential equation, and hence considers the boundary value problem as solved, and then examines whether this solution really yields an extremum, *D. Hilbert*[65] addresses the problem of proving in certain cases at the outset the existence

[61] *Hamel 1901; 1903.*

[62] For the case of an ordinary differential equation of second order, this theory is worked out in *Darboux 1894*, no. 604–606, which, among other things, contains the proof of the theorem that for any equation $y'' = \varphi(x, y, y')$ it is possible to find infinitely many integrals $\int f(x, y, y')\,dx$ whose extremals are the solutions of the equation (on this, see also *Kürschák 1903*, p. 163).—Furthermore, *A. Hirsch* (*1897*) proves that if the linear differential expression $\sum_{k=0}^{n} \frac{\partial F}{\partial y^{(k)}} u^{(k)}$ derived from the equation $F\big(x, y, y', \ldots, y^{(2n)}\big) = 0$ is "adjoint to itself" (*Vessiot 1900*, no. **26**), then one can find by quadratures an integral $\int f(x, y, y', \ldots, y^{(n)})\,dx$ whose extremals are given by the equation $F = 0$. Furthermore, the validity of an analogous theorem for partial differential expressions of second order with two or three independent variables is proved; the conjecture that this is valid for any number of independent variables was confirmed by *W. Hertz* (*1903*) for $n = 4, 5$. Further literature on this subject is given by *Hamel* [[cf. fn. 61]].—In this connection, we should also mention the problem, treated by *A. Guldberg* (*1902*), of finding the most general integral $\int f(x, y, y')\,dx$ that remains invariant under a continuous group of transformations. It turns out that for such integrals the *Lagrange* equation can be reduced to quadratures or to a first-order equation.

[63] *Yoshiye 1903.*

[64] *Hedrick 1903*, p. 141, 157.

[65] *Hilbert 1900b.*

Kurve oder Fläche von grösstem oder kleinstem Integralwerte zu beweisen, welche dann auf Grund dieser Eigenschaft die Differentialgleichung unter den gegebenen Randbedingungen erfüllen muss und damit einen Schluss auf die Lösbarkeit des Randwertproblemes gestattet.

639 | Seinen Gedankengang entwickelt *Hilbert* zunächst an dem Beispiel der kürzesten Linien auf Flächen. Da die Bogenlängen aller zwei gegebene Punkte der Fläche verbindenden Kurven eine untere Grenze l besitzen, so existiert eine unendliche Folge von Kurven, deren Bogenlängen diesen Wert l zur Grenze haben. Diese Kurven konvergieren aber gegen eine Grenzkurve, von der sich beweisen lässt[59], dass sie differentiierbar ist und die minimale Bogenlänge l besitzt. Dieselbe Schlussweise überträgt *Ch. A. Noble*[60] unter gewissen einschränkenden Voraussetzungen auf einfache Integrale, die eine oder zwei unbekannte Funktionen und deren erste Ableitungen enthalten.

Hilbert[61] selbst verwendet seine Methode zu einer strengen Begründung der als „*Dirichlet*'sches Prinzip" bekannten Beweismethode (II A 7 b, Nr. **24** u. **25**, *Burkhardt* u. *Meyer*), auf Grund deren *Riemann* die Existenz der überall endlichen Integrale auf einer vorgelegten *Riemann*'schen Fläche erschloss (II B 2, Nr. **12**, *Wirtinger*). Es handelt sich hier darum, die Existenz einer auf der ganzen *Riemann*'schen Fläche regulären Potentialfunktion $u(x, y)$ nachzuweisen, welche entlang einer geschlossenen, die Fläche nicht zerstückelnden Kurve C, die man immer aus geraden, den Koordinatenaxen parallelen Stücken bestehend annehmen kann, den Sprung 1 erleidet. Bildet man für alle beliebigen dieser Randbedingung genügenden Funktionen das über die ganze Fläche erstreckte „*Dirichlet*'sche Integral":

$$\iint \left\{ \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 \right\} dx\, dy \ ,$$

so haben die Integralwerte eine untere Grenze d. Es sei nun $U_1, U_2, U_3, \ldots$ eine Folge von Funktionen, deren *Dirichlet*'sche Integrale die Grenze d ergeben, und die etwa so gewählt seien, dass die über ein gewisses Stück der
640 x-Axe erstreckten Integrale $\int U_n\, dx$ sämtlich den | Wert Null haben. Dann lässt sich aus ihnen eine andere Folge u_1, u_2, u_3, ... so auswählen, dass für

[59] Eine Ausführung des Beweises giebt *Ch. A. Noble* (Diss. Gött. 1901, p. 10) auf Grund der Thatsache, dass zwei genügend nahe Punkte der Fläche sich immer durch eine geodätische Linie verbinden lassen, die auch die kürzeste ist. Ein hiervon unabhängiger Beweis liesse sich mit den von *Hilbert* zum Beweise des *Dirichlet*'schen Prinzips verwendeten Methoden führen. — Aus der Differentiierbarkeit der gefundenen Kurve folgt unmittelbar, dass sie eine geodätische Linie sein muss, und somit auf einer analytischen Fläche selbst analytisch ist.

[60] Diss. Gött. 1901.

[61] Über das *Dirichlet*'sche Prinzip, Gött. Festschrift, Berlin 1901. Bezüglich der Anwendung auf die gewöhnliche Randwertaufgabe der Potentialtheorie vgl. *D. Hilbert*, Deutsche Math.-Ver. 8 (1900), p. 186. Doch ist die dort skizzierte Methode etwas verschieden von der im Texte angegebenen. Eine etwas weitere Ausführung dieser älteren *Hilbert*'schen Methode bei *E. R. Hedrick*, Diss. Gött. 1901, p. 69 ff.

of a curve or surface of greatest or smallest integral value which then, by virtue of this property, must satisfy the differential equation under the given boundary conditions, thereby licensing the inference to the solvability of the boundary value problem.

At first, *Hilbert* develops his line of thought using the example of the shortest lines on surfaces. Since the arc lengths of all curves joining two given points of the surface possess a lower limit l, there exists an infinite sequence of curves whose arc lengths have this value l as a limit. But these curves converge to a limit curve of which it is possible to prove[66] that it is differentiable and possesses the minimal arc length l. *Ch. A. Noble*[67] adapts this line of reasoning under certain restricting assumptions to simple integrals containing one or two unknown functions and their first derivatives.

Hilbert[68] himself uses his method for a rigorous foundation of the proof method known as "*Dirichlet*'s principle" (*Burkhardt and Meyer 1900*, no. **24** a. **25**), on the basis of which *Riemann* showed the existence of the everywhere finite integrals on a given *Riemann* surface (*Wirtinger 1901*, no. **12**). This is a matter of showing the existence of a potential function $u(x, y)$ that is regular on the entire *Riemann* surface and that suffers the jump discontinuity 1 along a closed curve C which does not cut up the surface and which we may always assume to be composed of straight segments parallel to the coordinate axes. If we form for any functions satisfying this boundary condition the "*Dirichlet* integral" extended over the entire surface:

$$\iint \left\{ \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 \right\} dx\, dy \,,$$

then the integral values have a lower limit d. Now suppose that $U_1, U_2, U_3, \ldots$ is a sequence of functions whose *Dirichlet* integrals yield the limit d and which are chosen so that the integrals $\int U_n\, dx$ extended over a certain segment of the x-axis all have the value zero. Then it is possible to select from them

[66] *Ch. A. Noble* (*1901*, p. 10) carries out this proof by relying on the fact that two sufficiently close points of the surface can always be joined by a geodesic line that is also the shortest line. It is possible to provide a proof independent of this fact by means of the methods used by *Hilbert* in his proof of *Dirichlet*'s principle.—It immediately follows from the differentiability of the curve found that it must be a geodesic line, and hence that, on an analytic surface, it itself is analytic.

[67] *Noble 1901.*

[68] *Hilbert 1901b.* As for the application to the ordinary boundary value problem of potential theory, see *Hilbert 1900b*, p. 186. But the method outlined there is somewhat different from the one presented in the text. A somewhat broader outline of this older method by *Hilbert* can be found in *Hedrick 1901b*, p. 69 ff.

beliebige (a, b) der Grenzwert:

$$v(x,y) = \lim_{n=\infty} \int_a^x \int_b^y u_n \, dx \, dy$$

existiert und eine stetige Funktion von x und y darstellt, solange das Integrationsgebiet keinen Verzweigungspunkt enthält und von der Kurve C nicht geschnitten wird. Der von (a, b) unabhängige Ausdruck:

$$u(x,y) = \frac{\partial^2 v}{\partial x \partial y} = \lim_{\varepsilon=0,\ \eta=0} \lim_{n=\infty} \frac{1}{\varepsilon\eta} \int_x^{x+\varepsilon} \int_y^{y+\eta} u_n \, dx \, dy$$

ist dann die gesuchte Potentialfunktion, wie folgendermassen bewiesen wird. Für eine willkürliche Funktion ζ und ein beliebiges Rechteck R muss die Gleichung bestehen:

$$\lim_{n=\infty} \iint_{(R)} \left(\frac{\partial \zeta}{\partial x} \frac{\partial u_n}{\partial x} + \frac{\partial \zeta}{\partial y} \frac{\partial u_n}{\partial y} \right) dx \, dy = 0 \, ,$$

welche hier dieselbe Rolle spielt, wie das Verschwinden der ersten Variation in der klassischen Variationsrechnung. Durch partielle Integrationen und Ausführung des Grenzüberganges unter den Integralzeichen erhält man den Satz, dass eine Funktion w, für welche $\dfrac{\partial^4 w}{\partial x^2 \partial y^2} = v(x, y)$ ist, für willkürliche ζ der Gleichung genügen muss:

$$\iint_{(R)} \frac{\partial^6 \zeta}{\partial x^3 \partial y^3} \Delta w \, dx \, dy = 0 \, ,$$

und hieraus folgt[62], dass Δw die Form haben muss:

$$\Delta w = Xy^2 + X'y + X'' + Yx^2 + Y'x + Y'' \, ,$$

wo X, X', X'' nur von x abhängen, Y, Y', Y'' nur von y. Durch Hinzufügung leicht zu bildender Integrale kann man daher aus w eine Funktion z ableiten, die der Gleichung $\Delta z = 0$ genügt und mit u und v durch die Beziehung verbunden ist:

$$\frac{\partial^6 z}{\partial x^3 \partial y^3} = \frac{\partial^2 v}{\partial x \partial y} = u(x,y) \, ,$$

wodurch die Potentialeigenschaft von u erwiesen ist.

[62] l. c. § 4. Dieses Beweisverfahren zeigt eine große Analogie mit dem von *P. du Bois-Reymond* für einfache Integrale angewandten (siehe Fussn. 39), geht aber wesentlich weiter, da es auch die Existenz erster Derivierter der Minimalfunktion zu erschliessen gestattet.

a different sequence u_1, u_2, u_3, ... so that for any (a, b) the limit

$$v(x,y) = \lim_{n=\infty} \int_a^x \int_b^y u_n \, dx \, dy$$

exists and constitutes a continuous function of x and y as long as the domain of integration contains no branching point and is not intersected by the curve C. The expression

$$u(x,y) = \frac{\partial^2 v}{\partial x \partial y} = \lim_{\varepsilon=0,\ \eta=0} \lim_{n=\infty} \frac{1}{\varepsilon\eta} \int_x^{x+\varepsilon} \int_y^{y+\eta} u_n \, dx \, dy$$

independent of (a, b) is then the sought potential function, which is proved as follows. For any function ζ and any rectangle R the equation

$$\lim_{n=\infty} \iint_{(R)} \left(\frac{\partial \zeta}{\partial x} \frac{\partial u_n}{\partial x} + \frac{\partial \zeta}{\partial y} \frac{\partial u_n}{\partial y} \right) dx \, dy = 0$$

must hold, which plays the same part here as the vanishing of the first variation in the classic calculus of variations. By using partial integration and taking the limit under the integral sign, one obtains the theorem that a function w for which $\dfrac{\partial^4 w}{\partial x^2 \partial y^2} = v(x,y)$ must satisfy the equation

$$\iint_{(R)} \frac{\partial^6 \zeta}{\partial x^3 \partial y^3} \Delta w \, dx \, dy = 0$$

for any ζ, and from this it follows[69] that Δw must have the form

$$\Delta w = Xy^2 + X'y + X'' + Yx^2 + Y'x + Y'' \, ,$$

where X, X', X'' only depend on x and Y, Y', Y'' only on y. By adding readily formed integrals it is thus possible to derive a function z from w that satisfies the equation $\Delta z = 0$ and is connected with u and v by the relation

$$\frac{\partial^6 z}{\partial x^3 \partial y^3} = \frac{\partial^2 v}{\partial x \partial y} = u(x,y) \, ,$$

whereby the potential property of u is demonstrated.

[69] l. c. § 4. This proof procedure exhibits a strong analogy with the one used by *P. du Bois-Reymond* for simple integrals (see footn. 43), but goes significantly beyond it in making it also possible to show the existence of first derivatives of the minimal function.

641 Nach *Hilbert*'schen Prinzipien behandelt *E. R. Hedrick*[63] die | Randwertaufgabe für die *Liouville*'sche Gleichung $\Delta u = e^u$ (II A 7c; Nr. **12**, *Sommerfeld*). *Ch. M. Mason*[64] gelangt durch Betrachtung des einfachen Integrales $\int_a^b y'^2\,dx$ mit der Nebenbedingung:

$$\int_a^b A(x)y^2\,dx = 1 \qquad (A(x) > 0)$$

zu einer Behandlung der gewöhnlichen linearen Differentialgleichung zweiter Ordnung $y'' + \lambda A(x)y = 0$ (II A 7 a, *Bôcher*). Er bestätigt die in der Physik aus blossen Analogiegründen zugelassene Annahme, dass zu jeder „homogenen" Randbedingung unendlich viele „ausgezeichnete" Werte von λ gehören, und insbesondere, dass es bei periodischem $A(x)$ unendlich viele Werte von λ giebt, für welche die Gleichung periodische Lösungen besitzt[65].

(Abgeschlossen im Januar 1904.)

[63] Diss. Gött. 1901, p. 64 ff.

[64] Diss. Gött. 1903, p. 29 ff. Diese Beweismethode ist nachgebildet der von *H. Weber* (Math. Ann. 1 (1869), p. 1) auf die partielle Differentialgleichung der schwingenden Membran (II A 7 c, Nr. **9**, *Sommerfeld*) angewandten Betrachtungsweise. Ihre Durchführung bei *Mason* ist aber nicht überall einwandfrei.

[65] l. c. p. 52 ff.

E. R. Hedrick[70] treats the boundary value problem for the *Liouville* equation $\Delta u = e^u$ (*Sommerfeld 1900b*, no. **12**) in accordance with *Hilbert*'s principles. *Ch. M. Mason*[71] approaches the treatment of the ordinary linear differential equation of second order $y''+\lambda A(x)y = 0$ (*Bôcher 1900*) by considering the simple integral $\int_a^b y'^2\,dx$ with the ancillary condition

$$\int_a^b A(x)y^2\,dx = 1 \qquad (A(x) > 0)\ .$$

He confirms the assumption, which, in physics, is admissible for reasons of analogy only, that to every "homogeneous" boundary condition there belong infinitely many "distinguished" values of λ, and in particular, that for periodic $A(x)$ there are infinitely many values of λ for which the equation has periodic solutions.[72]

(Completed in January 1904.)

[70] *Hedrick 1901b*, p. 64 ff.

[71] *Mason 1903*, p. 29 ff. This proof procedure is modeled on the approach applied by *H. Weber* (*1869*) to the partial differential equation of the oscillating membrane (*Sommerfeld 1900b*, no. **9**). Its execution in *Mason* is not always without flaw.

[72] l. c. p. 52 ff.

Introductory note to 1906

Jos Uffink

Josiah Willard Gibbs's book *1902* provided an approach to the foundations of statistical physics that differs in many respects from Boltzmann's. Perhaps the foremost difference, at least for present concerns, is that Gibbs aimed to provide what he called a "rational" foundation to the statistical approach, while refraining from "solving the mysteries of nature". Indeed, Gibbs's book reads like an exposition of the rigorous theorems of a theory that he baptized "statistical mechanics", assuming only the most general assumptions needed for such an account.[1] For example, the assumption that the systems under consideration are actually composed of a finite number of molecules is only introduced in the final chapter. And whenever Gibbs discusses the consequences that this theory would have for the reduction of thermodynamics, these discussions are embedded in thoughtful reservations, pointing out that they merely suggest certain analogies between the two theories—never a bold statement that he had shown that a law of thermodynamics could rigorously be reduced to a theorem in another theory, as Boltzmann did on various occasions. For example, Gibbs emphasized that the analogies in question are far from unique. His argument was rather that the existence of such analogies could make it seem plausible that thermodynamical relations are not incompatible with the statistical-mechanical approach. Roughly speaking, Gibbs was a cautious theoretical physicist who kept a keen eye on mathematical rigor, whereas Boltzmann, at times, boldly announced claims that overstated what he had actually proved. It seems to me that Zermelo must have been much more sympathetic to Gibbs's approach than to Boltzmann's, since Gibbs's goal seems to coincide with the goal stated in Zermelo's *1896a*, i.e. to state clearly what can and cannot be proved. More extensive discus-

[1] The distinction between "statistical mechanics" proper and the more general concept of "statistical physics" is somewhat subtle. It is typical for statistical mechanics that one assumes a mechanical phase space for the entire system considered (whether this be a gas, liquid or another macroscopic body), together with the Hamiltonian equations of motion, and a probability measure defined on this phase space. This stands in contrast to other approaches within statistical physics, for example the kinetic theory of gases, where probability considerations are applied on the level of a single molecule. Although Gibbs invented the term "statistical mechanics" and developed this theory systematically, this is not to say that he originated the approach. Indeed, several of Boltzmann's papers took the same approach, and so did Zermelo. However, Boltzmann never clearly distinguished the cases where he was pursuing statistical mechanics, as we now use this term, from cases where he was pursuing kinetic theory or some other set of assumptions within statistical physics.

H.-D. Ebbinghaus, A. Kanamori (Eds.), *Ernst Zermelo - Collected Works/ Gesammelte Werke II*, DOI 10.1007/978-3-540-70856-8_11, Schriften der Mathematisch-naturwissenschaftlichen Klasse der Heidelberger Akademie der Wissenschaften 23, © Springer-Verlag Berlin Heidelberg 2013
Introductory Note, © Springer-Verlag, Paper, © B.G. Teubner
English translation with kind permission by Springer-Verlag

sions of the difference between the Boltzmannian and Gibbsian approaches are offered by *Frigg 2008*.

Zermelo produced the first German translation of Gibbs's book. I do not know whether this translation was commissioned by the publisher or whether it was undertaken on Zermelo's own initiative. In the latter case, one would suppose that the effort expended would indicate a significant affinity that Zermelo must have felt towards Gibbs's work. But in any case, Zermelo's appreciation for Gibbs is evident from the words he wrote in the preface to the translation (Zermelo in *Gibbs 1905*, iii–iv):

> This is the first effort to develop rigorously and on a secure mathematical basis the statistical and probabilistic considerations in mechanics which are indispensable in various areas of physics, especially in the kinetic theory of gases. As such, this work is of paramount interest and of permanent value, even if not all the derivations by the author should turn out to be tenable. Which objections I have to make, from my side, in particular against the considerations in Chapter 12, in as far as these attribute to mechanical systems a tendency towards a state of statistical equilibrium, and on this basis, a complete analogy with thermodynamical systems, I intend to explicate on another occasion. [. . .]
> May this work in its new form contribute to raise the interest for and understanding of statistical mechanical problems amongst German physicists and mathematicians.[2]

The "another occasion" to which this quote refers is Zermelo's own review *1906* of Gibbs's book. After a clear statement of the problems and purpose of statistical mechanics as set out by Gibbs, Zermelo argues that similar arguments had been used earlier in the theory of gases, but that their application "was not always unobjectionable" (p. 234). This, of course, is a veiled reference to Boltzmann. By contrast, Gibbs is praised for developing

[2] "Es ist der erste Versuch, die statistischen und die Wahrscheinlichkeits-Betrachtungen in der Mechanik, wie sie auf verschiedenen Gebieten der Physik, namentlich aber in der kinetischen Gastheorie unentbehrlich sind, unabhängig von ihrem Anwendungsgebiet auf sicherer Grundlage mathematisch streng zu entwickeln, und ist als solcher von hervorragendem Interesse und bleibender Bedeutung, auch wenn sich nicht alle Deduktionen des Verfassers als haltbar erweisen sollten. Welche Einwendungen ich meinerseits namentlich gegen die Betrachtungen des zwölften Kapitels zu machen habe, soweit sie den mechanischen Systemen eine Tendenz nach einem Zustande statistischen Gleichgewichts und auf Grund dessen eine vollständige Analogie mit den thermodynamischen Systemen im Sinne des zweiten Wärmesatzes zuschreiben, gedenke ich demnächst an anderer Stelle auszuführen. [. . .]
Möge das Werk in seiner neuen Form dazu beitragen, das Interesse und Verständnis für statistisch-mechanische Probleme bei deutschen Physikern und Mathematikern zu fördern."

the foundations of the subject, independently of the intended domain of application, on a rigorous mathematical basis. Nevertheless, Zermelo has two main objections against Gibbs's work. The first has to do with the analogies that Gibbs aims to point out between relations holding in statistical mechanics and those in thermodynamics. The second deals with his treatment of irreversible processes.

In Gibbs's approach, as is well known, one considers a probability measure, say ρ, which may be represented by a density function P on the phase space Γ, such that

$$\rho(A) = \int_A P(x) d\mu(x)$$

represents the probability of finding the state of the system at some given initial time in a measurable set $A \subseteq \Gamma$. (Zermelo writes $dp_1 \cdots dq_n$ for $d\mu(x)$, the Lebesgue measure on Γ in the canonical coordinates.)

This probability is physically interpreted by reference to an *ensemble*, i.e. one imagines a great number N of replicas of the system under consideration with arbitrary initial states and supposes that the relative number of systems that have their initial state in some measurable region A converges, as $N \to \infty$, to $\int_A P(x) d\mu(x)$.

The average or expected value of any phase space function $f(x)$ over this ensemble is then given by

$$\overline{f} = \int_\Gamma f(x) P(x) d\mu(x)\,. \tag{1}$$

In general, the probability density function P will change with time, in accordance with the flow generated by the Hamiltonian equations of motion. But in certain cases, it can be invariant under time evolution. These cases Gibbs called *statistical equilibrium*. In general, this will happen if P depends only on phase space functions that are conserved by the Hamiltonian flow. And since the Hamiltonian H itself is always conserved under the Hamiltonian flow, any probability density that depends only on H will be in statistical equilibrium.

Gibbs devoted special attention to a particular family of probability density functions P in statistical equilibrium which he called *canonical*, which in modern notation would be written as:

$$P(x) = \frac{1}{Z(\beta)} e^{-\beta H(x)} \tag{2}$$

where H is the Hamiltonian, β an arbitrary parameter related to the temperature, and $Z(\beta)$ a normalization factor ensuring that P is indeed a probability density, i.e. $Z(\beta) := \int_\Gamma e^{-\beta H(x)} d\mu(x)$. In Gibbs's own notation, also used by Zermelo, this corresponds to

$$P(x) = e^{\frac{\psi - \epsilon}{\theta}}\,, \tag{3}$$

where ϵ denotes the Hamiltonian H, $\theta = \frac{1}{\beta}$, and $\psi = \ln Z(\beta)$.

For these canonical distributions (or the ensembles they represent), Gibbs argued that one could obtain relations which were analogous to results in equilibrium thermodynamics if we consider θ analogous to the system's absolute temperature, and

$$\overline{\eta} = \int P(x) \ln P(x) d\mu(x) \tag{4}$$

as analogous to (minus) its entropy.

In particular, assuming that the Hamiltonian depends smoothly on some parameters a_i $(i = 1, \ldots, n)$, and if $A_i := \dfrac{\partial A}{\partial a_i}$, Gibbs shows that

$$d\overline{H} = -\theta d\overline{\eta} - \sum_i \overline{A_i} da_i \tag{5}$$

which, at least typographically, is similar to Gibbs's own (*1875*) formulation of the fundamental equation of thermodynamics, which (in the special case of fluids) reads

$$dE = TdS - pdV\,. \tag{6}$$

So, if we consider θ to be analogous to temperature, $-\overline{\eta}$ to the entropy S, take the Hamiltonian to depend only on a single parameter, i.e. the volume V, and the pressure p to be analogous to $\overline{\partial H/\partial V}$, we recover this fundamental equation of thermodynamics in statistical mechanics, at least for this special case of a fluid described by a canonical ensemble.

However, Zermelo noticed serious problems with this proposed analogy (p. 237):

> But this analogy remains entirely superficial and cannot clarify the mechanical character of the thermodynamical equation. For the magnitude $\overline{\eta}$ [...] is not at all a property of the individual mechanical systems.

Indeed, while the thermodynamical relation (6) is supposed to relate the differential change of energy dE of an individual system to the changes in its entropy S and volume V, some quantities appearing in the suggested analogy (5), in particular θ and $\overline{\eta}$, are only defined by recourse to the envisaged ensemble. But any individual system can be regarded as being part of any ensemble whatsoever, and properties of the ensemble need not reflect properties of an individual system.

This remark by Zermelo is a penetrating one, and points out a deep problem in Gibbs's approach that still remains unsolved today. The point here is not just that there is a distinction between an individual system and an ensemble. For example, the energy of an individual system, as characterized by statistical mechanics, would be $H(x)$, where $x \in \Gamma$ is the state of the individual system. The putative analogy (5), instead, involves the ensemble average $\overline{H}$, which might very well be very different from the energy value $H(x)$

of the individual system. In general, an individual system need not resemble an average system. Similar remarks could be made about the pressure of the individual system $\partial H(x)/\partial V$ and the ensemble average $\overline{\partial H/\partial V}$.

But *this* point was considered by Gibbs: He showed that for systems with many constituents, the canonical probability density, considered as a function of energy, becomes increasingly peaked, and that the standard deviation in energy will approach zero if the number of constituents goes to infinity. So, if the system is large enough, the probability that (say) $H(x)$ deviates more than a tiny amount from $\overline{H}$ is negligible.

However, the problem that Zermelo pointed to is much deeper: His point, as I read it, is that θ, the putative candidate for temperature, is not even a function on phase space. θ is just a parameter in the canonical probability density (3) on Γ, and so for an individual system in a state $x \in \Gamma$, there is no way of determining what its value is. The quantity θ does not even refer to a property that the individual system might have. Something similar holds for $\eta(x)$ and $\overline{\eta}$. Here, perhaps, one might say that $\eta(x) = \ln P(x)$ actually does define a function on Γ and that each individual system in state x will have the corresponding value $\eta(x)$. But the phase space function η is not fixed by the mechanical constitution of the system, like the Hamiltonian H or the pressure $\partial H/\partial V$. It is only defined by recourse to the ensemble from which the individual system is taken to be a member. But any individual system can be considered as a member of many different ensembles, even if we just focus on canonical ensembles. Consequently, we cannot determine the value of $\eta(x)$ for an individual system in state x unless we are told which ensemble it is supposed to belong to, so that we can specify the form of P in $\eta(x) := \ln P(x)$. The value $\eta(x)$ is just not a mechanical property of an individual system in state x, i.e. a property that can be ascertained independently of the choice of ensemble. And, of course, a demonstration that, for canonical ensembles of systems whose number of constituents increases to infinity, the standard deviation in η goes to zero, as Gibbs provided, does nothing to alleviate this point.

Of course, I admit that it is not completely clear that Zermelo's briefly stated objection actually meant to point out this particular problem. However, if the above reading of what Zermelo was pointing out is even roughly correct, it shows a remarkably deep insight by Zermelo into the foundational problems of Gibbs's approach. I would not know of any other author in the early 20th century that expressed awareness of this problem. Whether Gibbs himself was aware of this problem is quite unclear. He expressed, as far as I can tell, no concerns about this issue, but one could imagine it might be one of the reasons why he was so cautious in calling the relationship between (6) and (5) a mere "analogy".

Gibbs's discussion of this topic might well be compared to Boltzmann's discussion in *1871b* of a similar issue, to which Gibbs refers several times in his book. Boltzmann, in effect, had already introduced in this paper what we

today (following Gibbs) call the canonical ensemble and applied it to study infinitesimal changes in the expected value of energy, to obtain a relation that was typographically identical to a relation in thermodynamics, similar to (6).[3] However, as mentioned before, the striking difference between Gibbs's discussion of this issue and *Boltzmann 1871b* is that Boltzmann presented his result boldly as an "analytical proof of the second law of thermodynamics", whereas Gibbs presented his result as merely establishing an "analogy".

The second objection that Zermelo raised in his *1906* is a return to the topic of his 1896–7 dispute with Boltzmann. Can statistical mechanics, rigorously developed, ever entail a description of irreversible processes? Gibbs discusses this point in his Chapter 12.

In this chapter, Gibbs provides a discussion of Poincaré's recurrence theorem and the objections it entails to the monotonic increase of entropy. Whether he learned about this theorem from studying Poincaré, or by reading the debate between Zermelo and Boltzmann, is unclear, since he gives no references. Gibbs reaches from this discussion more or less the same conclusions as Zermelo did about the feasibility of incorporating an irreversible approach to equilibrium in statistical mechanics, i.e., "we find by this method no approach to statistical equilibrium in the course of time (*1902*, 144)."

However, Gibbs did not draw the same devastating conclusions that Zermelo (*1896a*) did about the prospects of statistical mechanics to deal with irreversibility. Instead, he warned the reader to withhold judgment: "Yet we must here exercise extreme caution (ibid.)." Gibbs's discussion continues by pointing to an analogy between the mechanical motion in phase space and the stirring of two differently colored dyes, conceived of as incompressible fluids. When two such dyes, say black and white, are mixed by an external stirring device, the total volumes occupied by the black and white dyes remain invariant. Even so, these volumes can become so deformed and intermingled by the mixing operation that, if one does not look too closely, one may regard the resulting mixture as being uniformly gray.

Gibbs suggests that something similar might apply to mechanical motion in phase space (in this case, however, without external intervention). Thus, even if we start out with a probability density function on the phase space Γ which does not represent statistical equilibrium (say it is zero outside some region $A \subseteq \Gamma$ and a constant inside that region, and the constant and zero are taken as analogous to black and white), the autonomous evolution of that region under Hamiltonian mechanics might be such that this region develops into ever more branching and fibrillating tentacles that will in some sense cover the entire phase space uniformly, at least if we do not look too closely.

Gibbs showed that the mechanical evolution will necessarily conserve the value of $\overline{\eta}$ for any non-equilibrium probability density. Thus, at first sight,

[3] Actually, in *1971b* Boltzmann does not propose a concrete expression analogous to entropy, as Gibbs does in *1902*.

the expression he intended to serve as an analogy of thermodynamical entropy, $-\overline{\eta}$, could not increase in the course of time, even for a probability density that did not correspond to statistical equilibrium. But he did not conclude that his theory implies that entropy must be conserved under any Hamiltonian mechanical evolution. Instead, Gibbs argued that if the above picture applies, the probability density function that results after a long time in this process of fibrillation might be well approximated by a uniform, gray, distribution, and for this approximating uniform probability density, the expression $\overline{\eta}$ would actually be less than that of the probability density we started out with, corresponding to an increase in thermodynamical entropy. Accordingly, Gibbs diagnosed the problem of representing irreversible processes in statistical mechanics as a problem of the interchange of the order of two different limits in the representation of physical systems: one limit of taking a finer and finer coarse-graining, and the other limit of going to infinite time in the evolution of the density function.

Gibbs's suggestions here have been taken up by later authors in an attempt to fill in the details. The first ingredient is to distinguish between two kinds of entropy: "coarse-grained" versus "fine-grained", a distinction first proposed by Poincaré (*1906*) (as *entropie grossiere* and *entropie fine*). If we take the fine-grained Gibbs entropy as generally defined by (4),

$$-\overline{\eta}(P) = -\int P(x) \ln P(x) d\mu(x) \,,$$

the coarse-graining is introduced as an averaging of P over some partition of phase space: $\Gamma = \omega_1 \cup \ldots \cup \omega_m$ (with $\omega_1 \cap \omega_j = \emptyset$ if $i \neq j$). The coarse-grained version of P is given by

$$\mathcal{CG}P(x) = \sum_{i=1}^{m} \frac{1_{\omega_i}(x)}{\mu(\omega_i)} \int_{\omega_i} P(x) d\mu(x)$$

where 1_ω is the characteristic function of the set ω; and the coarse-grained entropy of P is then defined as the Gibbs entropy of the coarse-grained version of P:

$$\Sigma[P] = -\overline{\eta}(\mathcal{CG}P) \,.$$

One can show that, while the fine-grained entropy $-\overline{\eta}(P_t)$ is stationary under time evolution, the coarse-grained entropy can change its value under this evolution.

The other ingredient, stressed in particular in the 1940s by Nikolai Sergeevich Krylov (*1979*), is to introduce as an additional assumption that dynamical systems have the property known as "mixing". In the context of the theory of dynamical systems, the mixing property is the requirement that, for any two measurable subsets $A, B \subseteq \Gamma$:

$$\lim_{t \to \infty} \mu(T_t A \cap B) = \mu(A)\mu(B) \,. \tag{7}$$

This seems a plausible formalization of the dynamical property that Gibbs had in mind when he compared the evolution to the stirring of dyes.

However, when it comes to answering the question whether these ingredients would suffice to provide a statistical-mechanical analogue of the thermodynamical approach to equilibrium, the situation is somewhat diffuse. Thus, although it can be shown that if the mixing assumption holds, and for an appropriate choice of the partition, the coarse-grained entropy $\Sigma(P_t)$ of any continuous non-equilibrium probability density P_t will approach, for $t \to \infty$, the value of the coarse-grained entropy of statistical equilibrium, this approach need not be monotonic. Moreover, the same is true for the limit $t \to -\infty$.

To return from these later developments to Zermelo's discussion, it seems that Zermelo dismisses Gibbs's considerations almost immediately as being ad hoc. He argues that there is no reason to suppose that a system should have the mixing property. In a sense, Zermelo is right here. Yet to me his dismissal also seems a bit too easy. Of course, it is true if there are many dynamical systems, even Hamiltonian ones, that do not have the mixing property. So, if one takes the purpose of statistical mechanics to be the systematic and general exposition of rigorous results pertaining to all Hamiltonian systems without any regard to possible applications in the kinetic theory of gases, as Zermelo seems to understand Gibbs's aims to be, then the appeal to this mixing property might seem rather ad hoc.

However, there are also many examples of dynamical systems that do have the mixing property. In other words, there is nothing in the general formalism of statistical mechanics that makes this assumption incompatible. And so, it seems a perfectly viable strategy to argue that when it comes to explaining the observed tendency towards equilibrium—an issue that only arises in special applications like in the kinetic theory of gases—to appeal to the conjecture that those systems that display this tendency may happen to possess an additional property, in order to explain this particular type of behavior.[4]

It is another matter whether the strategy is also successful. Here, another objection raised by Zermelo is more to the point: even if the mixing property were to hold, it would not break the time-reversal invariance of the theory, and imply not only a tendency towards statistical equilibrium towards the future (i.e. the limit $t \to \infty$) but also towards the past, $t \to -\infty$. Hence, as Zermelo puts it, the property gives no basis to conclude that if we take two arbitrary times $t'' > t'$ that the Gibbs entropy at one instant should be less than at the other.

[4] The question whether the mixing property holds for the kind of systems of interest to the intended application, e.g. the hard spheres model, turns out to be extremely tedious (*Sinai and Chernov 1987, Szász 1996, Simányi and Szász 1999*).

Zermelo's papers on the Gibbs approach to the foundations of statistical mechanics are not well-known, but I hope to have made clear that he made a major contribution in making Gibbs's work accessible to the German-speaking community and, in his critical review of 1906, made a couple of

Besprechung von *Gibbs 1902* und *Gibbs 1905*

1906

Seit den frühesten Anfängen der Naturwissenschaft ist es ein Lieblingsgedanke der Forscher gewesen, die Mannigfaltigkeit der Naturvorgänge auf eine gemeinsame Grundform, auf die Bewegung ihrer kleinsten Teilchen, zurückzuführen. Namentlich für eine exakte quantitative Naturerkenntnis schien ei-
233 ne solche Zurückführung aussichtsvoll, nachdem durch *Galilei* und | *Newton* die Grundlage für eine mathematische Theorie der Bewegungserscheinungen geschaffen war. Es ist bekannt, daß *Newton* selbst die Optik in Gestalt der Emissionstheorie als einen Zweig der Mechanik behandelt hat, und wenn auch später diese Theorie zugunsten der Undulationstheorie aufgegeben werden mußte, so verstand man doch auch die letztere zunächst rein mechanisch, indem man sich den Äther wie eine elastische Flüssigkeit schwingend vorstellte. Eine Neubelebung erfuhr die mechanistische Vorstellungsweise, als mit der Entdeckung des Energiegesetzes die erkannte Äquivalenz von Wärme und Arbeit die Vermutung nahe legte, daß die Wärme eine reine Bewegungserscheinung und die Temperatur eines Körpers durch die lebendige Kraft der unregelmäßig bewegten Moleküle zu messen sei. Aber nur in der kinetischen Gastheorie, deren Hypothesen so besonders einfach sind, ist eine solche Zurückführung der thermischen Vorgänge auf mechanische bis zu einem gewissen Grade gelungen, bei den festen Körpern und den Flüssigkeiten ist sie bisher noch in ziemlich unvollkommenen Anfängen stecken geblieben.

Die Schwierigkeiten, die sich der exakten Durchführung der Molekularhypothese entgegenstellen, sind von zweierlei Art. Einmal entsteht die physikalische Frage: wie sollen wir uns die Beschaffenheit der kleinsten materiellen Teile vorstellen und welche Wirkungsgesetze den zwischen ihnen wirkenden Kräften zuschreiben, um einen bestimmten Ansatz für die Rechnung zu gewinnen? Es ist klar, daß wir uns hier mit ziemlich rohen Annäherungen, mit unvollkommenen Bildern begnügen können, sofern nur zu erwarten steht, daß die hierbei begangenen Fehler ohne wesentlichen Einfluß auf die Gesamterscheinung sein werden. So begnügt man sich in der Gastheorie meist mit der Vorstellung elastischer Stöße zwischen einfach geformten festen Körpern in der Annahme, daß das wahre Wirkungsgesetz der zwischen den Molekülen wirkenden Kräfte bei der kurzen Dauer der Einwirkung nur unwesentlich in

penetrating remarks, long before anyone else, pointing out that Gibbs's approach, while keeping close to the standard of a rigorous mathematical framework, did not provide a fully satisfactory explanation of the thermodynamical analogies, or the increase of entropy in isolated systems.

Review of *Gibbs 1902* and *Gibbs 1905*

1906

Since the earliest dawn of natural science researchers have been taken with the idea that the manifold of natural phenomena should be reduced to a common basic form, to the motion of their smallest particles. Such reduction seemed particularly promising for an exact quantitative knowledge of nature once *Galilei* and *Newton* had created the foundation for a mathematical theory of the phenomena of motion. *Newton* himself, as is well-known, had treated optics in the form of emission theory as a branch of mechanics. Even though it was later necessary to abandon this theory in favor of the undulation theory, understanding of the latter was purely mechanical at first as it was based on the idea that ether oscillates like an elastic liquid. The mechanistic conception was reanimated when the discovery of the law of energy led to the realization that heat and work are equivalent, which suggested that heat is but a phenomenon of motion and that the temperature of a body is to be measured as the living force of molecules in irregular motion. But only for the kinetic theory of gases, whose hypotheses are particularly simple, was it possible to reduce the thermal processes to mechanical ones to a certain extent, while similar attempts undertaken for solid bodies and for liquids have remained stalled in their early, inchoate beginnings.

The difficulties standing in the way of the rigorous implementation of the molecular hypothesis are of two different kinds. First, the physical question arises: How are we to conceive of the constitution of the smallest material parts, and which laws are we to attribute to the forces acting among them, in order to obtain a definitive starting point for calculation? It is evident that we can make do here with crude approximations and incomplete pictures, provided only that we can expect that the subsequent errors will have no substantive bearing on the phenomenon as a whole. Thus, in applying the theory of gases, we are mostly content to consider elastic collisions between simply formed solid bodies, assuming that the true laws of the forces acting between the molecules will not really be relevant considering the short

Betracht kommen werde. Aber jetzt beginnt eine weitere, die mathematische Aufgabe, auf Grund unserer physikalischen Hypothesen die zusammengesetzten Vorgänge zu berechnen, die in den gesamten Molekülsystemen, d. h. in den unserer Beobachtung zugänglichen Körpern unter den gemachten Annahmen stattfinden würden, damit wir in der Lage sind, die so berechneten Vorgänge mit der Erfahrung zu vergleichen und daraus Rückschlüsse auf die Berechtigung unserer Annahmen zu ziehen. Bedenkt man nun, welche außerordentlichen Schwierigkeiten schon der einfache Fall der drei nach dem *Newton*schen Gesetze bewegten Himmelskörper den Astronomen bietet, so erscheint es fast vermessen, die Bewegungsvorgänge in einem aus vielen Billionen von Molekülen zusammengesetzten Systeme berechnen zu wollen. Doch auch hier bietet sich eine bedeutende Erleichterung. Es ist ja weder notwendig noch auch wünschenswert, den Weg jedes einzelnen Teilchens zu verfolgen, sondern es genügt, die in ihrer Gesamtheit geltenden Gesetzmäßigkeiten zu erforschen. Wir brauchen weder den genauen Anfangszustand zu kennen, der uns durch messende Beobachtung doch niemals gegeben werden kann, noch auch den genauen Endzustand zu berechnen. Vielmehr kommt es nur darauf an, aus einem nur in ganz rohen Umrissen bestimmten Anfangszustande auf das Eintreten eines ebenso roh bestimmten Endzustandes zu schließen und dabei von allen singulären und wegen ihrer relativen Kleinheit und Seltenheit in der großen Masse | verschwindenden Einzelvorgängen abzusehen. Man 234
muß mit einem Worte die Mechanik der Moleküle nicht individuell sondern *statistisch* betreiben. Nicht die Bewegung eines einzelnen Teilchens ist zu bestimmen, sondern die Veränderung gewisser Funktionen, die von unzählig vielen solchen Teilchen abhängen, und auch hier nur die durchschnittliche oder wahrscheinliche Änderung für alle möglichen mit dem beobachtbaren physikalischen Anfangszustande verträglichen Positionen. Bei dieser statistischen Betrachtungsweise ist auch anzunehmen, daß unsere Hypothesen über Gestalt und Beschaffenheit der Moleküle und ihre Wirkungsgesetze nur teilweise zum Ausdruck kommen und daß speziellere Vorstellungen für die Beantwortung vieler Fragen sich als unnötig erweisen. Freilich müssen Grundlagen und Methoden dieser „statistischen Mechanik“ zuerst bekannt sein, es muß möglich sein, ohne willkürliche und zweifelhafte Hypothesen die Massenerscheinungen in mechanischen Systemen mit sehr viel Freiheitsgraden unter Beschränkung auf eine plausible Wahrscheinlichkeit ebenso sicher zu bestimmen wie vermöge der reinen Mechanik die Vorgänge in einem einzelnen System mit einer beschränkten Zahl von Freiheitsgraden.

Solcher Methoden hat man sich natürlich in der Gastheorie immer bedient. Aber die Art und Weise, wie man die in allen Richtungen durcheinander fliegenden Moleküle durch Gruppen von mittlerer Beschaffenheit ersetzte, wie man die Verteilung der Geschwindigkeiten nach dem Maxwellschen Gesetze zu begründen und die Wahrscheinlichkeitsrechnung auf die mittlere Anzahl und den Effekt der Zusammenstöße anzuwenden suchte, war nicht immer einwandsfrei. Auch bezogen sich alle diese Methoden meist auf den speziellen nur in der Gastheorie vorliegenden Fall eines Systems von geradlinig beweg-

duration of the action. But now a further, mathematical task awaits us: to calculate on the basis of our physical hypotheses the composite processes that would arise in the whole systems of molecules, i. e., in the bodies accessible to our observation, given the assumptions made, in order to be able to compare the processes thus calculated with experience and to draw conclusions from this about the extent to which our assumptions are justified. Considering what extraordinary difficulties even the simple case of three celestial bodies moving in accordance with *Newton*'s laws creates for astronomers, it appears almost presumptuous to try to calculate the processes of motion in a system composed of many trillions of molecules. But here, too, we are afforded significant relief. For it is neither necessary nor desirable to chart the path of each individual particle but, rather, sufficient to investigate the lawful regularities obtaining in their totality. We do not need to know the precise initial state, with which observation by measurement could never acquaint us anyway. Nor do we need to calculate the precise final state. Instead, all that matters is to deduce from a particular initial state given only in very rough outlines a final state determined in a similarly rough fashion while ignoring the singular individual processes that disappear in the great mass due to their relative small scale and their rarity. In a word, we must pursue the mechanics of the molecules not individually but *statistically.* Instead of determining the motion of individual particles, we must determine the changes of certain functions that depend on innumerably many such particles, and more specifically only the average or probable change for all possible positions compatible with the observable physical initial states. It is to be assumed that, in this statistical approach, our hypotheses about the form and constitution of the molecules and the laws governing them find only partial expression and that many questions need no more specific conceptions to receive an answer. Of course, we must first know what the basic ideas and methods of this "statistical mechanics" are. Restricting ourselves to a plausible probability, we must be able to determine, without relying on arbitrary and dubious hypotheses, the mass phenomena in mechanical systems with a great number of degrees of freedom with as great a certainty as we determine the processes in an individual system with a restricted number of degrees of freedom by means of pure mechanics.

Such methods have of course always been used in the theory of gases. But the manner in which the jumble of molecules moving in all directions was replaced by groups of average constitution, the way in which one tried to explain the distribution of velocities by means of *Maxwell*'s law and to apply the calculus of probabilities to the mean number and to the effect of the collisions was not always unobjectionable. Moreover, all these methods referred for the most part to the special case, exclusive to the theory of gases, of a sys-

ten Teilchen und waren somit auf die Theorie der Flüssigkeiten und festen Körper nicht ohne weiteres anwendbar. Es besteht also gewiß ein Bedürfnis, die Grundlagen der statistischen Mechanik unabhängig von ihrem Anwendungsgebiete auf streng mathematischer Grundlage zu entwickeln, und dies ist eben die Aufgabe, die der jüngst verstorbene große amerikanische Physiker sich in seinem hier vorliegenden letzten Werke gestellt hat, um hiermit der kinetischen Theorie einen ähnlichen Dienst zu leisten, wie er ihn in seinen grundlegenden „Thermodynamischen Studien“ der phänomenologischen Thermochemie erwiesen hat. Das behandelte Problem bietet in neuester Zeit noch ein erhöhtes Interesse: seitdem man begonnen hat, auch die Elektrizität atomistisch aufzufassen, wird man auch in der Theorie der bewegten „Elektronen“ statistische Methoden nicht entbehren können.

Ebenso wenig wie *Maxwell*[1] bei der Untersuchung der intramolekularen Bewegung legt *Gibbs* seiner Theorie eine speziellere Molekularvorstellung zugrunde, sondern ausschließlich die allgemeine Form der dynamischen Gleichungen, welche für ein System von n Freiheitsgraden unter der Einwirkung von inneren Potentialkräften in der *Hamilton*schen Form geschrieben werden können:

$$\dot{q}_i = \frac{dq_i}{dt} = \frac{\partial \varepsilon}{\partial p_i}\,, \quad \dot{p}_i = \frac{dp_i}{dt} = -\frac{\partial \varepsilon}{\partial q_i}\,, \tag{6}$$

wo $q_1, q_2 \ldots q_n$ die Koordinaten, $p_1, p_2 \ldots p_n$ die Impulse (momenta) und ε die
235 Gesamtenergie des Systemes bedeutet. Die Koordinaten für sich | bestimmen die Konfiguration, die Impulse den Geschwindigkeitszustand, beide zusammen die „Phase“ des Systems für einen variablen Zeitpunkt t. Durch die Anfangsphase zur Zeit $t = t_0$ ist nun die Phase zu jeder anderen Zeit t und damit die ganze Bewegung des Systems vollständig bestimmt. Läßt man also die Anfangsphase variieren innerhalb eines gewissen $2n$-dimensionalen Gebietes g_0, so variiert auch die Phase zur Zeit t_1 innerhalb eines entsprechenden Gebietes g_1, und nach einem bekannten Satze von *Liouville* hat das $2n$ fache Integral $\int \ldots \int dp_1 \ldots dp_n dq_1 \ldots dq_n$, das von der Wahl des Koordinatensystemes unabhängig ist und als „Phasenausdehnung“ (extension-in-phase) bezeichnet wird, für beide Gebiete g_0 und g_1 denselben Wert, d. h. es ist für bewegte Systeme in der Zeit konstant. Dieser Satz wird bezeichnet als der „Satz von der Erhaltung der Phasenausdehnung“ und ist für alles Folgende von grundlegender Bedeutung. Anstatt eines einfachen Systemes denken wir uns eine große Zahl von Systemen mit verschiedenen Anfangsphasen in der Weise verteilt, daß der Quotient aus der Anzahl der Systeme, die sich innerhalb eines Gebietes g befinden, durch die Ausdehnung dieses Gebietes bei entsprechender Verkleinerung einer Grenze D zustrebt, die im allgemeinen eine Funktion der Phase (p, q) ist und als „Phasendichte“ (density-in-phase) bezeichnet wird. Der Quotient P dieser Phasendichte durch die Anzahl N aller überhaupt betrachteten Systeme heißt der „Wahrscheinlich-

[1] *Maxwell*, Cambr. Trans. 12 (1879), p. 547; Papers 2, p. 713.

tem of particles in linear motion, and hence were not applicable to the theory of liquids and solid bodies without further ado. Thus there certainly exists a need for developing the foundations of statistical mechanics independently of its domain of application on the basis of a rigorous mathematical foundation, which is just the task the great, recently deceased American physicist has set himself in his last work under review here, in which he attempted to do for the kinetic theory what he had done for the phenomenological thermochemistry in his seminal *1892*. The problem under consideration has drawn increased attention recently: with the emergence of an atomistic conception of electricity, it will be impossible to do without statistical methods also in the theory of "electrons" in motion.

Like *Maxwell*[1] in the investigation of intramolecular motion, *Gibbs* does not base his theory upon a specific conception of molecules but only upon the general form of the dynamical equations, which, for a system of n degrees of freedom under the action of internal potential forces, can be written in *Hamilton*ian form:

$$\dot{q}_i = \frac{dq_i}{dt} = \frac{\partial \varepsilon}{\partial p_i}\,, \quad \dot{p}_i = \frac{dp_i}{dt} = -\frac{\partial \varepsilon}{\partial q_i}\,, \tag{6}$$

where $q_1, q_2 \ldots q_n$ denote the coordinates, $p_1, p_2 \ldots p_n$ the momenta, and ε the total energy of the system. The coordinates themselves determine the configuration, and the momenta determine the velocity state, and both of them together determine the "phase" of the system for a variable point in time t. The initial phase at time $t = t_0$ now completely determines the phase at any other time t, and hence the entire motion of the system. Thus, if we let the initial phase vary within a certain $2n$-dimensional region g_0, then the phase at time t_1 also varies within a corresponding region g_1, and, by a well-known theorem of *Liouville*, the $2n$-fold integral $\int \ldots \int dp_1 \ldots dp_n dq_1 \ldots dq_n$, which is independent of the choice of the coordinate system and is called "extension-in-phase", has the same value for both regions g_0 and g_1, i. e., it is constant over time for systems in motion. This theorem is called the "theorem of the preservation of the extension-in-phase" and is of basic significance for all that follows. Instead of a simple system, let us envision a great number of systems with different initial phases distributed so that the quotient from the number of systems lying within a region g, and the expansion of this region upon corresponding reduction, converges to a limit D, which, in general, is a function of the phase (p, q) and is called the "density-in-phase". The quotient P of this density-in-phase by the number N of all systems under

[1] *Maxwell 1879*; *Maxwell 1890*, pp. 713 ff.

keitskoeffizient“ (coefficient of probality), dessen natürlicher Logarithmus η ist der „Wahrscheinlichkeitsexponent“ (index of probality), und der Ausdruck $Pdp_1 \dots dq_n = e^{\eta} dp_1 \dots dq_n$ stellt den Wert der Wahrscheinlichkeit dar, daß ein gegebenes System sich zwischen den durch das Differentialprodukt angedeuteten Phasengrenzen befindet. Denken wir uns nun ein beliebiges System gleichzeitig mit den anderen Systemen, die von benachbarten Phasen ausgehen, gemäß den Bewegungsgleichungen verändert, so muß dabei mit der Phasenausdehnung und der Anzahl der Systeme auch die Phasendichte D und daher auch P und η ungeändert bleiben. Ist also die Phasendichte zu einer beliebigen Zeit t als Funktion der Phase (p, q) gegeben, so ist für die betrachtete „Gesamtheit“ (ensemble) die „Phasenverteilung“ (distribution-in-phase) für jede andere Zeit t bestimmt vermöge der partiellen Differentialgleichung

$$\Big(\frac{dD}{dt}\Big)_{p,q} + \sum \Big(\frac{\partial D}{\partial p_i}\dot{p}_i + \frac{\partial D}{\partial q_i}\dot{q}_i\Big) = 0 \tag{21}$$

oder

$$\frac{\partial D}{\partial t} = \sum \Big(\frac{\partial D}{\partial p_i}\frac{\partial \varepsilon}{\partial q_i} - \frac{\partial D}{\partial q_i}\frac{\partial \varepsilon}{\partial p_i}\Big),$$

die als die „Grundgleichung der statistischen Mechanik“ bezeichnet werden kann. Hier bedeutet $\dfrac{\partial D}{\partial t}$ oder genauer $\Big(\dfrac{dD}{dt}\Big)_{p,q}$ die zeitliche Veränderung von D für konstante p, q, d. h. an einer bestimmten Stelle des Phasenraumes. Die Phasendichte bleibt an jeder Stelle (p, q) ungeändert: $\Big(\dfrac{dD}{dt}\Big)_{p,q} = 0$, wenn die rechte Seite der letzten Gleichung verschwindet, und in diesem Falle sagen wir, unsere Gesamtheit sei in „statistischem Gleichgewicht“ (statistical equilibrium). Dies wird insbesondere dann eintreten, wenn die Phasendichte
236 konstant oder eine Funktion der Gesamt- | energie ist. Ob eine Gesamtheit in statistischem Gleichgewichte ist oder nicht, hängt demnach nicht von der Beschaffenheit der einzelnen Systeme selbst ab, sondern lediglich von der willkürlich angenommenen anfänglichen Phasenverteilung, d. h. von der Anzahl der Systeme, die wir von jedem Phasenelement aus ihre Bewegung beginnen lassen. Die in einem einzelnen Systeme stattfindenden Vorgänge sind daher ganz unabhängig davon, ob das betrachtete System einer in statistischem Gleichgewichte stehenden Gesamtheit angehört oder nicht. Zur Erleichterung der meisten Untersuchungen wird es zweckmäßig sein, Verteilungen in statistischem Gleichgewichte vorzugsweise zu betrachten, und es bietet sich eine weitere Vereinfachung, wenn wir die Phasendichte kontinuierlich nach 0 abnehmen lassen, während die Koordinaten und Impulse q, p ins Unendliche wachsen. Beiden Forderungen genügt eine Verteilung, in welcher der Wahrscheinlichkeitsexponent eine lineare Funktion der Energie ist:

$$\eta = \log P = \frac{\psi - \varepsilon}{\theta} \quad \text{oder} \quad P = e^{\eta} = e^{\frac{\psi-\varepsilon}{\theta}},$$

consideration is called the "coefficient of probability", whose natural logarithm η is the "index of probability", and the expression $P dp_1 \dots dq_n = e^{\eta} dp_1 \dots dq_n$ represents the value of the probability that a given system lies between the phase boundaries indicated by the differential product. Let us now consider any system together with the other systems starting from neighboring phases, all modified in accordance with the equations of motion. Then along with the extension-in-phase and the number of the systems the density-in-phase D, too, and hence also P and η, must remain unchanged. Thus, if the density-in-phase is given at any time t as function of the phase (p, q), then, for any other time t, the "distribution-in-phase" is determined for the considered "ensemble" by virtue of the partial differential equation

$$\Big(\frac{dD}{dt}\Big)_{p,q} + \sum \Big(\frac{\partial D}{\partial p_i}\dot{p}_i + \frac{\partial D}{\partial q_i}\dot{q}_i\Big) = 0 \tag{21}$$

or

$$\frac{\partial D}{\partial t} = \sum \Big(\frac{\partial D}{\partial p_i}\frac{\partial \varepsilon}{\partial q_i} - \frac{\partial D}{\partial q_i}\frac{\partial \varepsilon}{\partial p_i}\Big),$$

which can be called the "basic equation of statistical mechanics". Here, $\dfrac{\partial D}{\partial t}$, or more precisely, $\Big(\dfrac{dD}{dt}\Big)_{p,q}$, denotes the temporal change of D for constant p, q, i. e., at a particular point of the phase space. The density-in-phase remains unchanged at every point (p, q): $\Big(\dfrac{dD}{dt}\Big)_{p,q} = 0$, if the right side of the last equation vanishes, and in this case we say that our ensemble is in "statistical equilibrium". This will occur in particular when the density-in-phase is constant or a function of the total energy. Thus, whether or not an ensemble is in statistical equilibrium does not depend on the constitution of the individual systems themselves but only on the arbitrarily assumed initial distribution-in-phase, i. e., on the number of the systems whose motions we let begin from every phase element. Therefore, the processes occurring in an individual system are entirely independent of whether or not the considered system belongs to an ensemble in statistical equilibrium. To simplify most investigations, it will be convenient to preferably consider distributions in statistical equilibrium, and a further simplification is afforded by letting the density-in-phase continually decrease toward 0 while the coordinates and momenta q, p increase to infinity. Both requirements are satisfied by a distribution in which the index of probability[2] is a linear function of the energy:

$$\eta = \log P = \frac{\psi - \varepsilon}{\theta} \quad \text{or} \quad P = e^{\eta} = e^{\frac{\psi-\varepsilon}{\theta}},$$

[2] ⟦Zermelo's term "Wahrscheinlichkeitsexponent" is translated here and in the following by the Gibbsian term "index of probability".⟧

wo ψ und θ Konstanten sind und ψ als Funktion von θ so zu bestimmen ist, daß die Gesamtwahrscheinlichkeit, d. h. das Integral

$$\int e^{\frac{\psi-\varepsilon}{\theta}} dp_1 \dots dq_n \ ,$$

über alle möglichen Phasen bis ins Unendliche erstreckt, den Wert 1 hat. Eine solche Verteilung bezeichnet *Gibbs* als eine „kanonische“ und beschränkt seine weiteren Untersuchungen zumeist auf solche Gesamtheiten. Die Analogie einer kanonischen Verteilung mit dem *Maxwell*schen Verteilungsgesetze in der Gastheorie ist unmittelbar einleuchtend, nur handelt es sich hier nicht um verschiedene Teilchen desselben Systems, sondern um lauter voneinander unabhängige Systeme. Speziell für solche kanonischen Verteilungen werden nun „Durchschnittswerte“, d. h. Integrale der Form

$$\overline{u} = \int \dots \int u e^{\frac{\psi-\varepsilon}{\theta}} dp_1 \dots dq_n$$

für verschiedene Phasenfunktionen u sowie ihre „mittleren Fehler“ d. h. ihre durchschnittlichen Abweichungen vom Durchschnittswerte berechnet.

Die für eine kanonische Verteilung charakteristische Konstante, den „Verteilungsmodul“ θ, betrachtet *Gibbs* in seiner Analogie mit der Temperatur eines thermischen Systems. Werden nämlich zwei kanonische Gesamtheiten zu einer dritten Gesamtheit in der Weise vereinigt, daß der Wahrscheinlichkeitskoeffizient des Gesamtsystems durch das Produkt $P_1 P_2$ der beiden gegebenen Koeffizienten für dieselbe Phase dargestellt wird, so ist die entstehende Gesamtheit wieder eine kanonische, also in statistischem Gleichgewichte, wenn die Verteilungsmoduln θ_1 und θ_2 der beiden Gesamtheiten gleich sind, ebenso wie zwei thermische Systeme nur dann miteinander im Gleichgewichte sein können, wenn ihre Temperaturen gleich sind. Denkt man sich ferner den Verteilungsmodul θ einer kanonischen Gesamtheit variiert, betrachtet also eine kontinuierliche Schar solcher Gesamtheiten mit verschiedenen Moduln, so ändert sich auch ψ und ergibt für die Durchschnittswerte $\overline{\varepsilon}$ und $\overline{\eta}$ der Energie und des Wahrscheinlichkeitsexponenten die Beziehung

$$d\psi = \frac{\psi - \overline{\varepsilon}}{\theta} d\theta = \overline{\eta} d\theta \ .$$

237 | Werden gleichzeitig auch die „äußeren Koordinaten“ verändert, d. h. die Parameter $a_1, a_2, a_3 \dots$, welche in dem Ausdrucke ε der Gesamtenergie außer den Koordinaten und Impulsen p, q als Veränderliche auftreten, sodaß sie für alle Systeme einer einzigen Gesamtheit dieselben Werte, für jede veränderte Gesamtheit aber andere Werte haben, so kommen noch weitere Glieder hinzu, welche die Durchschnittswerte der „äußeren Kräfte“ $A_1 = -\frac{\partial \varepsilon}{\partial a_1}$, $A_2 = -\frac{\partial \varepsilon}{\partial a_2}, \dots$ darstellen und mit $\overline{A}_1, \overline{A}_2, \dots$ bezeichnet werden mögen, und

where ψ and θ are constants and ψ is to be determined as a function of θ so that the total probability, i. e., the integral

$$\int e^{\frac{\psi-\varepsilon}{\theta}} dp_1 \dots dq_n \ ,$$

extended over all possible phases to infinity, has the value 1. *Gibbs* refers to a distribution of this sort as "canonical" and restricts his further investigations for the most part to such ensembles. The analogy of a canonical distribution with *Maxwell*'s law of distribution in the theory of gases is immediately evident, but it is concerned here with mutually independent systems rather than different particles of the same system. Now, specifically for such canonical distributions "mean values" are calculated, i. e., integrals of the form

$$\overline{u} = \int \dots \int u e^{\frac{\psi-\varepsilon}{\theta}} dp_1 \dots dq_n$$

for different phase functions u and their "mean errors", i. e., their average deviations from the mean value.

Gibbs considers the constant characteristic of a canonical distribution, the "distribution module" θ, in his analogy with the temperature of a thermal system. For if two canonical ensembles are united with a third ensemble so that the coefficient of probability of the total system is represented by the product $P_1 P_2$ of the two given coefficients for the same phase, then the ensemble obtained is in turn a canonical ensemble, and hence in statistical equilibrium, if the distribution modules θ_1 and θ_2 of the two ensembles are identical, just like two thermal systems can be in equilibrium with one another only if their temperatures are identical. Furthermore, if we assume that the distribution module θ of a canonical ensemble varies, and hence if we consider a continuous family of such ensembles with different modules, then ψ, too, changes and, for the mean values $\overline{\varepsilon}$ and $\overline{\eta}$ of the energy and of the index of probability, we get the relation

$$d\psi = \frac{\psi - \overline{\varepsilon}}{\theta} d\theta = \overline{\eta} d\theta \ .$$

If we also alter the "external coordinates", i. e., the parameters $a_1, a_2, a_3 \dots$ occurring in the expression ε of the total energy besides the coordinates and momenta p, q as variables, so that they have the same values for all systems of a single ensemble but different values for each altered ensemble, then further terms are added which represent the mean values of the "external forces" $A_1 = -\frac{\partial \varepsilon}{\partial a_1}$, $A_2 = -\frac{\partial \varepsilon}{\partial a_2}, \dots$ and which we denote by $\overline{A}_1, \overline{A}_2, \dots$, and we

wir erhalten die Formeln

$$d\psi = \overline{\eta} d\theta - \overline{A}_1 da_1 - \overline{A}_2 da_2 - \cdots$$

oder wegen $d\psi - d\overline{\varepsilon} = d(\theta\overline{\eta})$ schließlich

$$d\overline{\varepsilon} = -\theta d\overline{\eta} - \overline{A}_1 da_1 - \overline{A}_2 da_2 - \cdots \tag{114}$$

ganz analog der Differentialbeziehung des zweiten Hauptsatzes, wenn θ der Temperatur, $\overline{\varepsilon}$ der Energie und $\overline{\eta}$ der negativen Entropie entspricht. Diese Analogie bleibt aber rein äußerlich und kann uns keine Aufklärung über den mechanischen Charakter der thermodynamischen Gleichung bieten. Denn die Größe $\overline{\eta}$, der Durchschnittswert des Wahrscheinlichkeitsexponenten, ist gar keine Eigenschaft der einzelnen mechanischen Systeme, auch keine durchschnittliche Eigenschaft, welche das allen Gemeinsame oder Vorwiegende zum Ausdruck bringen könnte, sondern sie hängt wie der Wahrscheinlichkeitsexponent selbst ausschließlich von der willkürlichen anfänglichen Phasenverteilung ab, während die Entropie eines Systems doch immer durch die Natur dieses einzelnen Systems selbst bestimmt sein muß. Ebensowenig kann das Glied $\theta d\overline{\eta}$ im eigentlichen Sinne die übergehende Wärme repräsentieren, da es sich lediglich auf eine Veränderung der Phasenverteilung, auf den Übergang zu einer neuen Gesamtheit, gar nicht auf eine Wechselwirkung der Systeme mit der Außenwelt bezieht. Will man in einwandsfreier Weise die Eigenschaften eines mechanischen Systems statistisch untersuchen, so muß man immer die Gesamtheit beibehalten und nur innerhalb dieser Gesamtheit die Änderungen der Durchschnittswerte bestimmen. Denn das einzelne System in seinem durchschnittlichen oder wahrscheinlichen Verhalten, nicht eine zwischen solchen Systemen willkürlich festgesetzte Beziehung sollte doch das mechanische Analogon eines physikalisch veränderten Naturkörpers darstellen.

Im 10. Kapitel untersucht der Verfasser eine andere Analogie mit Hilfe der „mikrokanonischen Phasenverteilung“, d. h. einer Phasenverteilung, in welcher alle Systeme die gleiche Energie haben, und die auch als ein Grenzfall einer kanonischen Gesamtheit aufgefaßt werden kann, indem man den Ausdruck $\eta = c - \dfrac{(\varepsilon - \varepsilon')^2}{\omega^2}$ für sehr große Werte von ω als Wahrscheinlichkeitsexponenten ansetzt. Jedem Energiewerte ε entspricht nun ein bestimmter Wert des im 8. Kapitel eingeführten Integrales

$$V = \int \ldots \int dp_1 \ldots dq_n$$

erstreckt über alle Phasen mit einer Energie $< \varepsilon$. Somit ist V, die Phasenausdehnung unter einer Energiegrenze ε, sowie auch der abgeleitete Wert $\varphi = \log \dfrac{\partial V}{\partial \varepsilon}$ eine Funktion von ε und den äußeren Koordinaten $a_1, a_2, \ldots,$
238 | also konstant in jeder mikrokanonischen Gesamtheit. Bezeichnet man nun

obtain the formulas

$$d\psi = \overline{\eta} d\theta - \overline{A}_1 da_1 - \overline{A}_2 da_2 - \cdots$$

or eventually, on account of $d\psi d\overline{\varepsilon} = d(\theta\overline{\eta})$,

$$d\overline{\varepsilon} = -\theta d\overline{\eta} - \overline{A}_1 da_1 - \overline{A}_2 da_2 - \cdots \tag{114}$$

in a manner entirely analogous to the differential relation of the second law, where θ corresponds to the temperature, $\overline{\varepsilon}$ to the energy and $\overline{\eta}$ to the negative entropy. But this analogy remains entirely superficial and cannot clarify the mechanical character of the thermodynamical equation. For the magnitude $\overline{\eta}$, the mean value of the index of probability, is not at all a property of the individual mechanical systems. Nor is it an average property capable of expressing what is common to or dominant in all. Rather, like the index of probability itself, it depends exclusively on the arbitrary initial distribution-in-phase, while the entropy of a system must always be determined by the nature of this individual system itself. Likewise, the term $\theta d\overline{\eta}$ cannot actually represent the transmitted heat, since it only refers to a change of the distribution-in-phase, to the transition to a new ensemble, and not at all to an interaction of the systems with the external world. Anyone wishing to statistically investigate in an unobjectionable manner the properties of a mechanical system must always retain the ensemble and only determine the changes of the mean values within this ensemble. For, after all, it is the individual system in its average or probable behavior, not a relation arbitrarily stipulated among such systems, that should be the mechanical analogy of a physically altered natural body.

In chapter 10, the author examines another analogy by means of the "microcanonical distribution-in-phase", i. e., a distribution-in-phase in which all systems have the same energy and which can also be regarded as a borderline case of a canonical ensemble by setting the expression $\eta = c - \dfrac{(\varepsilon - \varepsilon')^2}{\omega^2}$ as index of probability for very large values of ω. To every energy value ε there now corresponds a particular value of the integral introduced in chapter 8

$$V = \int \ldots \int dp_1 \ldots dq_n$$

extended over all phases with energy $< \varepsilon$. Thus, V, the extension-in-phase below some energy limit ε, as well as the derived value $\varphi = \log \dfrac{\partial V}{\partial \varepsilon}$, is a function of ε and the external coordinates $a_1, a_2, \ldots$, and hence constant in every

mit $\overline{A_1}\big|_\varepsilon$, $\overline{A_2}\big|_\varepsilon$ usw. die in der mikrokanonischen Gesamtheit genommenen Durchschnittswerte der äußeren Kräfte, so gelangt man zu der Gleichung

$$d\varepsilon = e^{-\varphi} V \, d\log V - \overline{A_1}\big|_\varepsilon \, da_1 - \overline{A_2}\big|_\varepsilon \, da_2 - \cdots \,, \tag{418}$$

welche wieder der Gleichung des zweiten thermodynamischen Haupsatzes entspricht, wenn $e^{-\varphi} V = V \Big/ \dfrac{\partial V}{\partial \varepsilon}$ die Temperatur und $\log V$ die Entropie repräsentiert. Diese zweite Analogie entspricht nun schon sehr viel besser einem wirklichen Naturvorgange, da alle hier vorkommenden Größen wirkliche Eigenschaften der einzelnen Systeme sind, bestimmt durch den Ausdruck, welcher die Energie als Funktion der äußeren und inneren Koordinaten und der Impulse darstellt. Bezeichnet man ferner mit ε_p die kinetische Energie des Systems, so gilt für ihren Durchschnittswert in der mikrokanonischen Gesamtheit die Beziehung

$$\frac{2}{n} \, \overline{\varepsilon_p}\big|_\varepsilon = e^{-\varphi} V \,, \tag{377}$$

es wird also wie in der kinetischen Gastheorie die lebendige Kraft der bewegten Teilchen ein Analogon der Temperatur. Dagegen gilt für solche mikrokanonischen Gesamtheiten nicht allgemein das Analogon des Satzes, daß zwei Körper von gleicher Temperatur vereinigt wieder ein System von derselben Temperatur ergeben.

Einen wesentlichen Bestandteil der Theorie bildet das 12. Kapitel, in welchem die Bewegung der Systeme und Gesamtheiten in langen Zeiten untersucht wird. Aus dem Gesetze von der Erhaltung der Phasenausdehnung wird der Satz von *Poincaré* hergeleitet, daß ein innerhalb endlicher Grenzen eingeschlossenes und sich selbst überlassenes, nur den *Hamilton*schen Gleichungen folgendes mechanisches System im *Poisson*schen Sinne *stabil* ist. Betrachten wir nämlich ein beliebig kleines Gebiet g von Anfangsphasen von endlicher Phasenausdehnung, so wird wenigstens ein Teil, sogar der größte Teil dieser Phasen nach Ablauf einer beliebig langen Zeit T, wenn auch nicht genau in die Anfangsphase, so doch in dasselbe Gebiet g zurückkehren, die Bewegung wird also für den überwiegenden Teil der Phasen einen quasiperiodischen Charakter haben. Somit steht nicht zu erwarten, daß sich bei Bewegungen dieser Art eine Tendenz nach irgend einem Endzustande geltend machen werde. Gleichwohl versucht der Verfasser ein solches Verhalten wenigstens wahrscheinlich, plausibel zu machen, indem er seinen Gesamtheiten, sofern sie nicht schon in statistischem Gleichgewichte sind, eine Tendenz nach einem Endzustande statistischen Gleichgewichtes zuschreiben möchte. Für den Fall des statistischen Gleichgewichtes ist nun, wie im elften Kapitel gezeigt wird, der durchschnittliche Wahrscheinlichkeitsexponent $\overline{\eta}$ *kleiner* als bei jeder anderen Verteilung. Es müßte sich also der Wert $\overline{\eta}$ mit unbegrenzt wachsender Zeit seiner unteren Grenze $\overline{\eta}_0$ beliebig nähern. Dem gegenüber steht aber die Tatsache, daß der Wahrscheinlichkeitsexponent, welcher der jeweiligen Phase eines bewegten Systemes entspricht, und damit auch der mittlere Wahr-

microcanonical ensemble. If we now denote the mean values of the external forces taken in the microcanonical ensemble by $\overline{A_1}\big|_\varepsilon$, $\overline{A_2}\big|_\varepsilon$ etc., then we obtain the equation

$$d\varepsilon = e^{-\varphi} V d \log V - \overline{A_1}\big|_\varepsilon \, da_1 - \overline{A_2}\big|_\varepsilon \, da_2 - \cdots \, , \tag{418}$$

which in turn corresponds to the equation of the second law of thermodynamics, assuming that $e^{-\varphi} V = V \Big/ \dfrac{\partial V}{\partial \varepsilon}$ represents the temperature, and $\log V$ the entropy. Now, this second analogy is already in much better accord with a real process in nature, since all magnitudes occurring here are real properties of the individual systems, determined by the expression representing the energy as a function of the external and internal coordinates and of the momenta. If, furthermore, we denote the kinetic energy of the system by ε_p, then for its mean value in the microcanonical ensemble the following relation holds:

$$\frac{2}{n} \overline{\varepsilon_p}\big|_\varepsilon = e^{-\varphi} V \, , \tag{377}$$

and hence, just as in the kinetic theory of gases, the living force of the particles in motion turns into an analogue of temperature. By contrast, an analogue to the theorem that uniting two bodies of the same temperature results in a system of the same temperature does not hold generally for such microcanonical ensembles.

An essential part of the theory is formed by chapter 12, which examines the motion of the systems and ensembles over extended periods of time. From the law of the preservation of the extension-in-phase *Gibbs* deduces *Poincaré*'s theorem that a mechanical system enclosed within finite boundaries, when left to its own devices, only following the *Hamilton*ian equations, is *stable* in *Poisson*'s sense of the word. For if we consider an arbitrarily small region g of initial phases of finite extension-in-phase, then, after an arbitrarily long period of time T has passed, at least a few, and even the majority, of these phases will return, if not exactly to the initial phase, at least into the same region g. Hence, the motion will have quasi-periodic character for the majority of phases. It is therefore not to be expected that a tendency towards any final state will assert itself for motions of this kind. Nonetheless, the author tries to render such behavior at least probable or plausible, by wishing to assign to his ensembles a tendency toward a final state of statistical equilibrium, provided they are not already in statistical equilibrium. As is shown in chapter 11, the mean index of probability $\overline{\eta}$ is *smaller* for the case of the statistical equilibrium than for any other distribution. Hence, the value $\overline{\eta}$ would have to get arbitrarily close to its lower boundary $\overline{\eta}_0$ at indefinitely increasing time. But this runs counter to the fact that the index of probability, which corresponds to the respective phase of a system in motion, and hence also the

scheinlichkeitsexponent der ganzen Gesamtheit auf Grund des Gesetzes von der Erhaltung der Phasendichte von der Zeit unabhängig, eine Konstante ist.
239 Hier sucht sich nun der Verfasser durch Umdeutung und Ana-|logie zu helfen. Er denkt sich in einem geschlossenen Raum zwei nicht-diffundierbare inkompressible Flüssigkeiten von gleicher hydrodynamischer Beschaffenheit, eine gefärbte und eine ungefärbte, durcheinander gerührt. Dann wird nach einiger Zeit, im allgemeinen wenigstens, eine nahezu homogene Mischung entstehen, obwohl jedes Flüssigkeitsteilchen nach wie vor entweder der einen oder der anderen Flüssigkeit angehören muß, die Dichtigkeit des Farbstoffes also in jedem kleinsten Teilchen dieselbe geblieben ist. Das mittlere Quadrat dieser Dichte, das im Falle wahrer Homogenität seinen Minimalwert annehmen müßte, wird also dasselbe sein wie am Anfang, wenn die Dichte in jedem Augenblicke durch unbegrenzte Verkleinerung des Volumenelementes bestimmt wird. Berechnet man dagegen die mittlere Dichte für endliche wenn auch beliebig kleine Volumenelemente, die nach einer gewissen Zeit eine große Menge verschiedenartiger Flüssigkeitsteilchen enthalten werden, so erscheint die so bestimmte Dichte im Endzustande nahezu homogen, und ihr mittleres Quadrat hat gegen den Anfangszustand abgenommen. In analoger Weise glaubt der Verfasser durch Vertauschung der beiden Grenzübergänge der Folgerung von der Konstanz des mittleren Wahrscheinlichkeitsexponenten entgehen zu können, indem er nicht ohne Berechtigung der betrachteten Bewegung in dem Sinne „Instabilität“ zuschreibt, daß im allgemeinen anfangs benachbarte Phasen sich im Laufe der Zeit immer weiter voneinander entfernen und anfangs getrennte sich nähern werden. Diese Analogie ist aber irreführend. In dem vorliegenden Beispiele handelt es sich wesentlich um einen Gegensatz zwischen kontinuierlicher und diskontinuierlicher Dichtigkeitsverteilung. Eine anfangs vorhandene Diskontinuitätsfläche von beschränkter Ausdehnung wird durch fortgesetzte Bewegung in eine schwammartig verwickelte Fläche von ungeheurer Ausdehnung auseinander gezogen und über den ganzen Raum verbreitet, sodaß jeder nicht allzu kleine Raumteil schroffe Übergänge von der einen in die andere Flüssigkeit in sich enthält, seine mittlere Dichtigkeit aber von der des benachbarten nur unbedeutend abweicht. Es wird also durch das Umrühren die Inhomogenität aus dem Großen ins Kleine übertragen, während ihre durch das mittlere Quadrat der Farbstoffdichte gemessene Gesamtgröße ungeändert bleibt. Von einem solchen Gegensatz zwischen kontinuierlichen und diskontinuierlichen Dichtigkeitsänderungen ist aber bei den von *Gibbs* betrachteten Gesamtheiten durchaus nicht die Rede. Vielmehr wird die Phasendichte von vornherein als eine stetige Funktion des Ortes im Phasenraume aufgefaßt, und es liegt kein Grund vor, anzunehmen, daß sie in den kleinsten Teilchen immer inhomogener werden sollte, um in den größeren Teilen im Mittel homogener werden zu können. Daß sich der Verfasser jede Phasenausdehnung bei der ursprünglichen Einführung der Phasendichte nicht stetig sondern durch eine endliche wenn auch sehr große Zahl von Systemen erfüllt denkt, tut nichts zur Sache, da dies anfangs so gut wie später gelten müßte, seine Formeln wenigstens haben es nur mit einer im allgemeinen stetigen

mean index of probability of the entire ensemble by virtue of the law of the preservation of the density-in-phase, is independent of time, is a constant. The author tries to find a way here by reinterpretation and analogy. He envisions two non-diffusible incompressible liquids of identical hydrodynamical constitution, one dyed and one undyed, blended in a closed space. Then, after some time, at least in general, an almost homogeneous mixture is formed, even though each liquid-particle must belong to either one or the other liquid, and hence the density of the colorant in every smallest particle has remained the same. Thus, the mean square of this density, which would have to assume its minimal value in the case of true homogeneity, will be the same as in the beginning, if the density is determined at every instant by unlimited decrease of the volume element. If, on the other hand, we calculate the mean density for finite, if arbitrarily small, volume elements that will, after a certain time, contain a great amount of various liquid-particles, then the density thus determined appears almost homogeneous in the final state, and its mean square has decreased toward the initial state. The author believes that it is possible to analogously avoid the conclusion about the constancy of the mean index of probability through a switch of the two passages to the limit by assigning, not without justification, "instability" to the considered motion so that, in general, initially neighboring phases will move ever further apart from one another over the course of time, and initially distant phases will approach one another. But this analogy is misleading. The present example hinges on a contrast between continuous and discontinuous density distributions. An initially given discontinuity surface of limited extension is stretched into a sponge-like entangled surface of immense extension by continued motion and is spread over the entire space so that each part of the space that is not too small contains jagged transitions from one liquid to the other while its mean density deviates from that of the neighboring part only insignificantly. Thus, the non-homogeneity is carried over from the large to the small scale by the stirring, while its total magnitude, which is measured by the mean square of the colorant density, remains unaltered. But, when considering his ensembles, *Gibbs* does not at all refer to such a contrast between continuous and discontinuous changes in density. Rather, the density-in-phase is conceived of at the outset as a continuous function of the position in the phase space, and there is no reason to assume that it should become ever more inhomogeneous in the smallest particles in order to be able to become on average more homogeneous in the greater parts. That the author assumes that each extension-in-phase is not continuous at the original introduction of the density-in-phase but satisfied by a finite if very large number of systems is irrelevant here, for this should hold true at the beginning just as at a later stage. At least his formulas only refer to a phase distribution, which, in

Phasenverteilung zu tun. Aus der oben zitierten Grundgleichung (21) der statistischen Mechanik folgt mathematisch streng, daß auch $\eta = \log \frac{D}{N}$ der Differentialgleichung genügt

$$\frac{\partial \eta}{\partial t} = \sum \left(\frac{\partial \eta}{\partial p_i} \frac{\partial \varepsilon}{\partial q_i} - \frac{\partial \eta}{\partial q_i} \frac{\partial \varepsilon}{\partial p_i} \right)$$

240 | und daher $\int .. \int \eta dp_1 .. \, dq_n$ eine Bewegungsinvariante ist, d. h. $\overline{\eta} = \text{const.}$, und dies gilt ganz unabhängig vom Koordinatensystem, mag man die Volumenelemente ursprünglich größer oder kleiner wählen, die Phasendichte durch den einen oder den anderen Grenzübergang bestimmen.

Wäre es aber auch gelungen, durch geeignete Umdeutung des Dichtigkeitsbegriffes das gewünschte Resultat zu erzielen, so stände es doch jedenfalls im Widerspruche mit den anfangs gegebenen Grundgleichungen, d. h. mit der gesamten vorher entwickelten Theorie. Man müßte geradezu von vorn anfangen und eine neue Theorie begründen, in welcher der Satz von der Herstellung des statistischen Gleichgewichtes gelten könnte. Ganz unstatthaft ist es aber, wie der Verfasser es im folgenden Kapitel tut, Formeln der alten und der neuen Theorie zu kombinieren, indem er z. B. in (462) die Annahme $\overline{\eta}''_{12} \leqq \overline{\eta}'_{12}$ zu den Folgerungen aus früheren Entwicklungen hinzufügt. Mit widersprechenden Voraussetzungen kann man freilich alles beweisen, was man will, beweist aber damit in Wirklichkeit nur die Unzulässigkeit eines solchen Verfahrens. Aber noch ein Weiteres hätte ich gegen die Formel (462) einzuwenden. Angenommen selbst, es wäre bewiesen, daß $\overline{\eta}$ mit der Zeit seinem Minimalwerte zustrebte, so müßte dies, wie der Verfasser selbst zugibt (*Gibbs*, p. 150, deutsche Bearbeitung S. 153) für abnehmende Zeiten ebenso gut gelten wie für wachsende, und die Grenzwerte für $t = +\infty$ und für $t = -\infty$ wären dieselben. Dann ist es aber völlig willkürlich, einer Zeit $t'' > t'$ einen Wert $\overline{\eta}'' < \overline{\eta}'$ zuzuschreiben, ebenso gut könnte man umgekehrt verfahren; denn welcher der beiden Zustände, der frühere oder der spätere, dem Grenzzustande näher liegt, ist allgemein nicht zu entscheiden. Es gelten eben hier dieselben Einwände wie gegen das *Boltzmann*sche H-Theorem, und es sei mir daher gestattet, auf die prinzipielle Seite dieser früher von mir erörterten Frage hier noch mit einigen Worten zurückzukommen.

Daß aus den Prinzipien der statistischen Mechanik, die auch ich für die einzig zuverlässige Grundlage solcher Wahrscheinlichkeitsbetrachtungen über dynamische Systeme halte, die beständige Zunahme irgend einer Phasenfunktion im Sinne des zweiten Hauptsatzes weder bewiesen noch auch wahrscheinlich gemacht werden kann, habe ich schon im Jahre 1896 mit Hilfe des *Liouville*schen und des *Poincaré*schen Satzes nachgewiesen[1] und diesen Beweis dann 1899 nach verschiedenen Richtungen ergänzt.[2] Es sei nämlich G irgend ein „invariantes Gebiet“ im Phasenraume, d. h. ein solches, aus welchem vermöge der Grundgleichungen weder Systeme austreten noch eintreten können,

[1] *Zermelo*, Ann. d. Phys. 57 (1896) S. 485.
[2] *Zermelo*, Phys. Ztschr. 1 (1900) S. 317.

general, is continuous. From the basic equation (21) of statistical mechanics cited above it follows with mathematical rigor that $\eta = \log \frac{D}{N}$, too, satisfies the differential equation

$$\frac{\partial \eta}{\partial t} = \sum \left(\frac{\partial \eta}{\partial p_i} \frac{\partial \varepsilon}{\partial q_i} - \frac{\partial \eta}{\partial q_i} \frac{\partial \varepsilon}{\partial p_i} \right)$$

and hence that $\int .. \int \eta dp_1 .. \, dq_n$ is a motion invariant, i. e., $\overline{\eta} = \text{const.}$, and this holds true entirely independently of the coordinate system, no matter whether we originally choose the volume elements greater or smaller, and determine the density-in-phase by means of one or the other passage to the limit.

But if a suitable reinterpretation of the concept of density had really led to the desired result, then it would certainly run counter to the basic equations given initially, i. e., to the entire theory previously developed. One would almost have to start from scratch and develop a new theory in which the theorem on the establishment of the statistical equilibrium could hold. But it is completely inadmissible to combine formulas of the old theory with the new theory, as the author does in the next chapter by adding, e. g., in (462) the assumption $\overline{\eta}''_{12} \leqq \overline{\eta}'_{12}$ to the consequences from previous developments. Of course, we can prove anything from contradictory assumptions, really proving only the inadmissibility of such a procedure. But I would like to raise a further objection against formula (462). Even if we assume that it is proved that $\overline{\eta}$ converges with time to its minimal value, then this would have to hold true for decreasing just as for increasing times, as is conceded by the author himself (*Gibbs 1902*, p. 150, *Gibbs 1905*, p. 153), and the boundary values for $t = +\infty$ and for $t = -\infty$ would be the same. But then, it is entirely arbitrary to assign to a time $t'' > t'$ a value $\overline{\eta}'' < \overline{\eta}'$, it is just as well to proceed conversely. For one cannot decide in general which of the two states, the earlier or the later one, is closer to the limit state. Here, the same objections apply as in the case of *Boltzmann*'s H-theorem, and I may still be permitted a few remarks on the basic aspects of this question, which I discussed on an earlier occasion.

Already in 1896, I showed, using *Liouville*'s and *Poincaré*'s theorems,[3] that it is impossible to prove, or even make probable, the continual increase of any phase function in the sense of the second law from the principles of statistical mechanics, which I too consider the only reliable foundation of such probabilistic considerations about dynamical systems. In 1899 I provided various supplements to this proof.[4] For let G be any "invariant region" in the phase space, i. e., one that, because of the basic equations, systems can neither

[3] *Zermelo 1896a.*
[4] *Zermelo 1900.*

z. B. eines, das durch irgend welche Energiegrenzen ε' und ε'' bestimmt ist, und es sei u irgend eine eindeutige Phasenfunktion. Dann wird der für das Gebiet G genommene Durchschnittswert in der *Gibbs*schen Bezeichnung

$$\overline{u} = \int uP\, dp_1 .. \, dq_n$$

und ist von der Zeit unabhängig. Somit erhalten wir, weil die Phasenwahrscheinlichkeit $Pdp_1 .. \, dq_n$ eine Bewegungsinvariante ist,

$$\frac{d\overline{u}}{dt} = \int \frac{du}{dt} P\, dp_1 .. \, dq_n = \overline{\frac{du}{dt}} = 0\,,$$

241 | d. h. die Funktion u wird im betrachteten Gebiete im Durchschnitt weder zunehmen noch abnehmen, sondern beides wird gleich wahrscheinlich sein. Dasselbe ergibt sich aber auch, wenn G kein invariantes Gebiet, sondern ein in zwei Grenzen u' und u'' der betrachteten Phasenfunktion u eingeschlossener Bereich ist.[1] Daß ich in meiner damaligen Darstellung der Einfachheit halber $P = 1$ angenommen hatte, ist für die Beweisführung selbst völlig unwesentlich. Meine Argumentation richtete sich nun auch gegen das *Boltzmann*sche H-Theorem, d. h. gegen die Behauptung, daß die von *Boltzmann* in die Gastheorie eingeführte Funktion H (*Boltzmann*, Vorlesungen über Gastheorie I § 5, 6), welche ebenso wie das *Gibbs*sche $\overline{\eta}$ im wesentlichen den Logarithmus der Phasenwahrscheinlichkeit darstellt und der Entropie entsprechen soll, beständig abnehmen müsse. In der sich hieraus entwickelnden Polemik[2] versuchte dann *Boltzmann* das H-Theorem zu retten, indem er seiner H-Kurve nicht mehr einen wesentlich absteigenden, sondern einen im allgemeinen gleichförmigen Verlauf parallel zur t-Achse zuschrieb, wobei die hin und wieder auftretenden „Buckel“ um so seltener sein sollen, je größer sie sind. Mag hiermit auch dem Minimalwerte von H entsprechend die überwiegende Wahrscheinlichkeit der *Maxwell*schen Verteilung plausibel gemacht werden, so kann ich doch in einem solchen nahezu symmetrischen Verlaufe der Kurve eine wirkliche Analogie zur Irreversibilität der Naturvorgänge keineswegs erblicken, und welchen Anspruch auf Allgemeingültigkeit dann die Ungleichung $\frac{dH}{dt} \leqq 0$ noch haben soll, ist mir unverständlich geblieben. Weder die *Boltzmann*schen noch die *Gibbs*schen Deduktionen haben meine Überzeugung erschüttern können, daß eine kinetische Wärmetheorie auch im Sinne der Wahrscheinlichkeitsrechnung sich nur dann mit dem zweiten Hauptsatze wird vereinigen lassen, wenn man sich entschließt, an Stelle der *Hamilton*schen Bewegungsgleichungen solche Differentialgleichungen zugrunde zu legen, welche das Prinzip der Irreversibilität bereits in sich enthalten.

Die Ausführungen des 13. Kapitels, in dem die Wirkung verschiedener äußerer Einflüsse auf Systeme und Gesamtheiten untersucht wird, beruhen

[1] *Zermelo*, Phys. Ztschr. 1 S. 319.

[2] *Boltzmann*, Ann. d. Phys. 57 (1896) S. 773; ibid. 60 (1897) S. 392. *Zermelo*, Ann. d. Phys. 59 (1896) S. 793, vgl. auch *Boltzmann*, Enzykl. d. math. Wiss. V. 8.

exit nor enter, such as one that is determined by any energy limits ε' and ε'', and let u be any single-valued phase function. Then, in *Gibbs*'s terminology, the mean value taken for the region G becomes

$$\overline{u} = \int uP\, dp_1 .. \, dq_n$$

and is independent of time. Thus, since the phase probability $P dp_1 .. \, dq_n$ is a motion invariant, we obtain

$$\frac{d\overline{u}}{dt} = \int \frac{du}{dt} P\, dp_1 .. \, dq_n = \overline{\frac{du}{dt}} = 0\,,$$

i. e., the function u will neither increase nor decrease on average in the region under consideration, but both cases will be equally probable. But we obtain the same result when G is not an invariant region but a domain enclosed within two boundaries u' and u'' of the phase function u under consideration.[5] That I simply assumed that $P = 1$ in my previous account for simplicity's sake is completely irrelevant to the proof itself. My line of argument was now also directed against *Boltzmann*'s H-theorem, i. e., against the assertion that the function H introduced into the theory of gases by *Boltzmann* (*1896a*, § 5, 6), which, like *Gibbs*'s $\overline{\eta}$, basically represents the logarithm of the phase probability and is supposed to correspond to the entropy, would have to decrease continually. In the ensuing polemic[6] *Boltzmann* then tried to salvage the H-theorem by attributing to his H-curve a course that is no longer essentially descending but generally uniform and parallel to the t-axis, where the occasionally occurring "bumps" are supposed to be the rarer the greater they are. While this may well make plausible the predominant probability of the *Maxwell* distribution in accordance with the minimal value of H, I certainly cannot recognize a genuine analogy to the irreversibility of the natural processes in such an almost symmetric course of the curve, and I fail to see how then the general validity of the inequality $\frac{dH}{dt} \leqq 0$ is supposed to be justified. Neither *Boltzmann*'s nor *Gibbs*'s deductions have been able to shake my conviction that it is possible to unite a kinetic theory of heat also in the sense of probability calculus with the second law only if we resolve to replace as a basis *Hamilton*'s equations of motion by differential equations already containing the principle of irreversibility in themselves.

The elaborations of chapter 13, which examines the effects of various external influences on systems and ensembles, rest for the most part on the

[5] *Zermelo 1900*, p. 319.

[6] *Boltzmann 1896*; *Boltzmann 1897*. *Zermelo 1896b*, see also *Boltzmann and Nabl 1907*.

nun großenteils auf den im 12. Kapitel gewonnenen Ergebnissen und sind daher nach den vorstehenden Ausführungen nicht frei von inneren Widersprüchen. Es wird angenommen, daß eine Gesamtheit, wenn einmal durch äußere Einflüsse ihr statistisches Gleichgewicht gestört ist, innerhalb einer gewissen Zeit wieder zu einem solchen Gleichgewichte zurückkehren werde, und hieraus werden dann in Verbindung mit den sonst abgeleiteten Beziehungen Ungleichungen gewonnen, die den thermodynamischen zu entsprechen scheinen. Im letzten Kapitel werden endlich die allgemeinen Betrachtungen spezialisirt auf Systeme von Molekülen verschiedener Gattungen, in welchen gegenseitige Umsetzungen zwischen den verschiedenen Gattungen stattfinden können. Auch solche Gesamtheiten denkt *Gibbs* sich kanonisch verteilt mit einem Wahrscheinlichkeitsexponenten der Form

$$\mathsf{H} = \frac{\Omega + \mu_1\nu_1 + \cdot\cdot + \mu_h\nu_h - \varepsilon}{\theta} \,, \tag{503}$$

242 | wo die Zahlen $\nu_1, \nu_2, .. \nu_h$ Anzahlen für die verschiedenen Partikelgattungen und $\Omega, \theta, \mu_1, \mu_2, .. \mu_h$ Konstanten bezeichnen. Aus diesem Ansatz ergeben sich dann wieder Gleichgewichtsbedingungen wie die folgende

$$d\Psi = \overline{\mathsf{H}}\, d\theta + \sum \mu_1 d\overline{\nu}_1 - \sum \overline{A}_1 da_1 \,, \tag{532}$$

die der thermodynamischen analog ist, wenn θ der Temperatur, $-\overline{\mathsf{H}}$ der Entropie und $\Psi = \overline{\varepsilon} + \theta\overline{\mathsf{H}}$ der freien Energie entspricht. Es ist aber physikalisch nicht recht einzusehen, wie denn das chemische Gleichgewicht der einzelnen Systeme mit dem statistischen Gleichgewichte ganzer Gesamtheiten zusammenhängen soll. Auch beruht die formale Analogie mit den thermodynamischen Gleichungen zum großen Teile nur auf der zu diesem Zwecke gewählten Form des Wahrscheinlichkeitsexponenten. Ich kann daher auch hier nicht zugeben, daß wir von solchen Analogien einen tieferen Einblick in das Wesen der chemischen Umsetzungen und ihrer Gleichgewichtsbedingungen zu erwarten hätten.

Einige weitere in dem besprochenen Werke mehr nebensächlich behandelte Gegenstände will ich zum Schlusse noch kurz erwähnen: im 2. Kapitel die Anwendung des Gesetzes von der Erhaltung der Phasenausdehnung auf die „Fehlertheorie“, wo aber im wesentlichen nur der Begriff, nicht die „Erhaltung“ der Phasenausdehnung benutzt wird, die Frage also nicht eigentlich der statistischen Mechanik angehört, und im 3. Kapitel die Anwendung desselben Prinzips auf die Integration der Bewegungsgleichungen mit Hilfe des „letzten Multiplikators“, für die aber statt *Jacobi*s „Vorlesungen über Dynamik“ eine viel spätere Arbeit von *Boltzmann* als Quelle zitiert wird. Besonders verdienstvoll sind die Untersuchungen des 7. Kapitels, wo die Berechnung einer großen Zahl von Durchschnittswerten ausgeführt und Methoden zu ihrer Fortsetzung angegeben werden. Hervorzuheben ist dabei die eingehende Berücksichtigung der „mittleren Fehler“, deren Abschätzung allein uns gelegentlich berechtigen kann, die in der Untersuchung vorkommenden Größen

results obtained in chapter 12, and hence are, according to the previous remarks, not devoid of inner contradictions. It is assumed that an ensemble, when once disturbed in its statistical equilibrium by external influences, returns to such an equilibrium within a certain period of time, and from this, in connection with the relations otherwise derived, inequalities are obtained that appear to correspond to the thermodynamical ones. Finally, in the last chapter, the general considerations are specifically applied to systems of molecules of various categories, in which mutual conversions among the various categories take place. *Gibbs* assumes that such ensembles, too, are canonically distributed with an index of probability of the form

$$\mathsf{H} = \frac{\Omega + \mu_1\nu_1 + \cdot\cdot + \mu_h\nu_h - \varepsilon}{\theta} \,, \tag{503}$$

where the numbers $\nu_1, \nu_2, .. \nu_h$ denote numbers for the various categories of particles, and $\Omega, \theta, \mu_1, \mu_2, .. \mu_h$ denote constants. From this ansatz in turn follow equilibrium conditions, such as

$$d\Psi = \overline{\mathsf{H}}\, d\theta + \sum \mu_1 d\overline{\nu}_1 - \sum \overline{A}_1 da_1 \,, \tag{532}$$

which is analogous to the thermodynamical one, if θ corresponds to the temperature, $-\overline{\mathsf{H}}$ to the entropy, and $\Psi = \overline{\varepsilon} + \theta\overline{\mathsf{H}}$ to the free energy. But from a physical perspective it is difficult to see how the chemical equilibrium of the individual systems should be related to the statistical equilibrium of entire ensembles. Moreover, the formal analogy with the thermodynamical equations rests for the most part only on the form of the index of probability chosen for this purpose. Thus here, too, I cannot concede that we should expect to gain deeper insights from such analogies into the nature of the chemical conversions and their equilibrium conditions.

In closing, I would like to mention some further matters which did not receive full attention in the work under review: the application of the law of the preservation of the extension-in-phase to the "error theory" in chapter 2, where, however, essentially only the concept, and not the "preservation", of the extension-in-phase is used and thus where the question does not really lie within the purview of statistical mechanics, and the application of the same principle to the integration of the equations of motion by means of the "last multiplier" in chapter 3, for which, however, instead of *Jacobi*'s "Vorlesungen über Dynamik", a much later work by *Boltzmann* is cited as a source. Particularly commendable are the investigations in chapter 7, where a great number of mean values are calculated and methods for their continuation are presented. I should highlight here the close attention paid to the "mean errors" whose estimation alone can, on occasion, justify the statistical replacement

statistisch durch ihre Durchschnittswerte zu ersetzen. Interessant ist auch das 11. Kapitel, in dem eine Reihe von Sätzen über Maximums- und Minimumseigenschaften verschiedener Verteilungen entwickelt werden. —

In seinem Vorworte bekennt der Verfasser, er versuche nicht, die Geheimnisse der Natur zu erklären, sondern wolle sich mit dem bescheideneren Ziele begnügen, die einfachsten Sätze einer Disziplin zu entwickeln, die einer künftigen mechanischen Naturerklärung als Grundlage dienen könne. Auf diesem Gebiete sei auch schon jetzt ein sicheres Fortschreiten möglich, während einer eingehenden Durchführung der Molekularhypothese zur Zeit noch unüberwindliche Schwierigkeiten entgegenständen. Hat ihn bei der Verfolgung dieses Zieles sein lebhaftes Interesse für die Sache auch gelegentlich verleitet, die selbstgesteckten Grenzen zu überschreiten und dem Irrtum zu verfallen, so kann doch das Unternehmen selbst, mit dem der hochverdiente Forscher seine Lebensarbeit beschlossen hat, nur als gelungen betrachtet werden. Wo immer man statistischer und wahrscheinlichkeitstheoretischer Betrachtungen in der Mechanik bedarf, wird man nicht umhin können, sich zunächst mit dem *Gibbs*schen Werke vertraut zu machen und auf den hier gelegten Grundlagen weiterzubauen.

Göttingen. *E. Zermelo.*

of the magnitudes occurring in the investigation by their mean values. Of further interest is chapter 11, where a series of theorems on maximum and minimum properties of various distributions is developed.—

In his preface, the author declares that he will not seek to explain the secrets of nature but content himself with the more modest goal of developing the simplest theorems of a discipline that may serve as the basis for a future mechanistic explanation of nature. ⟦He states that⟧ while, already today, it is possible to make sound progress in this field, a detailed implementation of the molecular hypothesis still faces unsurmountable difficulties. Even if in the pursuit of this goal his lively interest in the matter has, on occasions, led him to cross the self-imposed boundaries and to fall into error, the enterprise itself, which marks the conclusion of the life-work of this exemplary researcher, can only be considered a success. Anyone in need of statistical and probabilistic considerations in mechanics cannot fail to first turn to *Gibbs*'s work and to build on the foundations laid in it.

Göttingen. *E. Zermelo.*

Introductory note to Riesenfeld and Zermelo 1909

Hans-Georg Bartel and Rainer Bruggemann

The physical chemist Ernst Hermann Riesenfeld was the impetus for this paper.[1] Already at the Physicochemical Institute of the University of Göttingen Riesenfeld investigated electrochemical effects at the boundary of two solvents.[2] These investigations were part of his Ph. D. (*Riesenfeld 1902a*; *Nernst and Riesenfeld 1901, 1902* and *Riesenfeld 1901, 1902b, 1902c*). Fick's first law of diffusion (*Fick 1855*)[3] as well as Nernst's distribution law (*Nernst 1890, 1891*)[4] played a central role in these and later investigations.

At the Chemisches Universitätslaboratorium in Freiburg, Riesenfeld worked together with his student Bernhard Reinhold[5] on transference numbers (*Riesenfeld and Reinhold 1909, 1910a, 1910b*). Once again the boundary between two solvents was the main concern; hence this kind of research can be seen as a continuation and extension of the work in Göttingen (*Riesenfeld 1902a, 1902b*; *Nernst and Riesenfeld 1902*). In *Riesenfeld and Reinhold 1910b* a mathematical problem was identified which inspired the paper discussed here. In order to solve the mathematical-physical problems Riesenfeld may have asked Zermelo, whom he knew since his time in Göttingen, for a collaboration.[6]

[1] Ernst Hermann Riesenfeld was born in Brieg, Silesia (now Brzeg, Śląsk, Poland) on 25 October 1877 and died in Stockholm on 19 May 1957. He studied at the universities of Heidelberg, Breslau and Göttingen. In Göttingen he studied with Walther Nernst (1864–1941) and earned his doctorate under him in 1901 (*Riesenfeld 1902a*). Two years later Riesenfeld went as an assistant to the Chemisches Universitätslaboratorium at the University of Freiburg. From 1913 until 1920 he was director and Extraordinarius at the Technologisches Institut of the university. In 1920 he became an Extraordinarius at the Institut für Physikalische Chemie at the University of Berlin. Riesenfeld lost this position in 1934 and went to Stockholm where he worked at the Nobelinstitut för fysikalisk kemi until 1952. In 1953 he was appointed an honorary professor at the University of Hamburg. Besides physical chemistry Riesenfeld worked in inorganic chemistry.

[2] The two solvents in Riesenfeld's experiments were (water-saturated) phenol and water.

[3] The law got its name from Adolf Fick (1829–1901), professor of physiology in Zurich and Würzburg, who formulated it in 1855 as an assistant professor (*Privatdozent*) in Zurich.

[4] Nernst referred to earlier results (1872) of Marcelin Berthelot (1827–1907) and Émile Jungfleisch (1839–1916) (*Nernst 1890*, 404–407, *1891*, 113–115).

[5] In 1909 Bernhard Reinhold (1885–?) got his Ph. D. degree, based on his investigations of the dependence on transference numbers on the concentration of the solute and the type of the solvents (*Reinhold 1910*).

H.-D. Ebbinghaus, A. Kanamori (Eds.), *Ernst Zermelo - Collected Works/ Gesammelte Werke II*, DOI 10.1007/978-3-540-70856-8_12, Schriften der Mathematisch-naturwissenschaftlichen Klasse der Heidelberger Akademie der Wissenschaften 23, © Springer-Verlag Berlin Heidelberg 2013

The main topic of *Riesenfeld and Zermelo 1909* is the determination of the values and of the temporal behavior of concentrations of a solute at the boundary of two not or only slightly miscible solvents.[7] The authors point out that this problem is of current importance, as phenomena at boundary surfaces have been treated by several scholars because of their chemical and physiological importance.[8]

For solving the problem, the validity of Nernst's distribution law (*Nernst 1891*)[9] is assumed as well as the realization that a chemical equilibrium at the boundary of two phases will instantaneously appear and that it will be maintained in diffusion-controlled processes. With regard to this knowledge the authors refer to *Nernst 1897*[10] and to *Noyes and Whitney 1897*.[11]

[6] This conjecture is highly probable, because there were close connections between Walther Nernst and his team on the one hand and the mathematicians of the University of Göttingen on the other. In 1894 the 30-year old Nernst got invited to chair mathematical physics in Munich as successor of Ludwig Boltzmann (1844–1906). Not alone because of the initiative of the dean at this time, namely Felix Klein (1849–1925), Nernst remained in Göttingen, where he was appointed a full professor of physical chemistry at the end of 1894 and got his own institute (cf. *Bartel and Huebener 2007*, 83–88). Nernst was a good friend of David Hilbert (1862–1943), and together with Arthur Schoenflies (1853–1928) he edited the textbook *Nernst and Schoenflies 1895* about mathematics for natural scientists (cf. *Bartel and Huebener 2007*, 129–131).

[7] This property of the two solvents is not explicitly mentioned, but can be deduced from the text.

[8] In that context it may be possible that Nernst published his law of electrical nerve stimulus threshold (Reizschwellengesetz) of 1899 (*Nernst 1899*) in an extended and more detailed manner in 1908 (*Nernst 1908*). In these papers "concentration changes, which are generated by the corresponding current at the boundary between protoplasm and cell liquid" ("Konzentrationsänderungen [. . .], die durch den betreffenden Strom an der Grenze von Protoplasma und Zellsaft hervorgebracht werden") (*Nernst 1908*, 312) play a role.

[9] Nernst's law of distribution states that in the equilibrium and at a given temperature the quotient of the two activities (concentrations in the case of dilution) of a substance soluble in two immiscible solvents becomes a constant (distribution coefficient) which does not depend on other substances solved in the two solvents. Eduard Riecke (1845–1915), full professor (Ordinarius) of physics in Göttingen, derived the law of distribution strictly from thermodynamics (*Riecke 1891*, 108–110).

[10] In this publication Nernst is mainly investigating the appearance and dissolution of solid phases of metal mixtures. The boundary surface separates the solid or liquid phase (amalgam) from the aqueous solution. Nernst emphasizes "that the chemical equilibrium at the boundary surface is preserved at any moment in time" ("dass an der Grenzfläche das chemische Gleichgewicht in jedem Augenblicke gewahrt bleibt") (*Nernst 1897*, 541).

[11] The publication of Arthur Amos Noyes (1866–1936) and Willis Rodney Whitney (1868–1958) was written at the Massachusetts Institute of Technology in Boston. Both were Ph. D. students of Wilhelm Ostwald (1853–1932) in Leipzig.

To solve the mathematical problem a horizontal diffusion cylinder with a unit area of the cross-section is supposed[12] and the following is assumed:

Coordinate x	infinite length	$-\infty < x < 0$	$x = 0$	$0 < x < \infty$
	finite length	$-l_1 \leq x < 0$		$0 < x \leq l_2$
Phase		solvent L_1	boundary surface	solvent L_2

The paper may be divided into five sections schematized as follows:

	Number of substances	Length of cylinder	Task
(A)	1	infinite	Determination of the boundary concentrations u_1, u_2 under seven conditions (see below) and under the assumption that they are constant in time
(B)			Proof for the correctnes of the assumption about the boundary concentrations (see (A))
(C)	> 1		Generalization to several substances
(D)	1	finite	Determination of the concentration u_1 in L_1 and of u_2 in L_2 at all points x for $t \to \infty$
(E)			Proof of the constancy of the boundary concentrations u_1 and u_2 if $\frac{l_1}{l_2} = \sqrt{\frac{D_1}{D_2}}$ (D_1, D_2 the diffusion constants in L_1 and L_2, resp.)

The paper contains the experimental confirmation of the Noyes-Whitney equation that describes the rate of dissolution of a solid by $dx/dt = C(S - x)$ (x the concentration of the solute, t the time, S the concentration at saturation, C a constant). With respect to the boundary between the solid and the aqueous solution it is noted "that the solid substance is always surrounded by an infinitely thin layer of its saturated solution" ("dass die feste Substanz stets mit einer unendlich dünnen Schicht von ihrer gesättigten Lösung umgebenen ist") (*Noyes and Whitney 1897*, 690).—The results, mentioned in the two publications, were applied to the problem of "the boundary surface of different solvents". Therefore it seems to be justified to assume instantaneous equilibria as well as the validity of Nernst's law of distribution for the actually considered system, although the authors do not explicit mention these two assumptions (see below).

[12] The paper does not state that the temperature is assumed to be constant in time and space. This assumption is crucial because both the diffusion coefficient and the distribution coefficient depend on the temperature.

(**A**) In Table 1 the seven conditions are summarized; they are the basis for calculating the concentration $u = u(x,t)$ and especially $u(-0,t) = u_1$ and $u(+0,t) = u_2$.

	Coordinate x	Time t	Conditions (A)
(1)[13]	< 0	> 0	$\dfrac{\partial u}{\partial t} = D_1 \dfrac{\partial^2 u}{\partial x^2}$
(2)[13]	> 0		$\dfrac{\partial u}{\partial t} = D_2 \dfrac{\partial^2 u}{\partial x^2}$
(3)[14]	$= 0$		$u_1 - k u_2 = 0$
(4)	$= 0$		$D_1 \Big(\dfrac{\partial u}{\partial x}\Big)_1 - D_2 \Big(\dfrac{\partial u}{\partial x}\Big)_2 = 0$
(5)	< 0	$= 0$	$u = c_1$
	$\to -\infty$	> 0	
(6)	> 0	$= 0$	$u = c_2$
	$\to \infty$	> 0	
(7)	$x \neq 0$	> 0	u and $\dfrac{\partial u}{\partial x}$ are continuous

Table 1

Assuming that u_1 and u_2 are constant with respect to time and under the conditions (5) and (6), the solutions of the differential equations (1) and (2) resp., can be found in Riemann's textbook on partial differential equations of mathematical physics ("after Riemann's lectures" ("nach Riemann's Vorlesungen"), *Riemann and Weber 1900, 1901*):[15]

$$u = u_1 + (c_1 - u_1)\,\Theta\,(\xi_1)\ ,$$

$$\text{where} \quad \xi_1 = \frac{(-1)^i x}{2\sqrt{D_1 t}} \quad \text{and} \quad \Theta(\xi) = \frac{2}{\sqrt{\pi}} \int_0^{\xi} e^{-\beta^2} d\beta \quad (i = 1, 2)\,.$$

[13] Fick's law (D_1, D_2 the diffusion coefficients in L_1 and L_2, resp.).

[14] Nernst's distribution law ($k = \text{const}$ the distribution coefficient).

[15] Riesenfeld and Zermelo cite §37 of the second volume of Riemann's textbook (*Riemann and Weber 1901*, 94–96, actualized and newly edited (4th edition) by Heinrich Weber (1842–1913)). This (sub)section with the title "Begrenzter Körper" ("Finite body") is part of the section "Probleme der Wärmeleitung, die nur von einer Coordinate abhängig sind" ("Problems of heat transport depending on only one coordinate") and is the logical extension of the preceding section "[...] Unbegrenzter Körper" ("[...] Infinite body") (*Riemann and Weber 1901*, 91–94).

Taking into account conditions (3) and (4), the concentrations at the boundary surface (x = 0) can be formulated as[16]

$$u_i = \frac{k^{(2-i)}\left(c_1 + c_2\sqrt{\dfrac{D_2}{D_1}}\right)}{k + \sqrt{\dfrac{D_2}{D_1}}} \quad (i = 1, 2)\,. \tag{8}$$

(**B**) Riesenfeld and Zermelo show that equations (8) are also valid for all $t > 0$ in the case of an infinite cylinder. The basic idea of the proof is the following: Equations (8) are a solution of the partial differential equations, taking all seven conditions (A) (Table 1) into account. If under all these seven conditions two solutions $u = u^{(I)}$ and $u = u^{(II)}$ would exist, then the difference $v = u^{(I)} - u^{(II)}$ (with the values v_1 and v_2 at $x = 0$) would be a function which is not identically 0 and satisfies the seven conditions gained from those listed in Table 1 in the following way: (i) replace u by v in the equations (1) to (4) and (7); (ii) use $v = 0$ both for conditions (5) and (6).

However, it can be shown that the function

$$V = V(t) = \int_{-\infty}^{0} v^2 dx + k\int_{0}^{\infty} v^2 dx$$

is 0 under these conditions. Therefore, the difference v vanishes for all $t > 0$ and for all x. Hence the time-independent equations (8) are the unique solution for any time $t > 0$ in the case of an infinite diffusion cylinder.

In the case of a finite cylinder with $-l_1 \le x \le l_2$, uniqueness is proved similarly.

(**C**) Assuming dilute solutions and the independence of diffusion on other solutes, it is concluded that the equations (8) are valid for any single component of a solute system.

(**D**) Considerations are made for the presence of a finite diffusion cylinder. In this case, the vanishing of the concentration gradient by diffusion in a finite time will be noticeable, so that the initial boundary concentrations also will be changed in general.

The concentrations $u_1^{(\infty)}$ and $u_2^{(\infty)}$ are calculated, which prevail everywhere in the solvent L_1 and L_2, resp., after total equalization ($t \to \infty$) depending on the initial ones c_1 and c_2 ($t = 0$), resp., and on the length l_1 and l_2, resp. Since for the total amount of substances[17] $c_1 l_1 + c_2 l_2 = u_1^{(\infty)} l_1 + u_2^{(\infty)} l_2$

[16] The two equations (8) correspond to equations (10) and (11) in the original paper (*Riesenfeld and Zermelo 1909*, 959). In these equations, k is the distribution coefficient (see condition (3) in Table 1 and n. 16).

[17] The text speaks of "amounts of salt" ("Salzmengen") (*Riesenfeld and Zermelo 1909*, 960), possibly with a view to a later application.

must be valid, the following equations are found:

$$u_i^{(\infty)} = \frac{k^{(2-i)}\left(c_1 + c_2\frac{l_2}{l_1}\right)}{k + \frac{l_2}{l_1}} \quad (i = 1, 2)\,. \tag{9}$$

Comparing (8) with (9) shows that under the condition $\frac{l_1}{l_2} = \sqrt{\frac{D_1}{D_2}}$ the initial concentrations u_1, u_2 at the boundary surface are equal to those ones $u_1^{(\infty)}$, $u_2^{(\infty)}$ after full equalization.

(**E**) Finally it is shown that in the latter mentioned case, the concentrations at the boundary surface maintain their initial values (8) during the entire process.

In the final paragraph, an application of the relation (8) is announced in a "forthcoming" ("demnächst erscheinenden") publication. This is the paper *Riesenfeld and Reinhold 1910b* in the *Zeitschrift für physikalische Chemie*. As stated above, it in fact inspired the present paper. The editorial office of the *Zeitschrift* obtained the manuscript already on 29 September 1909 whereas the present paper was received by *Physikalische Zeitschrift* only later, namely on 12 October 1909. This may be the reason that the reference in *Riesenfeld and Reinhold 1910b* (p. 461, n. 1) to *Riesenfeld and Zermelo 1909* is given without page numbers.

In *Riesenfeld and Reinhold 1910b* the task is to determine transference numbers by measuring the electromotive force E in a solvent L_{II}, whose miscibility with water is only limited. For the implementation of the required theory, the following arrangement is considered, which was already used by Riesenfeld earlier in *Riesenfeld 1902a* and *1902c*:

electrode	L_{I}		L_{II}			L_{I}		electrode
	c_1	$\overline{c}_1$	$\overline{\gamma}_1$	γ	$\overline{\gamma}_2$	$\overline{c}_2$	c_2	
	A		B		C		D	

A, B, C, D denote the phase boundaries, c_1, c_2, γ the concentrations and $\overline{c}_1, \overline{\gamma}_1, \overline{\gamma}_2, \overline{c}_2$ the concentrations at the boundaries. Riesenfeld and Reinhold assume an instantaneous distribution equilibrium at B and C. They justify this assumption, referring to *Nernst 1897* and *Noyes and Whitney 1897*. Beyond this they refer to a publication of Emil Abel (*Abel 1906*), who describes a problem, similar to that of *Riesenfeld and Reinhold 1910b*.[18] Following

[18] Like Riesenfeld, Emil Abel (1875–1958) was graduated in 1901 with Walther Nernst in Göttingen. Later he was professor of physical chemistry and director of the Institute for Physical Chemistry at the Technische Hochschule in Vienna. In 1938 he emigrated to London.

Nernst's distribution law, they find

$$\frac{\overline{c}_1}{\overline{\gamma}_1} = \frac{\overline{c}_2}{\overline{\gamma}_2} = k.$$

Consider now a binary electrolyte with transference numbers n_I, n_{II} and diffusion coefficients D_I, D_{II} in L_I and L_{II}, resp. Then the electromotive force E of the arrangement shown above is

$$E = 2RTn_I \ln\frac{c_1}{c_2} + 2RT(n_{II} - n_I)\ln\frac{\overline{c}_1}{\overline{c}_2} \tag{10}$$

(R universal gas constant, T absolute temperature). "According to the theorem, which was established by one of us" ("Nach dem von einem von uns aufgestellten Satze", *Riesenfeld and Reinhold 1910b*, 461), the concentrations $\overline{c}_1$ and $\overline{c}_2$ can be calculated according to equations (8). Inserting these values

Die Einstellung der Grenzkonzentrationen an der Trennungsfläche zweier Lösungsmittel

Riesenfeld and Zermelo 1909

Werden zwei Lösungsmittel, in denen beliebig viele Stoffe gelöst sind, miteinander in Berührung gebracht, so gleichen sich durch Diffusion die Konzentrationen allmählich derart aus, daß im Gleichgewichtszustande jede Molekülart ein konstantes Teilungsverhältnis zwischen den einzelnen Lösungsmitteln hat, unabhängig von der Gegenwart anderer Molekülgattungen und davon, ob die betrachtete Molekülgattung mit den andern sich chemisch umsetzt oder nicht[1]. Es ist ferner bekannt, daß an der Grenzfläche verschiedener Lösungsmittel das chemische Gleichgewicht in jedem Augenblick gewahrt bleibt, da im Vergleich zu der Langsamkeit, mit der sich Diffusionsvorgänge abspielen, sich das Gleichgewicht an der Grenzfläche mit sehr großer Geschwindigkeit einstellt[2].

Unbeantwortet ist aber bis jetzt die Frage, wie groß die Konzentrationen sind, die sich an der Trennungsfläche einstellen, und in welcher Weise sie sich im Verlaufe des Diffusionsvorganges ändern. Diese Frage hat dadurch besondere Bedeutung gewonnen, daß die elektrischen und kapillaren Erscheinungen an Grenzflächen, zum Teil auch infolge ihrer chemischen und physiologischen Bedeutung, neuerdings mehrfach behandelt wurden.

[1] *Nernst*, Zeitschr. f. phys. Chem. **8**, 110, 1891.

[2] *Nernst*, Zeitschr. f. phys. Chem. **22**, 541, 1897; *Noyes* und *Whithney*, Zeitschr. f. phys. Chem. **23**, 689, 1897.

in (10) the following expression is obtained:

$$E = 2RTn_I \ln\frac{c_1}{c_2} + 2RT(n_{II} - n_I)\ln\frac{c_1 + \gamma\sqrt{\frac{D_{II}}{D_I}}}{c_2 + \gamma\sqrt{\frac{D_{II}}{D_I}}}. \tag{11}$$

Assuming $\gamma \ll \min(c_1, c_2)$ and $D_I \ll D_{II}$, which in general is justified, the expression $\gamma\sqrt{\frac{D_{II}}{D_I}}$ can be neglected in comparison to c_1 and c_2, resp. Then from (11) the equation

$$E = 2RTn_{II}\ln\frac{c_1}{c_2}$$

follows, which was some years earlier found by Riesenfeld (*Riesenfeld 1902a, 1902b*, 619) in order to determine the transference number n_{II}.

The settling of the boundary concentrations at the dividing surface of two solvents

Riesenfeld and Zermelo 1909

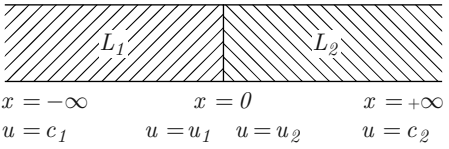
If two solvents in which any number of substances are dissolved are brought into contact with one another, then the concentrations slowly balance out by diffusion so that in the equilibrium state each kind of molecule has a constant division ratio between the individual solvents independently of the presence of other categories of molecules and of whether the category of molecules under consideration chemically reacts with the others or not.[1] Furthermore, it is well-known that the chemical equilibrium at the boundary surface of different solvents is preserved at any moment in time, since the equilibrium at the boundary surface sets in at a very great velocity, compared to the slow pace in which processes of diffusion unfold.[2]

But heretofore the question has remained unanswered as to how great the concentrations are that occur at the dividing surface and how they change during the process of diffusion. This question has become particularly significant due to the fact that the electric and capillary phenomena at boundary surfaces have been the subject of several discussions recently, in part also owing to their chemical and physiological relevance.

[1] *Nernst 1891.*

[2] *Nernst 1897*, 541; *Noyes and Whitney 1897.* ⟦Riesenfeld and Zermelo erroneously write "Whithney" instead of "Whitney".⟧

Der Einfachheit halber sei im folgenden angenommen, daß die Lösungsmittel L_1 und L_2 nur den Stoff S gelöst haben, und daß, solange die Lösungen noch getrennt sind, die Konzentrationen von S in L_1 überall c_1 und in L_2 überall c_2 betragen. Zur Zeit $t = 0$ werden L_1 und L_2 in einem horizontalen Diffusionszylinder von unendlicher Länge und dem Querschnitt 1 so aneinander geschichtet, daß sich bei $x = 0$ die Grenzfläche befindet. Für alle Werte $x < 0$ sei S in L_1 und für alle Werte $x > 0$ in L_2 gelöst; seine Konzentration während des Diffusionsvorganges sei mit u bezeichnet (s. die Fig.).



Die unmittelbar an der Grenzfläche in L_1 und L_2 herrschenden Konzentrationen seien u_1 und u_2 genannt. Die Diffusionskoeffizienten von S mögen in L_1 und L_2 die Werte D_1 und D_2 haben, und k sei der „Teilungskoeffizient" von S in L_1 und L_2. Unter dem Teilungskoeffizienten wird das Verhältnis der Konzentrationen von S in L_1 und L_2 im Gleichgewichtszustande verstanden, das, wie oben angeführt, bei konstanter Temperatur von den Konzentrationen unabhängig ist.

Dann gelten für u die folgenden Bedingungen. Nach dem *Fick*schen Diffusionsgesetz ist:

$$\frac{\partial u}{\partial t} = D_1 \frac{\partial^2 u}{\partial x^2} \quad (x < 0)\,, \tag{1}$$

$$\frac{\partial u}{\partial t} = D_2 \frac{\partial^2 u}{\partial x^2} \quad (x > 0)\,. \tag{2}$$

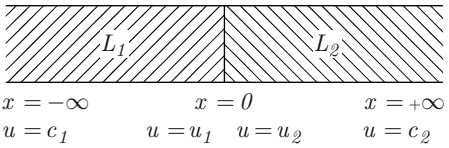
Nach dem *Nernst*schen Verteilungssatz ist

$$u_1 - k u_2 = 0 \quad (x = 0)\,. \tag{3}$$

Da die an der Grenzfläche aus einem Lösungsmittel hinausdiffundierende Substanzmenge gleich der in das andere Lösungsmittel hineindiffundierenden sein muß, so ist ferner bei $x = 0$

$$D_1 \Big(\frac{\partial u}{\partial x}\Big)_1 - D_2 \Big(\frac{\partial u}{\partial x}\Big)_2 = 0\,. \tag{4}$$

For simplicity's sake, we shall assume in what follows that the solvents L_1 and L_2 have only dissolved the substance S and that, as long as the solutions are still separated, the concentrations of S are c_1 everywhere in L_1 and c_2 everywhere in L_2. At time $t = 0$, L_1 and L_2 are layered in a horizontal diffusion cylinder of infinite length and with a cross sectional area of 1 so that the boundary surface lies at $x = 0$. Let us assume that S is dissolved in L_1 for all values $x < 0$ and dissolved in L_2 for all values $x > 0$; its concentration during the diffusion process shall be denoted by u (s. the fig.).



The concentrations directly obtaining at the boundary surfaces in L_1 and L_2 shall be denoted by u_1 and u_2. The diffusion coefficients of S in L_1 and L_2 shall have the values D_1 and D_2, and k shall be the "division coefficient" of S in L_1 and L_2. By the division coefficient we mean the ratio of the concentrations of S in L_1 and L_2 in the state of equilibrium, which at constant temperature is independent of the concentrations, as stated above.

Then the following conditions hold for u. By *Fick*'s law of diffusion we have

$$\frac{\partial u}{\partial t} = D_1 \frac{\partial^2 u}{\partial x^2} \quad (x < 0)\,, \tag{1}$$

$$\frac{\partial u}{\partial t} = D_2 \frac{\partial^2 u}{\partial x^2} \quad (x > 0)\,. \tag{2}$$

By *Nernst*'s law of distribution we have

$$u_1 - k u_2 = 0 \quad (x = 0)\,. \tag{3}$$

Since the amount of substance diffusing out of one solvent at the boundary surface must be equal to the amount of substance diffusing into the other solvent, we also have, for $x = 0$,

$$D_1 \Big(\frac{\partial u}{\partial x}\Big)_1 - D_2 \Big(\frac{\partial u}{\partial x}\Big)_2 = 0\,. \tag{4}$$

Endlich ist nach dem oben Gesagten

$$u = c_1 \text{ für } \begin{cases} x < 0 \\ t = 0 \end{cases} \quad \text{und für } \begin{cases} x = -\infty \\ t > 0 \end{cases} , \tag{5}$$

$$u = c_2 \text{ für } \begin{cases} x > 0 \\ t = 0 \end{cases} \quad \text{und für } \begin{cases} x = +\infty \\ t > 0 \end{cases} , \tag{6}$$

$$u \text{ und } \frac{\partial u}{\partial x} \text{ stetig an jeder Stelle } x \gtrless 0 . \tag{7}$$

Macht man zunächst die Voraussetzung, daß bei $x = 0$ von der Zeit unabhängig $u = u_1$ ist, so erhält man für (1) und (5) die Lösung[3]

$$u = u_1 + (c_1 - u_1)\, \Theta(\xi_1) ,$$

worin

$$\xi_1 = \frac{-x}{2\sqrt{D_1 t}}$$

ist. Hier bedeutet

$$\Theta(\xi) = \frac{2}{\sqrt{\pi}} \int_0^{\xi} e^{-\beta^2} d\beta ,$$

$$\Theta'(\xi) = \frac{2}{\sqrt{\pi}} e^{-\xi^2} .$$

Für $x = 0$ wird somit

$$\left(\frac{\partial u}{\partial x}\right)_1 = \frac{u_1 - c_1}{\sqrt{\pi D_1 t}} . \tag{8}$$

Ganz analog erhält man unter der Voraussetzung, daß bei $x = 0$ von der Zeit unabhängig $u = u_2$ ist, für (2) und (6) die Lösung

$$u = u_2 + (c_2 - u_2)\, \Theta(\xi_2) ,$$

wobei

$$\xi_2 = \frac{x}{2\sqrt{D_2 t}}$$

959 | ist. Für $x = 0$ wird dann

$$\left(\frac{\partial u}{\partial x}\right)_2 = \frac{c_2 - u_2}{\sqrt{\pi D_2 t}} . \tag{9}$$

[3] *Riemann-Weber*, Die partiellen Differentialgleichungen der mathematischen Physik, Bd. II, § 37.

Finally, by what has been said above,

$$u = c_1 \text{ for } \begin{cases} x < 0 \\ t = 0 \end{cases} \quad \text{and for } \begin{cases} x = -\infty \\ t > 0 \end{cases} , \tag{5}$$

$$u = c_2 \text{ for } \begin{cases} x > 0 \\ t = 0 \end{cases} \quad \text{and for } \begin{cases} x = +\infty \\ t > 0 \end{cases} , \tag{6}$$

$$u \text{ and } \frac{\partial u}{\partial x} \text{ are continuous at every point } x \gtrless 0 . \tag{7}$$

If we first assume that, for $x = 0$, we have $u = u_1$ independently of time, we then obtain for (1) and (5) the solution[3]

$$u = u_1 + (c_1 - u_1)\, \Theta\,(\xi_1) ,$$

wherein

$$\xi_1 = \frac{-x}{2\sqrt{D_1 t}} .$$

Here

$$\Theta(\xi) = \frac{2}{\sqrt{\pi}} \int_0^{\xi} e^{-\beta^2} d\beta ,$$

$$\Theta'(\xi) = \frac{2}{\sqrt{\pi}} e^{-\xi^2} .$$

Hence for $x = 0$

$$\left(\frac{\partial u}{\partial x}\right)_1 = \frac{u_1 - c_1}{\sqrt{\pi D_1 t}} . \tag{8}$$

Analogously, assuming that $u = u_2$ independently of time when $x = 0$, we obtain for (2) and (6) the solution

$$u = u_2 + (c_2 - u_2)\, \Theta\,(\xi_2) ,$$

where

$$\xi_2 = \frac{x}{2\sqrt{D_2 t}} .$$

Then for $x = 0$

$$\left(\frac{\partial u}{\partial x}\right)_2 = \frac{c_2 - u_2}{\sqrt{\pi D_2 t}} . \tag{9}$$

[3] *Riemann and Weber 1901*, § 37.

Setzt man diese Werte in (4) ein, so folgt

$$\sqrt{D_1}\,(u_1 - c_1) - \sqrt{D_2}\,(c_2 - u_2) = 0$$

oder

$$u_1\sqrt{D_1} + u_2\sqrt{D_2} = c_1\sqrt{D_1} + c_2\sqrt{D_2}$$

und hieraus unter Berücksichtigung von (3)

$$u_1 = \frac{k\left(c_1 + c_2\sqrt{\dfrac{D_2}{D_1}}\right)}{k + \sqrt{\dfrac{D_2}{D_1}}}\,, \tag{10}$$

$$u_2 = \frac{c_1 + c_2\sqrt{\dfrac{D_2}{D_1}}}{k + \sqrt{\dfrac{D_2}{D_1}}}\,. \tag{11}$$

Die so gefundenen Gleichungen (10) und (11) stellen eine Lösung der oben aufgestellten 7 Bedingungsgleichungen dar. Um zu beweisen, daß die bei ihrer Ableitung gemachte Voraussetzung berechtigt ist, daß nämlich die Grenzkonzentrationen für alle Zeiten konstant sind, ist noch zu zeigen, daß diese 7 Bedingungen den ganzen Vorgang für $t > 0$ *eindeutig* bestimmen.

Wären $u = u'$ und $u = u''$ *zwei* Lösungen derselben Gleichungen, so müßte die Differenz $v = u' - u''$, welche bei $x = 0$ die Werte v_1 und v_2 annehmen möge, folgenden Bedingungen genügen:

$$\frac{\partial v}{\partial t} = D_1\frac{\partial^2 v}{\partial x^2} \quad (x < 0) \tag{1}'$$

$$\frac{\partial v}{\partial t} = D_2\frac{\partial^2 v}{\partial x^2} \quad (x > 0) \tag{2}'$$

$$\left.\begin{aligned} v_1 - kv_2 &= 0 \qquad (3)' \\ D_1\Big(\frac{\partial v}{\partial x}\Big)_1 - D_2\Big(\frac{\partial v}{\partial x}\Big)_2 &= 0 \qquad (4)' \end{aligned}\right\} (x = 0\,,\ t > 0)\,,$$

$$v = 0 \text{ für } \begin{cases} t = 0 \\ x < 0 \end{cases} \text{ und für } \begin{cases} t > 0 \\ x = -\infty \end{cases}\,, \tag{5}'$$

$$v = 0 \text{ für } \begin{cases} t = 0 \\ x > 0 \end{cases} \text{ und für } \begin{cases} t > 0 \\ x = +\infty \end{cases}\,, \tag{6}'$$

$$v \text{ und } \frac{\partial v}{\partial x} \textit{ stetig} \text{ für alle } x \gtrless 0\,, \tag{7}'$$

und es ist zu zeigen, daß dann immer $v = 0$ sein muß.

If we now substitute these values in (4), it then follows that

$$\sqrt{D_1}\,(u_1 - c_1) - \sqrt{D_2}\,(c_2 - u_2) = 0$$

or

$$u_1\sqrt{D_1} + u_2\sqrt{D_2} = c_1\sqrt{D_1} + c_2\sqrt{D_2}\,,$$

and from this, considering (3),

$$u_1 = \frac{k\left(c_1 + c_2\sqrt{\dfrac{D_2}{D_1}}\right)}{k + \sqrt{\dfrac{D_2}{D_1}}}\,, \tag{10}$$

$$u_2 = \frac{c_1 + c_2\sqrt{\dfrac{D_2}{D_1}}}{k + \sqrt{\dfrac{D_2}{D_1}}}\,. \tag{11}$$

The equations (10) and (11) so found represent a solution of the 7 constraint equations identified above. In order to prove that the assumption made for their derivation is justified, namely that the boundary concentrations are constant for all times, we must also show that these 7 conditions *uniquely* determine the entire process for $t > 0$.

If $u = u'$ and $u = u''$ were *two* solutions to the same equations, then the difference $v = u' - u''$, which shall assume the values v_1 and v_2 for $x = 0$, would have to satisfy the following conditions:

$$\frac{\partial v}{\partial t} = D_1\frac{\partial^2 v}{\partial x^2} \quad (x < 0) \tag{1}'$$

$$\frac{\partial v}{\partial t} = D_2\frac{\partial^2 v}{\partial x^2} \quad (x > 0) \tag{2}'$$

$$\left.\begin{aligned} v_1 - kv_2 &= 0 \qquad (3)' \\ D_1\Big(\frac{\partial v}{\partial x}\Big)_1 - D_2\Big(\frac{\partial v}{\partial x}\Big)_2 &= 0 \qquad (4)' \end{aligned}\right\} (x = 0\,,\ t > 0)\,,$$

$$v = 0 \text{ for } \begin{cases} t = 0 \\ x < 0 \end{cases} \text{ and for } \begin{cases} t > 0 \\ x = -\infty \end{cases}\,, \tag{5}'$$

$$v = 0 \text{ for } \begin{cases} t = 0 \\ x > 0 \end{cases} \text{ and for } \begin{cases} t > 0 \\ x = +\infty \end{cases}\,, \tag{6}'$$

$$v \text{ and } \frac{\partial v}{\partial x} \textit{ are continuous} \text{ for all } x \gtrless 0\,, \tag{7}'$$

and it is to be shown that, in this case, always necessarily $v = 0$.

Zu diesem Zwecke betrachten wir die Funktion

$$V = V(t) = \int_{-\infty}^{0} v^2 dx + k \int_{0}^{\infty} v^2 dx$$

und finden durch Differentiation und Benutzung von $(1)'$ und $(2)'$

$$\frac{1}{2}\frac{dV}{dt} = \int_{-\infty}^{0} v\frac{\partial v}{\partial t}dx + k\int_{0}^{\infty} v\frac{\partial v}{\partial t}dx\,,$$

$$\frac{1}{2}\frac{dV}{dt} = D_1 \int_{-\infty}^{0} v\frac{\partial^2 v}{\partial x^2}dx + kD_2 \int_{0}^{\infty} v\frac{\partial^2 v}{\partial x^2}dx\,,$$

und weiter durch partielle Integration

$$\frac{1}{2}\frac{dV}{dt} = \left[D_1 v\frac{\partial v}{\partial x}\right]_{-\infty}^{-0} - D_1 \int_{-\infty}^{0} \left(\frac{\partial v}{\partial x}\right)^2 dx$$
$$+ \left[kD_2 v\frac{\partial v}{\partial x}\right]_{+0}^{\infty} - kD_2 \int_{0}^{\infty} \left(\frac{\partial v}{\partial x}\right)^2 dx$$

oder, da wegen $(5)'$ und $(6)'$ $\dfrac{\partial v}{\partial x}$ im Unendlichen verschwindet und bei $x = +0$ wegen $(3)'$ und $(4)'$ die Grenzglieder sich aufheben,

$$\frac{1}{2}\frac{dV}{dt} = -D_1 \int_{-\infty}^{0} \left(\frac{\partial v}{\partial x}\right)^2 dx - kD_2 \int_{0}^{\infty} \left(\frac{\partial v}{\partial x}\right)^2 dx \leqq 0\,.$$

Durch Integration folgt also für ein beliebiges positives t

$$V(t) - V(0) = \int_{0}^{t} \frac{dV}{dt}dt \leqq 0$$

oder, da nach $(5)'$ und $(6)'$ bei $t = 0$ mit v auch V verschwindet,

$$V(t) \leqq 0\,,$$

während doch V seiner Definition nach wesentlich positiv sein muß. Somit folgt $V = 0$ und hieraus weiter $v = 0$ für beliebige Werte von $t > 0$ und x.

Dieser Eindeutigkeitsbeweis läßt sich ohne weiteres auf den Fall eines *endlichen* Diffusionszylinders übertragen, an dessen Grenzen $x = -l_1$ und

To this end we consider the function

$$V = V(t) = \int_{-\infty}^{0} v^2 dx + k \int_{0}^{\infty} v^2 dx$$

and, through differentiation and use of $(1)'$ and $(2)'$, find

$$\frac{1}{2}\frac{dV}{dt} = \int_{-\infty}^{0} v\frac{\partial v}{\partial t} dx + k \int_{0}^{\infty} v\frac{\partial v}{\partial t} dx\,,$$

$$\frac{1}{2}\frac{dV}{dt} = D_1 \int_{-\infty}^{0} v\frac{\partial^2 v}{\partial x^2} dx + kD_2 \int_{0}^{\infty} v\frac{\partial^2 v}{\partial x^2} dx\,,$$

and furthermore, by partial integration,

$$\frac{1}{2}\frac{dV}{dt} = \left[D_1 v\frac{\partial v}{\partial x} \right]_{-\infty}^{-0} - D_1 \int_{-\infty}^{0} \left(\frac{\partial v}{\partial x}\right)^2 dx$$

$$+ \left[kD_2 v\frac{\partial v}{\partial x} \right]_{+0}^{\infty} - kD_2 \int_{0}^{\infty} \left(\frac{\partial v}{\partial x}\right)^2 dx$$

or, since, on account of $(5)'$ and $(6)'$, $\dfrac{\partial v}{\partial x}$ vanishes in the infinite and since the boundary terms cancel each other out for $x = +0$ on account of $(3)'$ and $(4)'$,

$$\frac{1}{2}\frac{dV}{dt} = -D_1 \int_{-\infty}^{0} \left(\frac{\partial v}{\partial x}\right)^2 dx - kD_2 \int_{0}^{\infty} \left(\frac{\partial v}{\partial x}\right)^2 dx \leqq 0\,.$$

Hence, for any positive t, it follows by integration that

$$V(t) - V(0) = \int_{0}^{t} \frac{dV}{dt} dt \leqq 0$$

or, since, by $(5)'$ and $(6)'$, V vanishes along with v for $t = 0$,

$$V(t) \leqq 0\,,$$

whereas, by its definition, V must essentially be positive. Thus, it follows that $V = 0$, and from this in turn that $v = 0$ for any values of $t > 0$ and x.

This uniqueness proof can be extended easily to the case of a *finite* diffusion cylinder at the boundaries $x = -l_1$ and $x = +l_2$ of which the condition

$x = +l_2$ die Bedingung $\dfrac{\partial u}{\partial x} = 0$ für jede Lösung erfüllt sein muß. Dann ergibt sich ebenso für die Differenz v zweier Lösungen

$$V = \int_{-l_1}^{0} v^2 dx + k \int_{0}^{l_2} v^2 dx = 0$$

und hieraus $v = 0$.

Bisher wurde stets angenommen, daß nur *ein* Stoff in beiden Lösungsmitteln gelöst ist. Sind mehrere Stoffe gelöst, so läßt sich für jeden einzelnen das gleiche ausführen, da sowohl die Verteilung eines Stoffes zwischen zwei Lösungsmitteln wie seine Diffusion von der Gegenwart anderer Molekülgattungen in verdünnten Lösungen unabhängig ist.

Es ist also hierdurch der allgemeine Satz gefunden: *Schichtet man zwei Lö-*
960 *sungsmittel übereinander, in denen beliebig viele | Stoffe gelöst sind, so stellt sich an der Trennungsfläche sofort für jeden der Stoffe eine Konzentration u_1 bzw. u_2 ein, die durch die Anfangskonzentrationen in beiden Lösungsmitteln c_1 und c_2 und die Teilungs- und Diffusionskoeffizienten k, D_1 und D_2 bestimmt ist. Sie beträgt:*

$$u_1 = \frac{k\left(c_1 + c_2\sqrt{\dfrac{D_2}{D_1}}\right)}{k + \sqrt{\dfrac{D_2}{D_1}}}, \quad u_2 = \frac{c_1 + c_2\sqrt{\dfrac{D_2}{D_1}}}{k + \sqrt{\dfrac{D_2}{D_1}}}.$$

Diese Konzentrationen bleiben auch während des Diffusionsverlaufes konstant, vorausgesetzt, daß die Diffusionszylinder genügend groß gemacht werden können, so daß keine Spiegelung an ihren Enden eintritt, was bei der Langsamkeit der Diffusionsvorgänge praktisch leicht zu erreichen ist.

Für das erste Stadium des Diffusionsvorganges kann man also stets die Konzentrationen an der Trennungsfläche als konstant ansehen und nach obigem leicht berechnen. Nach längerer Zeit aber müssen sich, wenn die Diffusionszylinder die *begrenzten* Längen l_1 und l_2 haben, die Konzentrationen allmählich ausgleichen, dann tritt also im allgemeinen eine Änderung der Grenzkonzentrationen ein. Der nach unendlich langer Zeit erreichte Wert ist auf folgende Weise leicht zu berechnen:

Die Summe der vor dem Zusammenbringen und nach dem völligen Ausgleich der Konzentrationen in beiden Lösungen vorhandenen Salzmengen muß gleich sein. Nennt man die an allen Orten zur Zeit $t = \infty$ in L_1 herrschende Konzentration u_1 und die in L_2 herrschende Konzentration u_2, so ist also:

$$c_1 l_1 + c_2 l_2 = u_1 l_1 + u_2 l_2 . \tag{12}$$

Setzt man in dieser Gleichung nach dem Verteilungssatz einmal

$$u_1 = k u_2$$

$\frac{\partial u}{\partial x} = 0$ must be satisfied for each solution. Likewise, we then obtain for the difference v of two solutions

$$V = \int_{-l_1}^{0} v^2 dx + k \int_{0}^{l_2} v^2 dx = 0$$

and, from this, $v = 0$.

So far, we have always assumed that only *one* substance is dissolved in the two solvents. If several substances are dissolved, then it is possible to proceed for each of them in identical fashion, since both the distribution of a substance among two solvents and its diffusion are independent of the presence of other categories of molecules in diluted solutions.

We have thus found the general theorem: *If two solvents in which any number of substances are dissolved are layered on top of one another, then, for each substance, a concentration u_1, and u_2 respectively, instantly occurs at the dividing surface that is determined by the initial concentrations in the two solvents c_1 and c_2 as well as by the division and diffusion coefficients k, D_1 and D_2. It is:*

$$u_1 = \frac{k\left(c_1 + c_2\sqrt{\frac{D_2}{D_1}}\right)}{k + \sqrt{\frac{D_2}{D_1}}}, \quad u_2 = \frac{c_1 + c_2\sqrt{\frac{D_2}{D_1}}}{k + \sqrt{\frac{D_2}{D_1}}}.$$

These concentrations also remain constant during the diffusion process, provided that the diffusion cylinders can be made sufficiently large so that no reflection occurs at their ends, which can be achieved easily, considering the slow pace of the diffusion processes.

Hence, we can always consider the concentrations at the dividing surface as being constant for the first phase of the diffusion process and easily calculate it according to what was said above. However, if the diffusion cylinders have the *limited* lengths l_1 and l_2, then the concentrations must gradually balance out after a while. Thus, generally, a change in the boundary concentrations occurs in this case. The value reached after an infinitely long period of time can easily be calculated as follows:

The sums of the amounts of salt present in both solutions before they are brought together and after the concentrations have completely balanced out must be equal to one another. If we denote the concentration obtaining in L_1 at all points at time $t = \infty$ by u_1, and the concentration obtaining in L_2 by u_2, we thus have:

$$c_1 l_1 + c_2 l_2 = u_1 l_1 + u_2 l_2 . \tag{12}$$

If we substitute in this equation according to the distribution theorem

$$u_1 = k u_2$$

und ein andermal

$$u_2 = \frac{u_1}{k}$$

ein, so folgt

$$u_1 = \frac{k\left(c_1 + c_2\frac{l_2}{l_1}\right)}{k + \frac{l_2}{l_1}} \tag{13}$$

$$u_2 = \frac{c_1 + c_2\frac{l_2}{l_1}}{k + \frac{l_2}{l_1}}\,. \tag{14}$$

Der Vergleich dieser beiden Formeln mit (10) und (11) zeigt, daß, *wenn sich die Längen der Diffusionszylinder wie die Quadratwurzeln aus den Diffusionskoeffizienten verhalten*, dann die während des ersten Teils der Diffusion und nach dem völligen Konzentrationsausgleich an den Trennungsflächen herrschenden Konzentrationen gleich sind.

Nun läßt sich aber zeigen, daß in dem hier betrachteten Falle

$$\frac{l_1}{l_2} = \frac{\sqrt{D_1}}{\sqrt{D_2}} \tag{15}$$

die Konzentrationen u_1 und u_2 während des ganzen Diffusionsvorganges *konstant* bleiben.

Bezeichnet man nämlich mit

$$\omega = F(\xi, \tau)$$

die (existierende und eindeutig bestimmte) Lösung der Differentialgleichung

$$\frac{\partial\omega}{\partial\tau} = \frac{\partial^2\omega}{\partial\xi^2}$$

mit den Grenzbedingungen

$$\begin{aligned} \omega &= 0 \quad \text{für} \quad \tau = 0\,, \\ \omega &= 1 \quad \text{für} \quad \xi = 0\,, \quad \tau > 0\,, \\ \frac{\partial\omega}{\partial\xi} &= 0 \quad \text{für} \quad \xi = 1\,, \quad \tau > 0\,, \end{aligned}$$

so genügt für konstante u_1 und u_2 ersichtlich der Ansatz

$$u = c_1 + (u_1 - c_1)\,F\Big(\frac{-x}{l_1}\,,\,\frac{D_1}{l_1^2}t\Big)$$
$$(-l_1 \leqq x \leqq 0)\,,$$
$$u = c_2 + (u_2 - c_2)\,F\Big(\frac{x}{l_2}\,,\,\frac{D_2}{l_2^2}t\Big)$$
$$(0 \leqq x \leqq l_2)\,,$$

on the one hand, and

$$u_2 = \frac{u_1}{k}$$

on the other hand, it then follows that

$$u_1 = \frac{k\left(c_1 + c_2\frac{l_2}{l_1}\right)}{k + \frac{l_2}{l_1}} \tag{13}$$

$$u_2 = \frac{c_1 + c_2\frac{l_2}{l_1}}{k + \frac{l_2}{l_1}}\,. \tag{14}$$

The comparison of these two formulas with (10) and (11) shows that *if the lengths of the diffusion cylinders are proportional to the square roots of the diffusion coefficients,* then the concentrations obtaining at the dividing surfaces during the first part of the diffusion and after the completed adjustment of the concentrations are equal to one another.

But it is now possible to show that in the case under consideration

$$\frac{l_1}{l_2} = \frac{\sqrt{D_1}}{\sqrt{D_2}} \tag{15}$$

the concentrations u_1 and u_2 remain *constant* during the entire process of diffusion.

For if by

$$\omega = F(\xi, \tau)$$

we denote the (existing and uniquely determined) solution of the differential equation

$$\frac{\partial\omega}{\partial\tau} = \frac{\partial^2\omega}{\partial\xi^2}$$

with the boundary conditions

$$\begin{aligned} \omega &= 0 \quad \text{for} \quad \tau = 0\,, \\ \omega &= 1 \quad \text{for} \quad \xi = 0\,, \quad \tau > 0\,, \\ \frac{\partial\omega}{\partial\xi} &= 0 \quad \text{for} \quad \xi = 1\,, \quad \tau > 0\,, \end{aligned}$$

then for constant u_1 and u_2 the ansatz

$$u = c_1 + (u_1 - c_1)\,F\Big(\frac{-x}{l_1}\,,\,\frac{D_1}{l_1^2}t\Big)$$
$$(-l_1 \leqq x \leqq 0)\,,$$
$$u = c_2 + (u_2 - c_2)\,F\Big(\frac{x}{l_2}\,,\,\frac{D_2}{l_2^2}t\Big)$$
$$(0 \leqq x \leqq l_2)\,,$$

den Differentialgleichungen (1) und (2) sowie den modifizierten Grenzbedingungen

$$u = c_1 \quad \text{für} \quad x < 0\,, \quad t = 0\,, \tag{5a}$$
$$u = c_2 \quad \text{für} \quad x > 0\,, \quad t = 0 \tag{6a}$$

und

$$\frac{\partial u}{\partial x} = 0 \quad \text{für} \quad t > 0\,, \quad x = -l_1 \tag{5, 6)''}$$
$$\text{sowie} \qquad \text{für} \quad t > 0\,, \quad x = l_2\,.$$

Um nun auch die Bedingung (4) bei $x = 0$ zu erfüllen, haben wir

$$D_1 \frac{u_1 - c_1}{l_1} F_\xi\Big(0, \frac{D_1}{l_1^2} t\Big) + D_2 \frac{u_2 - c_2}{l_2} F_\xi\Big(0, \frac{D_2}{l_2^2} t\Big) = 0\,,$$

wo F_ξ die partielle Ableitung von F nach dem ersten Argumente bedeutet. Ist die Bedingung (15) erfüllt, so reduziert sich diese Gleichung auf

$$l_1 u_1 + l_2 u_2 = l_1 c_1 + l_2 c_2 \tag{12}$$

und liefert mit (3) zusammen für u_1 und u_2 die in (13) und (14) gegebenen Ausdrücke. Aus der *Eindeutigkeit* der Lösung unter den gewählten Bedingungen[1] folgt demnach streng die Konstanz von u_1 und u_2.

961 | *Wenn sich also die Längen der Diffusionszylinder wie die Quadratwurzeln aus den Diffusionskoeffizienten verhalten, so bleiben während des ganzen Verlaufs des Diffusionsvorganges die Konzentrationen an der Grenzfläche konstant.*

Eine praktische Anwendung der hier abgeleiteten Formeln wird von dem einen von uns in einer demnächst erscheinenden Arbeit über die elektromotorischen Kräfte an Trennungsflächen gemacht werden.

Freiburg und Göttingen, im September 1909.

(Eingegangen 12. Oktober 1909.)

[1] Vgl. oben.

obviously satisfies the differential equations (1) and (2) as well as the modified boundary conditions

$$u = c_1 \quad \text{for} \quad x < 0\,, \quad t = 0\,, \tag{5a}$$

$$u = c_2 \quad \text{for} \quad x > 0\,, \quad t = 0 \tag{6a}$$

and

$$\frac{\partial u}{\partial x} = 0 \quad \text{for} \quad t > 0\,, \quad x = -l_1 \tag{5, 6)''}$$

$$\text{as well as} \quad \text{for} \quad t > 0\,, \quad x = l_2\,.$$

Now in order to also satisfy condition (4) for $x = 0$, we have

$$D_1 \frac{u_1 - c_1}{l_1} F_\xi\Big(0, \frac{D_1}{l_1^2} t\Big) + D_2 \frac{u_2 - c_2}{l_2} F_\xi\Big(0, \frac{D_2}{l_2^2} t\Big) = 0\,,$$

where F_ξ denotes the partial derivative of F with respect to the first argument. If condition (15) is satisfied, then this equation reduces to

$$l_1 u_1 + l_2 u_2 = l_1 c_1 + l_2 c_2 \tag{12}$$

and yields, together with (3), for u_1 and u_2 the expression given in (13) and (14). Thus, the constancy of u_1 and u_2 follows strictly from the *uniqueness* of the solution, given the chosen conditions.[4]

Hence, if the lengths of the diffusion cylinders are proportional to the square roots from the diffusion coefficients, then the concentrations remain constant at the boundary surface during the entire course of the diffusion process.

One of the authors will provide a practical application of the formulas derived here in a forthcoming paper on the electromagnetic forces at dividing surfaces.

Freiburg and Göttingen, September 1909.

(Received October 12, 1909.)

[4] See above.

*Introductory note to 1928 (= 1929)**

Mark E. Glickman

Zermelo 1928 is the second of two papers that Zermelo wrote concerning the game of chess during the period between that of his initial foundational work and his later return to foundations. Both *Zermelo 1913* and *Zermelo 1928* would remain initially obscure but later would come to be considered as classic, ground-breaking contributions to emergent subjects. However, while *Zermelo 1913* only has fragmentary and precursory results (cf. its introductory note in volume I of this Zermelo edition), *Zermelo 1928* has substantive results resonant with modern work.

Zermelo's *1928* focuses on measuring participants' playing strengths in chess tournaments, and is a remarkable work in the history of paired comparison modeling. Apart from several contemporary papers by Thurstone (*1927a, 1927b, 1927c*), Zermelo's paper was an isolated excursion into paired comparison methods that was far ahead of its time. After this paper, the field remained mostly dormant for about 25 years until seminal publications by Bradley and Terry (*1952*) and Mosteller (*1951a, 1951b, 1951c*) led to increased interest in paired comparison methodology. Subsequently, work in paired comparison methods has remained an active area of research since the 1950's, and with paired comparison applications to large data sets including marketing research problems involving choice preference among consumers, the field continues to be active.

Scholarly work in this area arguably led to popularizing paired comparison methods in the form of competitor rating systems. For example, in the United States a rating system due to Elo (*1978*) for chess players was implemented in the early 1960s (supplanting a fairly crude system developed in the 1950s) which had as a basis the model appearing in Zermelo's paper. The system was adapted to ensure computational simplicity for the analysis of large collections of players whose abilities are changing over time. Elo's system was eventually adopted by the World Chess Federation in the early 1970s and is still used today to rate chess players in international tournaments. Given that the basis of Elo's system can be traced back to Zermelo, tournament chess players both in the United States and elsewhere are therefore indirect inheritors of Zermelo's work on rating chess players.

* The paper is in volume 29 of *Mathematische Zeitschrift* which was completed in 1929. Actually, the paper is in issue 2/3 of that volume which appeared already in 1928, and Zermelo in his personal bibliographies gave that year, a mode which has been followed so far in this edition. The additional "1929" is given here for the benefit of those readers who are familiar with the year 1929 of appearance attributed to the paper in the current literature. —A first draft of the paper was written already in 1919 under the title "Eine neue Methode zur Berechnung eines Turniers" ("A new method for calculating a tournament"). It is preserved in the Universitätsarchiv Freiburg, Zermelo *Nachlass*, under signature C129/262.

H.-D. Ebbinghaus, A. Kanamori (Eds.), *Ernst Zermelo - Collected Works/ Gesammelte Werke II*, DOI 10.1007/978-3-540-70856-8_13, Schriften der Mathematisch-naturwissenschaftlichen Klasse der Heidelberger Akademie der Wissenschaften 23,

Interestingly, despite sharing the same model and developing nearly identical numerical algorithms, the prominent paired comparison researchers from the 1950s, as well as later authors including Elo, may not have known about Zermelo's paper, as they did not cite it. The earliest citation in paired comparison literature that I could find was by Good (*1955*) who limits mention to a property of his own model being shared by that of Zermelo. Only in the monograph *David 1988* documenting the history of paired comparison methods do we begin to get a sense for the profundity of Zermelo's work.

Zermelo's paper is mostly concerned with estimating the relative playing strengths of chess players in imbalanced designs, that is, tournaments in which each player does not necessarily compete against every other the same number of times. To address the problem, Zermelo introduces a probability model for game outcomes as a function of the players' unknown strengths. The model assumes that chess games result in only wins and losses.[1] For a tournament with players $A_1, A_2, \ldots, A_n$, the model is given by

$$\mathrm{P}(A_r \text{ defeats } A_s) = \frac{u_r}{u_r + u_s} \tag{1}$$

where u_r and u_s are the unknown playing strengths of A_r and A_s, respectively, that are to be estimated. Perhaps unfairly to Zermelo, the model in (1) is now commonly called the Bradley-Terry model, a naming convention that gained popularity following the publication of *Bradley and Terry 1952*. Ebbinghaus (*2007*) in fact suggests referring to (1) as the Zermelo-Bradley-Terry model. There is no known evidence that this model for playing strength has been published prior to Zermelo's paper.

These days, it is more convenient and common to work with a reparameterized version of this model. Setting $v_r = \ln u_r$ for player A_r, the model can be rewritten as

$$\mathrm{P}(A_r \text{ defeats } A_s) = \frac{\exp(v_r)}{\exp(v_r) + \exp(v_s)} = \frac{1}{1 + \exp(-(v_r - v_s))} \tag{2}$$

which is the standard logistic cumulative distribution function evaluated at $v_r - v_s$. In fact, letting $\boldsymbol{x} = (x_1, \ldots, x_n)$ be a vector of length n with $x_r = 1$, $x_s = -1$, the remaining elements of $\boldsymbol{x}$ set to 0, and $\boldsymbol{v} = (v_1, \ldots, v_n)$, the model can be rewritten as

$$\operatorname{logit} \mathrm{P}(A_r \text{ defeats } A_s) = \boldsymbol{x}'\boldsymbol{v} \tag{3}$$

[1] Zermelo does not explicitly consider a draw (tie) as an outcome in his model, but in footnote 2 he implies that he can reframe the estimation problem by doubling the number of wins and losses and then counting draws as one win and one loss. A similar approach was advocated by Glickman (*1999*). It is worth pointing out that direct application of Zermelo's approach will result in overly optimistic standard error estimates of the playing strengths, as the sample size has doubled.

where $\boldsymbol{x}'\boldsymbol{v} = \sum_{r=1}^{n} x_r v_r$,[2] and logit $p = \ln\left(\frac{p}{1-p}\right)$. Equation (3) is simply a logistic regression model with a linear predictor containing unknown coefficients $\boldsymbol{v}$, a member of the class of generalized linear models (*McCullagh and Nelder 1989*). Understanding this connection provides greater insights into the model properties and numerical methods in Zermelo's paper. It is interesting that Zermelo simply asserts the model in (1) as the *de facto* choice, as the contemporaneous model of Thurstone assumed a Gaussian cumulative distribution function rather than Zermelo's logistic distribution function.

The optimization problem on which Zermelo focuses involves maximizing the joint probability of the tournament game results as a function of the u_r. This is precisely the method of maximum likelihood estimation, which is arguably still the most common approach in the 21st century to fitting probability models. It is likely that Zermelo's familiarity with statistical mechanics and its reliance on the principle of maximum entropy (which is intricately connected to maximum likelihood estimation) may have guided his choice. At the time of Zermelo's paper, maximum likelihood estimation was in its infancy, with the formal development appearing in a series of papers by R. A. Fisher between 1912 and 1925. While Zermelo's statement of the optimization problem does not rely on the details of Fisher's development, it is unclear whether Zermelo was aware of Fisher's foundational work.

The focus of the first half of Zermelo's paper is the decomposition of tournaments into disjoint collections of players to establish the optimization of the likelihood function (that is, Zermelo's function $\Phi(u_1, \ldots, u_n)$) under the constraints that the u_r are non-negative and that $\sum_{r=1}^{n} u_r \leq 1$. The key point is that tournaments can be decomposed uniquely into a set of irreducible partial tournaments, or "prime" tournaments. A group C of players constitutes a prime tournament if the other players can be partitioned into two groups P and Q such that: (a) no player in P has defeated any player in $C \cup Q$; (b) no player in C has defeated any player in Q; and (c) C cannot be partitioned into two groups of players C' and C'' such that no player in C' has defeated any player in C''. Zermelo then supposes that there is an ordering $C_1 < C_2 < \ldots < C_t$ of the prime tournaments such that for $1 \leq j < t$, there has been at least one game between a player in C_j and a player in C_{j+1}, and every game between a player in C_j and a player in C_{j+1} has resulted in a loss for the C_j player. With this, the main theorem of the paper is that maximizing the likelihood involves essentially maximizing factors in the likelihood product corresponding to the prime tournaments. The second part of the theorem connects optima across the prime tournaments: The ratio of optimized strength parameters between players A_r and A_s in two different ordered prime tournaments C_j and C_k where players in C_j are dominated by the players in C_k goes to 0 in the limit. It is important to note that the theorem only applies to tournament decompositions in which prime tournaments can be ordered as described above, which is not always

[2] This notation for the inner product is conventional in statistics literature.

guaranteed. As Zermelo acknowledges, tournaments can exist in which prime tournaments are incomparable; for example, when the players of C_j have not competed against any player in C_k, and (say) both C_j and C_k are dominated by every other prime tournament.

Zermelo's analysis, particularly with the comparison of strengths of players between prime tournaments, is more detailed than that of later authors. *Ford 1957*, an important paper that re-derives a subset of Zermelo's results without ostensibly being aware of his paper (and, for that matter, of *Bradley and Terry 1952*), focuses on estimating playing strengths in tournaments whose decomposition results in a single prime tournament. Ford asserts the condition for the irreducibility of a tournament in a clean way: In every possible partition of players into two non-empty subsets, some player in the second set has defeated at least one player in the first set. Note that this definition of irreducibility coincides with Zermelo's for prime tournaments containing at least two players. The paper then demonstrates that this condition is sufficient for a unique optimum of Zermelo's likelihood function. By contrast to Ford's paper, Zermelo considers more general tournaments that can be decomposed into many prime tournaments. In practice, however, optimization only makes sense when analyzing each prime tournament separately.

The second half of Zermelo's paper concentrates on numerically solving the optimization problem. He first proves that, for a balanced design (in which each player competes against every other the same number of times) in a tournament that cannot be decomposed into more than one prime tournament, the total score for each player is in the same rank order as the optimized playing strengths. Ford (*1957*) independently derives this result, and Bühlmann and Huber (*1963*) demonstrate that in fact the model assumed by Zermelo is the only linear paired comparison model (that is, linear in the $v_r = \ln u_r$) for which ranking according to the total score is equivalent to ranking according to the maximum likelihood estimates. Zermelo then derives an iterative numerical algorithm to solve the optimization problem for such irreducible, not necessarily balanced tournaments. The method is the same as the one described by Bradley and Terry (*1952*), though Zermelo deserves the credit as the method's originator. The algorithm, however, has been shown to have slow convergence (*Dykstra 1956*), particularly with poorly chosen initial values. The modern treatment for optimizing Zermelo's model is to recognize it as a logistic regression and use Fisher scoring (see *McCullagh and Nelder 1989*, 42) to carry out the optimization, which is quite fast.

One interesting detail of Zermelo's development is how he addresses the non-identifiability of the player strengths. The probability model in (1) implies that if $(u_1^*, \ldots, u_n^*)$ is a solution to the optimization problem, then so is $(au_1^*, \ldots, au_n^*)$ for $a > 0$, provided that $\sum_{r=1}^n au_r \leq 1$. Thus, without an additional constraint, the solution is not unique. As Zermelo demonstrates, this indeterminacy does not affect the decomposition theorem, but the actual optimization requires a norm restriction on the u_r. Zermelo's extra constraint is to fix $\sum_{r=1}^n u_r$ at a constant (unity, in the development, and then 100

in his example). The modern way to add a constraint is to reparameterize the model as in (3) and to assume a constraint on $\sum_{r=1}^{n} v_r$; in other words, to fix the product of the u_r rather than the sum of the u_r. For example, it is conventional to assume that $\sum_{r=1}^{n} v_r = 0$ and then to estimate only $v_1, v_2, \ldots, v_{n-1}$, in which case the logistic regression model in (3) changes by the deletion of a term in the linear predictor. Specifically, recognizing that $v_n = -(v_1 + \cdots + v_{n-1})$, we have

$$\begin{aligned} \text{logit P}(A_r \text{ defeats } A_s) = \boldsymbol{x}'\boldsymbol{v} &= \left(\sum_{r=1}^{n-1} x_r v_r\right) + x_n v_n \\ &= \left(\sum_{r=1}^{n-1} x_r v_r\right) + x_n(-v_1 - \cdots - v_{n-1}) \\ &= \tilde{\boldsymbol{x}}'\tilde{\boldsymbol{v}} \end{aligned} \tag{4}$$

where $\tilde{\boldsymbol{v}} = (v_1, \ldots, v_{n-1})$, and $\tilde{\boldsymbol{x}} = (x_1 - x_n, x_2 - x_n, \ldots, x_{n-1} - x_n)$. This logistic regression can then be fit via Fisher scoring to obtain unique maximum likelihood estimates, assuming the tournament is irreducible in the sense of *Ford 1957*.

Zermelo, at the end of his paper, applies the numerical algorithm to the analysis of playing strengths of participants in the New York Masters' Tournament of 1924, which was won by Emanuel Lasker (the world chess champion from 1894 through 1921). Lasker, who was also an accomplished mathematician, wrote a letter to Zermelo in 1929 expressing appreciation of his work on measuring chess strength, pointing out that this paper was the first to use probability theory for this problem (see *Ebbinghaus 2007*, 149). Lasker's performance in the 1924 tournament was arguably the last great chess performance of his career before he effectively retired from playing.

The development of models for estimating competitor playing strengths has come a long way since Zermelo's paper, but Zermelo's work was remarkable in the level of detail in which it laid the foundation for paired comparison methods. Once Zermelo's approach caught on some 25 years later mostly through the work of Bradley and Terry and their contemporaries, a variety

of interesting extensions were explored that are noteworthy. For example, models were developed for outcomes that include ties and other degrees of partial preferences. An extension to the model of Zermelo that has become standard is the incorporation of an (unknown) advantage for playing white in chess, which is more generally known as an "order" effect (in the sense that the probability of a preference between two objects may depend on the order in which they are presented). A detailed synopsis of recent developments in paired comparison modeling appears in the monograph *David 1988*.

It is worth pointing out that recent approaches via Bayesian modeling (see, for example, *Leonard 1977* and *Glickman 1999*) in which a proper prior distribution is assumed for the player strengths avoid the difficulties connected with optimizing over Zermelo's prime decomposition of tournaments. In the Bayesian framework, the results of a tournament simply revise the prior distribution of playing strengths to a posterior distribution, regardless of whether subsets of players have not competed against each other, or if any player has won or lost all of his games. These occurrences are problematic for Zermelo's (and in general the maximum likelihood estimation) approach. The Bayesian approach, however, still has important connections to the groundwork laid out by Zermelo in his analysis of the likelihood function, so that much of Zermelo's development is unquestionably relevant today.

While seeming to have little influence immediately following its publication, Zermelo's paper has had a long-lasting impact. His work is now regularly cited in papers on paired comparison models, and his name is now immortalized in connection with rating competitors in games and sports. For example, statistician David Marcus, who constructed a rating system applicable for tournament table tennis (*Marcus 2001*), paid tribute by developing Windows software called "Zermelo" that organizes and runs table tennis tournaments. Even American football rankings[3] have been connected to Zermelo. The resurgence of interest in Zermelo's paper on measuring playing strength over recent years is the appropriate recognition for an impressive piece of work.

[3] See the web site `http://www.ezfootballrankings.com`, where the "ez" in the domain name stands for "Ernst Zermelo."

Die Berechnung der Turnier-Ergebnisse als ein Maximumproblem der Wahrscheinlichkeitsrechnung

1928

Schachturniere werden neuerdings immer in der Weise ausgeführt, daß jeder Teilnehmer mit jedem anderen eine bestimmte Anzahl k von Partien (gewöhnlich zwei mit wechselndem Anzuge) zu spielen hat und dann die Reihenfolge der Sieger nach der Anzahl ihrer Gewinnpartien bestimmt wird, wobei eine unentschiedene („Remis"-) Partie jedem der beiden Partner immer als halbe Gewinnpartie angerechnet wird. Dieses Berechnungsverfahren, das alle gespielten Partien unabhängig von der Reihenfolge gleichmäßig berücksichtigt und dadurch dem Zufall möglichst wenig Einfluß gestattet, hat sich in der Praxis anscheinend auch durchaus bewährt, soweit eben nur die *Reihenfolge* in Betracht kommt, versagt aber völlig bei *abgebrochenen* Turnieren (in welchen die Anzahl der gespielten Partien nicht bei jedem Teilnehmer die gleiche ist) und liefert auch bei durchgeführten keine befriedigende *quantitative* Bestimmung der *relativen Spielstärken*, da es ganz augenscheinlich der Mittelmäßigkeit zu gute käme. So würde etwa bei $k = 1$ ein Meister gegen n Stümper doch nur n Partien gewinnen, diese alle zusammen aber $\frac{n(n-1)}{2}$, jeder von ihnen also durchschnittlich $\frac{n-1}{2}$, und der Quotient der berechneten Werte $n\colon \frac{n-1}{2} = \frac{2n}{n-1}$ würde sich mit wachsendem n der 2 nähern, der Sieger also nur etwa doppelt so hoch bewertet werden wie jeder andere. Zur Behebung dieses Übelstandes und zur quantitativen Verfeinerung des üblichen Verfahrens sind in der Schachliteratur verschiedene Vorschläge gemacht
437 und diskutiert worden, u. a. von E. Landau[1], die aber, so inter- | essant sie auch in mathematischer Beziehung waren, gelegentlich zu paradoxen Ergebnissen führten und bisher keine befriedigende Lösung des Problems ergaben. Im folgenden soll nun ein neues Verfahren der Turnier-Berechnung entwickelt werden, das von den bisherigen Bedenken frei zu sein scheint und namentlich auch bei „abgebrochenen" und „kombinierten" Turnieren zu einem mathematisch bestimmten und allen vernünftigen Anforderungen entsprechenden Ergebnisse führt.

Unser Verfahren kommt darauf hinaus, daß die relativen Spielstärken als Wahrscheinlichkeiten aufgefaßt und so bestimmt werden, daß die Wahrscheinlichkeit für das Eintreten des beobachteten Turnier-Ergebnisses eine möglichst große wird.

[1] E. Landau, Über Preisverteilung bei Spielturnieren. Ztschr. f. Math. u. Phys. **63** (1914), S. 192.

The calculation of the results of a tournament as a maximum problem in the calculus of probabilities

1928

Of late, chess tournaments have always been conducted so that each participant must play a certain number k of games against every other participant (usually a pair with alternate first move) and so that the order of the winners is then determined according to the number of the games they have won, where an undecided game ("draw") is always counted as half a game won for each player. This method of calculation, which equally considers all played games independently of the order, and hence allows chance as little influence as possible, seems to have proved itself in practice, insofar as only the *order* is concerned. But it completely fails for *aborted* tournaments (in which the number of played games is not the same for all participants) and also provides no satisfactory *quantitative* determination of the *relative playing strengths* for completed tournaments, since, as is very obvious, it would favor mediocrity. Thus for, say, $k = 1$, a champion player would win only n games against n dilettante players, but all of these together would win $\frac{n(n-1)}{2}$, and hence each of them $\frac{n-1}{2}$ on average, and the quotient of the calculated values $n\colon \frac{n-1}{2} = \frac{2n}{n-1}$ would approach 2 as n increases. Hence, the winner would only be rated about twice as high as any other player. In order to remedy this deficiency and to refine quantitatively the usual method, several suggestions, such as one by E. Landau,[1] have been made and discussed in the literature on chess, which, their intrinsic mathematical interest notwithstanding, have still led to paradoxical results on occasion and have not yet produced a satisfactory solution to the problem. In what follows we shall now develop a new method of tournament calculation that does not seem to be encumbered by the usual concerns and that, in particular also for "aborted" and "combined" tournaments leads to mathematically definite results meeting all reasonable requirements.

Our method is based on the idea that the relative playing strengths are conceived of as probabilities and determined so that the probability of the occurrence of the observed tournament result is as great as possible.

[1] *Landau 1914.*

§ 1.

Ansatz und Reduktion des Problems.

Jedem der n an einem Turnier teilnehmenden Spieler A_r denken wir uns eine positive Zahl u_r als „Spielstärke“ zugeordnet und nehmen an, daß diese Zahlen $u_1, u_2, \ldots, u_n$ sich wie ihre relativen Gewinnchancen verhalten. Die Wahrscheinlichkeit dafür, daß in einer bestimmten Partie A_r gegen A_s gewinnt, wird also gegeben durch den Bruch:

$$u_{rs} = \frac{u_r}{u_r + u_s} \,. \tag{1}$$

Ist dann k_{rs} die Anzahl der zwischen A_r und A_s gespielten Partien, von denen wir hier[2] annehmen wollen, daß sie alle zur Entscheidung kommen, g_{rs} die dabei von A_r und $g_{sr} = k_{rs} - g_{rs}$ die von A_s gewonnenen, so wird, wenn wir die verschiedenen Partien als „unabhängige“ Ereignisse betrachten, die kombinierte Wahrscheinlichkeit des auf A_r und A_s bezüglichen Teilergebnisses, daß nämlich A_r gerade g_{rs} und A_s entsprechend g_{sr} Partien gegen den andern gewinnt, bis auf einen von u_r, u_s unabhängigen Zahlenfaktor $\sigma_{rs} = \binom{k_{rs}}{g_{rs}}$, auf den mich Herr Prof. Ostrowski freundlichst aufmerksam machte,

$$w_{rs} = u_{rs}^{g_{rs}} u_{sr}^{g_{sr}} = \frac{u_r^{g_{rs}} u_s^{g_{sr}}}{(u_r + u_s)^{k_{rs}}} \,, \tag{2}$$

438 | und somit die Wahrscheinlichkeit des durch die Matrix g_{rs} bestimmten Gesamtresultates das Produkt aller dieser Wahrscheinlichkeiten

$$w = \prod_{r,\,s}{}' w_{r,s} = \frac{u_1^{g_1} u_2^{g_2} \ldots u_n^{g_n}}{\prod_{r,\,s}{}'(u_r + u_s)^{k_r s}} = \Phi(u_1, \ldots, u_n) \,, \tag{3}$$

wo

$$g_r = \sum_{s}^{s \neq r} g_{rs} \tag{4}$$

immer die Gesamtzahl der von A_r gewonnenen Partien bezeichnet und das Produkt $\prod'$ über alle Kombinationen r, s mit $r \neq s$ zu erstrecken ist.

Die (nicht negativen) Zahlen $u_1, \ldots, u_n$ sollen nun so bestimmt werden, daß $\Phi(u_1, \ldots, u_n)$ möglichst groß und damit das beobachtete, durch

[2] Bei allen praktischen Anwendungen berücksichtigen wir die etwa vorkommenden unentschiedenen („Remis“-) Partien durch die Fiktion, daß doppelt so viel Partien gespielt seien, indem wir jedem Spieler jede tatsächlich gewonnene Partie als *doppelten* und jede unentschiedene als *einfachen* Gewinn anrechnen. Dies entspricht genau der im Eingang erwähnten herkömmlichen Zählung der Remis-Partien als halbe Gewinnpartien.

§ 1.

Approach to and reduction of the problem.

Suppose that there is a positive number u_r associated with each of the n players A_r participating in a tournament that designates the "playing strength" and assume that these numbers $u_1, u_2, \ldots, u_n$ are proportional to their relative winning chances. Hence, the probability that, in a certain game, A_r wins against A_s is given by the fraction

$$u_{rs} = \frac{u_r}{u_r + u_s} \,. \tag{1}$$

Let k_{rs} be the number of games played between A_r and A_s, of which we shall assume here[2] that they all are decided, and let g_{rs} be the number of games won by A_r, and $g_{sr} = k_{rs} - g_{rs}$ that of games won by A_s. Then, if we conceive of the different games as "independent" events, the combined probability of the partial result with respect to A_r and A_s, namely that A_r wins exactly g_{rs}, and correspondingly A_s g_{sr}, games against the other player, except for a numeric factor $\sigma_{rs} = \binom{k_{rs}}{g_{rs}}$, to which Prof. Ostrowski was kind enough to draw my attention, becomes

$$w_{rs} = u_{rs}^{g_{rs}} u_{sr}^{g_{sr}} = \frac{u_r^{g_{rs}} u_s^{g_{sr}}}{(u_r + u_s)^{k_{rs}}} \,, \tag{2}$$

and hence the probability of the total result determined by the matrix g_{rs} becomes the product of all these probabilities

$$w = \prod_{r,\,s}{}' w_{r,s} = \frac{u_1^{g_1} u_2^{g_2} \ldots u_n^{g_n}}{\prod_{r,\,s}{}'(u_r + u_s)^{k_r s}} = \Phi(u_1, \ldots, u_n) \,, \tag{3}$$

where

$$g_r = \sum_{s}^{s \neq r} g_{rs} \tag{4}$$

always denotes the total number of games won by A_r, and where the product $\prod'$ must be extended over all combinations r, s for $r \neq s$.

The (non-negative) numbers $u_1, \ldots, u_n$ shall now be determined so that $\Phi(u_1, \ldots, u_n)$ is as great as possible, and hence the observed result of the

[2] For all practical applications, we account for any undecided games ("draws") through the fiction that twice as many games are played by crediting each player *twice* for each game actually won and *once* for each undecided game. This precisely corresponds to the way of counting the draws as half a game won, which was mentioned at the beginning.

die Zahlen $g_1, \ldots, g_n$ bestimmte Turnier-Ergebnis möglichst wahrscheinlich wird[3].

Da Φ in bezug auf die Variablen u_r homogen von der nullten Dimension ist, also nur von ihren Verhältnissen abhängt, so können wir o. B. d. A. annehmen, daß

$$\sum_{r=1}^{n} u_r \leqq 1 \tag{5a}$$

ist, während gleichzeitig für alle r

$$u_r \geqq 0 \qquad (r = 1, 2, \ldots, n) \tag{5b}$$

vorausgesetzt wird.

Der durch (5a) und (5b) bestimmte Bereich $\mathfrak{B}$ der u_r ist endlich und abgeschlossen; in ihm ist die durch (3) als Wahrscheinlichkeit definierte Funktion Φ überall $\leqq 1$, besitzt also eine endliche obere Grenze $\overline{w} \leqq 1$; dabei muß

$$\overline{w} \geqq \Phi(1, \ldots, 1) = \left(\frac{1}{2}\right)^N > 0 \tag{6}$$

sein, wo

$$N = \sum_{r=1}^{n} g_r = \sum_{r,\,s}^{r \neq s} k_{rs} \tag{7}$$

die Anzahl aller überhaupt gespielten Partien ist. Nach Weierstraß gibt es in $\mathfrak{B}$ mindestens einen Punkt $\overline{\mathfrak{q}}$ mit den Koordinaten $\overline{u}_1, \ldots, \overline{u}_n$, in dessen Umgebung der Wert $\overline{w}$ von Φ approximiert wird.

Wir müssen jetzt die beiden Fälle unterscheiden, daß $\overline{\mathfrak{q}}$ im Innern oder auf dem Rande von $\mathfrak{B}$ liegt.

439 | *Fall 1.* $\overline{\mathfrak{q}}$ liege im Innern von $\mathfrak{B}$, d. h.

$$\sum_{r=1}^{n} \overline{u}_r < 1 \quad \text{und} \quad \overline{u}_r > 0 \qquad (r = 1, 2, \ldots, n)\,.$$

Dann ist Φ *stetig* an der Stelle $\overline{\mathfrak{q}}$ und es wird

$$\Phi(\overline{\mathfrak{q}}) = \overline{w}$$

das *Maximum* von Φ in $\mathfrak{B}$. Da Φ dann in $\overline{\mathfrak{q}}$ zugleich auch nach allen Variablen differentiierbar und von Null verschieden ist, so müssen auch die partiellen logarithmischen Ableitungen sämtlich in $\overline{q}$ verschwinden, d. h.

$$\frac{g_r}{\overline{u}_r} - \sum_{t}{}' \frac{k_{rt}}{\overline{u}_r + \overline{u}_t} = 0$$

[3] Dieser Ansatz ist einem insbesondere in der Gastheorie häufig angewendeten Verfahren analog.

tournament, which is determined by the numbers $g_1, \ldots, g_n$, becomes as likely as possible.[3]

Since Φ is homogeneous of the zeroth dimension with respect to the variables u_r, and hence only depends on their relations, we can assume without loss of generality that

$$\sum_{r=1}^{n} u_r \leqq 1 \tag{5a}$$

while, at the same time, we assume

$$u_r \geqq 0 \qquad (r = 1, 2, \ldots, n) \tag{5b}$$

for all r.

The domain $\mathfrak{B}$ of the u_r determined by (5a) and (5b) is finite and closed; in it, the function Φ, which is defined as probability by (3), is $\leqq 1$ everywhere, and hence possesses a finite upper limit $\overline{w} \leqq 1$; here necessarily

$$\overline{w} \geqq \Phi(1, \ldots, 1) = \left(\frac{1}{2}\right)^N > 0\,, \tag{6}$$

where

$$N = \sum_{r=1}^{n} g_r = \sum_{r,\,s}^{r \neq s} k_{rs} \tag{7}$$

is the total number of all games played. According to Weierstrass, there is at least one point $\overline{\mathfrak{q}}$ in $\mathfrak{B}$ with coordinates $\overline{u}_1, \ldots, \overline{u}_n$ in whose neighborhood Φ approximates the value of $\overline{w}$.

We now have to distinguish between the case where $\overline{\mathfrak{q}}$ lies in the interior of $\mathfrak{B}$ and the case where it lies on its boundary.

Case 1. $\overline{\mathfrak{q}}$ lies in the interior of $\mathfrak{B}$, i. e.,

$$\sum_{r=1}^{n} \overline{u}_r < 1 \quad \text{and} \quad \overline{u}_r > 0 \qquad (r = 1, 2, \ldots, n)\,.$$

Then Φ is *continuous* at point $\overline{\mathfrak{q}}$ and

$$\Phi(\overline{\mathfrak{q}}) = \overline{w}$$

becomes the *maximum* of Φ in $\mathfrak{B}$. Since, in this case, Φ is also differentiable in $\overline{\mathfrak{q}}$ with respect to all variables and different from zero, the partial logarithmic derivatives, too, must all vanish in $\overline{q}$, i. e.,

$$\frac{g_r}{\overline{u}_r} - \sum_{t}{}' \frac{k_{rt}}{\overline{u}_r + \overline{u}_t} = 0$$

[3] This approach is analogous to a method often used in particular in the theory of gases.

oder

$$\sum_t{}' k_{rt}\frac{\overline{u}_r}{\overline{u}_r+\overline{u}_t}=g_r \quad \text{für} \quad r=1,\ldots,n\,, \tag{8}$$

wobei die Summe $\sum'$ über alle $t\neq r$ zu nehmen ist. Wir erhalten also ein System von n homogenen algebraischen Gleichungen zur Bestimmung von $\overline{u}_1,\ldots,\overline{u}_n$ oder vielmehr ihrer Verhältnisse $\overline{u}_r:\overline{u}_s$.

Durch Addition der n Gleichungen (8) erhält man wegen

$$\frac{\overline{u}_r}{\overline{u}_r+\overline{u}_t}+\frac{\overline{u}_t}{\overline{u}_r+\overline{u}_t}=1$$

die Identität (7) als Ausdruck der zwischen den Gleichungen des Systems bestehenden Abhängigkeit.

Da $k_{rt}\dfrac{u_r}{u_r+u_t}$ die „wahrscheinliche Anzahl“ der von A_r gegen A_t gewonnenen Partien darstellt, so bedeutet das Gleichungssystem (8), daß bei der unserm Ansatze entsprechenden Bewertung der Spieler *die wahrscheinliche Anzahl der von einem Spieler gegen alle übrigen Spieler zu gewinnenden Partien gleich der Anzahl der von ihm wirklich gewonnenen ist.*

Fall 2. $\overline{\mathfrak{q}}$ liege auf dem *Rande* von $\mathfrak{B}$ derart, daß alle

$$\overline{u}_r>0 \qquad (r=1,2,\ldots,n)\,,$$

aber

$$\sum_r \overline{u}_r=1$$

ist. Da nun aber die Funktion Φ in $u_1,\ldots,u_n$ homogen von 0-ter Dimension ist

$$\Phi(\lambda u_1,\ldots,\lambda u_n)=\Phi(u_1,\ldots,u_u)\,, \tag{9}$$

so wird der Grenzwert $\overline{w}$ dann für $0<\lambda<1$ auch im *Innern* von $\mathfrak{B}$ approximiert und damit der Fall 2 auf den Fall 1 zurückgeführt.

440 | *Fall 3.* Der Punkt $\overline{\mathfrak{q}}$ liege auf dem Rande von $\mathfrak{B}$ derart, daß

$$\prod \overline{u}_r=\overline{u}_1\overline{u}_2\ldots\overline{u}_n=0 \tag{10}$$

wird, d. h. daß mindestens ein $\overline{u}_r=0$ wird.

Ist z. B. $\overline{u}_1=0$, während alle übrigen $\overline{u}_r>0$ sind, so ist der Nenner der Funktion Φ an der Stelle $\overline{\mathfrak{q}}$ von Null verschieden. Also muß auch $g_1=0$ sein, da nach (3) sonst Φ eine Potenz $u_1^{g_1}$ von u_1 als Faktor enthielte und daher w an der Stelle $\overline{\mathfrak{q}}$ anstatt des Wertes $\overline{w}>0$ die Null approximieren würde. Ist umgekehrt $g_1=0$, so erscheint u_1 in (3) nur im Nenner, und zwar in allen Faktoren u_1+u_s als positiver Summand, der den Quotienten verkleinert,

or

$$\sum_t{}' k_{rt} \frac{\overline{u}_r}{\overline{u}_r + \overline{u}_t} = g_r \quad \text{for} \quad r = 1, \ldots, n \,, \tag{8}$$

where the sum $\sum'$ is to be taken over all $t \neq r$. Thus, we obtain a system of n homogeneous algebraic equations for the determination of $\overline{u}_1, \ldots, \overline{u}_n$, or rather their ratios $\overline{u}_r : \overline{u}_s$.

By addition of the n equations (8) we obtain, on account of

$$\frac{\overline{u}_r}{\overline{u}_r + \overline{u}_t} + \frac{\overline{u}_t}{\overline{u}_r + \overline{u}_t} = 1 \,,$$

the identity (7) as an expression of the dependence holding among the equations of the system.

Since $k_{rt} \dfrac{u_r}{u_r + u_t}$ represents the "probable number" of the games won by A_r against A_t, the system of equations (8) denotes the fact that, for the evaluation of the players corresponding to our approach, *the probable number of games to be won by one player against all other players is equal to the number of games he actually won.*

Case 2. $\overline{\mathfrak{q}}$ lies on the *boundary* of $\mathfrak{B}$ so that all

$$\overline{u}_r > 0 \qquad (r = 1, 2, \ldots, n) \,,$$

but

$$\sum_r \overline{u}_r = 1 \,.$$

However, since the function Φ in $u_1, \ldots, u_n$ is homogeneous of 0th dimension

$$\Phi(\lambda u_1, \ldots, \lambda u_n) = \Phi(u_1, \ldots, u_u) \,, \tag{9}$$

the limit $\overline{w}$ is also approximated in the *interior* of $\mathfrak{B}$ for $0 < \lambda < 1$, and hence case 2 is reduced to case 1.

Case 3. The point $\overline{\mathfrak{q}}$ lies on the boundary of $\mathfrak{B}$ so that

$$\prod \overline{u}_r = \overline{u}_1 \overline{u}_2 \ldots \overline{u}_n = 0 \,, \tag{10}$$

i. e., at least one $\overline{u}_r$ becomes $= 0$.

If, e. g., $\overline{u}_1 = 0$, while all other $\overline{u}_r$ are > 0, then the denominator of the function Φ is different from zero at point $\overline{\mathfrak{q}}$. Thus, we must also have $g_1 = 0$, since otherwise, by (3), Φ would contain a power $u_1^{g_1}$ of u_1 as factor, and hence w would approximates zero instead of the value $\overline{w} > 0$ at point $\overline{\mathfrak{q}}$. If, conversely, $g_1 = 0$, then u_1 appears in (3) only in the denominator, and in particular in all factors $u_1 + u_s$ as a positive summand diminishing the

muß sich also bei der Approximation der oberen Grenze $\overline{w}$ seiner eigenen unteren Schranke 0 nähern, d. h. $\overline{u}_1 = 0$. Nach unserem Ansatz wird also ein einzelner Spieler mit der Spielstärke 0 bewertet dann und nur dann, wenn er als einziger im Turnier alle seine Partien verliert[4].

Seien jetzt allgemein[5] bei geeignet gewählter Reihenfolge

$$\overline{u}_1 = \overline{u}_2 = \ldots = \overline{u}_p = 0\,; \quad \overline{u}_{p+1}\,,\ \overline{u}_{p+2}\,,\ \ldots\,,\ \overline{u}_n > 0\,.$$

Dann folgt wie im Spezialfall, daß Φ in (3) keinen Faktor der Form $u_{rs} = \dfrac{u_r}{u_r + u_s}$ für $r \leqq p < s$ enthalten darf, weil sich sonst dieser und mit ihm Φ bei der Approximation an die Grenzstelle $\overline{\mathfrak{q}}$ der Null nähern müßte, was wegen (6) $\overline{w} > 0$ unmöglich ist. In diesem Falle müssen also alle $g_{rs} = 0$ sein für $r \leqq p < s$, d. h. wir haben eine *Einteilung aller Spieler in zwei Klassen derart, daß jeder Spieler A_r der ersten Klasse gegen keinen Spieler A_s der zweiten Klasse in dem Turnier eine Partie gewinnt* (sondern alle etwa gespielten Partien verliert).

Hat umgekehrt das Ergebnis des Turniers zu einer solchen Klassenbildung $\{A_r, A_s\}$ mit $g_{rs} = 0$ geführt, *so müssen bei der Approximation von $\overline{w}$ durch Φ alle u_r, die den Spielern A_r der Unterklasse zugeordnet sind und für wenigstens ein $s > p$ die Bedingung $k_{rs} > 0$ erfüllen* (d. h. sofern die A_r überhaupt gegen die A_s gespielt haben), *sich der Null nähern:* $\overline{u}_r = 0$.

Um dies zu beweisen, zerlegen wir die durch (3) definierte Funktion Φ in drei Faktoren

$$\Phi(u) = \Phi_1(u_r)\Phi_2(u_s)\Phi_{1,2}(u_r, u_s)\,, \tag{11}$$

441 | wo

$$\Phi_1(u_r) = \prod_{1 \leqq r,\, r' \leqq p}^{r \neq r'} \frac{u_r^{g_{rr'}}\, u_{r'}^{g_{r'r}}}{(u_r + u_{r'})^{k_{rr'}}}\,, \quad \Phi_2(u_s) = \prod_{p+1 \leqq s,\, s' \leqq n}^{s \neq s'} \frac{u_s^{g_{ss'}}\, u_{s'}^{g_{s's}}}{(u_s + u_{s'})^{k_{ss'}}}\,,$$

$$\Phi_{1,2}(u_r, u_s) = \prod_{r=1}^{p} \prod_{s=p+1}^{n} \Big(\frac{u_s}{u_r + u_s}\Big)^{k_{rs}}\,,$$

da nach der gemachten Annahme wegen $g_{rs} = 0$ Faktoren u_{rs} gar nicht auftreten.

Ersetzen wir jetzt alle u_r mit $r \leqq p$ durch λu_r, wo $0 < \lambda < 1$ ist, so bleiben wegen der Homogenität Φ_1 und Φ_2 ungeändert, während $\Phi_{1,2}$ übergeht in

$$\Phi_{1,2}^{(\lambda)} = \prod_{r=1}^{p} \prod_{s=p+1}^{n} \Big(\frac{u_s}{\lambda u_r + u_s}\Big)^{k_{rs}} = \Phi_{1,2}\,(\lambda u_r; u_s)\ . \tag{12}$$

[4] Dieses Ergebnis ist nur scheinbar paradox, da kein Spieler bei noch so geringer Spielstärke schlechter abschneiden könnte. Vgl. E. Landau a. a. O. S. 201.

[5] Der Fall, daß alle $\overline{u}_r = 0$ sind, können wir wegen der Homogenität (9) von Φ ausschließen.

quotient, and hence it must approach its own lower bound 0 as the upper boundary $\overline{w}$ is approximated, i. e., $\overline{u}_1 = 0$. Thus, our approach rates an individual player at playing strength 0 if and only if he is the only one in the tournament losing all his games.[4]

Let us now assume generally[5] for a suitable chosen order

$$\overline{u}_1 = \overline{u}_2 = \ldots = \overline{u}_p = 0\,; \quad \overline{u}_{p+1}\,,\ \overline{u}_{p+2}\,,\ \ldots\,,\ \overline{u}_n > 0\,.$$

As in the special case, it then follows that Φ in (3) must not contain a factor of the form $u_{rs} = \dfrac{u_r}{u_r + u_s}$ for $r \leqq p < s$, since otherwise the latter, and with it Φ, would have to approach zero as the limit point $\overline{\mathfrak{q}}$ is approximated, which, on account of (6) $\overline{w} > 0$, is impossible. Thus, in this case, all the g_{rs} must be $= 0$ for $r \leqq p < s$, i. e., we have a *partition of all players into two classes so that each player A_r of the first class does not win a game in the tournament against any player A_s of the second class* (but loses all games he may have played).

If, conversely, the result of the tournament has led to such a partition $\{A_r, A_s\}$ with $g_{rs} = 0$, *then, as Φ approximates $\overline{w}$, all u_r that are associated with the players A_r of the subclass and satisfy the condition $k_{rs} > 0$ for at least one $s > p$* (i. e., provided the A_r's have played against the A_s's at all), *must approach zero:* $\overline{u}_r = 0$.

To prove this we decompose the function Φ defined by (3) into three factors

$$\Phi(u) = \Phi_1(u_r)\Phi_2(u_s)\Phi_{1,2}(u_r, u_s)\,, \tag{11}$$

where

$$\Phi_1(u_r) = \prod_{1 \leqq r,\, r' \leqq p}^{r \neq r'} \frac{u_r^{g_{rr'}} u_{r'}^{g_{r'r}}}{(u_r + u_{r'})^{k_{rr'}}}\,, \quad \Phi_2(u_s) = \prod_{p+1 \leqq s,\, s' \leqq n}^{s \neq s'} \frac{u_s^{g_{ss'}} u_{s'}^{g_{s's}}}{(u_s + u_{s'})^{k_{ss'}}}\,,$$

$$\Phi_{1,2}(u_r, u_s) = \prod_{r=1}^{p} \prod_{s=p+1}^{n} \Big(\frac{u_s}{u_r + u_s}\Big)^{k_{rs}}\,,$$

since, by our assumption, the factors u_{rs} do not occur at all on account of $g_{rs} = 0$.

If we now replace all u_r's with $r \leqq p$ by λu_r's, where $0 < \lambda < 1$, then Φ_1 and Φ_2 remain unaltered on account of the homogeneity, while $\Phi_{1,2}$ is transformed into

$$\Phi_{1,2}^{(\lambda)} = \prod_{r=1}^{p} \prod_{s=p+1}^{n} \Big(\frac{u_s}{\lambda u_r + u_s}\Big)^{k_{rs}} = \Phi_{1,2}\,(\lambda u_r; u_s)\,. \tag{12}$$

[4] The paradoxical air of this result is due to the fact that no player, however low his playing strength, could possibly do worse. See *Landau 1914*, p. 201.

[5] We can exclude the case where all $\overline{u}_r$ are $= 0$ on account of the homogeneity (9) of Φ.

Angenommen, an einer Approximationsstelle $\overline{\mathfrak{q}}$ von $\overline{w}$ sei eine der Größen $\overline{u}_r > 0$ $(r \leqq p)$ mit $k_{rs} > 0$, dann wäre wegen $\lambda > 0$

$$\frac{\Phi^{(\lambda)}_{1,\,2}}{\Phi_{1,\,2}} \geqq \Big(\frac{u_r + u_s}{\lambda u_r + u_s}\Big)^{k_{rs}} ,$$

wo der Quotient rechts bei der Annäherung an $\overline{\mathfrak{q}}$ sich wegen $\overline{u}_r > 0$, $\lambda < 1$, und $k_{rs} > 0$ dem Werte:

$$\Big(\frac{\overline{u}_r + \overline{u}_s}{\lambda \overline{u}_r + \overline{u}_s}\Big)^{k_{rs}} > 1$$

nähert; es wäre also in $\overline{\mathfrak{q}}$

$$\Phi^{(\lambda)}_{1,2} > \Phi_{1,2} ,$$

und demnach

$$\lim \Phi\,(\lambda u_1, \ldots, \lambda u_p; u_{p+1}, \ldots, u^n) > \lim \Phi\,(u_1, \ldots, u_n) = \overline{w} \qquad (13)$$

im Widerspruch mit der Definition von $\overline{w}$.

Es folgt also, daß für $r \leqq p$ alle $\overline{u}_r = 0$ sein müssen, für die mindestens ein $k_{rs} > 0$ ist.

Bei der Approximation von $\overline{\mathfrak{q}}$ durch $\mathfrak{q}$ strebt dann $\Phi_{1,\,2}$ gegen 1, da hier nur solche u_r mit $k_{rs} > 0$, für die $\overline{u}_r = 0$ wird, wirklich auftreten. Demnach wird $\overline{w}$ von Φ dann und nur dann approximiert, wenn $\Phi_1(u_r)$ und $\Phi_2(u_s)$ *unabhängig voneinander möglichst groß* werden.

Um also die verhältnismäßigen Spielstärken der Spieler einer Klasse untereinander zu ermitteln, genügt es, unseren Ansatz auf das Teilturnier anzuwenden, welches aus dem Gesamtturnier durch Beschränkung auf die Spieler dieser einen Klasse entsteht.

Damit wird unser Problem auf solche mit kleinerer Spielerzahl zurückgeführt, die wieder entsprechenden Reduktionen unterliegen, bis man schließlich
442 zu „irreduziblen“ Problemen gelangt, für die dann alle $\overline{u}_r > 0$ sind | und daher die durch Gleichung (8) ausgedrückte Lösung des Falles 1 Gültigkeit hat.

Wir zeigen zunächst, daß die Zerlegung des „Gesamtturniers“ T in „irreduzible“ Teilturniere oder *„Primturniere“* $C_1, C_2, C_3, \ldots$ *eindeutig* ist. Die einem solchen „Primturnier“ angehörende Spielergruppe C ist nämlich charakterisiert durch die Eigenschaft, daß alle übrigen zu T gehörenden Spieler in zwei Gruppen P und Q zerfallen,

$$T = P + C + Q , \qquad (14)$$

wobei *kein* Spieler von P gegen einen von C oder Q und *kein* Spieler von C gegen einen von Q eine Partie gewinnt, während *innerhalb* C *keine* Einteilung in zwei Klassen C', C'' von der angegebenen Beschaffenheit möglich ist. Ist nun

$$T = T' + T''$$

Suppose that, at an approximation point $\overline{\mathfrak{q}}$ of $\overline{w}$, one of the magnitudes $\overline{u}_r$ were > 0 ($r \leqq p$) with $k_{rs} > 0$. Then, on account of $\lambda > 0$, we would have

$$\frac{\Phi_{1,2}^{(\lambda)}}{\Phi_{1,2}} \geqq \left(\frac{u_r + u_s}{\lambda u_r + u_s}\right)^{k_{rs}} ,$$

where, as $\overline{\mathfrak{q}}$ is approached, the quotient on the right side approaches the value

$$\left(\frac{\overline{u}_r + \overline{u}_s}{\lambda \overline{u}_r + \overline{u}_s}\right)^{k_{rs}} > 1$$

on account of $\overline{u}_r > 0$, $\lambda < 1$, and $k_{rs} > 0$; thus we would have

$$\Phi_{1,2}^{(\lambda)} > \Phi_{1,2}$$

in $\overline{\mathfrak{q}}$, and hence

$$\lim \Phi(\lambda u_1, \ldots, \lambda u_p; u_{p+1}, \ldots, u^n) > \lim \Phi(u_1, \ldots, u_n) = \overline{w} , \qquad (13)$$

contrary to the definition of $\overline{w}$.

Hence, it follows that when $r \leqq p$ all those $\overline{u}_r$ must be $= 0$ for which at least one k_{rs} is > 0.

As $\mathfrak{q}$ approximates $\overline{\mathfrak{q}}$, $\Phi_{1,2}$ tends toward 1, since here only the u_r with $k_{rs} > 0$, for which $\overline{u}_r$ becomes $= 0$, actually occur. Thus, $\overline{w}$ is approximated by Φ if and only if $\Phi_1(u_r)$ and $\Phi_2(u_s)$ become *as large as possible independently of one another*.

Hence, in order to determine the relative playing strengths of the players of a class it is sufficient to apply our approach to the partial tournament obtained from the total tournament by restriction to the players of this one class.

Our problem is thus reduced to ones with a smaller number of players, which are in turn subject to further reductions, so that we eventually arrive at "irreducible" problems for which all $\overline{u}_r$ are thus > 0, and hence the solution, expressed by equation (8), of case 1 is valid.

We first show that the decomposition of the "total tournament" T into "irreducible" partial tournaments, or "*prime tournaments*", $C_1, C_2, C_3, \ldots$ is *unique*. For the group of players C belonging to such a "prime tournament" is characterized by the property that all other players belonging to T fall into two groups P and Q,

$$T = P + C + Q , \qquad (14)$$

where *no* player from P wins a game against a player from C or Q, and *no* player from C wins a game against a player from Q, while *no* such partition into two classes C', C'' is possible *within* C. If now

$$T = T' + T''$$

irgendeine Klasseneinteilung des Gesamtturniers, bei der kein Spieler von T' gegen einen von T'' gewinnt, so zerfallen dabei auch P und Q in je zwei Klassen

$$P = P' + P'' , \quad Q = Q' + Q'' ,$$

wo P' und Q' zu T', P'' und Q'' zu T'' gehören, während C *ganz* zu T' oder ganz zu T'' gehört, da sonst

$$C = C' + C''$$

zerlegbar wäre gegen die Annahme. Wird also bei fortgesetzter Zerlegung des Turniers immer derjenige Bestandteil weiter zerlegt, der mit C auch nur ein Element gemein hat, so muß der schließlich übrigbleibende unzerlegbare Rest C^* das ganze C enthalten, und ebenso umgekehrt C wieder C^*, d. h. $C = C^*$, und die Zerlegung (in Primturniere) ist somit *eindeutig*.

Ist C_1 irgendein von C verschiedenes Primturnier, so liegt es entweder *ganz* in P oder *ganz* in Q, da es sonst zerlegbar wäre, oder *beliebig* in einem von beiden, sofern nämlich C und C_1 gar nicht miteinander gespielt haben.

§ 2.

Die eindeutige Lösung des irreduziblen Problems.

Wir beschränken uns zunächst auf den „irreduziblen“ Fall, setzen also voraus, daß eine Einteilung aller Spieler in zwei Klassen *unmöglich* ist, bei der kein Spieler A_r der Unterklasse gegen einen A_s der Oberklasse auch nur eine Partie gewinnt, bei der also für jede solche Kombination r, s immer $g_{rs} = 0$
443 wäre. Dann kann, wie aus den Betrachtungen des § 1 | hervorgeht, bei *keiner* Approximation der oberen Grenze $\overline{w}$ irgendeiner der Quotienten $\frac{u_r}{u_s}$ sich der Null nähern (wobei einige, aber nicht alle $\overline{u}_r = 0$ würden), sondern es kann nur einer der Fälle 1 und 2 vorliegen, wo $\overline{w}$ als das *wahre Maximum* der Funktion Φ für ein System *positiver* Werte $u_t = \overline{u}_t > 0$ tatsächlich angenommen wird. Für solche Werte u_t gelten also auch die Gleichungen (8), die wir jetzt in der Form schreiben:

$$\sum_t{}' k_{rt} \frac{u_r}{u_r + u_t} = g_r \quad \text{für} \quad r = 1, 2, \ldots, n , \tag{8a}$$

und die *Existenz* einer *positiven* Lösung $u_r = \overline{u}_r > 0$ ist für den „irreduziblen“ Fall damit bewiesen.

Um aber auch die *Eindeutigkeit* dieser Lösung zu zeigen, bedienen wir uns einer leicht herleitbaren Hilfsformel. Durch Summation der Gleichun-

is any partition of the total tournament for which no player from T' wins against a player from T'', then P and Q, too, decompose into two classes

$$P = P' + P'' , \quad Q = Q' + Q'' ,$$

where P' and Q' belong to T', P'' and Q'' to T'', whereas C belongs *entirely* to T' or entirely to T'', since otherwise

$$C = C' + C''$$

would be decomposable, contrary to the assumption. Hence, if that part which shares even only one element with C is always further decomposed as the tournament is gradually decomposed, then the indecomposable remainder C^* left over in the end must contain the entire C, and, conversely, C also C^*, i. e., $C = C^*$, and the decomposition (into prime tournaments) is thus *unique*.

If C_1 is any prime tournament different from C, then it lies either *entirely* in P or *entirely* in Q, since otherwise it would be decomposable, or *arbitrarily* in one of them, namely in case C and C_1 have not played each other at all.[6]

§ 2.

The unique solution of the irreducible problem.

First, we restrict ourselves to the "irreducible" case, and hence assume that it is *impossible* to divide all players into two classes so that no player A_r of the lower class wins even only one game against a player A_s of the upper class, and hence so that for each such combination r, s always $g_{rs} = 0$. Then, as is obvious from the considerations of § 1, *none* of the quotients $\frac{u_r}{u_s}$ can approach zero for *any* approximation of the upper limit $\overline{w}$ (while a few but not all $\overline{u}_r$ would become $= 0$). Rather, only one of the cases 1 and 2 can occur, where, in fact, $\overline{w}$ is taken as the *true maximum* of the function Φ for a system of *positive* values $u_t = \overline{u}_t > 0$. Hence, equations (8), which we now write in the form

$$\sum_t{}' k_{rt} \frac{u_r}{u_r + u_t} = g_r \quad \text{for} \quad r = 1, 2, \ldots, n , \tag{8a}$$

also hold for such values u_t, and the *existence* of a *positive* solution $u_r = \overline{u}_r > 0$ is thus proved for the "irreducible" case.

But in order to also demonstrate the *uniqueness* of this solution we use an auxiliary formula that is readily derived. For by summation of the equations

[6] ⟦Here and later Zermelo writes in an abbreviated form about tournaments playing each other when he evidently has in mind players in tournaments playing each other.⟧

gen (8a) über $r = 1, 2, \ldots, p$, wo $1 \leqq p \leqq n$ ist, erhalten wir nämlich unter Berücksichtigung der Identität

$$k_{rt}\frac{u_r}{u_r+u_t} + k_{tr}\frac{u_t}{u_t+u_r} = k_{rt} = k_{tr}$$

die Formel

$$\sum_{r=1}^{p}\sum_{t=p+1}^{n} k_{rt}\frac{u_r}{u_r+u_t} = \sum_{r=1}^{p} g_r - {\sum_{1\leqq r,\,t\leqq p}}' k_{rt} = G_p - K_p = \gamma_p \geqq 0\ , \qquad (15)$$

in der die rechte Seite den Überschuß γ_p der von den Spielern $A_1, A_2, \ldots, A_p$ überhaupt erzielten Gewinne über die Anzahl K_p der zwischen ihnen gespielten Partien, also die Anzahl der von ihnen gegen die *übrigen* Spieler A_s gewonnenen Partien ausdrückt, die im „irreduziblen" Falle, wo eine Klasseneinteilung mit $g_{rs} = 0$ ausgeschlossen ist, immer *positiv* ausfällt.

Angenommen nun, es seien u_t und u'_t zwei *wesentlich verschiedene* positive Lösungen des homogenen Gleichungssystems (8a), d. h. solche, die sich nicht nur durch einen gemeinsamen Multiplikator ϱ unterscheiden, so können wir sie beide gleichzeitig auf die Summe 1 normiert denken

$$\sum_{t=1}^{n} u_t = \sum_{t=1}^{n} u'_t = 1 \qquad (16)$$

und die Reihenfolge der u_t so wählen, daß die ersten p dieser Größen *kleiner*, alle übrigen aber größer oder gleich den entsprechenden Werten u'_t ausfallen, also

$$u_r < u'_r \quad \text{und} \quad u_s \geqq u'_s \quad \text{für} \quad 1 \leqq r \leqq p < s \leqq n \qquad (17)$$

wird. Dann folgt für jede solche Kombination r, s

$$\frac{u_r}{u_s} < \frac{u'_r}{u'_s} \qquad (17a)$$

444 | sowie

$$\frac{u_r}{u_r+u_s} < \frac{u'_r}{u'_r+u'_s} \qquad (17b)$$

und daher auch

$$\sum_{r=1}^{p}\sum_{s=p+1}^{n} k_{rs}\frac{u_r}{u_r+u_s} < \sum_{r=1}^{p}\sum_{s=p+1}^{n} k_{rs}\frac{u'_r}{u'_r+u'_s} \qquad (18)$$

im Widerspruch mit (15), sofern auch nur *eine* der Zahlen $k_{rs} > 0$ ist, d. h. sofern die Spieler A_r mit den andern A_s überhaupt gespielt haben, was wegen $k_{rs} \geqq g_{rs} > 0$ im irreduziblen Falle immer zutrifft.

(8a) over $r = 1, 2, \ldots, p$, where $1 \leqq p \leqq n$ and by taking into account the identity

$$k_{rt}\frac{u_r}{u_r + u_t} + k_{tr}\frac{u_t}{u_t + u_r} = k_{rt} = k_{tr}$$

we obtain the formula

$$\sum_{r=1}^{p}\sum_{t=p+1}^{n} k_{rt}\frac{u_r}{u_r + u_t} = \sum_{r=1}^{p} g_r - \sum_{1 \leqq r,\, t \leqq p}{}' k_{rt} = G_p - K_p = \gamma_p \geqq 0\,, \quad (15)$$

in which the right side expresses the surplus γ_p of the total gains attained by the players $A_1, A_2, \ldots, A_p$ over the number K_p of the games played between them, and hence it expresses the number of games won by them against the *remaining* players A_s, which is always *positive* in the "irreducible" case, where a partition into classes with $g_{rs} = 0$ is excluded.

Let us now assume that u_t and u'_t are two *essentially different* positive solutions of the homogeneous system of equations (8a), i. e., ones that do not only differ by a common multiplier ϱ. Then we can consider both of them simultaneously normed to the sum 1,

$$\sum_{t=1}^{n} u_t = \sum_{t=1}^{n} u'_t = 1\,, \quad (16)$$

and we can choose the order of the u_t's so that the first p's of these magnitudes are *smaller* than, but all remaining magnitudes greater than or equal to, the corresponding values u'_t, and hence

$$u_r < u'_r \quad \text{and} \quad u_s \geqq u'_s \quad \text{for} \quad 1 \leqq r \leqq p < s \leqq n\,. \quad (17)$$

It then follows for every such combination r, s that

$$\frac{u_r}{u_s} < \frac{u'_r}{u'_s} \quad (17a)$$

and also

$$\frac{u_r}{u_r + u_s} < \frac{u'_r}{u'_r + u'_s}\,, \quad (17b)$$

and hence also

$$\sum_{r=1}^{p}\sum_{s=p+1}^{n} k_{rs}\frac{u_r}{u_r + u_s} < \sum_{r=1}^{p}\sum_{s=p+1}^{n} k_{rs}\frac{u'_r}{u'_r + u'_s} \quad (18)$$

contrary to (15), provided that even *one* of the numbers k_{rs} is > 0, i.e., provided that the players A_r have played the other players A_s at all, which, on account of $k_{rs} \geqq g_{rs} > 0$, is always true in the irreducible case.

Durch die Gleichungen (8a) sind also im betrachteten Falle die *Verhältnisse* der (positiven) Größen u_t *eindeutig* bestimmt und *jede* positive Lösung des Gleichungssystems muß als die *einzige* Lösung des *irreduziblen* Problems auch das Maximum $\overline{w}$ der durch (3) definierten Funktion $\Phi(u)$ liefern.

Mit Hilfe der Formel (15) läßt sich aber noch eine wichtige Eigenschaft unseres Lösungssystems herleiten. Wir betrachten wieder zwei verschiedene positive Lösungen u_t und u'_t der Gleichungen (8a), die wir uns gleichfalls gemäß (16) normiert denken, die aber zu zwei *verschiedenen* Verteilungen der Gewinnzahlen g_t und g'_t, d. h. zu verschiedenen Turnier-Ergebnissen gehören mögen. Auch hier denken wir uns wieder die Reihenfolge so gewählt, daß für $r \leqq p < s$ die Beziehungen (17), (17a) und (17b) gelten[6]. Dann erhalten wir, da nach unserer Annahme immer $k_{rt} = k'_{rt}$ die Anzahl der gespielten Partien beidemal die gleiche sein soll, wieder die Ungleichheit (18) und somit auch wegen (15)

$$G_p = g_1 + g_2 + \ldots + g_p < g'_1 + g'_2 + \ldots + g'_p = G'_p \,. \tag{19}$$

Es muß also bei jedem Übergang von einem Turnier-Ergebnis g_{rt} zum andern g'_{rt} die Gesamtsumme aller Gewinne für diejenige Spielergruppe A_r wachsen, deren normierte Spielstärken u_r zunehmen. Angenommen nun, bei einem solchen Übergange sei g_q die *einzige* zunehmende $g_q < g'_q$, während sonst immer $g_t \geqq g'_t$ sein soll. Dann muß g_q notwendig in der Teilsumme G_q vorkommen, d. h. $q \leqq p$ und $u_q < u'_q$ sein und *die relative* (*normierte*) *Spielstärke eines Spielers A_r wird dadurch notwendig vergrößert, daß er bei sonst ungeänderten Ergebnissen eine einzelne Partie gewinnt, anstatt sie zu verlieren*[7].

445 |

§ 3.

Die Lösung des allgemeinen Problems.

Nachdem für den „irreduziblen“ Fall die Existenz einer eindeutig bestimmten Lösung erwiesen ist, handelt es sich nun um die Frage, was sich daraus für den allgemeinen Fall ergibt. In § 1 haben wir gesehen, wie das Gesamtturnier T *eindeutig* in irreduzible Teilturniere $C_1, C_2, \ldots$ zerlegt werden kann. Für die Spieler jedes solchen „Primturniers“ C sind also nach § 2 die gesuchten „Spielstärken“ als Verhältnisse positiver Zahlen bestimmt. Es seien C_1, C_2 zwei verschiedene irreduzible Teilturniere derart, daß kein Spieler von C_1 gegen einen von C_2 gewonnen, aber wenigstens einer A_r von C_1 mit einem A_s

[6] Den Fall, wo alle $u_t = u'_t$ sind, können wir ausschließen, weil wegen (8a) sonst auch alle $g_t = g'_t$ wären, gegen die Annahme.

[7] Die Wichtigkeit dieser an eine brauchbare Bewertungsmethode zu stellenden Forderung betont E. Landau a. a. O. § 3, wo er auch zeigt, daß sie in den von ihm behandelten bisherigen Lösungsversuchen des Problems nicht durchweg erfüllt ist.

Thus, equations (8a) *uniquely* determine the *ratios* of the (positive) magnitudes u_t in the case under consideration, and *every* positive solution of the system of equations, being the *only* solution of the *irreducible* problem, must also yield the maximum $\overline{w}$ of the function $\Phi(u)$ defined by (3).

By means of formula (15) we can derive yet another important property of our system of solutions. Once again, we consider two different positive solutions u_t and u'_t of equations (8a) which, in accordance with (16), we assume to be normed as well but which may belong to two *different* distributions of numbers of victories g_t and g'_t, i. e., to different tournament results. Here, again, we assume that the order is chosen so that for $r \leqq p < s$ the relations (17), (17a) and (17b) hold.[7] Then, since by our assumption always $k_{rt} = k'_{rt}$, the number of games played is supposed to be the same both times, we again obtain inequality (18), and hence, on account of (15), also

$$G_p = g_1 + g_2 + \ldots + g_p < g'_1 + g'_2 + \ldots + g'_p = G'_p \,. \tag{19}$$

Thus, for each transition from one tournament result g_{rt} to the other g'_{rt} the sum total of all gains must increase for that group of players A_r whose normed playing strengths u_r increase. Let us now suppose that for such a transition g_q is the *only* increasing $g_q < g'_q$, while otherwise we shall always have $g_t \geqq g'_t$. Then g_q must necessarily occur in the partial sum G_q, i. e., we must have $q \leqq p$ and $u_q < u'_q$, and *the relative* (*normed*) *playing strength of a player* A_r *is necessarily increased due to the fact that for otherwise unaltered results he wins a particular game instead of losing it.*[8]

§ 3.

The solution of the general problem.

Having demonstrated the existence of a uniquely determined solution for the "irreducible" case, we now turn to the question of its import for the general case. In § 1 we learned how to decompose the total tournament T *uniquely* into irreducible partial tournaments $C_1, C_2, \ldots$. Hence, by § 2, the sought "playing strengths" are determined as ratios of positive numbers for the players of every such "prime tournament" C. Let C_1, C_2 be two different irreducible partial tournaments so that no player from C_1 has won against one from C_2 but so that at least one A_r from C_1 has played with (and hence

[7] We can exclude the case where all u_t's are $= u'_t$, since otherwise, on account of (8a), all g_t's would also be $= g'_t$, contrary to the assumption.

[8] The significance of this requirement, which any useful evaluation method must meet, is stressed in *Landau 1914*, § 3, where Landau also demonstrates that it is not always satisfied by his previous attempts to solve the problem.

von C_2 gespielt (und also verloren) hat. Dann muß, wie wir § 1 gezeigt haben, die Spielstärke von A_r im Verhältnis zu der von A_s verschwinden, und das gleiche muß auch für *alle* Spieler von C_1 gegenüber *allen* von C_2 gelten, da die Quotienten der Spielstärken in jedem irreduziblen Teilturnier endlich und von 0 verschieden sind. Wir schreiben dann $C_1 < C_2$.

Ist nun

$$T = P + C_1 + Q \tag{20}$$

die für das Primturnier C_1 charakteristische Zerlegung des Gesamtturniers von der Form (14), und ist $C_1 < C_2$, wobei Spieler von C_1 mit solchen von C_2 wirklich gespielt (und verloren) haben, so muß C_2 in Q enthalten sein, da keiner von P gegen einen von C_1 gewonnen hat. Ist $C_2 < C_3$ im Sinne unserer Definition, so kann C_3 nicht in P enthalten sein, weil *kein* Spieler von P gegen einen von $Q = C_2 + Q'$ gewonnen hat und doch C_3 gegen C_2 gewonnen haben soll. Vielmehr liegt C_3 ganz in Q', also in Q, und kein Spieler von C_1 kann gegen einen von C_3 gewonnen haben. Bildet man weiter von C_1 ausgehend eine Kette von Primturnieren $C_1, C_2, \ldots, C_t$, für die sukzessive

$$C_1 < C_2\ ,\ C_2 < C_3\ , \ldots , C_{t-1} < C_t$$

sein möge, derart, daß jedes gegen das nachfolgende gespielt und verloren hat, so ergibt sich durch vollständige Induktion, daß *alle* auf C_1 folgenden in Q enthalten sind, also kein Spieler von C_1 gegen einen von C_t gewonnen haben und mithin sicher *nicht* $C_t < C_1$ sein kann. Da also zyklische Ketten dieser Art ausgeschlossen sind, so kann man festsetzen, daß jedesmal, wo $C_i < C_k$ und $C_k < C_l$ ist, auch $C_i < C_l$ sein soll, ohne daß bei dieser erweiterten Definition jemals *gleichzeitig* $C_i < C_k$ und $C_k < C_i$ werden könnte. Durch diese Festsetzung wird also die Menge der „primitiven Teilturniere“
446 oder „Primturniere“ C_i im Sinne von | Hausdorff[8] „teilweise geordnet“, und jedesmal, wo $C_i < C_k$ ist, muß bei der Approximation der oberen Grenze $\overline{w}$ jeder Quotient u_r/u_s, wo der Spieler A_r zu C_i und A_s zu C_k gehört, nach Null konvergieren, d. h. alle Spieler von C_i sind unendlich gering zu bewerten gegen alle von C_k. „Vergleichbar“ in dem Sinne, daß entweder $C_i < C_k$ oder $C_k < C_i$ ausfällt, sind also zwei primitive Teilturniere C_i und C_k einmal dann, wenn zwischen ihnen überhaupt Spiele stattgefunden haben, weiter aber auch in dem Falle, wo zwischen ihnen eine lückenlose Kette von primitiven Teilturnieren $C_1, C_2, \ldots$ eingeschaltet werden kann, in welcher je zwei aufeinanderfolgende miteinander gespielt und jedes gegen das vorangehende gewonnen haben:

$$C_1 < C_2 < C_3 < \ldots < C_t\ .$$

Als *„unvergleichbar“* gelten dagegen nur solche Paare von primitiven Teilturnieren, für die eine solche Kette nicht existiert. So wären z. B. im Falle dreier

[8] F. Hausdorff, Grundzüge der Mengenlehre, Kap. VI, § 1.

lost to) an A_s from C_2. Then, as we have shown in § 1, the playing strength of A_r must disappear in relation to the one of A_s, and the same must be true of *all* players from C_1 with respect to *all* players from C_2, since the quotients of the playing strengths are finite in every irreducible partial tournament and different from 0. We then write $C_1 < C_2$.

Now if

$$T = P + C_1 + Q \tag{20}$$

is the decomposition of the total tournament that is characteristic for the prime tournament C_1 and of the form (14), and if $C_1 < C_2$, where players from C_1 have actually played with (and lost to) those from C_2, then C_2 must be contained in Q, since no player from P has won against any from C_1. If $C_2 < C_3$ in the sense of our definition, then C_3 cannot be contained in P since *no* player from P has won against any from $Q = C_2 + Q'$ even though C_3 is supposed to have won against C_2. Rather, C_3 entirely lies in Q', and hence in Q, and no player from C_1 can have won against any from C_3. Furthermore, if, starting from C_1, we form a chain of prime tournaments $C_1, C_2, \ldots, C_t$ for which successively

$$C_1 < C_2 \ , \ C_2 < C_3 \ , \ldots , C_{t-1} < C_t \ ,$$

so that each has played against and lost to its successor, then, by mathematical induction, it follows that *all* successors of C_1 are contained in Q, and hence that no player from C_1 can have won against any from C_t, and, consequently, that it is certainly *not* possible that $C_t < C_1$. Thus, since cyclic chains of this sort are excluded, it is possible to stipulate that whenever $C_i < C_k$ and $C_k < C_l$, we are also supposed to have $C_i < C_l$ without there ever being the possibility that we *simultaneously* have $C_i < C_k$ and $C_k < C_i$ for this expanded definition. Hence, the set of "primitive partial tournaments" or "prime tournaments" C_i is "partially ordered" by this stipulation in Hausdorff's sense,[9] and whenever $C_i < C_k$ every quotient u_r/u_s must converge to zero as the upper limit $\overline{w}$ is approximated, where the player A_r belongs to C_i and A_s to C_k, i. e., all players from C_i must be rated infinitely low against all from C_k. Therefore, two primitive tournaments C_i and C_k are "comparable" in the sense that either $C_i < C_k$ or $C_k < C_i$ if, first of all, there are games that have been played between them, and if, furthermore, it is possible to interpolate between them a closed chain of primitive partial tournaments $C_1, C_2, \ldots$ in which any two consecutive ones have played with one another and each has won against its predecessor:

$$C_1 < C_2 < C_3 < \ldots < C_t \ .$$

By contrast, only those pairs of primitive partial tournaments are considered *"incomparable"* for which no such chain exists. E. g., for three prime

[9] *Hausdorff 1914*, chap. VI, § 1.

Primturniere C_1, C_2, C_3 die beiden ersten miteinander unvergleichbar, wenn sie beide gegen C_3 gespielt und gewonnen oder *beide* gegen C_3 verloren, aber unter sich gar nicht gespielt haben sollten. Von dem so charakterisierten Falle der „Unvergleichbarkeit“ abgesehen, der durch die Natur der Sache gegeben ist, liefert uns aber unser Wahrscheinlichkeitsansatz stets eine eindeutig bestimmte Lösung.

Wir beweisen nämlich folgenden

Satz. *Im allgemeinsten Falle eines in die irreduziblen Teilturniere C_1, C_2, C_3, …, C_m zerfallenden Turnieres T wird die obere Grenze $\overline{w}$ der Turnier-Wahrscheinlichkeit $\Phi(u)$ approximiert, indem man* 1. *die verhältnismäßigen Spielstärken $u_r : u_t$ der zu jedem Primturnier C_i gehörenden Spieler so wählt, daß die auf C_i bezogene Teilfunktion $\Phi_i(u)$, also die Wahrscheinlichkeit des auf C_i beschränkten Teilergebnisses, ihren Maximalwert $\overline{w}_i$ annimmt, und* 2. *für je zwei Primturniere C_i und C_j, welche der Bedingung $C_i < C_j$ genügen, alle Quotienten u_r/u_s der Spielstärken eines Spielers A_r aus C_i und eines Spielers A_s aus C_j nach Null abnehmen läßt.*

Zum Beweise benutzen wir die Tatsache, daß jeder Faktor des Produktes $\Phi(u)$ von der Form ist:

$$w_{rs} = \frac{u_r^{g_{rs}} u_s^{g_{sr}}}{(u_r + u_s)^{k_{rs}}} , \tag{2}$$

wobei die entsprechenden Spieler A_r und A_s höchstens *zwei* verschiedenen Primturnieren C_i und C_j angehören können. Faßt man nun alle diejenigen Faktoren w_{rs} zu einem Teilprodukte $\Phi_i(u)$ zusammen, für welche bei-
447 de | Spieler *demselben* Primturniere C_i angehören, so erhält man genau die Wahrscheinlichkeit w_i für das auf C_i bezogene Teilergebnis, und diese besitzt dann, weil C_1 irreduzibel ist, bei geeigneter Wahl der in Φ_i vorkommenden relativen Spielstärken $u_r : u_s$ ein wahres Maximum $\Phi_i(\overline{u}) = \overline{w}_i$.

Außerdem enthält aber $\Phi(u)$ noch solche Faktoren w_{rs}, für welche die Spieler A_r und A_s zwei *verschiedenen* Primturnieren C_i und C_j angehören. Faßt man also für jede Kombination i, j die entsprechenden Faktoren w_{rs} wieder zu einem Teilprodukte $\Phi_{ij}(u)$ zusammen, so ergibt sich für die Gesamtfunktion folgende Zerlegung:

$$\Phi(u) = \prod_{i=1}^{m} \Phi_i(u) \cdot \prod_{i,\,j}^{i \neq j} \Phi_{ij}(u) . \tag{21}$$

Ist nun $u_t = \overline{u}_t$ irgendeine, etwa die auf die Summe $\sum_t \overline{u}_t = 1$ der vorkommenden Variablen normierte Lösung des auf das i-te Primturnier C_i reduzierten Maximumsproblems

$$\Phi_i(\overline{u}) = \overline{w}_i \geqq \Phi_i(u) ,$$

tournaments C_1, C_2, C_3, the first two would be mutually incomparable if they both had played and won against C_3 or *both* lost to C_3 but not played one another. Barring the case of "incomparability" thus characterized, which is given by the nature of things, our probabilistic approach always yields a uniquely determined solution.

For we prove the following

Theorem. *In the most general case of a tournament T that decomposes into the irreducible partial tournaments C_1, C_2, C_3, ..., C_m, the upper limit $\overline{w}$ of the tournament probability $\Phi(u)$ is approximated* 1. *by choosing the relative playing strength $u_r : u_t$ of the players belonging to each prime tournament C_i so that the partial function $\Phi_i(u)$ referring to C_i, and hence the probability of the partial result restricted to C_i, assumes its maximal value $\overline{w}_i$, and* 2. *by letting all quotients u_r/u_s of the playing strengths of a player A_r from C_i and a player A_s from C_j decrease toward zero for any two prime tournaments C_i and C_j satisfying the condition $C_i < C_j$.*

For a proof we use the fact that every factor of the product $\Phi(u)$ is of the form

$$w_{rs} = \frac{u_r^{g_{rs}} u_s^{g_{sr}}}{(u_r + u_s)^{k_{rs}}} , \tag{2}$$

where the corresponding players A_r and A_s can belong to at most *two* different prime tournaments C_i and C_j. If we now bring together all those factors w_{rs} into a partial product $\Phi_i(u)$ for which the two players belong to *the same* prime tournament C_i, we obtain precisely the probability w_i for the partial result referring to C_i, which then possesses a true maximum $\Phi_i(\overline{u}) = \overline{w}_i$ for a suitable choice of the relative playing strengths $u_r : u_s$ occurring in Φ_i, since C_1 is irreducible.

But, in addition, $\Phi(u)$ still contains those factors w_{rs} for which the players A_r and A_s belong to two *different* prime tournaments C_i and C_j. Thus, if we again bring together for every combination i, j the corresponding factors w_{rs} into a partial product $\Phi_{ij}(u)$, then we obtain the following decomposition for the total function:

$$\Phi(u) = \prod_{i=1}^{m} \Phi_i(u) \cdot \prod_{i,j}^{i \neq j} \Phi_{ij}(u) . \tag{21}$$

Now if $u_t = \overline{u}_t$ is any solution—say, one that is normed to the sum $\sum_t \overline{u}_t = 1$ of the occurring variables—of the maximum problem

$$\Phi_i(\overline{u}) = \overline{w}_i \geqq \Phi_i(u)$$

so entsteht die allgemeinste Lösung desselben Teilproblems durch Hinzufügung eines beliebigen positiven Faktors ϱ_i, da auch $\Phi_i(u)$ in bezug auf die vorkommenden u_t wieder homogen von 0-ter Dimension ist:

$$\Phi_i(\varrho\overline{u}) = \Phi_i(\overline{u}) = \overline{w}_i \ .$$

Für ein beliebiges System der ϱ_i wird also das über $i = 1, 2, \ldots, m$ gewonnene einfache Produkt in (21) durch den Ansatz $u_t = \varrho_i \overline{u}_t$ seinen Maximalwert $\overline{w}_1 \overline{w}_2 \ldots \overline{w}_n$ erreichen:

$$\prod_{i=1}^{m} \Phi_i\,(\varrho_i \overline{u}) = \overline{w}_1 \overline{w}_2 \ldots \overline{w}_m = w^* \geqq \prod_{i=1}^{m} \Phi_i(u)\ , \tag{22}$$

während das über i, j genommene Doppelprodukt in (21) durch Wahl der Multiplikatoren ϱ_i, wie wir zeigen wollen, seiner oberen Grenze 1 beliebig genähert werden kann. Ist nämlich $\Phi_{ij}(u)$ ein Faktor dieses Produktes, so sind nur folgende drei Fälle möglich:

1. Zwischen C_i und C_j ist überhaupt nicht gespielt worden, alle $k_{rs} = 0$, wenn A_r und A_s zu C_i bzw. C_j gehören, dann sind auch die entsprechenden g_{rs} und g_{sr} sämtlich $= 0$ und die Funktion $\Phi_{ij}(u)$ reduziert sich auf den konstanten Wert 1. Solche Faktoren Φ_{ij} können also ganz außer Betracht bleiben.

2. Mindestens ein Spieler A_r von C_i hat gegen einen A_s von C_j gespielt und *verloren.* Dann muß in der durch C_i gemäß (14) bestimmten Zerlegung

$$S = P_i + C_i + Q_i$$

448 | C_j notwendig in Q_i enthalten sein, d. h. alle $g_{rs} = 0$, sofern A_r zu C_i und A_s zu C_j gehört, und $C_i < C_j$ in der im Anfang des § eingeführten Schreibweise. Dann erscheinen aber alle Variablen u_r, die zu C_i gehören, ausschließlich im *Nenner* von $\Phi_{ij}(u)$, und der Ausdruck

$$\Phi_{ij}(u) = \prod_{r,\,s} \frac{u_s^{g_{sr}}}{(u_r + u_s)^{k_{rs}}} = \prod_{r,\,s} \Big(\frac{u_s}{u_r + u_s}\Big)^{k_{rs}} \tag{23}$$

wird nach der Substitution $u_r = \varrho_i \overline{u}_r$, $u_s = \varrho_j \overline{u}_s$, wo r alle auf C_i und s alle auf C_j bezüglichen Indizes durchläuft,

$$\Phi_{ij}(u) = \prod_{r,\,s} \Big(\frac{\lambda_j \overline{u}_s}{\lambda_i \overline{u}_r + \lambda_j \overline{u}_s}\Big)^{k_{rs}} = \prod_{r,\,s} \Big(\frac{\overline{u}_s}{\overline{u}_s + \frac{\lambda_i}{\lambda_j}\overline{u}_r}\Big)^{k_{rs}}\ , \tag{24}$$

sich seiner oberen Grenze 1 nähern, wenn der Quotient $\dfrac{\lambda_i}{\lambda_j}$ der Grenze 0 zustrebt.

3. Mindestens ein Spieler A_r von C_i hat gegen einen A_s von C_j *gewonnen*, $g_{rs} > 0$. Dann gehört C_j notwendig zu P_i und alle $g_{sr} = 0$, sowie $C_j < C_i$. Hier gilt wieder das gleiche für C_j und C_i wie im Fall 2 für C_i und C_j, und die

that is reduced to the i-th partial tournament C_i, then the most general solution of the same partial problem arises by adding any positive factor ϱ_i, since also $\Phi_i(u)$ is again homogeneous of 0-th dimension with respect to the occurring u_t:

$$\Phi_i(\varrho\overline{u}) = \Phi_i(\overline{u}) = \overline{w}_i \ .$$

Hence, for any system of the ϱ_i, the simple product over $i = 1, 2, \ldots, m$ in (21) will reach its maximal value $\overline{w}_1\overline{w}_2 \ldots \overline{w}_n$ by means of the ansatz $u_t = \varrho_i\overline{u}_t$:

$$\prod_{i=1}^{m} \Phi_i\,(\varrho_i\overline{u}) = \overline{w}_1\overline{w}_2 \ldots \overline{w}_m = w^* \geqq \prod_{i=1}^{m} \Phi_i(u) \ , \tag{22}$$

whereas it is possible to make the double product in (21) taken over i, j approach its upper limit 1 as close as we wish, as we shall show, by choice of the multipliers ϱ_i. For if $\Phi_{ij}(u)$ is a factor of this product, then only the following three cases are possible:

1. No game at all has been played between C_i and C_j, all k_{rs}'s are $= 0$, if A_r and A_s belong to C_i and C_j respectively, then the corresponding g_{rs}'s and g_{sr}'s are also all $= 0$, and the function $\Phi_{ij}(u)$ reduces to the constant value 1. Hence, such factors Φ_{ij} need not be considered at all.

2. At least one player A_r from C_i has played against and *lost* to an A_s from C_j. Then C_j must necessarily be contained in Q_i in the decomposition

$$S = P_i + C_i + Q_i$$

that is determined by C_i in accordance with (14), i. e., all g_{rs}'s are $= 0$, provided A_r belongs to C_i, and A_s to C_j, and $C_i < C_j$ in the notion introduced at the beginning of the §. But then all variables u_r belonging to C_i exclusively appear in the *denominator* of $\Phi_{ij}(u)$, and the expression

$$\Phi_{ij}(u) = \prod_{r,\,s} \frac{u_s^{g_{sr}}}{(u_r + u_s)^{k_{rs}}} = \prod_{r,\,s} \Big(\frac{u_s}{u_r + u_s}\Big)^{k_{rs}} \ , \tag{23}$$

upon the substitution $u_r = \varrho_i\overline{u}_r$, $u_s = \varrho_j\overline{u}_s$, where r runs through all indices referring to C_i and s through all indices referring to C_j,

$$\Phi_{ij}(u) = \prod_{r,\,s} \Big(\frac{\lambda_j\overline{u}_s}{\lambda_i\overline{u}_r + \lambda_j\overline{u}_s}\Big)^{k_{rs}} = \prod_{r,\,s} \Big(\frac{\overline{u}_s}{\overline{u}_s + \frac{\lambda_i}{\lambda_j}\overline{u}_r}\Big)^{k_{rs}} \ , \tag{24}$$

will approach its upper limit 1 if the quotient $\dfrac{\lambda_i}{\lambda_j}$ tends toward the limit 0.

3. At least one player A_r from C_i has *won* against a player A_s from C_j, $g_{rs} > 0$. Then C_j necessarily belongs to P_i, and all g_{sr}'s are $= 0$, and also $C_j < C_i$. In this case, what holds for C_i and C_j in case 2 also holds for C_j

Funktion $\Phi_{ij}(u) = \Phi_{ij}\,(\lambda_i\overline{u}, \lambda_j\overline{u})$ nähert sich der Eins, wenn der Quotient $\frac{\lambda_j}{\lambda_i}$ der Null zustrebt.

Sind nun $\Phi_{ij}(u)$ und $\Phi_{jl}(u)$ zwei (nicht konstante) Faktoren von $\Phi(u)$, für welche gleichzeitig $C_i < C_j$ und $C_j < C_l$ ausfällt, so daß Spieler von C_i gegen solche von C_j, sowie Spieler von C_j gegen solche von C_l gespielt und verloren haben, so werden *beide* Funktionen Φ_{ij} und Φ_{jl} sich der Grenze 1 nähern, wenn gleichzeitig die beiden Quotienten

$$\frac{\varrho_i}{\varrho_j}\,,\,\frac{\varrho_j}{\varrho_l} \quad \text{und damit auch ihr Produkt } \frac{\varrho_i}{\varrho_l}$$

der Null zustrebt, und das Analoge gilt für jede Kette von Primturnieren, für welche der Reihe nach

$$C_i < C_j < C_l < \ldots < C_z$$

ist und von denen je zwei aufeinanderfolgende gegeneinander gespielt haben. Daraus ergibt sich aber, daß auch nach unserer allgemeinen Festsetzung *jedesmal*, wo $C_i < C_j$ ist, auch der Quotient $\frac{\varrho_i}{\varrho_j}$ bei der betrachteten Approximation sich der Null nähert. Da nun diese „teilweise Ordnung“ der Primturniere, wie wir S. 445–446 gesehen haben, widerspruchsfrei möglich ist, so können wir durch Befolgung dieser Regel

$$\lim \frac{\varrho_i}{\varrho_j} = 0 \quad \text{für} \quad C_i < C_j$$

449 | gleichzeitig *alle* in (21) vorkommenden Faktoren $\Phi_{ij}(u)$ und damit auch ihr Produkt sich der Eins nähern lassen und erhalten demgemäß

$$\lim \Phi(\varrho_i\overline{u}_i) \begin{array}{l} = \overline{w}_1\overline{w}_2\ldots\overline{w}_m \lim \prod_{i,\,j} \Phi_{ij}\,(\varrho_i\overline{u}_r, \varrho_j\overline{u}_s)\ , \\ = \overline{w}_1\overline{w}_2\ldots\overline{w}_m \geqq \prod_i \Phi_i(u) \geqq \Phi(u)\ , \end{array} \tag{25}$$

d. h. es ist $w^* = \overline{w}$ in der That die obere Grenze von $\Phi(u)$, die wir auf diese Weise approximieren.

§ 4.

Diskussion der Lösung für reguläre Turniere.

In dem praktisch wichtigen Fall, den wir als den „regulären“ bezeichnen wollen, wo alle

$$k_{rs} = k$$

and C_i, and the function $\Phi_{ij}(u) = \Phi_{ij}(\lambda_i \overline{u}, \lambda_j \overline{u})$ approaches 1 if the quotient $\frac{\lambda_j}{\lambda_i}$ tends toward zero.

If now $\Phi_{ij}(u)$ and $\Phi_{jl}(u)$ are two (non-constant) factors of $\Phi(u)$, for which simultaneously $C_i < C_j$ and $C_j < C_l$ so that players from C_i and C_j have played against and lost to players from C_j and C_l, respectively, then *both* functions Φ_{ij} and Φ_{jl} will approach the limit 1 if simultaneously approaching zero are the two quotients

$$\frac{\varrho_i}{\varrho_j}\,,\ \frac{\varrho_j}{\varrho_l} \quad \text{and hence also their product } \frac{\varrho_i}{\varrho_l}\,,$$

and the analog holds for every chain of prime tournaments for which successively

$$C_i < C_j < C_l < \ldots < C_z$$

and of which any two prime tournaments immediately succeeding one another have played each other. But from this it follows also in accordance with our general stipulation that *whenever* $C_i < C_j$, the quotient $\frac{\varrho_i}{\varrho_j}$, too, approaches zero for the approximation under consideration. Since, as we have seen on pp. 445–446, this "partial order" of the prime tournaments is possible without contradiction, we can, by following the rule

$$\lim \frac{\varrho_i}{\varrho_j} = 0 \quad \text{for} \quad C_i < C_j\ ,$$

let *all* factors $\Phi_{ij}(u)$ occurring in (21), and hence also their product, simultaneously approach 1, and thus obtain

$$\lim \Phi(\varrho_i \overline{u}_i) \begin{array}{l} = \overline{w}_1 \overline{w}_2 \ldots \overline{w}_m \lim \prod_{i,\,j} \Phi_{ij}\,(\varrho_i \overline{u}_r, \varrho_j \overline{u}_s)\ , \\ = \overline{w}_1 \overline{w}_2 \ldots \overline{w}_m \geqq \prod_i \Phi_i(u) \geqq \Phi(u)\ , \end{array} \tag{25}$$

i. e., $w^* = \overline{w}$ is really the upper limit of $\Phi(u)$, which we approximate in this way.

§ 4.

Discussion of the solution for regular tournaments.

In the practically significant case, which we shall call the "regular" one, where all

$$k_{rs} = k$$

einander gleich sind, wo also jeder Spieler mit jedem die gleiche Anzahl Partien gespielt hat, geht (8a) über in

$$g_r^* \equiv \frac{g_r}{k} = u_r \sum_{t=1}^{n}{}' \frac{1}{u_r + u_t} \quad \text{für} \quad r = 1, \ldots, n \geqq 2 \,. \tag{8b}$$

Sind in der Reihe $u_1, \ldots, u_n$ zwei Größen einander gleich: $u_r = u_s$, so muß auch

$$g_r = g_s$$

sein; ist aber

$$u_r < u_s \,, \tag{26}$$

so ist auch

$$\frac{1}{u_r + u_t} > \frac{1}{u_s + u_t} \tag{27}$$

für alle $r \neq t \neq s$ und

$$\frac{u_r}{u_r + u_t} < \frac{u_s}{u_s + u_t} \,, \tag{28}$$

also

$$g_r < g_s \,,$$

aber, sofern $n > 2$ ist,

$$\frac{g_r}{u_r} > \frac{g_s}{u_s}$$

oder

$$\frac{u_r}{u_s} < \frac{g_r}{g_s} < 1 \,. \tag{29}$$

Umgekehrt folgt aus $g_r \leqq g_s$ wieder $u_r \leqq u_s$.

Wir haben also den

Satz. *Bei einem Turnier, in welchem jeder Spieler mit jedem die gleiche*
450 *Anzahl Partien spielt, haben — im irreduziblen Falle — die be-| zeichneten Spielstärken u_t dieselbe Reihenfolge wie die Gewinnzahlen g_t, aber für $n > 2$ in quantitativ größeren Abstandsverhältnissen.*

In unserem Falle $k_{rs} = k$ wird, wenn man die g_r (und damit zugleich die u_r) nach ansteigender Größe ordnet,

$$1 \leqq g_1 \leqq g_2 \leqq \ldots \leqq g_n \,,$$

are equal to one another, and hence where every player has played every ⟦other⟧ player the same number of games, (8a) is transformed into

$$g_r^* \equiv \frac{g_r}{k} = u_r \sum_{t=1}^{n}{}' \frac{1}{u_r + u_t} \quad \text{for} \quad r = 1, \ldots, n \geqq 2 \,. \tag{8b}$$

If two magnitudes are equal to one another in the series $u_1, \ldots, u_n$: $u_r = u_s$, then also necessarily

$$g_r = g_s \,;$$

but if

$$u_r < u_s \,, \tag{26}$$

then also

$$\frac{1}{u_r + u_t} > \frac{1}{u_s + u_t} \tag{27}$$

for all $r \neq t \neq s$ and

$$\frac{u_r}{u_r + u_t} < \frac{u_s}{u_s + u_t} \,, \tag{28}$$

and hence

$$g_r < g_s \,,$$

but, provided that $n > 2$,

$$\frac{g_r}{u_r} > \frac{g_s}{u_s}$$

or

$$\frac{u_r}{u_s} < \frac{g_r}{g_s} < 1 \,. \tag{29}$$

Conversely, it again follows from $g_r \leqq g_s$ that $u_r \leqq u_s$.

We thus have the

Theorem. *For a tournament in which every player has played every* ⟦*other*⟧ *player the same number of games, the designated playing strengths* u_t *have—in the irreducible case—the same order as the number of victories* g_t*, but in quantitatively greater ratios of distances when* $n > 2$*.*

If, in our case $k_{rs} = k$, we arrange the g_r's (and hence also the u_r's) according to their increasing magnitude,

$$1 \leqq g_1 \leqq g_2 \leqq \ldots \leqq g_n \,,$$

jede p-te Teilsumme

$$g_1 + g_2 + \ldots + g_p \qquad (p = 1, 2, \ldots, n-1)$$

wegen (15) größer ausfallen als die entsprechende Teilsumme K_p

$$G_p = g_1 + g_2 + \ldots + g_p > K_p = \sum_{1 \leqq r,\, t \leqq p}{}' k_{rt} = k\frac{p(p-1)}{2} , \tag{30}$$

und nur für $p = n$ ergibt sich

$$g_1 + g_2 + \ldots + g_n = k\frac{n(n-1)}{2} = N . \tag{7a}$$

Durch Zerlegung der ganzen Zahl N in n Summanden erhält man also alle im betrachteten Falle möglichen durch die Gewinnzahlen g_t ausdrückbaren Turnier-Ergebnisse.

So haben wir z. B. im Falle $k = 1$, $n = 2$, $N = 1 = 0 + 1$ nur eine einzige mögliche Verteilung

$$g_1 = 0 , \quad g_2 = 1 ,$$

die dem „reduziblen“ Falle angehört und der Bewertung $\frac{u_1}{u_2} = 0$ entspricht.

Für $k = 2$, $n = 2$, $N = 2 = 0 + 2 = 1 + 1$ noch die irreduzible Verteilung $g_1 = g_2 = 1$, $u_1 = u_2$.

Für $k = 3$, $n = 2$, $N = 3 = 1 + 2$ ergibt sich die irreduzible Verteilung $g_1 = 1$, $g_2 = 2$ mit den zugehörigen Bestimmungsgleichungen (8)

$$\frac{3u_1}{u_1 + u_2} = 1 , \quad \frac{3u_2}{u_1 + u_2} = 2$$

und dem Ergebnis $u_1 : u_2 = 1 : 2$.

Allgemein ist für $n = 2$, $N = k = g_1 + g_2$

$$g_1 = k\frac{u_1}{u_1 + u_2} , \quad g_2 = k\frac{u_2}{u_1 + u_2} ,$$

also $u_1 : u_2 = g_1 : g_2$, d. h. die berechneten Spielstärken verhalten sich wie die Gewinnzahlen, während für $n > 2$ nach (29) für $r < s$ immer

$$\frac{u_r}{u_s} < \frac{g_r}{g_s} < 1$$

ausfällt.

Für $n = 3$, $k = 1$, $N = 3 = g_1 + g_2 + g_3 = 1 + 1 + 1$ haben wir nur einen einzigen irreduziblen Typus

$$g_1 = g_2 = g_3 = 1 , \quad u_1 = u_2 = u_3 ,$$

then every p-th partial sum

$$g_1 + g_2 + \ldots + g_p \qquad (p = 1, 2, \ldots, n-1)$$

will, on account of (15), be greater than the corresponding partial sum K_p

$$G_p = g_1 + g_2 + \ldots + g_p > K_p = \sum_{1 \leqq r,\, t \leqq p}{}' k_{rt} = k\frac{p(p-1)}{2} , \tag{30}$$

and only for $p = n$ we have

$$g_1 + g_2 + \ldots + g_n = k\frac{n(n-1)}{2} = N . \tag{7a}$$

By decomposing the whole number N into n summands we thus obtain all tournament results possible in the case under consideration that can be expressed by the numbers of victories g_t.

Thus, e. g., in the case $k = 1$, $n = 2$, $N = 1 = 0 + 1$ there is only one possible distribution

$$g_1 = 0 , \quad g_2 = 1$$

that belongs to the "reducible" case and corresponds to the evaluation $\frac{u_1}{u_2} = 0$.

For $k = 2$, $n = 2$, $N = 2 = 0 + 2 = 1 + 1$ there is also the irreducible distribution $g_1 = g_2 = 1$, $u_1 = u_2$.

For $k = 3$, $n = 2$, $N = 3 = 1 + 2$, we obtain the irreducible distribution $g_1 = 1$, $g_2 = 2$ together with the associated determining equations (8)

$$\frac{3u_1}{u_1 + u_2} = 1 , \quad \frac{3u_2}{u_1 + u_2} = 2$$

and the result $u_1 : u_2 = 1 : 2$.

For $n = 2$, $N = k = g_1 + g_2$, generally,

$$g_1 = k\frac{u_1}{u_1 + u_2} , \quad g_2 = k\frac{u_2}{u_1 + u_2} ,$$

and hence $u_1 : u_2 = g_1 : g_2$, i. e., the calculated playing strengths are proportional to the playing numbers, whereas for $n > 2$, by (29), always

$$\frac{u_r}{u_s} < \frac{g_r}{g_s} < 1$$

when $r < s$.

For $n = 3$, $k = 1$, $N = 3 = g_1 + g_2 + g_3 = 1 + 1 + 1$ there is only one irreducible type

$$g_1 = g_2 = g_3 = 1 , \quad u_1 = u_2 = u_3 ,$$

451 | während dem Falle $n = 3$, $k = 2$, $N = 6$ die drei Zerlegungen

$$6 = 1 + 1 + 4 = 1 + 2 + 3 = 2 + 2 + 2$$

und damit drei verschiedene Verteilungen und Bewertungen entsprechen. Von diesen ist aber die erste *reduzibel* wegen $g_1 + g_2 = 2 = 2\dfrac{2(2-1)}{2}$ entsprechend der Matrix $g_{12} = g_{21} = 1$, $g_{13} = g_{23} = 0$, $g_{31} = g_{32} = 2$, und würde sich auch ergeben, wenn jeder mit jedem nur *eine* Partie gespielt und dabei die beiden ersten miteinander Remis gemacht und gegen den dritten verloren hätten. Es bleibt also außer dem trivialen Typus $g_1 = g_2 = g_3$, $u_1 = u_2 = u_3$ nur noch *ein* irreduzibler übrig $g_1 = 1$, $g_2 = 2$, $g_3 = 3$. Hier ergeben sich, wenn man

$$\frac{u_1}{u_3} = x, \quad \frac{u_2}{u_3} = y$$

setzt, für x und y die Gleichungen

$$\frac{x}{x+y} + \frac{x}{x+1} = \frac{1}{2}\,, \quad \frac{y}{x+y} + \frac{y}{y+1} = 1\,,$$

und hieraus weiter

$$x = y^2\,, \quad 1 = y + y^2 + 3y^3\,.$$

Die kubische Gleichung hat als einzige positive Lösung

$$y = 0{,}46940\,, \quad x = y^2 = 0{,}22034\,,$$

woraus sich, auf die Summe 1 normiert, als Spielstärken ergeben

$$\begin{aligned} u_1 &= 0{,}1304\,, & g_1 &= 1 = \quad 1 + 0 \quad = \quad 0 + 1\,,\\ u_2 &= 0{,}2778\,, & g_2 &= 2 = 1 \qquad + 1 = 2 \quad + 0\,,\\ u_3 &= 0{,}5918\,, & g_3 &= 3 = 2 + 1 \qquad = 1 + 2\,. \end{aligned}$$

Diese Verteilung ergibt sich insbesondere, wenn jeder der 3 Spieler mit jedem anderen je *eine* Partie gespielt hat und dabei

1. A_1 gegen A_2 Remis gemacht und gegen A_3 verloren, A_2 gegen A_3 Remis gemacht,

oder

2. A_1 gegen A_2 verloren und gegen A_3 Remis gemacht, A_2 gegen A_3 verloren hat.

Für $n = 4$, $k = 1$ haben wir wieder nur eine einzige „irreduzible“ Verteilung entsprechend der Zerlegung

$$N = 6 = 1 + 1 + 2 + 2$$

mit dem Ergebnis

$$u_1 : u_2 : u_3 : u_4 = 1 : 1 : 3 : 3\,.$$

whereas to the case $n = 3$, $k = 2$, $N = 6$ there correspond the three decompositions

$$6 = 1 + 1 + 4 = 1 + 2 + 3 = 2 + 2 + 2 ,$$

and hence three different distributions and evaluations. But of these the first one is *reducible* on account of $g_1 + g_2 = 2 = 2\dfrac{2(2-1)}{2}$ according to the matrix $g_{12} = g_{21} = 1$, $g_{13} = g_{23} = 0$, $g_{31} = g_{32} = 2$, and we would also obtain it if everyone had played only *one* game against everyone ⟦else⟧ and the first two players had played a draw against one another and lost to the third player. Therefore, apart from the trivial type $g_1 = g_2 = g_3$, $u_1 = u_2 = u_3$, there only remains *one* irreducible one, $g_1 = 1$, $g_2 = 2$, $g_3 = 3$. Here, we obtain, if we set

$$\frac{u_1}{u_3} = x, \quad \frac{u_2}{u_3} = y ,$$

for x and y the equations

$$\frac{x}{x+y} + \frac{x}{x+1} = \frac{1}{2} , \quad \frac{y}{x+y} + \frac{y}{y+1} = 1 ,$$

and from this furthermore

$$x = y^2 , \quad 1 = y + y^2 + 3y^3 .$$

The only positive solution to the cubic equation is

$$y = 0.46940 , \quad x = y^2 = 0.22034 ,$$

from which, when normed to the sum 1, the playing strengths

$$\begin{array}{lll} u_1 = 0.1304 , & g_1 = 1 = & 1 + 0 \quad = \quad 0 + 1 , \\ u_2 = 0.2778 , & g_2 = 2 = 1 & + 1 = 2 \quad + 0 , \\ u_3 = 0.5918 , & g_3 = 3 = 2 + 1 & = 1 + 2 \end{array}$$

arise. This distribution arises in particular if each of the 3 players played every other *once* and

1. A_1 played a draw against A_2 and lost to A_3, A_2 played a draw against A_3,

or

2. A_1 lost to A_2 and played a draw against A_3, A_2 lost to A_3.

For $n = 4$, $k = 1$, there is again only one "irreducible" distribution according to the decomposition

$$N = 6 = 1 + 1 + 2 + 2$$

with the result

$$u_1 : u_2 : u_3 : u_4 = 1 : 1 : 3 : 3 .$$

452 | Für $n = 4$, $k = 2$ dagegen ergeben sich bereits 7 irreduzible Verteilungen:

$$\begin{aligned} N = 12 &= 1 + 2 + 4 + 5 \\ &= 1 + 3 + 3 + 5 \\ &= 1 + 3 + 4 + 4 \\ &= 2 + 2 + 3 + 5 \\ &= 2 + 2 + 4 + 4 \\ &= 2 + 3 + 3 + 4 \\ &= 3 + 3 + 3 + 3 \,, \end{aligned}$$

welche mit Ausnahme der letzten (mit der trivialen Lösung $u_1 = u_2 = u_3 = u_4$) auf Gleichungen dritten und höheren Grades führen. Die zweite z. B. liefert:

$$\frac{u_1}{u_4} = x = y^2 \,, \quad \frac{u_2}{u_4} = \frac{u_3}{u_4} = y \,, \quad 5y^3 + y^2 + 3y - 1 = 0 \,.$$

Für $n = 5$, $k = 1$ erhalten wir 3 irreduzible Zerlegungen

$$\begin{aligned} N = 10 &= 1 + 1 + 2 + 3 + 3 \\ &= 1 + 2 + 2 + 2 + 3 \\ &= 2 + 2 + 2 + 2 + 2 \,, \end{aligned}$$

von denen die beiden ersten wieder auf kubische Gleichungen führen.

Die Tatsache, daß für $k_{rs} = k$ die Spielstärken u_t gleichgeordnet sind den entsprechenden Gewinnzahlen g_t, gilt aber auch im *reduziblen* Falle. Haben wir nämlich eine Klasseneinteilung aller Spieler in solche A_r und A_s, daß für $1 \leqq r \leqq p < s \leqq r$ immer $g_{rs} = 0$ wird, so gilt für jedes g_r und jedes g_s

$$g_r \leqq k(p-1) < kp \leqq g_s \,, \tag{31}$$

da ein Spieler A_r der ersten Klasse höchstens gegen $p - 1$ Spieler derselben Klasse je k Partien und andererseits jeder Spieler A_s der zweiten Klasse mindestens kp Partien gegen alle der ersten Klasse gewinnt. Jede Gewinnzahl der unteren Klasse ist also kleiner als jede der oberen, während gleichzeitig jede zur unteren Klasse gehörende Spielstärke u_r gegen jede zur oberen gehörende u_s verschwindet. In dem hier betrachteten Falle erfolgt also, wenn man die g_r nach aufsteigender Größe ordnet, jede Klassentrennung immer zwischen zwei aufeinanderfolgenden g_p, g_{p+1} und zwar immer an der Stelle, wo

$$G_p = g_1 + g_2 + \ldots + g_p = k\frac{p(p-1)}{2} = k + 2k + \ldots + (p-1)k \tag{32}$$

ausfällt, während wegen (30) sonst immer

$$G_p = g_1 + g_2 + \ldots + g_p > k\frac{p(p-1)}{2}$$

For $n = 4$, $k = 2$, by contrast, there are already 7 irreducible decompositions:

$$\begin{aligned} N = 12 &= 1 + 2 + 4 + 5 \\ &= 1 + 3 + 3 + 5 \\ &= 1 + 3 + 4 + 4 \\ &= 2 + 2 + 3 + 5 \\ &= 2 + 2 + 4 + 4 \\ &= 2 + 3 + 3 + 4 \\ &= 3 + 3 + 3 + 3 \,, \end{aligned}$$

which, except for the last (with the trivial solution $u_1 = u_2 = u_3 = u_4$) lead to equations of third and higher degree. The second, e. g., yields

$$\frac{u_1}{u_4} = x = y^2 \,, \quad \frac{u_2}{u_4} = \frac{u_3}{u_4} = y \,, \quad 5y^3 + y^2 + 3y - 1 = 0 \,.$$

For $n = 5$, $k = 1$, we obtain 3 irreducible distributions

$$\begin{aligned} N = 10 &= 1 + 1 + 2 + 3 + 3 \\ &= 1 + 2 + 2 + 2 + 3 \\ &= 2 + 2 + 2 + 2 + 2 \,, \end{aligned}$$

of which the first two again lead to cubic equations.

But the fact that for $k_{rs} = k$ the playing strengths u_t are coordinated with the corresponding numbers of victories g_t also holds for the *reducible* case. For if we have a partition of all players into classes A_r and A_s so that for $1 \leqq r \leqq p < s \leqq n$ [10] always $g_{rs} = 0$, then for each g_r and each g_s

$$g_r \leqq k(p-1) < kp \leqq g_s \,, \tag{31}$$

since a player A_r of the first class can win against at most $p - 1$ players of the same class in k games each and, on the other hand, each player A_s of the second class wins at least kp games against all of the first class. Each number of victories of the lower class is therefore smaller than each of the upper class, whereas, at the same time, each playing strength u_r belonging to the lower class vanishes with respect to each u_s belonging to the upper class. Hence, in the case under consideration, each partition into classes always occurs between two g_p, g_{p+1} immediately succeeding one another if the g_r's are ordered according to increasing magnitudes, and in particular at positions where

$$G_p = g_1 + g_2 + \ldots + g_p = k\frac{p(p-1)}{2} = k + 2k + \ldots + (p-1)k \,, \tag{32}$$

whereas otherwise, on account of (30), always necessarily

$$G_p = g_1 + g_2 + \ldots + g_p > k\frac{p(p-1)}{2} \,.$$

[10] ⟦Zermelo erroneously writes "$s \leq r$" instead of "$s \leq n$".⟧

453 | sein muß. Es genügt also, die sukzessiven Teilsummen der nach der Größe geordneten Gewinnzahlen $g_1 \leqq g_2 \leqq g_3 \leqq \ldots \leqq g_n$ mit den entsprechenden Teilsummen der Reihe $0, k, 2k, \ldots, (n-1)k$ zu vergleichen, um sofort die vollständige Zerlegung in „Primturniere“ zu erhalten.

§ 5.

Die numerische Auflösung der Gleichungen durch sukzessive Approximation.

Für größere Werte von n, wo die Berechnung durch algebraische Elimination nicht mehr angängig ist, empfiehlt sich ein Approximationsverfahren, das, wie wir zeigen wollen, im irreduziblen Falle stets zum Ziele führt.

Da wir die Gleichungen (8) des Problems zur Berechnung der jetzt einfach mit u_r bezeichneten relativen Spielstärken $\overline{u}_r$ in der Form schreiben können

$$\frac{g_r}{u_r} = \sum_t{}' \frac{k_{rt}}{u_r + u_t} \qquad \begin{pmatrix} r, t = 1, 2, \ldots, n \\ t \neq r \end{pmatrix}, \tag{8}'$$

so liegt es nahe, gegebene (positive) Annäherungswerte x_r der u_r sukzessive zu verbessern, indem wir setzen

$$\frac{g_r}{x'_r} = \sum_t{}' \frac{k_{rt}}{x_r + x_t} = \frac{g_r}{f_r(x)}, \quad \text{also} \quad x'_r = f_r(x) \tag{33}$$

und so von einem System (x_r) der Reihe nach übergehen zu weiteren Systemen $(x'_r), (x''_r), \ldots, \left(x_r^{(i)}\right)$, die im Falle der Konvergenz die gesuchte Lösung liefern werden. Denn ist etwa

$$\overline{x}_r = \lim_{i \to \infty} x_r^{(i)} \qquad (r = 1, \ldots, n),$$

so ergibt sich aus (33) analog

$$\frac{g_r}{x_r^{(i+1)}} = \sum_t{}' \frac{k_{rt}}{x_r^{(i)} + x_t^{(i)}} \tag{33}_i$$

und durch Grenzübergang

$$\frac{g_r}{\overline{x}_r} = \sum_t{}' \frac{k_{rt}}{x_r + \overline{x}_t}, \tag{34}$$

d. h. die Grenzwerte $\overline{x}_r$ bilden ein Lösungssystem von (8).

Hence, it suffices to compare the successive partial sums of the numbers of victories $g_1 \leqq g_2 \leqq g_3 \leqq \ldots \leqq g_n$ ordered according to magnitude with the corresponding partial sums of the series $0, k, 2k, \ldots, (n-1)k$ in order to obtain instantly the complete decomposition into "prime tournaments".

§ 5.

The numeric solution of the equations through successive approximation.

For larger values of n where the calculation through algebraic elimination is no longer possible it is advisable to use an approximation method, which, as we shall show, always leads to the desired result in the irreducible case.

Since, in order to calculate the relative playing strengths $\overline{u}_r$ now simply designated by u_r, we can write the equations (8) of the problem in the form

$$\frac{g_r}{u_r} = \sum_t{}' \frac{k_{rt}}{u_r + u_t} \qquad \begin{pmatrix} r, t = 1, 2, \ldots, n \\ t \neq r \end{pmatrix}, \tag{8}'$$

it seems obvious that we can successively improve the given (positive) approximation values x_r of the u_r's by setting

$$\frac{g_r}{x'_r} = \sum_t{}' \frac{k_{rt}}{x_r + x_t} = \frac{g_r}{f_r(x)}, \quad \text{hence} \quad x'_r = f_r(x) \tag{33}$$

and thus successively proceed from a system (x_r) to further systems $(x'_r), (x''_r)$, $\ldots, \left(x_r^{(i)}\right)$ that will provide the sought solution in the case of convergence. For if, say,

$$\overline{x}_r = \lim_{i \to \infty} x_r^{(i)} \qquad (r = 1, \ldots, n),$$

then it analogously follows from (33) that

$$\frac{g_r}{x_r^{(i+1)}} = \sum_t{}' \frac{k_{rt}}{x_r^{(i)} + x_t^{(i)}} \tag{33}_i$$

and, taking the limit,[11]

$$\frac{g_r}{\overline{x}_r} = \sum_t{}' \frac{k_{rt}}{\overline{x}_r + \overline{x}_t}, \tag{34}$$

i. e., the limits $\overline{x}_r$ form a system of solutions of (8).

[11] ⟦In the denominator on the right side of the following formula, Zermelo erroneously writes "x_r" instead of "$\overline{x}_r$".⟧

Um die tatsächliche Konvergenz im irreduziblen Falle unseres Problems zu beweisen, behandeln wir ein allgemeineres Iterationsproblem, dessen Voraussetzungen, wie leicht zu zeigen, in unserem Falle zutreffen.

454 | Theorem. *Es sei gegeben eine Substitution von n Variablen x_r in der Form*

$$x'_r = f_r(x) \equiv f_r(x_1, x_2, \ldots, x_n) \quad (r = 1, 2, \ldots, n) \,, \tag{35}$$

wo die (reellen) Funktionen f_r folgende Eigenschaften besitzen:

1. *Die Funktionen f_r sind* homogen *von der Dimension* 1, *d. h. es ist für jedes (positive) λ stets*

$$f_r(\lambda x) = \lambda f_r(x) \,. \tag{36}$$

2. *Die Funktionen sind „positiv beschränkt", d. h. ihre Funktionswerte liegen zwischen* positiven (*von* 0 *verschiedenen*) *endlichen Schranken in jedem Variabilitätsbereiche, in welchem die sämtlichen Variablen x_r selbst zwischen positiven endlichen Schranken gelegen sind.*

3. *Die Funktionen f_r sind* stetig differentiierbar *und auch ihre sämtlichen* partiellen Ableitungen $f_{rs}(x) = \dfrac{\partial}{\partial x_s} f_r(x)$ *sind „positiv beschränkte" Funktionen, soweit die Variablen x_s „wirklich vorkommen", d. h. die Ableitungen f_{rs} nicht identisch verschwinden. Die Funktionen f_r sind also in den wirklich vorkommenden Variablen x_s auch* monoton.

4. *Jede der Funktionen f_r enthält „eigentlich" die ihr entsprechende Variable x_r, so daß $f_{rr} \neq 0$ ist in jedem positiven Bereiche der Veränderlichen; und bei* jeder *Einteilung der Zahlen $1, 2, \ldots, n$ in zwei Klassen (r, s) gibt es immer mindestens ein Zahlenpaar r, s, wo r der einen und s der anderen angehört, für welches $f_{rs} \neq 0$ ist.*

5. *Die Gleichungen*

$$u_r = f_r(u) \qquad (r = 1, 2, \ldots, n) \tag{37}$$

besitzen mindestens eine positive *Lösung $u_r > 0$.*

Dann konvergieren für jedes positive System von Werten $x_r > 0$ die sukzessive gebildeten Werte

$$x'_r = f_r(x),\ x''_r = f_r(x'), \ldots, x_r^{(i+1)} = f_r\big(x^{(i)}\big), \ldots \tag{35$_i$}$$

mit wachsendem i nach positiven Grenzwerten

$$\overline{x}_r = \lim_{i \to \infty} x_r^{(i)} \qquad (r = 1, 2, \ldots, n) \,, \tag{38}$$

welche dem Gleichungssystem (37) *genügen und sich von den vorausgesetzten Lösungen u_r höchstens durch einen gemeinsamen positiven Faktor unterscheiden*

$$\overline{x}_r = \lambda^* u_r \,, \tag{39}$$

wodurch zugleich die Eindeutigkeit *der positiven Lösung gesichert ist.*

In order to prove the actual convergence in the irreducible case of our problem we examine a more general problem of iteration whose assumptions apply to our case, as can be readily shown.

Theorem. *Suppose there is a substitution of n variables* x_r *in the form*

$$x'_r = f_r(x) \equiv f_r(x_1, x_2, \ldots, x_n) \quad (r = 1, 2, \ldots, n) \,, \tag{35}$$

where the (*real*) *functions* f_r *possess the following properties*:

1. *The functions* f_r *are* homogeneous *of dimension* 1, *i. e., for every* (*positive*) λ *always*

$$f_r(\lambda x) = \lambda f_r(x) \,. \tag{36}$$

2. *The functions are "'positively bounded", i. e., their function values lie between* positive *finite bounds* (*different from* 0) *in every domain of variability in which the variables* x_r *themselves all lie between positive finite bounds.*

3. *The functions* f_r *are* continuously differentiable *and their* partial derivatives $f_{rs}(x) = \dfrac{\partial}{\partial x_s} f_r(x)$, *too, are all "positively bounded" functions, provided that the variables* x_s *"are really occurring", i. e., provided that the derivatives* f_{rs} *do not vanish identically. Hence, the functions* f_r *are also* monotonous *in the really occurring variables* x_s.

4. *Each of the functions* f_r *"actually" contains the variable* x_r *corresponding to it so that* $f_{rr} \neq 0$ *in every positive domain of the variable; and for* every *partition of the numbers* $1, 2, \ldots, n$ *into two classes* (r, s) *there is always at least one pair of numbers* r, s *where* r *belongs to the one and* s *to the other for which* $f_{rs} \neq 0$.

5. *The equations*

$$u_r = f_r(u) \qquad (r = 1, 2, \ldots, n) \tag{37}$$

have at least one positive *solution* $u_r > 0$.

Then for each positive system of values $x_r > 0$ *the successively formed values*

$$x'_r = f_r(x), \; x''_r = f_r(x'), \ldots, x_r^{(i+1)} = f_r\big(x^{(i)}\big), \ldots \tag{35}_i$$

converge with increasing i *to positive limits*

$$\overline{x}_r = \lim_{i \to \infty} x_r^{(i)} \qquad (r = 1, 2, \ldots, n) \,, \tag{38}$$

that satisfy the system of equations (37) *and that differ from the assumed solutions* u_r *at most by a common positive factor*

$$\overline{x}_r = \lambda^* u_r \,, \tag{39}$$

whereby, at the same time, the uniqueness *of the positive solution is secured.*

Bevor wir dieses allgemeine Theorem beweisen, überzeugen wir uns zunächst, daß die über die Funktionen f_r gemachten Voraussetzungen für den von uns behandelten speziellen Fall auch erfüllt sind.

455 | 1. Die Homogeneität erhellt unmittelbar aus der Form der Gleichungen (8).

2. Im „irreduziblen“ Falle sind nach § 1 alle $g_r > 0$ und für jedes r mindestens ein $k_{rt} > 0$, und die Funktionen f_r sind also nicht verschwindende rationale Funktionen mit lauter *positiven* Koeffizienten. Solche Funktionen sind aber immer „positiv beschränkt“, weil, wie leicht ersichtlich, die Summe, das Produkt und der Quotient „positiv beschränkter“ ⟦Funktionen⟧ stets wieder die gleiche Eigenschaft besitzen. Allgemein sind weiter alle positiv beschränkten Funktionen von positiv beschränkten Funktionen wieder positiv beschränkt.

3. Durch partielle Differentiation ergibt sich aus (33)

$$g_r \frac{\partial f_r(x)}{\partial x_r} = f_r(x)^2 \sum_t{}' \frac{k_{rt}}{(x_r + x_t)^2} \tag{40}$$

und

$$g_r \frac{\partial f_r(x)}{\partial x_s} = f_r(x)^2 \frac{k_{rs}}{(x_r + x_s)^2} \quad \text{für} \quad s \neq r\,; \tag{41}$$

also sind auch hier alle $f_{rs}(x)$, soweit sie nicht identisch verschwinden, stetige und „positiv beschränkte“ Funktionen.

4. Aus den letzten Formeln ergibt sich unmittelbar, daß stets $f_{rr}(x) > 0$ und $f_{rs}(x) = 0$ nur im Falle $k_{rs} = 0$, wo der r-te Spieler mit dem s-ten überhaupt nicht gespielt hat. Gäbe es nun eine Klasseneinteilung in Spieler A_r und A_s, bei welcher alle $k_{rs} = 0$ wären, so wären auch alle $g_{rs} = 0$, was im „irreduziblen“ Falle ausgeschlossen ist.

5. Die *Existenz* einer positiven Lösung des Gleichungssystems ist für den irreduziblen Fall durch unsere Betrachtungen im § 1 mit der Existenz des Maximums von Φ nach dem Weierstraßschen Satze gesichert. Dagegen braucht die *Eindeutigkeit* der Lösung hier *nicht* vorausgesetzt zu werden, sie wird vielmehr — unabhängig von unserem in § 2 gegebenen Beweise — als spezielle Folgerung des allgemeinen Satzes gleichzeitig *mitbewiesen*.

Beweis des Konvergenztheorems. Für $t = 1, 2, \ldots, n$ sei $x_t > 0$ ein beliebiges positives Wertsystem und $u_t > 0$ eine positive Lösung der Gleichungen (37). Bezeichnen wir hier mit λ den kleinsten und mit μ den größten Quotienten aus der Reihe $\dfrac{x_t}{u_t}$,

$$\lambda u_t \leqq x_t \leqq \mu u_t \qquad (t = 1, 2, \ldots, n)\,, \tag{42}$$

so folgt aus 3., da die Funktionen f_r monoton sind,

$$\lambda f_r(u) \leqq f_r(x) \leqq \mu f_r(u) \qquad (r = 1, 2, \ldots, n) \tag{43}$$

Before proving this general theorem we first make sure that the assumptions made about the functions f_r are indeed satisfied in our special case.

1. The homogeneity is immediately evident from the form of the equations (8).

2. In the "irreducible" case all g_r's are, by § 1, > 0 and for each r there is at least one $k_{rt} > 0$, and hence the functions f_r are non-vanishing rational functions all of whose coefficients are *positive*. But such functions are always "positively bounded" since, as can be readily seen, the sum, the product and the quotient of "positively bounded" ⟦functions⟧ always possess the same property in turn. Furthermore, all positively bounded functions of positively bounded functions are, in general, positively bounded in turn.

3. By partial differentiation we obtain from (33)

$$g_r \frac{\partial f_r(x)}{\partial x_r} = f_r(x)^2 \sum_t{}' \frac{k_{rt}}{(x_r + x_t)^2} \tag{40}$$

and

$$g_r \frac{\partial f_r(x)}{\partial x_s} = f_r(x)^2 \frac{k_{rs}}{(x_r + x_s)^2} \quad \text{for} \quad s \neq r\,; \tag{41}$$

hence here, too, all $f_{rs}(x)$'s are continuous and "positively bounded" functions, provided they do not vanish identically.

4. The last formulas immediately lead to the fact that $f_{rr}(x)$ is always > 0 and that $f_{rs}(x) = 0$ only in the case $k_{rs} = 0$, where the r-th player has never played the s-th player. Now, if there existed a partition into players A_r and A_s for which all k_{rs}'s were $= 0$, then all g_{rs}'s would also be $= 0$, which is impossible in the "irreducible" case.

5. The *existence* of a positive solution of the system of equations is secured for the irreducible case according to our considerations in § 1 by the existence of the maximum of Φ in accordance with *Weierstrass*'s theorem. By contrast, the *uniqueness* of the solution does *not* have to be assumed here. Rather, it is simultaneously *proved along with* the general theorem as a special corollary—independently of our proof given in § 2.

Proof of the convergence theorem. For $t = 1, 2, \ldots, n$ let $x_t > 0$ be any positive system of values and $u_t > 0$ a positive solution of equations (37). If we denote by λ the least, and by μ the greatest, quotient from the series $\frac{x_t}{u_t}$,

$$\lambda u_t \leqq x_t \leqq \mu u_t \qquad (t = 1, 2, \ldots, n)\,, \tag{42}$$

it then follows from 3. that

$$\lambda f_r(u) \leqq f_r(x) \leqq \mu f_r(u) \qquad (r = 1, 2, \ldots, n) \tag{43}$$

und wegen (37) auch

$$\lambda u_r \leqq x'_r \leqq \mu u_t \ ,$$

456 | sowie für alle folgenden Approximationen $x_t^{(i)}$

$$\lambda u_t \leqq x_t^{(i)} \leqq \mu u_t \qquad (t = 1, 2, \ldots, n) \ . \qquad (42)_i$$

Ebenso ist aber auch für jedes i, wenn wieder $\lambda^{(i)}$ den kleinsten und $\mu^{(i)}$ den größten Quotienten $\dfrac{x_t^{(i)}}{u_t}$ bezeichnet,

$$\lambda^{(i)} u_t \leqq x_t^{(i)} \leqq \mu^{(i)} u_t$$

und

$$\lambda^{(i)} u_t \leqq x_t^{(i+1)} \leqq \mu^{(i)} u_t \ ,$$

also

$$\lambda^{(i)} \leqq \lambda^{(i+1)} \leqq \mu^{(i+1)} \leqq \mu^{(i)} \ ,$$

d. h. die Werte $\lambda^{(i)}$ bilden eine niemals fallende, die $\mu^{(i)}$ eine niemals steigende Reihe, wobei die Intervalle $(\lambda^{(i)}, \mu^{(i)})$ sukzessive ineinander liegen und alle im Gesamtintervall (λ, μ) enthalten sind. Die $\lambda^{(i)}$ besitzen also eine *obere* Grenze $\overline{\lambda}$, die $\mu^{(i)}$ eine *untere* Grenze $\overline{\mu}$, und es ist

$$\lambda \leqq \lambda^{(i)} \leqq \overline{\lambda} \leqq \overline{\mu} \leqq \mu^{(i)} \leqq \mu \quad \text{für} \quad i = 1, 2, 3, \ldots \qquad (43)$$

sowie

$$\overline{\lambda} = \lim_{i \to \infty} \lambda^{(i)} \ , \quad \overline{\mu} = \lim_{i \to \infty} \mu^{(i)} \ . \qquad (44)$$

Läßt sich nun zeigen, daß beide Grenzwerte übereinstimmen,

$$\overline{\lambda} = \overline{\mu} = \varrho \ , \qquad (45)$$

so wird auch

$$\lim_{i \to \infty} x_t^{(i)} = \varrho u_t \qquad (t = 1, 2, \ldots, n) \ , \qquad (46)$$

wie zu beweisen.

Um aber diesen Nachweis zu führen, müssen wir zwei Fälle unterscheiden:

Fall I. Alle $f_{rs} > 0$, d. h. beim Turnierproblem $k_{rs} > 0$ für alle $r \neq s$: jeder Spieler hat mit jedem andern gespielt. Dann sind nach 3. auch sämtliche Ableitungen f_{rs} „positiv beschränkt“, und für alle den Bedingungen

$$\lambda u_t \leqq x_t \leqq \mu u_t \qquad (t = 1, 2, \ldots, n) \qquad (42)$$

since the functions f_r are monotonic, and, on account of (37), also

$$\lambda u_r \leqq x'_r \leqq \mu u_t \ ,$$

and also for all subsequent approximations $x_t^{(i)}$

$$\lambda u_t \leqq x_t^{(i)} \leqq \mu u_t \qquad (t = 1, 2, \ldots, n) \ . \qquad (42)_i$$

But likewise, if $\lambda^{(i)}$ again denotes the least, and $\mu^{(i)}$ the greatest, quotient $\frac{x_t^{(i)}}{u_t}$, for each i

$$\lambda^{(i)} u_t \leqq x_t^{(i)} \leqq \mu^{(i)} u_t$$

and

$$\lambda^{(i)} u_t \leqq x_t^{(i+1)} \leqq \mu^{(i)} u_t \ ,$$

hence

$$\lambda^{(i)} \leqq \lambda^{(i+1)} \leqq \mu^{(i+1)} \leqq \mu^{(i)} \ ,$$

i. e., the values $\lambda^{(i)}$ form a never-decreasing, the $\mu^{(i)}$'s a never-increasing, series, where the intervals $(\lambda^{(i)}, \mu^{(i)})$ are successively nested and all contained in the total interval (λ, μ). Thus, the $\lambda^{(i)}$'s have an *upper* limit $\overline{\lambda}$, the $\mu^{(i)}$'s a *lower* limit $\overline{\mu}$, and

$$\lambda \leqq \lambda^{(i)} \leqq \overline{\lambda} \leqq \overline{\mu} \leqq \mu^{(i)} \leqq \mu \quad \text{for} \quad i = 1, 2, 3, \ldots \qquad (43)$$

and also

$$\overline{\lambda} = \lim_{i \to \infty} \lambda^{(i)} \ , \quad \overline{\mu} = \lim_{i \to \infty} \mu^{(i)} \ . \qquad (44)$$

Now, if we can show that the two limits correspond,

$$\overline{\lambda} = \overline{\mu} = \varrho \ , \qquad (45)$$

then also

$$\lim_{i \to \infty} x_t^{(i)} = \varrho u_t \qquad (t = 1, 2, \ldots, n) \ , \qquad (46)$$

which is what is to be proved.

But in order to conduct this proof we need to distinguish two cases:

Case I. All $f_{rs} > 0$, i. e., for the tournament problem $k_{rs} > 0$ for all $r \neq s$: each player has played every ⟦other⟧ player. Then, by 3., all derivatives f_{rs} are "positively bounded", and for all systems of values x_t satisfying the conditions

$$\lambda u_t \leqq x_t \leqq \mu u_t \qquad (t = 1, 2, \ldots, n) \ , \qquad (42)$$

genügenden Wertsysteme x_t, zu denen nach $(42)_i$ auch alle $x_t^{(i)}$ gehören, besitzen sie positive untere Schranken:

$$f_{rs}(x) \geqq \gamma_{rs} > 0\,. \tag{47}$$

Nach dem Mittelwertsatze ist aber wegen (37)

$$x_r^{(i+1)} - \lambda^{(i)} u_r = f_r\big(x^{(i)}\big) - f_r\big(\lambda^{(i)} u\big) = \sum_{t=1}^{n} f_{rt}(\tilde{x})\big(x_t^{(i)} - \lambda^{(i)} u_t\big)\,, \tag{48}$$

457 | wo

$$\tilde{x}_t = (1-\vartheta_r)\,\lambda^{(i)} u_t + \vartheta_r x_t^{(i)} \qquad (0 \leqq \vartheta_r \leqq 1;\ t = 1, 2, \ldots, n)$$

einen Punkt auf der Verbindungsstrecke zwischen den Punkten $\big(x_t^{(i)}\big)$ und $\big(\lambda^{(i)} u_t\big)$ kennzeichnet, der gleichfalls den Beziehungen (42) genügt. Daher ist nach (47) für jedes $s \leqq n$

$$x_r^{(i+1)} - \lambda^{(i)} u_r \geqq \gamma_{rs}\big(x_s^{(i)} - \lambda^{(i)} u_s\big)\,. \tag{49}$$

Insbesondere gilt dies auch für die Maximal- bzw. Minimalwerte der Quotienten, wo

$$x_r^{(i+1)} = \lambda^{(i+1)} u_r\,, \quad x_s^{(i)} = \mu^{(i)} u_s\,,$$

so daß wir erhalten

$$\lambda^{(i+1)} - \lambda^{(i)} \geqq \gamma_{rs}\frac{u_s}{u_r}\big(\mu^{(i)} - \lambda^{(i)}\big) \geqq \varkappa\,\big(\mu^{(i)} - \lambda^{(i)}\big)\,, \tag{50}$$

wenn $\varkappa$ den kleinsten der positiven Werte $\varkappa_{rs} = \gamma_{rs}\dfrac{u_s}{u_r}$ bedeutet, woraus folgt:

$$\mu^{(i)} - \lambda^{(i)} \leqq \tfrac{1}{\varkappa}\big(\lambda^{(i+1)} - \lambda^{(i)}\big) \quad \text{und}$$

$$\overline{\mu} - \lambda = \lim_{i \to \infty}\big(\mu^{(i)} - \lambda^{(i)}\big) = 0\,, \tag{51}$$

w. z. b. w.

Um nun auch den Grad der Konvergenz abzuschätzen, bedienen wir uns noch einer zweiten, ganz analog wie (50) aus (47) abgeleiteten Formel: Es ist

$$\mu^{(i)} u_r - x_r^{(i+1)} = f_r\big(\mu^{(i)} u\big) - f_r\big(x^{(i)}\big) = \sum_{t=1}^{n} f_{rt}(\tilde{x})\big(\mu^{(i)} u_t - x_t^{(i)}\big)$$

$$\geqq \gamma_{rs}\big(\mu^{(i)} u_s - x_s^{(i)}\big) \quad \text{für beliebige } r, s,$$

also für $x_r^{(i+1)} = \mu^{(i+1)} u_r$, $x_s^{(i)} = \lambda^{(i)} u_s$ weiter

$$\mu^{(i)} - \mu^{(i+1)} \geqq \varkappa\,\big(\mu^{(i)} - \lambda^{(i)}\big)\,. \tag{52}$$

to which, by $(42)_i$, also belong all $x_t^{(i)}$'s, they have positive lower bounds

$$f_{rs}(x) \geqq \gamma_{rs} > 0 \,. \tag{47}$$

But according to the mean value theorem we have on account of (37)

$$x_r^{(i+1)} - \lambda^{(i)} u_r = f_r\big(x^{(i)}\big) - f_r\big(\lambda^{(i)} u\big) = \sum_{t=1}^{n} f_{rt}(\tilde{x})\big(x_t^{(i)} - \lambda^{(i)} u_t\big) \,, \tag{48}$$

where

$$\tilde{x}_t = (1 - \vartheta_r)\,\lambda^{(i)} u_t + \vartheta_r x_t^{(i)} \qquad (0 \leqq \vartheta_r \leqq 1;\ t = 1, 2, \ldots, n)$$

denotes a point along the line joining the points $\big(x_t^{(i)}\big)$ and $\big(\lambda^{(i)} u_t\big)$ that also satisfies conditions (42). Hence, by (47), for each $s \leqq n$

$$x_r^{(i+1)} - \lambda^{(i)} u_r \geqq \gamma_{rs}\big(x_s^{(i)} - \lambda^{(i)} u_s\big) \,. \tag{49}$$

This holds in particular for the maximal and minimal values of the quotients where

$$x_r^{(i+1)} = \lambda^{(i+1)} u_r \,, \quad x_s^{(i)} = \mu^{(i)} u_s \,,$$

so that we obtain

$$\lambda^{(i+1)} - \lambda^{(i)} \geqq \gamma_{rs} \frac{u_s}{u_r}\big(\mu^{(i)} - \lambda^{(i)}\big) \geqq \varkappa\,\big(\mu^{(i)} - \lambda^{(i)}\big) \,, \tag{50}$$

if $\varkappa$ denotes the least of the positive values $\varkappa_{rs} = \gamma_{rs}\dfrac{u_s}{u_r}$, from which follows

$$\begin{aligned} \mu^{(i)} - \lambda^{(i)} &\leqq \tfrac{1}{\varkappa}\big(\lambda^{(i+1)} - \lambda^{(i)}\big) \quad \text{and} \\ \overline{\mu} - \lambda &= \lim_{i \to \infty} \big(\mu^{(i)} - \lambda^{(i)}\big) = 0 \,, \end{aligned} \tag{51}$$

q. e. d.

Now in order to estimate also the degree of convergence we use yet another formula derived entirely analogously to (50) from (47): We have

$$\begin{aligned} \mu^{(i)} u_r - x_r^{(i+1)} &= f_r\big(\mu^{(i)} u\big) - f_r\big(x^{(i)}\big) = \sum_{t=1}^{n} f_{rt}(\tilde{x})\big(\mu^{(i)} u_t - x_t^{(i)}\big) \\ &\geqq \gamma_{rs}\big(\mu^{(i)} u_s - x_s^{(i)}\big) \quad \text{for any } r, s, \end{aligned}$$

and hence for $x_r^{(i+1)} = \mu^{(i+1)} u_r$, $x_s^{(i)} = \lambda^{(i)} u_s$ furthermore

$$\mu^{(i)} - \mu^{(i+1)} \geqq \varkappa\,\big(\mu^{(i)} - \lambda^{(i)}\big) \,. \tag{52}$$

Durch Addition von (50) und (52) erhält man dann

$$\mu^{(i)} - \lambda^{(i)} - \big(\mu^{(i+1)} - \lambda^{(i+1)}\big) \geqq 2\varkappa\,\big(\mu^{(i)} - \lambda^{(i)}\big)\ .$$

Also

$$\mu^{(i+1)} - \lambda^{(i+1)} \leqq (1 - 2\varkappa)\big(\mu^{(i)} - \lambda^{(i)}\big) \tag{53}$$

und somit

$$\mu^{(i)} - \lambda^{(i)} \leqq (\mu - \lambda)(1 - 2\varkappa)^i\ . \tag{54}$$

Die Approximationen konvergieren also mindestens wie eine geometrische Reihe.

Fall II. Nicht alle f_{rs} sind $\neq 0$, nicht alle f_r enthalten „eigentlich“ sämtli-
458 che Variablen, dagegen gilt die Eigenschaft 4. Dieser Fall läßt sich | nun, wie wir zeigen wollen, auf den Fall I zurückführen, indem man die Substitutionen (35) iteriert und setzt:

$$\begin{aligned}
f_r^{(1)}(x) &\equiv f_r(f) && \equiv f_r\big(f_1(x), f_2(x) && \ldots f_n(x)\big)\\
f_r^{(2)}(x) &\equiv f_r(f^{(1)}) && \equiv f_r\big(f_1^{(1)}(x), f_2^{(1)}(x) && \ldots f_n^{(1)}(x)\big)\\
&\ .\ \ .\ \ . && .\ \ .\ \ .\ \ .\ && .\ \ .\ \ .\\
f_r^{(i+1)}(x) &\equiv f_r(f^{(i)}) && \equiv f_r\big(f_1^{(i)}(x) && \ldots f_n^{(i)}(x)\big)
\end{aligned} \tag{55}$$

$$(r = 1, 2, \ldots, n;\ i = 1, 2, \ldots)\ .$$

Diese durch Iteration entstehenden Funktionen $f_r^{(i)}(x)$ besitzen nämlich, wie leicht ersichtlich, die gleichen Eigenschaften 1–3 wie die ursprünglichen $f_r(x)$, da sich diese Eigenschaften stets auf die zusammengesetzten Funktionen übertragen. Insbesondere ist auch für jede solche Funktion F in jedem positiv beschränkten Bereiche

$$\frac{\partial}{\partial x_r} F\,(f_1, f_2, \ldots) = \sum_t \frac{\partial F}{\partial f_t}\frac{\partial f_t}{\partial x_r} \geqq \frac{\partial F}{\partial f_s}\frac{\partial f_s}{\partial x_r} > 0\ , \tag{56}$$

wenn auch nur für *ein* s die nach f_s genommene Ableitung von F sowie die nach x_r genommene von f_s nicht verschwindet. Hieraus folgt zunächst, daß mit $f_{rr} \neq 0$ (vermöge 4) auch

$$f_{rr}^{(1)}(x) > 0\,,\ f_{rr}^{(2)}(x) > 0\,,\ \ldots,\ f_{rr}^{(i)}(x) > 0\,,$$

daß also sämtliche $f_r^{(i)}$ *monotone* Funktionen von x_r sind.

Angenommen nun, die der Reihe der Funktionen f_t entnommene Funktion $F(x) = F\,(x_1, x_2, \ldots, x_n)$ enthalte „eigentlich“ die mit x_r bezeichneten Variablen aus der Reihe $x_1 \ldots x_n$, so daß alle $\dfrac{\partial F}{\partial x_r} > 0$ sind, während für alle

By addition of (50) and (52) we then obtain

$$\mu^{(i)} - \lambda^{(i)} - \left(\mu^{(i+1)} - \lambda^{(i+1)}\right) \geqq 2\varkappa \left(\mu^{(i)} - \lambda^{(i)}\right) .$$

Thus

$$\mu^{(i+1)} - \lambda^{(i+1)} \leqq (1 - 2\varkappa)\left(\mu^{(i)} - \lambda^{(i)}\right) , \tag{53}$$

and hence

$$\mu^{(i)} - \lambda^{(i)} \leqq (\mu - \lambda)(1 - 2\varkappa)^i . \tag{54}$$

The approximations thus converge at least like a geometric series.

Case II. Not all of the f_{rs}'s are $\neq 0$, not all of the f_r's "really" contain all of the variables. By contrast, property 4. holds. Now it is possible, as we will show, to reduce this case to case I by iterating the substitutions (35) and by setting

$$\begin{aligned}
f_r^{(1)}(x) &\equiv f_r(f) && \equiv f_r\big(f_1(x), f_2(x) && \ldots f_n(x)\big) \\
f_r^{(2)}(x) &\equiv f_r(f^{(1)}) && \equiv f_r\big(f_1^{(1)}(x), f_2^{(1)}(x) && \ldots f_n^{(1)}(x)\big) \\
&\ \ .\quad .\quad . && \ \ .\quad .\quad .\quad . && .\quad .\quad . \\
f_r^{(i+1)}(x) &\equiv f_r(f^{(i)}) && \equiv f_r\big(f_1^{(i)}(x) && \ldots f_n^{(i)}(x)\big)
\end{aligned} \tag{55}$$

$$(r = 1, 2, \ldots, n;\ i = 1, 2, \ldots) .$$

For these functions $f_r^{(i)}(x)$ obtained by iteration possess, as is readily seen, the same properties 1.–3. as the original $f_r(x)$, since these properties are always passed on to the composite functions. In particular, for each such function F in every positively bounded domain

$$\frac{\partial}{\partial x_r} F\,(f_1, f_2, \ldots) = \sum_t \frac{\partial F}{\partial f_t}\frac{\partial f_t}{\partial x_r} \geqq \frac{\partial F}{\partial f_s}\frac{\partial f_s}{\partial x_r} > 0 , \tag{56}$$

if for even only *one* s the derivative of F taken with respect to f_s and the derivative of f_s taken with respect to x_r do not vanish. From this it first follows that along with $f_{rr} \neq 0$ (by virtue of 4.) also

$$f_{rr}^{(1)}(x) > 0 ,\ f_{rr}^{(2)}(x) > 0 , \ldots ,\ f_{rr}^{(i)}(x) > 0 ,$$

and hence that all of the $f_r^{(i)}$'s are *monotonic* functions of x_r.

Suppose now that the function $F(x) = F\,(x_1, x_2, \ldots, x_n)$ taken from the series of functions f_t "really" contains the variables denoted by x_r from the series $x_1 \ldots x_n$ so that all $\dfrac{\partial F}{\partial x_r}$ are > 0, whereas for all remaining ones (de-

übrigen (hier mit x_s bezeichneten) immer $\dfrac{\partial F}{\partial x_s} = 0$ sei, so gibt es nach 4. ein Wertepaar r, s, für welches $f_{rs} \neq 0$ ist, und wir haben wegen (56) einmal

$$\frac{\partial}{\partial x_r} F(f) > 0 \quad \text{für alle } r\,, \text{ für die } \quad \frac{\partial F(x)}{\partial x_r} > 0 \quad \text{ist}\,, \tag{57}$$

sowie

$$\frac{\partial}{\partial x_s} F(f) \geqq \frac{\partial F}{\partial f_r} \frac{\partial f_r}{\partial x_s} > 0\,, \tag{58}$$

d. h. die Funktion $F(f) = F^{(1)}(x)$ enthält eigentlich *alle* x_r, welche $F(x)$ eigentlich enthält, und *außerdem* noch mindestens eine weitere Variable x_s. Somit enthält in der Reihe

$$F(x)\,,\ F^{(1)}(x) = F(f)\,,\ F^{(2)}(x) = F^{(1)}(f)\,,\ \ldots,\ F^{(i+1)}(x) = F^{(i)}(f)$$

459 | bereits $F^{(n-1)}(x)$ sämtliche n Variablen $x_1, \ldots, x_n$, und insbesondere gibt es auch eine Zahl $h \leqq n-1$ von der Beschaffenheit, daß die n Funktionen

$$f_1^{(h)}(x),\ f_2^{(h)}(x), \ldots,\ f_n^{(h)}(x)\,,$$

welche durch h-fache Iteration der Substitution (35) entstehen, die sämtlichen Variablen $x_1, \ldots, x_n$ eigentlich enthalten. Diese Funktionen $f_r^{(h)}$ genügen den sämtlichen Bedingungen 1 bis 5, da ersichtlich auch für jedes positive Lösungssystem (u) von (37)

$$u_r = f_r^{(h)}(u) \qquad (r = 1, 2, \ldots, n) \tag{$37)_h$}$$

ist, und damit sind alle Voraussetzungen der Konvergenz im Falle I gegeben. Es werden also für ein beliebiges positives System (x) von Anfangswerten die Iterationswerte[9]

$$x_r^{(h)} = f_r^{(h)}(x),\ x_r^{(2h)} = f_r^{(2h)}(x), \ldots,\ x_r^{(ih)} = f_r^{(ih)}(x)$$

mit wachsendem i einer Lösung von (37) zustreben, indem die entsprechenden Differenzen $\mu^{(ih)} - \lambda^{(ih)}$ sich der 0 nähern, und somit gilt auch allgemein

$$\overline{\mu} - \lambda = \lim_{j=\infty} \left(\mu^{(j)} - \lambda^{(j)}\right) = 0\,,$$

d. h. auch in dem durch die Bedingungen 1 bis 5 gegebenen *allgemeinen* Falle führt unser Verfahren der sukzessiven Approximation zur numerischen Lösung des irreduziblen Problems. Der Unterschied der beiden Fälle I und II macht sich nur insofern geltend, daß im Falle II in der Folge der Approximationen $x^{(i)}$ gelegentlich auch

$$\mu^{(i+1)} - \lambda^{(i+1)} = \mu^{(i)} - \lambda^{(i)}$$

[9] Es wird hier von der Gültigkeit des Assoziationsgesetzes für die Zusammensetzung von Substitutionen Gebrauch gemacht.

noted here by x_s) always $\dfrac{\partial F}{\partial x_s} = 0$. Then, by 4., there is a pair of values r, s for which $f_{rs} \neq 0$, and we have, on account of (56),

$$\frac{\partial}{\partial x_r} F(f) > 0 \quad \text{for all } r\,, \text{ for which } \quad \frac{\partial F(x)}{\partial x_r} \quad \text{is} > 0\,, \tag{57}$$

and also

$$\frac{\partial}{\partial x_s} F(f) \geqq \frac{\partial F}{\partial f_r} \frac{\partial f_r}{\partial x_s} > 0\,, \tag{58}$$

i. e., the function $F(f) = F^{(1)}(x)$ really contains *all* the x_r's that $F(x)$ really contains, and *in addition* also at least one further variable x_s. Thus, in the series

$$F(x)\,,\ F^{(1)}(x) = F(f)\,,\ F^{(2)}(x) = F^{(1)}(f)\,,\ \ldots,\ F^{(i+1)}(x) = F^{(i)}(f)\,,$$

$F^{(n-1)}(x)$ already contains all of the n variables $x_1, \ldots, x_n$, and in particular there is also a number $h \leqq n-1$ so constituted that the n functions

$$f_1^{(h)}(x),\ f_2^{(h)}(x), \ldots,\ f_n^{(h)}(x)\,,$$

which arise through h-fold iteration of substitution (35), really contain all of the variables $x_1, \ldots, x_n$. These functions $f_r^{(h)}$ satisfy all of the conditions from 1. to 5., since, as is readily seen, also for every positive system of solutions (u) from (37)

$$u_r = f_r^{(h)}(u) \qquad (r = 1, 2, \ldots, n)\,, \tag{37}_h$$

and hence all presuppositions for the convergence in case I are satisfied. Thus, for any positive system (x) of initial values, the iteration values[12]

$$x_r^{(h)} = f_r^{(h)}(x),\ x_r^{(2h)} = f_r^{(2h)}(x), \ldots,\ x_r^{(ih)} = f_r^{(ih)}(x)$$

converge to a solution of (37) as i increases, while the corresponding differences $\mu^{(ih)} - \lambda^{(ih)}$ approach 0, and hence also in general[13]

$$\overline{\mu} - \overline{\lambda} = \lim_{j=\infty} \left(\mu^{(j)} - \lambda^{(j)}\right) = 0\,,$$

i. e., our method of successive approximation leads to the numeric solution of the irreducible problem also in the *general* case given by the conditions from 1. to 5. The difference between the two cases I and II is relevant only insofar as in case II

$$\mu^{(i+1)} - \lambda^{(i+1)} = \mu^{(i)} - \lambda^{(i)}$$

[12] Here we use the validity of the law of associativity for the composition of substitutions.

[13] [[In the following formula, Zermelo erroneously writes "λ" instead of "$\overline{\lambda}$".]]

ausfallen kann, während doch immer, wie aus unserer Reduktion vermöge $f_r^{(h)}$ hervorgeht,

$$\mu^{(i+h)} - \lambda^{(i+h)} < \mu^{(i)} - \lambda^{(i)}$$

sein muß.

Unsere Auflösungsmethode ist also, weit entfernt, nur auf „reguläre“ Turniere (wo $k_{rt} = k$ konstant ist) beschränkt zu sein, wie meines Wissens alle bisherigen Methoden der Turnierberechnung, im „irreduziblen“ Falle auch ohne weiteres auf „abgebrochene“ und auf „kombinierte“ Turniere (die aus verschiedenen Einzelturnieren mit teilweise identischen Spielern zusammengesetzt sind) anwendbar und liefert auch in „reduziblen“ Fällen stets eine das vorliegende Beobachtungsmaterial gleichmäßig berücksichtigende und allen vernünftigen Forderungen genügende vollständige oder teilweise Bestimmung der relativen Spielstärken.

460 | Bei der praktischen Ausführung der Rechnung wird man in der Regel mit dem Ansatz $x_1 = x_2 = \ldots = x_n = 1$ beginnen und hieraus mit Hilfe von Reziprokentafeln gemäß (33) die sukzessiven Annäherungen $x_r', x_r'', \ldots$ bestimmen und so lange damit fortfahren, bis je zwei aufeinanderfolgende sich nicht mehr merklich unterscheiden, das Gleichungssystem (8) also erfüllt ist. Vorher muß man sich aber erst von der „Irreduzibilität“ des Turniers überzeugt bzw. die erforderliche Zerlegung in „Primturniere“ vorgenommen haben, was im „regulären“ Falle ($k_{rt} = k$) am bequemsten geschieht nach dem am Schluß von § 4 angegebenen Verfahren, d. h. durch Anordnung der g_r nach aufsteigender Größe und Vergleichung ihrer Teilsummen G_r mit den entsprechenden Zahlen $\frac{1}{2}kr(r-1)$ gemäß (32).

Als Beispiel für die Auswertung eines größeren (regulären) Turniers gebe ich zum Schluß eine kurze Tabelle, die nach dem New-Yorker Meisterturnier von 1924 für $n = 11$ Teilnehmer (darunter E. Lasker als erster Preisträger) nach dem angegebenen Verfahren berechnet wurde, wobei die berechneten „Spielstärken“ u_r auf die Summe 100 normiert sind und γ_r wegen $k = 2$ die Differenz $G_r - r(r-1)$ bedeutet.

r	u_r	g_r	G_r	$r(r-1)$	γ_r
1	2,43	5	5	0	5
2	3,44	6,5	11,5	2	9,5
3	3,84	7	18,5	6	12,5
4	4,72	8	26,5	12	14,5
5	6,40	9,5	36	20	16
6	7,12	10	46	30	16
7	7,84	10,5	56,5	42	14,5
8	8,71	11	67,5	56	11,5
9	10,7	12	79,5	72	7,5
10	18,4	14,5	94	90	4
11	26,4	16	110	110	0

(Eingegangen am 28. Mai 1928.)

may occasionally drop out in the sequence of approximations $x^{(i)}$, whereas always necessarily

$$\mu^{(i+h)} - \lambda^{(i+h)} < \mu^{(i)} - \lambda^{(i)} ,$$

as follows from our reduction by virtue of $f_r^{(h)}$.

Our method of resolution is thus far from being restricted only to "regular" tournaments (where $k_{rt} = k$ is constant), as all previous methods of the calculation of tournaments have been, at least to my knowledge. In the "irreducible" case it is also applicable without further ado to "aborted" and to "combined" tournaments (composed of different individual tournaments with in part identical players) and in the "reducible" cases it also always yields a complete or partial determination of the relative playing strengths that pays due attention to the available observational data and meets all reasonable requirements.

When actually carrying out the calculation one usually begins with the ansatz $x_1 = x_2 = \ldots = x_n = 1$ on the basis of which one determines by means of reciprocal tables in accordance with (33) the successive approximations $x_r', x_r'', \ldots$. One continues with this until any two consecutive approximations no longer differ noticeably from one another, and hence the system of equations (8) is satisfied. But first one must convince oneself of the "irreducibility" of the tournament or one must have carried out the required decomposition into "prime tournaments", which, in the "regular" case ($k_{rt} = k$), is easiest to do by using the method specified at the end of § 4, i. e., by arranging the g_r's according to increasing magnitude and by comparison of their partial sums G_r with the corresponding numbers $\frac{1}{2}kr(r-1)$ in accordance with (32).

I conclude with an example of the evaluation of a larger (regular) tournament: a short table that was calculated subsequent to the New York Masters' Tournament of 1924 for $n = 11$ participants (among them E. Lasker as the first prize winner) in accordance with the specified method, where the calculated "playing strengths" u_r are normed to the sum 100 and γ_r denotes the difference $G_r - r(r-1)$ on account of $k = 2$.

r	u_r	g_r	G_r	$r(r-1)$	γ_r
1	2.43	5	5	0	5
2	3.44	6.5	11.5	2	9.5
3	3.84	7	18.5	6	12.5
4	4.72	8	26.5	12	14.5
5	6.40	9.5	36	20	16
6	7.12	10	46	30	16
7	7.84	10.5	56.5	42	14.5
8	8.71	11	67.5	56	11.5
9	10.7	12	79.5	72	7.5
10	18.4	14.5	94	90	4
11	26.4	16	110	110	0

(Received May 28, 1928.)

Introductory note to 1930c and 1931a

Craig G. Fraser

According to Heinz-Dieter Ebbinghaus (*2007*, 150) the navigation problem was suggested to Zermelo by the circumnavigation of the globe by the airship "Graf Zeppelin" in 1929. The problem concerns a blimp or plane that moves with a given velocity relative to the air, travelling between two points on the earth. Because of the action of wind, the motion of the airship over land is modified. Suppose that the strength and direction of the wind are given as a function of position and time. The problem is to find the trajectory followed by the airship and the corresponding steering angle such that the airship completes its journey in the least time. Following the Hindenburg disaster of 1937, transportation by dirigibles or zeppelins became less common. In later formulations of the problem the airship was often replaced by a boat and the wind by current, and the problem was one of navigation along water.

In his two papers Zermelo formulated the problem mathematically as follows. (Our account is restricted to the two-dimensional case, although it should be noted that Zermelo extended his analysis to three dimensions.) The ship moves with velocity k relative to the surrounding medium, the latter being air in the case of an airship, and water in the case of a boat. The ship must go from A to B. There is a wind (in the case of a airship) or a current (in the case of a boat) that affects the motion of the ship. We consider an $x-y$ coordinate system with A at the origin. Assume at time t that the ship moves relative to the medium at an angle $\varphi = \varphi(t)$ to the x-axis. This "steering angle" is the direction in which we power the ship, knowing that the actual direction followed by the ship will be modified as the result of the action of the wind or current. The magnitude of the wind or current at time t and position (x, y) is given in terms of its x and y components, respectively $u(x, y, t)$ and $v(x, y, t)$. If (x, y) are the coordinates of the ship then the equations that describe its motion are

$$\begin{aligned} \frac{dx}{dt} &= u(x, y, t) + k\cos\varphi\,, \\ \frac{dy}{dt} &= v(x, y, t) + k\sin\varphi\,. \end{aligned} \tag{1}$$

It is necessary to find the track and steering angle for which the ship travels from A to B in least time. The partial derivatives of u with respect to x and y are denoted u_x and u_y respectively, with similar expressions for the partials of v. Zermelo derived a formula that describes this optimal motion in terms

H.-D. Ebbinghaus, A. Kanamori (Eds.), *Ernst Zermelo - Collected Works/ Gesammelte Werke II*, DOI 10.1007/978-3-540-70856-8_14, Schriften der Mathematisch-naturwissenschaftlichen Klasse der Heidelberger Akademie der Wissenschaften 23,

of the steering angle φ:

$$\frac{d\varphi}{dt} = v_x \sin^2 \varphi + (u_x - v_y) \cos \varphi \sin \varphi - u_y \cos^2 \varphi \,. \tag{2}$$

Zermelo's derivation of (2) was complicated and carried out directly from first principles for the special case of this problem. To the knowledge of the present writer no subsequent investigator adopted Zermelo's solution. Immediately following its appearance Tullio Levi-Civita (*1931*) published an article in which he stated, "Zermelo's direct and elegant treatment is very interesting." However, Levi-Civita did not make use of this solution and showed that the result follows from standard results in the calculus of variations. A few years later Carathéodory (*1935*, 378) asserted that the navigation problem "was posed by Zermelo and completely solved by an extraordinarily ingenious method" (translation from the English edition (*1967*) of Carathéodory's book). Carathéodory also did not adopt this method and followed Levi-Civita in treating the problem by means of standard methods in the calculus of variations. Carathéodory's investigation was very complete and remains today the most detailed analysis of the navigation problem. Among other things, Carathéodory examined the question of sufficiency and extended and systematized Zermelo's discussion of this point. Other mathematicians of the period who worked on the navigation problem included Basilio Manià (*1937*), and Magnus R. Hestenes (*1937*). The problem also attracted the attention of physicists, and was the subject of papers by Richard von Mises (*1931*) and Philipp Frank (*1933*). The latter researchers did not use the calculus of variations at all, but made the basis of their investigation an analogy between the motion of the airship and the passage of a light ray through a medium of variable refractive index.

The first of Zermelo's two papers is rather difficult to follow because it is really only an extended abstract and some details are omitted. Zermelo was motivated in part to continue work on the problem by an error that Levi-Civita found in the 1930 paper which he communicated to Zermelo in a letter. (See *Ebbinghaus 2007*, 151, where some other details of the genesis and reception of Zermelo's two papers may be found.) While the second paper provides a fuller account of the navigation problem, the analysis remains difficult to follow, requiring steps that are suited to the problem at hand but not otherwise part of a general theory. In his investigation of the problem Zermelo was returning to a branch of mathematics that he had worked on thirty years earlier. While he may have been somewhat out of touch with contemporary research in the subject, his formidable mathematical powers led him to interesting and useful results. Although his analysis was of a personal and even singular character, his papers are nonetheless worth studying and contemplating for their own sake and mathematical interest.

We now give a derivation of Zermelo's navigation formula using the idea that originated with Levi-Civita (*1931*). (Our derivation is based on the account given in *Funk 1962*, 282–284.) The following theorem was originally

derived by Adolph Mayer (*1895*) and was presented by Oskar Bolza in his authoritative textbook *Vorlesungen über die Variationsrechnung* (*1909*, 572–574). We have a variational problem in which x is the independent variable and y_0, y_1 and y_2 are the dependent variables. The derivative of y_i with respect to x is denoted y_i'. The variables x and y_0, y_1, y_2 are connected by an equation of the form

$$g(x, y_0, y_1, y_2, y_0', y_1', y_2') = 0\,, \tag{3}$$

on the interval $[x_1, x_2]$. The variables y_1 and y_2 have prescribed values at the endpoints x_1 and x_2. The variable y_0 has a prescribed value at x_1. It is supposed that the value of y_0 at x_2 is a maximum or a minimum. A necessary condition for this to be the case is that there exists a multiplier function $\lambda = \lambda(x)$ defined on $[x_1, x_2]$ such that the following Euler equations are valid:

$$\begin{aligned} \frac{\partial(\lambda g)}{\partial y_1} - \frac{d}{dx}\Big(\frac{\partial(\lambda g)}{\partial y_1'}\Big) &= 0\,, \\ \frac{\partial(\lambda g)}{\partial y_2} - \frac{d}{dx}\Big(\frac{\partial(\lambda g)}{\partial y_2'}\Big) &= 0\,. \end{aligned} \tag{4}$$

The variational problem as formulated is known as a Mayer problem. Equations (4) are derived from a general multiplier rule for problems with constraints in the form of differential equations.

Consider now Zermelo's navigation problem. Because the time t itself is a variable that must be minimized we express t as a function of an independent variable τ, where τ lies on the interval $[\tau_1, \tau_2]$. The variables x and y are now to be regarded as functions of τ. From (1) we immediately obtain the following equation:

$$\Big(\frac{dx}{dt} - u\Big)^2 + \Big(\frac{dy}{dt} - v\Big)^2 - k^2 = 0\,. \tag{5}$$

Expressed in terms of the independent variable τ (5) becomes

$$\Big(\frac{x'}{t'} - u\Big)^2 + \Big(\frac{y'}{t'} - v\Big)^2 - k^2 = 0\,, \tag{6}$$

where the prime notation indicates differentiation with respect to τ. We now apply Mayer's result to this problem. In the application of this result the variables τ, t, x, y take the place of x, y_0, y_1, y_2 above. Here

$$g(\tau, t, x, y, t', x', y') = \Big(\frac{x'}{t'} - u\Big)^2 + \Big(\frac{y'}{t'} - v\Big)^2 - k^2 = 0\,. \tag{7}$$

Then there exists a multiplier function $\lambda = \lambda(\tau)$ such that

$$\begin{aligned} \frac{\partial(\lambda g)}{\partial x} - \frac{d}{d\tau}\Big(\frac{\partial(\lambda g)}{\partial x'}\Big) &= 0\,, \\ \frac{\partial(\lambda g)}{\partial y} - \frac{d}{d\tau}\Big(\frac{\partial(\lambda g)}{\partial y'}\Big) &= 0\,. \end{aligned} \tag{8}$$

These equations reduce to

$$\begin{aligned}\lambda\Big((\frac{x'}{t'}-u)u_x+(\frac{y'}{t'}-v)v_x\Big)+\frac{d}{d\tau}\Big(\frac{\lambda}{t'}(\frac{x'}{t'}-u)\Big) &= 0\,,\\ \lambda\Big((\frac{x'}{t'}-u)u_y+(\frac{y'}{t'}-v)v_y\Big)+\frac{d}{d\tau}\Big(\frac{\lambda}{t'}(\frac{y'}{t'}-v)\Big) &= 0\,.\end{aligned} \tag{9}$$

From (1) we have $(\frac{x'}{t'}-u)=(\frac{dx}{dt}-u)=k\cos\varphi$ and $(\frac{y'}{t'}-v)=(\frac{dy}{dt}-v)=k\sin\varphi$. Introduce the new multiplier function $\mu=\dfrac{\lambda}{t'}$. Note that for any variable z we have $\dfrac{dz}{d\tau}=t'\dfrac{dz}{dt}$. Then equations (9) simplify to

$$\begin{aligned}\mu\cos\varphi u_x+\mu\sin\varphi v_x+\frac{d}{dt}(\mu\cos\varphi) &= 0\,,\\ \mu\cos\varphi u_y+\mu\sin\varphi v_y+\frac{d}{dt}(\mu\sin\varphi) &= 0\,.\end{aligned} \tag{10}$$

Elimination of μ and $\dfrac{d\mu}{dt}$ from these two equations leads to Zermelo's navigation formula (2).

It is useful to consider some examples. We will examine two given by Zermelo (*1930*), although we present them in more detail than Zermelo did himself. For the sake of discussion, we will illustrate his formula for the case of a ship or power boat crossing a river. In the first example we are given a river that is straight with constant width a lying parallel to the x-axis, the latter being oriented in a west-east direction. The beginning point A on the south side of the river is taken to be the origin of the coordinate system; the y axis runs in a positive direction northwards, the x axis runs in a positive direction eastwards. The destination is the point B on the opposite side of the river with coordinates (b,a). The speed of the boat relative to the water (its speed over land in still water) is k. The current is constant in time and place, flowing from east to west. Its value at each point (x,y) is $-c$, where $k>c\geq 0$. The goal is to find the so-called "steering direction" $\varphi=\varphi(t)$. This is the direction in which we power the ship, knowing that the actual direction followed by the ship will be modified as the result of the action of the current. Equations (1) here become

$$\begin{aligned}&\text{(a)}\quad \frac{dx}{dt} &&= -c+k\cos\varphi\,,\\ &\text{(b)}\quad \frac{dy}{dt} &&= k\sin\varphi\,.\end{aligned} \tag{11}$$

The navigation formula (2) becomes $\dfrac{d\varphi}{dt}=0$, or $\varphi=\text{const}$. It is apparent from (11) that $\dfrac{dy}{dx}=\text{const.}$, and the path followed by the ship in least time

is the straight line joining A and B. If we let $a = b$ then $\dfrac{dy}{dx} = 1$. Dividing (11a) by (11b) we have $\cos\varphi - \sin\varphi = \dfrac{c}{k}$. Solving for the steering angle φ we obtain

$$\varphi = \arccos\left(\frac{\dfrac{c}{k} + \sqrt{2 - \dfrac{c^2}{k^2}}}{2}\right). \tag{12}$$

For example, if $k = 10$ knots and $c = 3$ knots then $\varphi = \arccos(.841) = 32.8°$. Hence we would place the ship on a steering bearing of $57.2°$ $(90° - 32.8°)$ with respect to true north for it to follow the track $y = x$ of least time from A to B.

Zermelo (*1930*, 48) refers to the second example as "the simplest non-trivial example of our theory." In this example the current in the river increases as a linear function of y, being 0 at A and reaching its maximum value at B. Assume again that the river flows from east to west. Equations (1) here are

$$\begin{aligned} &\text{(a)} \quad \frac{dx}{dt} = -y + \cos\varphi\,, \\ &\text{(b)} \quad \frac{dy}{dt} = \sin\varphi\,, \end{aligned} \tag{13}$$

where the units have been adjusted so that $k = 1$. The navigation formula gives $\dfrac{d\varphi}{dt} = \cos^2\varphi$. We integrate this to obtain $\tan\varphi = t + \tan\varphi_0$, which gives the steering angle as a function of time. To find the optimal track we begin by noting that from (13b) we have $\dfrac{dy}{d\varphi} \cdot \dfrac{d\varphi}{dt} = \sin\varphi$, or $\dfrac{dy}{d\varphi}\cos^2\varphi = \sin\varphi$. Hence $y = \dfrac{1}{\cos\varphi} + \text{const.}$, or

$$y = \frac{1}{\cos\varphi} - \frac{1}{\cos\varphi_0}\,. \tag{14}$$

We now divide(13a) by (13b) and use (14) to obtain

$$\frac{dx}{dy} = \frac{1 - y(y+f)}{\sqrt{(y+f)^2 - 1}}\,, \tag{15}$$

where $f = \sec\varphi_0$. (15) describes the tracks followed by the ship in least time from A to points B on the opposite bank. Fig. 1 provides graphs of solutions to (15) giving the optimal curves for values of f from 1.0 to 1.45 in increments of 0.05.

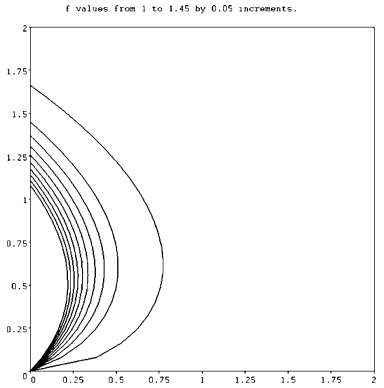


Fig. 1

Zermelo's navigation problem appears in several later textbooks on the calculus of variations. *Funk 1962* and *Klötzler 1970* are representatives from the German literature. Funk's treatment is very clear but relatively brief. A notable characteristic of Klötzler's account is its abstruseness, a result of the author's decision to develop the calculus of variations as part of functional analysis. English-language authors seem united in presenting the problem as one in nautical navigation. Some authors of note here are Fox (*1950*), Pars (*1962*), Young (*1969*), and Sagan (*1969*). With the exception of Funk, Zermelo's navigation formula is derived in none of these books. The only thing the authors seem to owe to Zermelo's original analysis is the formulation of the problem as one in the calculus of variations. In the English literature the navigation problem is presented for a ship moving through water against a current. There is no recognition that the problem was originally formulated for an aircraft or dirigible. Young states the problem but provides no solution. The treatment of Fox is very clear and derives from *Carathéodory 1935*. Fox (p.152) deduces the interesting fact that the ship can be steered along its optimal track by blind reckoning, that is, by a knowledge only of the time independent of any information about its external position. Pars's exposition is essentially a recapitulation of Fox's account. Of these later treatments, Sagan's is the simplest in its approach. This author solves for the time t and then uses the standard Euler equations to obtain a description of the motion of the boat.

It should be noted that the canonical problems of the calculus of variations—the isoperimetric problem, the hanging chain, the brachistochrone—go back centuries and appear at an early stage in the history of the subject. The navigation problem is somewhat unusual in providing a simple and signature example of very recent vintage, arising from technological developments of the twentieth century.

In the theory of control Zermelo's navigation problem appears as a standard problem of interest. Indeed, in modern mathematical science this subject would probably be the most likely place one would encounter the problem. For a sample of this literature see *Hsu and Meyer 1968, Bryson and Ho 1975*, and *Lewis 1986*. Writers refer to the Zermelo problem as "classic"; it is invariably presented for the case of a power boat moving against a current. In a recent book on optimal control and aerospace applications Ben-Asher (*2010*, 101) writes: "This problem was proposed by the German mathematician Ernest [*sic*] Zermelo in 1913 [*sic*]. Although formulated for a boat, it can also describe an aircraft flying in wind, assuming a fast response to heading changes. Therefore, it will be used extensively in this textbook." Historical misinformation is also present in the *Wikipedia* article on Zermelo, where we find the following: "Proposed in 1931, the Zermelo's navigation problem is a classic optimal control problem. The problem deals with a boat navigating on a body of water, originating from a point O to a destination point D. The boat is capable of a certain

Über die Navigation in der Luft als Problem der Variationsrechnung

1930c

(Auszug. Eine ausführliche Darstellung erfolgt demnächst in der „Zeitschrift für angewandte Mathematik".)

Das hier behandelte Problem ist das folgende. *In einer unbegrenzten Ebene, in welcher die Windverteilung durch ein Vektorfeld u, v als Funktion von Ort und Zeit gegeben ist, bewegt sich ein Luftschiff oder Flugzeug mit der konstanten Eigengeschwindigkeit k relativ zur umgebenden Luftmasse. Wie muß das Fahrzeug gesteuert werden, um in kürzester Zeit von einem Punkte P_0 zu einem anderen P_1 zu gelangen?* Wird mit φ der Winkel bezeichnet, den die „Steuerrichtung" des Fahrzeuges, d. h. der Vektor seiner Eigenbewegung mit der x-Achse bildet, so ergeben sich unmittelbar durch Vektoraddition für die Geschwindigkeitskomponenten der Gesamtbewegung die „Steuergleichungen"

$$\left.\begin{aligned} \frac{dx}{dt} &= k\cos\varphi + u(x,y,t) \\ \frac{dy}{dt} &= k\sin\varphi + v(x,y,t)\ , \end{aligned}\right\} \tag{1}$$

welche die Bewegung vollständig bestimmen, wenn φ als Funktion der Zeit gegeben ist. In Wirklichkeit soll aber diese Funktion so gewählt werden, daß die gestellte Minimumsbedingung erfüllt ist. Es handelt sich also um ein Va-

maximum speed, and we want to derive the best possible control to reach D in the least possible time" (retrieved November 1, 2012).

The term "Zermelo navigation" appears in the title of an article by David Bao, Colleen Robles and Zhongmin Shen (*2004*). The context of this research is different from the mathematical work that has been discussed up to now, since it does not concern the calculus of variations but rather differential geometry in which certain ideas about geodesics come into play. Although the paper bears only a limited relation to Zermelo's original investigation, it does retain the idea of motion (in a very abstract setting) that is modified by the action of continuous constraints in the form of differential equations. It is perhaps indicative of the seminal character of Zermelo's ideas that they would lend themselves to such abstract and general development.

Acknowledgment. The author would like to thank Bruce Petrie who used Maple to produce the graphical solution to equation (15), given in Fig. 1.

On navigation in the air as a problem in the calculus of variations

1930c

(Excerpt. A detailed account will soon appear in the "Zeitschrift für angewandte Mathematik".)

The present paper addresses the following problem. *In an unbounded plane in which the wind distribution is given by a vector field u, v as a function of space and time an airship or airplane moves at the constant airspeed k relative to the surrounding air mass. How is the vehicle to be steered in order to reach a point P_1 from another point P_0 in the shortest time possible?* If the angle formed by the "steering direction" of the vehicle, i. e., the vector of its proper motion, and by the x-axis is denoted by φ, then vector addition for the velocity components of the total motion immediately yields the "steering equations"

$$\left.\begin{aligned} \frac{dx}{dt} &= k\cos\varphi + u(x,y,t) \\ \frac{dy}{dt} &= k\sin\varphi + v(x,y,t)\ , \end{aligned}\right\} \tag{1}$$

which completely determine the motion, provided φ is given as a function of time. But in reality this function shall be chosen so that the specified

45 riationsproblem mit drei unbekannten Funktionen x, y, φ der | unabhängigen Veränderlichen t, zwischen denen zwei Differentialgleichungen erster Ordnung bestehen, und in welchem das Intervall $T = t_1 - t_0$ der Variablen t möglichst klein werden soll, innerhalb dessen das Wertepaar x, y von vorgeschriebenen Anfangswerten x_0, y_0 zu vorgeschriebenen Endwerten x_1, y_1 übergeht.

Gehen wir, wie in der Variationsrechnung üblich, von vorgelegten Funktionen $x(t)$, $y(t)$, $\varphi(t)$ zu einparametrigen Scharen $x(t,\alpha)$, $y(t,\alpha)$, $\varphi(t,\alpha)$ über, so erhalten wir durch „Variation", d. h. durch partielle Differentiation nach dem von t unabhängigen Parameter α aus (1)

$$\left.\begin{aligned} \frac{d\delta x}{dt} = \delta\frac{dx}{dt} &= -k\sin\varphi\delta\varphi + u_x\delta x + u_y\delta y \\ \frac{d\delta y}{dt} = \delta\frac{dy}{dt} &= k\cos\varphi\delta\varphi + v_x\delta x + v_y\delta y \end{aligned}\right\} \tag{2}$$

und weiter durch Einführung der noch zu bestimmenden Multiplikatoren $\lambda(t)$ und $\mu(t)$

$$\begin{aligned} \frac{d}{dt}(\lambda\delta x + \mu\delta y) &= k\delta\varphi(-\lambda\sin\varphi + \mu\cos\varphi) \\ &\quad + \delta x\Big(\frac{d\lambda}{dt} + u_x\lambda + v_x\mu\Big) + \delta y\Big(\frac{d\mu}{dt} + u_y\lambda + v_y\mu\Big)\,. \end{aligned} \tag{3}$$

Hier können wir die Glieder mit δx und δy auf der rechten Seite zum Verschwinden bringen vermöge der linearen und homogenen Differentialgleichungen

$$\left.\begin{aligned} \frac{d\lambda}{dt} + u_x\lambda + v_x\mu &= 0 \\ \frac{d\mu}{dt} + u_y\lambda + v_y\mu &= 0\,, \end{aligned}\right\} \tag{4}$$

welche, wenn x, y, φ als Funktionen von t gegeben sind, λ und μ bestimmen in der Form

$$\lambda = c_1\lambda_1 + c_2\lambda_2\,, \quad \mu = c_1\mu_1 + c_2\mu_2$$

mit den willkürlichen Konstanten c_1 und c_2. Diese Konstanten können nun so gewählt werden, daß der Ausdruck

$$\omega = -\lambda\sin\varphi + \mu\cos\varphi = c_1w_1 + c_2w_2 \tag{5}$$

eine „erlaubte Variation" $\delta\varphi$ von φ darstellt, die den Grenzbedingungen bei $t = t_0$ und $t = t_1$ genügt, nach denen alle Bahnkurven der Schar durch die Endpunkte $P_0(x_0, y_0)$ und $P_1(x_1, y_1)$ gehen sollen, d. h. es soll sein

$$\delta x_0 = \delta y_0 = 0 \qquad \text{und} \tag{6}$$

$$\Big(\frac{dy}{dt}\Big)_{(1)}\delta x_1 - \Big(\frac{dx}{dt}\Big)_{(1)}\delta y_1 = 0\,, \tag{7}$$

minimum condition is satisfied. Thus we have a variational problem with three unknown functions x, y, φ of the independent variable t between which two differential equations of first order obtain and in which the interval $T = t_1 - t_0$ of the variable t is supposed to become as small as possible, within which the pair of values x, y passes from prescribed initial values x_0, y_0 to prescribed final values x_1, y_1.

If we pass, as is customary in the calculus of variations, from given functions $x(t)$, $y(t)$, $\varphi(t)$ to one-parameter families $x(t,\alpha)$, $y(t,\alpha)$, $\varphi(t,\alpha)$, then we obtain by "variation", i. e., by partial differentiation with respect to the parameter α independent of t from (1)

$$\left.\begin{aligned} \frac{d\delta x}{dt} &= \delta\frac{dx}{dt} = -k\sin\varphi\delta\varphi + u_x\delta x + u_y\delta y \\ \frac{d\delta y}{dt} &= \delta\frac{dy}{dt} = \quad k\cos\varphi\delta\varphi + v_x\delta x + v_y\delta y \end{aligned}\right\} \tag{2}$$

and, furthermore, by introduction of the multipliers $\lambda(t)$ and $\mu(t)$, yet to be determined,

$$\begin{aligned} \frac{d}{dt}(\lambda\delta x + \mu\delta y) &= k\delta\varphi(-\lambda\sin\varphi + \mu\cos\varphi) \\ &\quad + \delta x\Big(\frac{d\lambda}{dt} + u_x\lambda + v_x\mu\Big) + \delta y\Big(\frac{d\mu}{dt} + u_y\lambda + v_y\mu\Big)\,. \end{aligned} \tag{3}$$

We can make the terms with δx and δy on the right side vanish by dint of the linear and homogeneous differential equations

$$\left.\begin{aligned} \frac{d\lambda}{dt} + u_x\lambda + v_x\mu &= 0 \\ \frac{d\mu}{dt} + u_y\lambda + v_y\mu &= 0\,, \end{aligned}\right\} \tag{4}$$

which, if x, y, φ are given as functions of t, determine λ and μ in the form

$$\lambda = c_1\lambda_1 + c_2\lambda_2\,, \quad \mu = c_1\mu_1 + c_2\mu_2$$

with the arbitrary constants c_1 and c_2. It is now possible to choose these constants so that the expression

$$\omega = -\lambda\sin\varphi + \mu\cos\varphi = c_1w_1 + c_2w_2 \tag{5}$$

represents a "permissible variation" $\delta\varphi$ of φ satisfying the boundary conditions at $t = t_0$ and $t = t_1$ according to which all paths of the family are supposed to pass through the endpoints $P_0(x_0, y_0)$ and $P_1(x_1, y_1)$, i. e., we are supposed to have

$$\delta x_0 = \delta y_0 = 0 \qquad \text{and} \tag{6}$$

$$\Big(\frac{dy}{dt}\Big)_{(1)}\delta x_1 - \Big(\frac{dx}{dt}\Big)_{(1)}\delta y_1 = 0\,, \tag{7}$$

da der feste Endpunkt x_1, y_1 in den variierten Bahnkurven zu verschiedenen Zeiten erreicht wird. Aus den linearen und unhomogenen Differentialgleichungen (2) bestimmen sich nämlich δx und δy wegen (6) eindeutig durch $\delta\varphi$ in der Form

$$\delta x = c_1\xi_1 + c_2\xi_2 \ , \quad \delta y = c_1\eta_1 + c_2\eta_2 \ ,$$

46 | wenn $\delta x = \xi_1$, $\delta y = \eta_1$ zu $\delta\varphi = \omega_1$ und entsprechend ξ_2, η_2 zu $\delta\varphi = \omega_2$ gehören. Durch Einsetzung in (7) erhält man dann den Quotienten von c_1 und c_2. Für jedes dieser Bedingung entsprechende Wertepaar c_1, c_2 stellt aber, wie näher gezeigt werden könnte, der Ausdruck (3) in der Tat eine erlaubte Variation von φ dar, und (3) reduziert sich auf

$$\frac{d}{dt}(\lambda\delta x + \mu\delta y) = k(-\lambda\sin\varphi + \mu\cos\varphi)^2 \geqq 0 \ , \tag{3a}$$

also wegen (6) durch Integration zwischen t_0 und t_1

$$\lambda(t_1)\,\delta x_1 + \mu(t_1)\,\delta y_1 > 0 \ ,$$

sofern nicht der Integrand an allen Stetigkeitsstellen des Intervalles verschwindet. Nun ist aber in Übereinstimmung mit (7)

$$\delta x_1 = -x_1'\delta t_1 \ , \quad \delta y_1 = -y_1'\delta t_1 \ ,$$

wo im Falle des Minimums $\delta t_1 = \dfrac{dt_1}{d\alpha} = 0$ sein muß. Es folgt also für das ganze Intervall die Gleichung

$$-\lambda\sin\varphi + \mu\cos\varphi = 0 \ , \tag{8}$$

welche zusammen mit den Differentialgleichungen (1) und (4) die 5 Funktionen x, y, φ, λ, μ als Funktionen von t und den Anfangswerten bestimmt. Aus (8) folgt aber

$$\operatorname{tg}\varphi = \frac{\mu}{\lambda}$$

und hieraus durch Differentiation mit Hilfe von (4) die Gleichung

$$\frac{d\varphi}{dt} = \frac{\lambda\dfrac{d\mu}{dt} - \mu\dfrac{d\lambda}{dt}}{\lambda^2+\mu^2} = v_x\sin^2\varphi + (u_x - v_y)\sin\varphi\cos\varphi - u_y\cos^2\varphi \ , \tag{9}$$

die als „Navigationsformel“ zusammen mit (1) die „extremale Bewegung“ bestimmt.

Wir erhalten also für die drei unbekannten Funktionen x, y, φ von t drei simultane Differentialgleichungen erster Ordnung, und ihre Integrale in der Form

$$x = x(t\,|\,x_0, y_0, \varphi_0) \ , \quad y = y(t\,|\,x_0, y_0, \varphi_0) \ , \quad \varphi = \varphi(t\,|\,x_0, y_0, \varphi_0) \ ,$$

since the fixed endpoint x_1, y_1 is reached along the varied paths at different times. For δx and δy are uniquely determined on the basis of the linear and non-homogeneous differential equations (2) on account of (6) by $\delta\varphi$ in the form

$$\delta x = c_1\xi_1 + c_2\xi_2 \ , \quad \delta y = c_1\eta_1 + c_2\eta_2 \ ,$$

if $\delta x = \xi_1$, $\delta y = \eta_1$ belong to $\delta\varphi = \omega_1$ and, accordingly, ξ_2, η_2 to $\delta\varphi = \omega_2$. By substitution in (7) we then obtain the quotient of c_1 and c_2. But, as could be shown in more detail, for each pair of values c_1, c_2 satisfying this condition the expression (3) is really a permissible variation of φ, and (3) reduces to

$$\frac{d}{dt}(\lambda\delta x + \mu\delta y) = k(-\lambda\sin\varphi + \mu\cos\varphi)^2 \geqq 0 \ , \tag{3a}$$

and hence, on account of (6), by integration between t_0 and t_1

$$\lambda(t_1)\,\delta x_1 + \mu(t_1)\,\delta y_1 > 0 \ ,$$

provided the integrand does not vanish at all continuity points of the interval. But now, in accordance with (7),

$$\delta x_1 = -x_1'\delta t_1 \ , \quad \delta y_1 = -y_1'\delta t_1 \ ,$$

where in the case of the minimum we must have $\delta t_1 = \dfrac{dt_1}{d\alpha} = 0$. Hence, for the entire interval there follows the equation

$$-\lambda\sin\varphi + \mu\cos\varphi = 0 \ , \tag{8}$$

which, together with the differential equations (1) and (4), determines the 5 functions x, y, φ, λ, μ as functions of t and the initial values. But from (8) follows

$$\operatorname{tg}\varphi = \frac{\mu}{\lambda}$$

and from this by differentiation by means of (4) the equation

$$\frac{d\varphi}{dt} = \frac{\lambda\dfrac{d\mu}{dt} - \mu\dfrac{d\lambda}{dt}}{\lambda^2 + \mu^2} = v_x\sin^2\varphi + (u_x - v_y)\sin\varphi\cos\varphi - u_y\cos^2\varphi \ , \tag{9}$$

which, as a "navigation formula", together with (1) determines the "extremal motion".

Thus we obtain for the three unknown functions x, y, φ of t three simultaneous differential equations of first order and their integrals in the form

$$x = x(t\,|\,x_0, y_0, \varphi_0) \ , \quad y = y(t\,|\,x_0, y_0, \varphi_0) \ , \quad \varphi = \varphi(t\,|\,x_0, y_0, \varphi_0) \ ,$$

soweit die vorkommenden Funktionen u, v, u_x, u_y, v_x, v_y den *Lipschitz*schen Bedingungen genügen.

Durch jeden Anfangspunkt P_0 geht also ein Büschel von Extremalen oder „Navigationskurven", dessen Parameter φ_0 (die anfängliche Steuerrichtung) an der Hand der gegenwärtigen und zu erwartenden Windkarte so gewählt werden muß, daß die zugehörige Extremale durch den vorgeschriebenen Zielpunkt P_1 geht.

Im einfachsten Falle, wo die Windkomponenten u, v in der Ebene *konstant* sind, ergibt sich aus (9) $\dfrac{d\varphi}{dt} = 0$, also $\varphi = \text{const.}$ und somit auch $\dfrac{dx}{dt}$ und $\dfrac{dy}{dt}$ konstant, d. h. die extremale Bewegung erfolgt *geradlinig* und *gleichförmig* zwischen P_0 und P_1 mit der durch (1) bestimmten Steuerrichtung $\varphi = \varphi_0$.

47 | Um nun auch *hinreichende* Bedingungen für das Bestehen eines Extremums zu finden, betrachten wir eine gleichzeitig bei $t = 0$ von einem Punkte 0 ausgehende Schar von extremalen Bahnkurven, d. h. von Funktionen

$$x = \overline{x}(t, \alpha)\ , \quad y = \overline{y}(t, \alpha)\ , \quad \varphi = \overline{\varphi}(t, \alpha)\ , \quad \lambda = \lambda(t, \alpha)\ , \quad \mu = \mu(t, \alpha)\ ,$$

die den Bedingungen (1), (4), (8) und (9) genügen, und erhalten aus (3) für variable t und α

$$\frac{\partial}{\partial t}\Big(\lambda\frac{\partial\overline{x}}{\partial\alpha} + \mu\frac{\partial\overline{y}}{\partial\alpha}\Big) = 0 \qquad \text{und hieraus}$$

$$\lambda\frac{\partial\overline{x}}{\partial\alpha} + \mu\frac{\partial\overline{y}}{\partial\alpha} = \text{const.} = 0\ , \tag{10}$$

da im Anfangspunkt $t = 0$ immer $\dfrac{\partial\overline{x}}{\partial\alpha} = \dfrac{\partial\overline{y}}{\partial\alpha} = 0$ ist. Durchläuft nun der variable Endpunkt P_1 dieser Extremalen selbst eine Bahnkurve $x_1 = x_1(\alpha)$, $y_1 = y_1(\alpha)$, wo α als Funktion von t und umgekehrt zu betrachten ist, so wird für jedes α an der Stelle $t = \overline{t}_1$, $x = x_1$, $y = y_1$

$$\Big(\frac{dx_1}{d\alpha_1}\Big)_1 = \Big(\frac{\partial\overline{x}}{\partial\alpha}\Big)_1 + \Big(\frac{\partial\overline{x}}{\partial t}\Big)_1\frac{d\overline{t}_1}{d\alpha}\ , \quad \Big(\frac{dy_1}{d\alpha_1}\Big)_1 = \Big(\frac{\partial\overline{y}}{\partial\alpha}\Big)_1 + \Big(\frac{\partial\overline{y}}{\partial t}\Big)\frac{d\overline{t}}{d\alpha}\ ,$$

und daher wegen (10)

$$\lambda\frac{dx_1}{d\alpha} + \mu\frac{dy_1}{d\alpha} = \Big(\lambda\frac{\partial\overline{x}}{\partial t} + \mu\frac{\partial\overline{y}}{\partial t}\Big)_1\frac{d\overline{t}_1}{d\alpha} = (\lambda\overline{x}'_1 + \mu\overline{y}'_1)\,\frac{d\overline{t}_1}{d\alpha}\ ,$$

also wegen (8) und (1)

$$\frac{d\overline{t}_1}{d\alpha} = \frac{\cos\overline{\varphi}_1\dfrac{dx_1}{d\alpha} + \sin\varphi_1\dfrac{dy_1}{d\alpha}}{\cos\overline{\varphi}_1\overline{x}'_1 + \sin\overline{\varphi}_1\overline{y}'_1} = \frac{\cos\overline{\varphi}_1\dfrac{dx_1}{d\alpha} + \sin\overline{\varphi}_1\dfrac{dy_1}{d\alpha}}{k + u_1\cos\overline{\varphi}_1 + v_1\sin\overline{\varphi}_1}\ ,$$

wo mit $\overline{\varphi}_1$ die auf zur Extremalen gehörende Steuerrichtung im Punkte x_1, y_1 und mit u_1, v_1 die dortigen Windkomponenten bezeichnet sind. Nun neh-

provided the occurring functions u, v, u_x, u_y, v_x, v_y satisfy the *Lipschitz* conditions.

Hence, through each initial point P_0 there passes a cluster of extremals, or "navigation curves", whose parameter φ_0 (the initial steering direction) is to be chosen on the basis of the present wind map and the wind map to be expected so that the associated extremal passes through the prescribed point of destination P_1.

In the simplest case where the wind components u, v are *constant* in the plane, we obtain $\dfrac{d\varphi}{dt} = 0$ from (9), and hence $\varphi = \text{const.}$; therefore $\dfrac{dx}{dt}$ and $\dfrac{dy}{dt}$, too, are constant, i. e., the extremal motion is *rectilinear* and *uniform* between P_0 and P_1 with the steering direction $\varphi = \varphi_0$ determined by (1).

Now, in order to also find *sufficient* conditions for the existence of an extremum, we consider a family of extremal curves simultaneously starting from a point 0 at $t = 0$, i. e., of functions

$$x = \overline{x}(t, \alpha)\ , \quad y = \overline{y}(t, \alpha)\ , \quad \varphi = \overline{\varphi}(t, \alpha)\ , \quad \lambda = \lambda(t, \alpha)\ , \quad \mu = \mu(t, \alpha)\ ,$$

satisfying conditions (1), (4), (8) and (9), and we obtain from (3) for variable t and α

$$\frac{\partial}{\partial t}\Big(\lambda\frac{\partial\overline{x}}{\partial\alpha} + \mu\frac{\partial\overline{y}}{\partial\alpha}\Big) = 0 \qquad \text{and from this}$$

$$\lambda\frac{\partial\overline{x}}{\partial\alpha} + \mu\frac{\partial\overline{y}}{\partial\alpha} = \text{const.} = 0\ , \tag{10}$$

since we always have $\dfrac{\partial\overline{x}}{\partial\alpha} = \dfrac{\partial\overline{y}}{\partial\alpha} = 0$ in the initial point $t = 0$. Now, if the variable endpoint P_1 of this extremal itself runs through a path $x_1 = x_1(\alpha)$, $y_1 = y_1(\alpha)$, where α is to be considered a function of t, and vice versa, then for each α at point $t = \overline{t}_1$, $x = x_1$, $y = y_1$

$$\Big(\frac{dx_1}{d\alpha_1}\Big)_1 = \Big(\frac{\partial\overline{x}}{\partial\alpha}\Big)_1 + \Big(\frac{\partial\overline{x}}{\partial t}\Big)_1\frac{d\overline{t}_1}{d\alpha}\ , \quad \Big(\frac{dy_1}{d\alpha_1}\Big)_1 = \Big(\frac{\partial\overline{y}}{\partial\alpha}\Big)_1 + \Big(\frac{\partial\overline{y}}{\partial t}\Big)\frac{d\overline{t}}{d\alpha}\ ,$$

and hence, on account of (10),

$$\lambda\frac{dx_1}{d\alpha} + \mu\frac{dy_1}{d\alpha} = \Big(\lambda\frac{\partial\overline{x}}{\partial t} + \mu\frac{\partial\overline{y}}{\partial t}\Big)_1\frac{d\overline{t}_1}{d\alpha} = (\lambda\overline{x}'_1 + \mu\overline{y}'_1)\,\frac{d\overline{t}_1}{d\alpha}\ ,$$

and hence, on account of (8) and (1),

$$\frac{d\overline{t}_1}{d\alpha} = \frac{\cos\overline{\varphi}_1\dfrac{dx_1}{d\alpha} + \sin\varphi_1\dfrac{dy_1}{d\alpha}}{\cos\overline{\varphi}_1\overline{x}'_1 + \sin\overline{\varphi}_1\overline{y}'_1} = \frac{\cos\overline{\varphi}_1\dfrac{dx_1}{d\alpha} + \sin\overline{\varphi}_1\dfrac{dy_1}{d\alpha}}{k + u_1\cos\overline{\varphi}_1 + v_1\sin\overline{\varphi}_1}\ ,$$

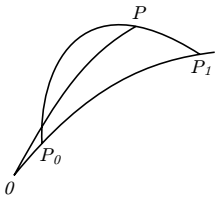
where $\overline{\varphi}_1$ denotes the steering direction belonging to the extremal at point x_1, y_1, and u_1, v_1 denote the associated wind components at that point.

men wir weiter an, der variable Punkt P_1 durchlaufe selbst eine nach der Formel (1) gesteuerte Bahn $\alpha = \alpha(t)$, so ergibt sich mit Weglassung des Index 1

$$\frac{d\overline{t}}{dt} = \frac{\cos\overline{\varphi}\dfrac{dx}{dt} + \sin\overline{\varphi}\dfrac{dy}{dt}}{k + u\cos\overline{\varphi} + v\sin\overline{\varphi}} \tag{11}$$

$$= \frac{k\cos(\varphi - \overline{\varphi}) + u\cos\overline{\varphi} + v\sin\overline{\varphi}}{k + u\cos\overline{\varphi} + v\sin\overline{\varphi}} ,$$

wo t die zur Durchlaufung der Kurve $\mathfrak{C}$ und $\overline{t}$ die zur Durchlaufung der



durch P gehenden Feldextremalen erforderliche Zeit bezeichnet. Liegen nun beide Endpunkte P_0 und P_1 selbst auf einer Extremalen des Feldes, so ergibt sich durch Integration

$$\overline{T} = \overline{t}_1 - \overline{t}_0 = \int_{t_0}^{t_1} \frac{k\cos(\varphi - \overline{\varphi}) + u\cos\overline{\varphi} + v\sin\overline{\varphi}}{k + u\cos\overline{\varphi} + v\sin\overline{\varphi}} dt , \qquad \text{also}$$

48 | $$T - \overline{T} = \int_{t_0}^{t_1} \frac{k(1 - \cos(\varphi - \overline{\varphi}))}{k + u\cos\overline{\varphi} + v\sin\overline{\varphi}} dt = \int_{t_0}^{t_1} E(x, y \,|\, \overline{\varphi}, \varphi)\, dt \geqq 0 , \tag{12}$$

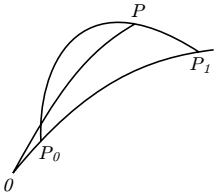
soweit der Nenner unseres Integranden E selbst positiv, d. h. soweit *die gegen die Steuerrichtung der Feldextremalen gerichtete Windkomponente kleiner ist als die Eigengeschwindigkeit* k.

Damit ist die *Weierstraß*sche *hinreichende* Bedingung für die Existenz des Extremums innerhalb des hier konstruierten „Feldes“ von Extremalen gefunden. Sie geht in die *Legendre*sche *notwendige* Bedingung über, wenn sie anstatt auf das Feld auf die in Frage stehende *Einzelextremale* angewendet wird. Die natürliche *Begrenzung* des Feldes liefert endlich die bekannte *Jacobi*sche Bedingung, wonach sich nicht zwei benachbarte Extremalen *zweimal* innerhalb des betrachteten Intervalles schneiden dürfen, eine Bedingung, die durch die aufeinander folgenden Nullstellen einer durch Variation aus der ursprünglichen ableitbaren *linearen Differentialgleichung* ausgedrückt werden kann.

If, furthermore, we assume that the variable point P_1 itself runs through a path $\alpha = \alpha(t)$ steered in accordance with formula (1), then, by dropping the index 1, we obtain

$$\begin{aligned} \frac{d\overline{t}}{dt} &= \frac{\cos\overline{\varphi}\dfrac{dx}{dt} + \sin\overline{\varphi}\dfrac{dy}{dt}}{k + u\cos\overline{\varphi} + v\sin\overline{\varphi}} \\ &= \frac{k\cos(\varphi - \overline{\varphi}) + u\cos\overline{\varphi} + v\sin\overline{\varphi}}{k + u\cos\overline{\varphi} + v\sin\overline{\varphi}} \,, \end{aligned} \tag{11}$$

where t denotes the time required to run through the curve $\mathfrak{C}$, and $\overline{t}$ denotes



the time required to run through the field extremal passing through P. If the two endpoints P_0 and P_1 themselves lie on an extremal of the field, then integration yields

$$\overline{T} = \overline{t}_1 - \overline{t}_0 = \int_{t_0}^{t_1} \frac{k\cos(\varphi - \overline{\varphi}) + u\cos\overline{\varphi} + v\sin\overline{\varphi}}{k + u\cos\overline{\varphi} + v\sin\overline{\varphi}} dt \,, \qquad \text{and hence}$$

$$T - \overline{T} = \int_{t_0}^{t_1} \frac{k(1 - \cos(\varphi - \overline{\varphi}))}{k + u\cos\overline{\varphi} + v\sin\overline{\varphi}} dt = \int_{t_0}^{t_1} E(x, y \,|\, \overline{\varphi}, \varphi)\, dt \geqq 0 \,, \tag{12}$$

provided the denominator of our integrand E itself is positive, i. e., provided *the wind component directed against the steering direction of the field extremal is smaller than the airspeed* k.

Thus we have found *Weierstrass*'s *sufficient* condition for the existence of the extremum within the "field" of extremals constructed here. It is transformed into *Legendre*'s *necessary* condition when, instead of being applied to the field, it is applied to the *individual extremal* in question. Finally, the natural *delineation* of the field yields the well-known *Jacobi* condition, according to which two neighboring extremals must not intersect one another *twice* in the interior of the considered interval, a condition that can be expressed by the successive roots of a *linear differential equation* derivable from the original one by variation.

Das einfachste nicht-triviale *Beispiel* unserer Theorie liefert der Fall einer *linearen* Windverteilung, in welcher etwa auf der nördlichen Halbebene *Ostwind* von wechselnder Stärke, auf der südlichen entsprechender *Westwind* angenommen wird: $u = -y$, $v = 0$. Hier ergibt sich $\frac{d\varphi}{dt} = \cos^2\varphi$, also

$$t = \operatorname{tg}\varphi + c = \operatorname{tg}\varphi - \operatorname{tg}\varphi_0\ , \quad y = \frac{k}{\cos\varphi} + c' = k\Big(\frac{1}{\cos\varphi} - \frac{1}{\cos\varphi_0}\Big)\ ,$$

und die bei $t = 0$ von einem Punkt des Äquators (wo die Windgeschwindigkeit 0 ist) ausgehenden Extremalen bedecken *einfach* die ganze Ebene, wobei auch die *Weierstraß*sche Bedingung überall erfüllt ist. Denn hier ist $k + u\cos\varphi = k\frac{\cos\varphi}{\cos\varphi_0} > 0$, weil $\cos\varphi$ sein Vorzeichen nicht wechseln kann, ohne daß t unendlich wird. In diesem Falle kann also auch bei *beliebig kleiner* Eigengeschwindigkeit *jeder* Punkt der Ebene auf eindeutig bestimmter Bahn in minimaler Zeit erreicht werden.

Über das Navigationsproblem bei ruhender oder veränderlicher Windverteilung

1931a

In meinem Prager Vortrage „Über die Navigation in der Luft als Problem der Variationsrechnung“ (Deutsche Mathematiker-Vrgg. Bd. 39, S. 44–48) habe ich das folgende Problem gestellt und gelöst: *In einer unbegrenzten Ebene, in welcher die Windverteilung durch ein Vektorfeld als Funktion von Ort und Zeit gegeben ist, bewegt sich ein Fahrzeug mit konstanter Eigengeschwindigkeit relativ zur umgebenden Luftmasse. Wie muß das Fahrzeug gesteuert werden, um in kürzester Zeit von einem Ausgangspunkte zu einem gegebenen Ziel zu gelangen?*

Da in der angegebenen Mitteilung die Darstellung etwas knapp gehalten war und in einigen Punkten noch der Ergänzung und Berichtigung bedarf, so möchte ich hier die ganze Entwicklung in erweiterter und verbesserter Form wiederholen, um das Problem in voller Allgemeinheit zu behandeln.

§ 1. Die Differentialgleichungen des Problems.

Bezeichnen wir mit t die Zeit, mit x, y die rechtwinkligen Koordinaten des bewegten Fahrzeugs, mit u, v die an der Eigengeschwindigkeit gemessenen

The simplest non-trivial *example* of our theory is provided by the case of a *linear* wind distribution in which we assume, say, an *easterly wind* of varying strength on the northern half-plane and a corresponding *westerly wind* on the southern half-plane: $u = -y$, $v = 0$. In this case, we have $\frac{d\varphi}{dt} = \cos^2\varphi$, and hence

$$t = \operatorname{tg}\varphi + c = \operatorname{tg}\varphi - \operatorname{tg}\varphi_0\ , \quad y = \frac{k}{\cos\varphi} + c' = k\Big(\frac{1}{\cos\varphi} - \frac{1}{\cos\varphi_0}\Big)\ ,$$

and the extremals starting at $t = 0$ from a point of the equator (where the wind velocity is 0) *simply* cover the entire plane, where also *Weierstrass*'s condition is satisfied everywhere. For in this case $k + u\cos\varphi = k\frac{\cos\varphi}{\cos\varphi_0} > 0$, since $\cos\varphi$ cannot change its sign without t becoming infinite. Hence, in this case, it is also possible to reach *every* point of the plane along a uniquely determined path in minimal time at an airspeed that is *arbitrarily small*.

On the navigation problem for a calm or variable wind distribution

1931a

⟦The introductory note just before *1930c* also addresses *1931a*.⟧

In my Prague lecture "On navigation in the wind as a problem in the calculus of variations" (*1930c*), I posed and solved the following problem: *In an unbounded plane in which the wind distribution is given by a vector field as a function of space and time a vehicle moves at constant airspeed relative to the surrounding air mass. How are we to steer the vehicle in order to reach a given destination from a starting point in the shortest time possible?*

Since the account I gave in this note is somewhat brief and still in need of amendment and correction in several points, I would like to return to the problem in the present expanded and revised account in order to treat it in full generality.

§ 1. The differential equations of the problem.

Let us denote time by t, the orthogonal coordinates of the vehicle in motion by x, y, the corresponding wind components measured against the

entsprechenden Windkomponenten, die als Funktionen von x, y, t gegeben sein sollen, und mit φ die „Steuerrichtung“, d. h. den Winkel, den der Vektor der Relativbewegung mit der x-Richtung bildet, so ergeben sich zunächst für jede „steuerbare“ Bahn die immer gültigen „Steuergleichungen“

$$\frac{dx}{dt} = u + \cos\varphi\ , \quad \frac{dy}{dt} = v + \sin\varphi\ . \tag{1}$$

Diese Differentialgleichungen bestimmen eindeutig die Bewegung von einem Ausgangspunkt des Windfeldes, wenn die (willkürliche) Steuerrichtung φ als Funktion der Zeit gegeben ist, und die Bestimmung dieser Funktion $\varphi = \varphi(t)$ ist die eigentliche Aufgabe des Variationsproblems. Gemäß den Methoden der
115 Variationsrechnung gehen wir von einer Bahn $x = x(t),\ |\ y = y(t),\ \varphi = \varphi(t)$ zu einer durch einen Parameter α charakterisierten Schar von „steuerbaren“ Bahnen über

$$x = x(t,\alpha)\ , \quad y = y(t,\alpha)\ , \quad \varphi = \varphi(t,\alpha)$$

und erhalten aus (1) durch „Variation“, d. h. durch partielle Differentiation nach α

$$\left.\begin{aligned} \frac{d}{dt}\frac{\partial x}{\partial\alpha} = \frac{\partial}{\partial\alpha}\frac{dx}{dt} &= u_x\frac{\partial x}{\partial\alpha} + u_y\frac{\partial y}{\partial\alpha} - \sin\varphi\frac{\partial\varphi}{\partial\alpha} \\ \frac{d}{dt}\frac{\partial y}{\partial\alpha} = \frac{\partial}{\partial\alpha}\frac{dy}{dt} &= v_x\frac{\partial x}{\partial\alpha} + v_y\frac{\partial y}{\partial\alpha} + \cos\varphi\frac{\partial\varphi}{\partial\alpha}\ , \end{aligned}\right\} \tag{2}$$

wo mit dem Zeichen d immer die *partielle* Differentiation nach t (bei konstantem α) bezeichnet wird. Durch Einführung zweier noch zu bestimmender Multiplikatoren $\lambda = \lambda(t)$, $\mu = \mu(t)$ erhalten wir weiter

$$\left.\begin{aligned} \frac{d}{dt}\Big(\lambda\frac{\partial x}{\partial\alpha} + \mu\frac{\partial y}{\partial\alpha}\Big) &= \Big(\frac{d\lambda}{dt} + u_x\lambda + v_x\mu\Big)\frac{\partial x}{\partial\alpha} + \Big(\frac{d\mu}{dt} + u_y\lambda + v_y\mu\Big)\frac{\partial y}{\partial\alpha} \\ &\quad + (-\lambda\sin\varphi + \mu\cos\varphi)\,\frac{\partial\varphi}{\partial\alpha} \\ &= (-\lambda\sin\varphi + \mu\cos\varphi)\,\frac{\partial\varphi}{\partial\alpha}\ , \end{aligned}\right\} \tag{3}$$

wenn wir die Funktionen $\lambda(t)$, $\mu(t)$ so wählen, daß

$$\left.\begin{aligned} \frac{d\lambda}{dt} + u_x\lambda + v_x\mu &= 0 \\ \frac{d\mu}{dt} + u_y\lambda + v_y\mu &= 0\ . \end{aligned}\right\} \tag{4}$$

Diese beiden linearen und homogenen Differentialgleichungen bestimmen λ und μ, wenn x, y und φ als Funktionen der Zeit gegeben sind, mit zwei willkürlichen Konstanten in der Form

$$\lambda = c_1\lambda_1 + c_2\lambda_2\ , \quad \mu = c_1\mu_1 + c_2\mu_2\ . \tag{5}$$

airspeed by u, v, which shall be given as functions of x, y, t, and the "steering direction" by φ, i. e., the angle formed by the vector of the relative motion and the x-direction. Then we obtain, first, for every "steerable" path the "steering equations" that always hold,

$$\frac{dx}{dt} = u + \cos\varphi\ , \quad \frac{dy}{dt} = v + \sin\varphi\ . \tag{1}$$

These differential equations uniquely determine the motion from a starting point of the wind field, if the (arbitrary) steering direction φ is given as a function of time, and the actual task of the variational problem is to determine this function $\varphi = \varphi(t)$. In accordance with the methods of the calculus of variations we proceed from a path $x = x(t)$, $y = y(t)$, $\varphi = \varphi(t)$ to a family of "steerable" paths which is characterized by a parameter α,

$$x = x(t, \alpha)\ , \quad y = y(t, \alpha)\ , \quad \varphi = \varphi(t, \alpha)\ ,$$

and obtain from (1) by "variation", i. e., by partial differentiation with respect to α,

$$\left.\begin{aligned} \frac{d}{dt}\frac{\partial x}{\partial \alpha} = \frac{\partial}{\partial \alpha}\frac{dx}{dt} &= u_x\frac{\partial x}{\partial \alpha} + u_y\frac{\partial y}{\partial \alpha} - \sin\varphi\frac{\partial \varphi}{\partial \alpha} \\ \frac{d}{dt}\frac{\partial y}{\partial \alpha} = \frac{\partial}{\partial \alpha}\frac{dy}{dt} &= v_x\frac{\partial x}{\partial \alpha} + v_y\frac{\partial y}{\partial \alpha} + \cos\varphi\frac{\partial \varphi}{\partial \alpha}\ , \end{aligned}\right\} \tag{2}$$

where the sign d always denotes the *partial* differentiation with respect to t (at constant α). By adding two multipliers $\lambda = \lambda(t)$, $\mu = \mu(t)$, which are yet to be determined, we furthermore obtain

$$\left.\begin{aligned} \frac{d}{dt}\left(\lambda\frac{\partial x}{\partial \alpha} + \mu\frac{\partial y}{\partial \alpha}\right) &= \left(\frac{d\lambda}{dt} + u_x\lambda + v_x\mu\right)\frac{\partial x}{\partial \alpha} + \left(\frac{d\mu}{dt} + u_y\lambda + v_y\mu\right)\frac{\partial y}{\partial \alpha} \\ &\quad + \left(-\lambda\sin\varphi + \mu\cos\varphi\right)\frac{\partial \varphi}{\partial \alpha} \\ &= \left(-\lambda\sin\varphi + \mu\cos\varphi\right)\frac{\partial \varphi}{\partial \alpha}\ , \end{aligned}\right\} \tag{3}$$

if we choose the functions $\lambda(t)$, $\mu(t)$ so that

$$\left.\begin{aligned} \frac{d\lambda}{dt} + u_x\lambda + v_x\mu &= 0 \\ \frac{d\mu}{dt} + u_y\lambda + v_y\mu &= 0\ . \end{aligned}\right\} \tag{4}$$

These two linear and homogeneous differential equations determine λ and μ, if x, y and φ are given as functions of time, with two arbitrary constants in the form

$$\lambda = c_1\lambda_1 + c_2\lambda_2\ , \quad \mu = c_1\mu_1 + c_2\mu_2\ . \tag{5}$$

Diese Konstanten wollen wir nun so wählen, daß die rechte Seite der Gl. (3) ein Quadrat wird, indem wir setzen

$$\frac{\partial\varphi}{\partial\alpha} = -\lambda\sin\varphi + \mu\cos\varphi = \omega\ . \tag{6}$$

Wie muß aber eine Funktion $\omega(t)$ beschaffen sein, um unter den gegebenen Bedingungen eine „erlaubte Variation“ $\dfrac{\partial\varphi}{\partial\alpha}$ der Steuerrichtung für eine bestimmte „Grundbahn“ $\alpha = \overline{\alpha}$, $x = \overline{x}(t)$, $y = \overline{y}(t)$, $\varphi = \overline{\varphi}(t)$ darzustellen? Die variierenden Bahnen der Schar α sollen alle *gleichzeitig* mit der Grundbahn für $t = t_1 = \text{const}$ vom Punkte 1 ausgehen, um dann (früher oder später) zur Zeit $t = \tau(\alpha)$ im Zielpunkte 2 einzumünden, wobei sie in ihrem ganzen Verlaufe überall den „Steuergleichungen“ (1) genügen müssen. Durch Differentiation nach α finden wir

$$\frac{\partial x}{\partial\alpha} = \frac{\partial y}{\partial\alpha} = 0 \quad \text{für} \quad t = t_1 \tag{7}$$

und

$$\frac{\partial x}{\partial\alpha} + \frac{\partial x}{\partial t}\cdot\frac{d\tau}{d\alpha} = \frac{\partial y}{\partial\alpha} + \frac{\partial y}{\partial t}\cdot\frac{d\tau}{d\alpha} = 0 \quad \text{für} \quad t = \tau(\alpha)\ , \tag{8}$$

und dies gilt für alle Werte α, insbesondere auch für $\alpha = \overline{\alpha}$, der unserer Grundbahn entspricht. Hier sind die Ausdrücke $\dfrac{\partial x}{\partial t} = \overline{x}'$, $\dfrac{\partial y}{\partial t} = \overline{y}'$ die (tatsächlichen) Geschwindigkeitskomponenten der Grundbahn, während die „Variationen“ $\dfrac{\partial x}{\partial\alpha}$, $\dfrac{\partial y}{\partial\alpha}$ für $\alpha = \overline{\alpha}$ zur Abkürzung mit $\xi(t)$, $\eta(t)$ bezeichnet werden sollen. Dann erhalten wir als „Grenzbedingungen“ für ξ und η aus (7) und (8) die folgenden

$$\xi(t_1) = \eta(t_1) = 0\ , \tag{7'}$$

$$\xi(t_2) + \overline{x}_2'\frac{d\tau}{d\alpha} = \eta(t_2) + \overline{y}_2'\frac{d\tau}{d\alpha} = 0 \tag{8'}$$

und hieraus

$$\overline{y}_2'\xi(t_2) - \overline{x}_2'\eta(t_2) = 0\ . \tag{8''}$$

116 | Außerdem müssen die Funktionen $\xi(t)$, $\eta(t)$ zusammen mit $\omega = \overline{\dfrac{\partial\varphi}{\partial\alpha}}$ (für $\alpha = \overline{\alpha}$) im ganzen Intervall (t_1, t_2) den linearen, nicht homogenen Differentialgleichungen (2) genügen in der Form

$$\left.\begin{aligned} \frac{d\xi}{dt} - \overline{u}_x\xi - \overline{u}_y\eta &= -\omega\sin\overline{\varphi} \\ \frac{d\eta}{dt} - \overline{v}_x\xi - \overline{v}_y\eta &= \omega\cos\overline{\varphi} \end{aligned}\right\} \tag{2'}$$

We shall now choose these constants so that the right side of eq⟦uation⟧ (3) becomes a square by setting

$$\frac{\partial\varphi}{\partial\alpha} = -\lambda\sin\varphi + \mu\cos\varphi = \omega\ . \tag{6}$$

But how must a function $\omega(t)$ be constituted in order to represent, under the given conditions, a "permissible variation" $\dfrac{\partial\varphi}{\partial\alpha}$ of the steering direction for a particular "basic path" $\alpha = \overline{\alpha}$, $x = \overline{x}(t)$, $y = \overline{y}(t)$, $\varphi = \overline{\varphi}(t)$? The varying paths of the family α are all supposed to start from the point 1 *simultaneously* with the basic path for $t = t_1 = \text{const}$ in order to lead (sooner or later) to the destination point 2 at time $t = \tau(\alpha)$, while having to satisfy everywhere the "steering equations" (1) over their entire course. By differentiation with respect to α we find

$$\frac{\partial x}{\partial\alpha} = \frac{\partial y}{\partial\alpha} = 0 \quad \text{for} \quad t = t_1 \tag{7}$$

and

$$\frac{\partial x}{\partial\alpha} + \frac{\partial x}{\partial t}\cdot\frac{d\tau}{d\alpha} = \frac{\partial y}{\partial\alpha} + \frac{\partial y}{\partial t}\cdot\frac{d\tau}{d\alpha} = 0 \quad \text{for} \quad t = \tau(\alpha)\ , \tag{8}$$

and this is valid for *all* values α, in particular also for $\alpha = \overline{\alpha}$, which corresponds to our basic path. Here, the expressions $\dfrac{\partial x}{\partial t} = \overline{x}'$, $\dfrac{\partial y}{\partial t} = \overline{y}'$ are the (actual) velocity components of the basic path, while the "variations" $\dfrac{\partial x}{\partial\alpha}$, $\dfrac{\partial y}{\partial\alpha}$ for $\alpha = \overline{\alpha}$ shall be denoted by $\xi(t)$, $\eta(t)$ for short. We then obtain as "boundary conditions" for ξ and η from (7) and (8) the following

$$\xi(t_1) = \eta(t_1) = 0\ , \tag{7'}$$

$$\xi(t_2) + \overline{x}'_2\frac{d\tau}{d\alpha} = \eta(t_2) + \overline{y}'_2\frac{d\tau}{d\alpha} = 0 \tag{8'}$$

and from this

$$\overline{y}'_2\xi(t_2) - \overline{x}'_2\eta(t_2) = 0\ . \tag{8''}$$

Furthermore, the functions $\xi(t)$, $\eta(t)$ along with $\omega = \overline{\dfrac{\partial\varphi}{\partial\alpha}}$ (for $\alpha = \overline{\alpha}$) must satisfy in the entire interval (t_1, t_2) the linear, non-homogeneous differential equations (2) in the form

$$\left.\begin{aligned} \frac{d\xi}{dt} - \overline{u}_x\xi - \overline{u}_y\eta &= -\omega\sin\overline{\varphi} \\ \frac{d\eta}{dt} - \overline{v}_x\xi - \overline{v}_y\eta &= \omega\cos\overline{\varphi} \end{aligned}\right\} \tag{2'}$$

and are uniquely determined on the basis of them by ω with the initial conditions (7'). They do not, however, satisfy the further boundary condition (8'')

und bestimmen sich eindeutig aus ihnen durch ω mit den Anfangsbedingungen $(7')$. Sie erfüllen aber nicht von selbst die weitere Grenzbedingung $(8'')$, sondern nur, wenn ω passend gewählt ist. Setzen wir aber gemäß (6) und (5)

$$\begin{aligned}\omega &= -\lambda \sin\overline{\varphi} + \mu\cos\overline{\varphi} \\ &= c_1(-\lambda_1 \sin\overline{\varphi} + \mu_1\cos\overline{\varphi}) + c_2(-\lambda_2\sin\overline{\varphi} + \mu_2\cos\overline{\varphi}) = c_1\omega_1 + c_2\omega_2\ ,\end{aligned}$$

so ergeben sich entsprechend

$$\xi = c_1\xi_1 + c_2\xi_2\ , \quad \eta = c_1\eta_1 + c_2\eta_2\ ,$$

und $(8'')$ kann durch geeignete nicht verschwindende c_1, c_2 erfüllt werden. Diese Bedingungen $(2')$, $(7')$, $(8'')$ sind aber, wie aus den Existenzsätzen für partielle Differentialgleichungen geschlossen werden kann, auch *hinreichend* für die Existenz einer Schar $x = x(t,\alpha)$, $y = y(t,\alpha)$, die den Differentialgleichungen (1) nebst den Grenzbedingungen genügt und sich für $\alpha = \overline{\alpha}$ auf die Grundbahn $x = \overline{x}(t)$, $y = \overline{y}(t)$ reduziert. So ergibt sich aus (3) und (6) mit Hilfe von $(7')$ und $(8')$

$$\begin{aligned}\int_{t_1}^{t_2} (-\lambda\sin\overline{\varphi} + \mu\cos\overline{\varphi})^2\, dt &= \left[\lambda\frac{\partial x}{\partial\alpha} + \mu\frac{\partial y}{\partial\alpha}\right]_{t_1}^{t_2} \\ &= -\frac{d\tau}{d\alpha}\left(\lambda(t_2)\overline{x}_2' + \mu(t_2)\overline{y}_2'\right) = 0\ ,\end{aligned} \tag{9}$$

falls die untersuchte Grundbahn $\overline{x}(t)$, $\overline{y}(t)$, $\overline{\varphi}(t)$ wirklich ein Extremum der Fahrtdauer liefert und daher $\dfrac{d\tau}{d\alpha} = 0$ ist. In diesem Falle muß aber im ganzen Intervall gelten

$$-\lambda\sin\varphi + \mu\cos\varphi = 0 \tag{10}$$

(mit Weglassung der jetzt unnötigen Querstriche), oder

$$\lambda = \varrho\cos\varphi\,,\ \mu = \varrho\sin\varphi\,,\ \frac{\mu}{\lambda} = \operatorname{tg}\varphi\,,\ \varrho = \lambda\cos\varphi + \mu\sin\varphi\,. \tag{11}$$

Durch Differentiation von (10) erhalten wir dann wegen (4)

$$\begin{aligned}\varrho\frac{d\varphi}{dt} &= (\lambda\cos\varphi + \mu\sin\varphi)\frac{d\varphi}{dt} = -\frac{d\lambda}{dt}\sin\varphi + \frac{d\mu}{dt}\cos\varphi \\ &= (u_x\lambda + v_x\mu)\sin\varphi - (u_y\lambda + v_y\mu)\cos\varphi \\ &= \varrho\left(-u_y\cos^2\varphi + (u_x - v_y)\cos\varphi\sin\varphi + v_x\sin^2\varphi\right)\end{aligned}$$

und damit als die bestimmende Differentialgleichung unseres Variationsproblems die „Navigationsformel“

$$\frac{d\varphi}{dt} = -u_y\cos^2\varphi + (u_x - v_y)\cos\varphi\sin\varphi + v_x\sin^2\varphi\ , \tag{12}$$

on their own but only for a suitable choice of ω. However, if in accordance with (6) and (5) we set

$$\begin{aligned}\omega &= -\lambda \sin\overline{\varphi} + \mu\cos\overline{\varphi} \\ &= c_1(-\lambda_1 \sin\overline{\varphi} + \mu_1\cos\overline{\varphi}) + c_2(-\lambda_2\sin\overline{\varphi} + \mu_2\cos\overline{\varphi}) = c_1\omega_1 + c_2\omega_2\ ,\end{aligned}$$

we accordingly obtain

$$\xi = c_1\xi_1 + c_2\xi_2\ , \quad \eta = c_1\eta_1 + c_2\eta_2\ ,$$

and (8″) can be satisfied by suitable, non-vanishing c_1, c_2. But, as can be deduced from the existence theorems for partial differential equations, these conditions (2′), (7′), (8″) are also *sufficient* for the existence of a family $x = x(t,\alpha)$, $y = y(t,\alpha)$ which satisfies the differential equations (1) along with the boundary conditions and which is reduced for $\alpha = \overline{\alpha}$ to the basic path $x = \overline{x}(t)$, $y = \overline{y}(t)$. We thus obtain from (3) and (6) by means of (7′) and (8′)

$$\begin{aligned}\int_{t_1}^{t_2} (-\lambda\sin\overline{\varphi} + \mu\cos\overline{\varphi})^2\, dt &= \left[\lambda\frac{\partial x}{\partial\alpha} + \mu\frac{\partial y}{\partial\alpha}\right]_{t_1}^{t_2} \\ &= -\frac{d\tau}{d\alpha}\,(\lambda(t_2)\overline{x}_2' + \mu(t_2)\overline{y}_2') = 0\ ,\end{aligned} \tag{9}$$

if the examined basic path $\overline{x}(t)$, $\overline{y}(t)$, $\overline{\varphi}(t)$ really yields an extremum of the travel duration and hence $\dfrac{d\tau}{d\alpha} = 0$. In this case, however, we must have in the entire interval

$$-\lambda\sin\varphi + \mu\cos\varphi = 0 \tag{10}$$

(dropping the bars, which are now superfluous), or

$$\lambda = \varrho\cos\varphi\,,\ \mu = \varrho\sin\varphi\,,\ \frac{\mu}{\lambda} = \operatorname{tg}\varphi\,,\ \varrho = \lambda\cos\varphi + \mu\sin\varphi\,. \tag{11}$$

By differentiation of (10) we then obtain on account of (4)

$$\begin{aligned}\varrho\frac{d\varphi}{dt} &= (\lambda\cos\varphi + \mu\sin\varphi)\,\frac{d\varphi}{dt} = -\frac{d\lambda}{dt}\sin\varphi + \frac{d\mu}{dt}\cos\varphi \\ &= (u_x\lambda + v_x\mu)\sin\varphi - (u_y\lambda + v_y\mu)\cos\varphi \\ &= \varrho\left(-u_y\cos^2\varphi + (u_x - v_y)\cos\varphi\sin\varphi + v_x\sin^2\varphi\right)\end{aligned}$$

and thus as the determining differential equation of our variational problem the "navigation formula"

$$\frac{d\varphi}{dt} = -u_y\cos^2\varphi + (u_x - v_y)\cos\varphi\sin\varphi + v_x\sin^2\varphi\ , \tag{12}$$

welche sich im Falle $\varphi = 0$ (d. h. wenn wir vorübergehend die Abszissenachse in die augenblickliche Steuerrichtung legen) auf $\frac{d\varphi}{dt} = -u_y$ reduziert, also: *Das Steuer muß immer nach der Seite gedreht werden, nach welcher die gegen die Steuerrichtung wirkende Windkomponente* größer *wird.*

Die rechte Seite der Gleichung enthält nur Glieder mit partiellen Ableitungen nach den Raumkoordinaten, verschwindet also in dem Spezialfalle eines *gleichförmigen Windfeldes*, wo u, v von x, y unabhängig als Funktionen der *Zeit* gegeben sind. In diesem Falle ist die *Steuerrichtung* φ *auf jeder Extremalen konstant*, aber *Geraden* werden wegen (1) alle Extremalen *nur* dann, wenn das gleichförmige Windfeld auch in der *Zeit* unverändert bleibt.

Die Differentialgleichungen (1) und (12) bestimmen zusammen eine 3-parametrige Schar von Extremalen, so daß etwa zur Zeit $t = t_0$ von einem Punkte x_0, y_0 des Feldes eine Extremale mit vorgeschriebener Steuerrichtung $\varphi = \varphi_0$, aus jedem Punkte also ein (in der Zeit veränderliches) *Büschel* von Extremalen ausgeht.

117 | Für spätere Zwecke brauchen wir noch die Bestimmung der in (11) eingeführten Funktion

$$\varrho = \lambda \cos\varphi + \mu \sin\varphi$$

und wir finden mit Hilfe von (4) und (10)

$$\left.\begin{aligned} \frac{d\varrho}{dt} &= \frac{d\lambda}{dt}\cos\varphi + \frac{d\mu}{dt}\sin\varphi + (-\lambda\sin\varphi + \mu\cos\varphi)\frac{d\varphi}{dt} \\ &= -(u_x\lambda + v_x\mu)\cos\varphi - (u_y\lambda + v_y\mu)\sin\varphi \\ &= -\varrho\left[u_x\cos^2\varphi + (u_y + v_x)\cos\varphi\sin\varphi + v_y\sin^2\varphi\right] = -\varrho W\ . \end{aligned}\right\} \quad (13)$$

Hieraus folgt weiter

$$\varrho = Ce^{-\int W dt} \quad (13')$$

und die Funktion $\varrho = \varrho(t)$ kann auf keinem Extremalenstück ihr *Vorzeichen wechseln*, auf welchem alle partiellen Ableitungen von u, v und damit auch die Funktion W endlich bleibt.

§ 2. Die Konstruktion des Feldes und die hinreichenden Bedingungen des Extremums.

Nach dem Vorgang von *Weierstraß* versuchen wir, ein „Feld" von Extremalen zu konstruieren, das uns auch die *hinreichenden* Bedingungen für das Bestehen eines Maximums oder Minimums der Fahrzeit zwischen zwei gegebenen Punkten 1 und 2 liefern soll. Dabei ist aber in dem Falle, wo die Windkomponenten u, v auch von der Zeit t abhängen, nicht nur auf die *Zeit-*
Fig. 1 *dauer*, sondern auch auf die *Zeitlage* Rücksicht zu nehmen. Es ist nicht das

which reduces to $\dfrac{d\varphi}{dt} = -u_y$ for $\varphi = 0$ (i. e., when we temporarily lay the axis of abscissas in the momentary steering direction), and hence: *The steering wheel must always be turned to the side in which the wind component acting against the steering direction* increases.

The right side of the equation only contains terms with partial derivatives with respect to the space coordinates, and hence vanishes in the special case of a *uniform wind field*, where u, v are given independently of x, y as functions of *time*. In this case, the *steering direction* φ *is constant on every extremal*, but all extremals are turned into *straight lines* on account of (1) *only* if the uniform wind field also remains unaltered over *time*.

The differential equations (1) and (12) together determine a three-parameter family of extremals so that, for instance, an extremal with prescribed steering direction $\varphi = \varphi_0$ starts from a point x_0, y_0 of the field at time $t = t_0$, and hence a *cluster* of extremals (which is variable over time) starts from every point.

For later purposes, we also need to determine the function

$$\varrho = \lambda \cos\varphi + \mu \sin\varphi \,,$$

which was introduced in (11), and with the aid of (4) and (10), we find

$$\left.\begin{aligned} \frac{d\varrho}{dt} &= \frac{d\lambda}{dt}\cos\varphi + \frac{d\mu}{dt}\sin\varphi + (-\lambda\sin\varphi + \mu\cos\varphi)\,\frac{d\varphi}{dt} \\ &= -(u_x\lambda + v_x\mu)\cos\varphi - (u_y\lambda + v_y\mu)\sin\varphi \\ &= -\varrho\left[u_x\cos^2\varphi + (u_y + v_x)\cos\varphi\sin\varphi + v_y\sin^2\varphi\right] = -\varrho W\,. \end{aligned}\right\} \tag{13}$$

From this it in turn follows that

$$\varrho = Ce^{-\int W\,dt} \tag{13'}$$

and the function $\varrho = \varrho(t)$ cannot *change its sign* on any extremal segment on which all partial derivatives of u, v, and hence also the function W, remain finite.

§ 2. The construction of the field and the sufficient conditions for the extremum.

Following the procedure of *Weierstrass*, we try to construct a "field" of extremals that is supposed to yield also the *sufficient* conditions for the existence of a maximum or minimum of the travel time between two given points 1 and 2. But in the case where the wind components u, v also depend on time t we have to take into account not only the *temporal duration* but also the *temporal position*. It is not the same whether we, starting from 1,

Fig. 1

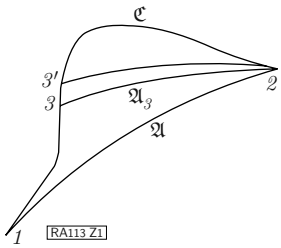


Abb. 1.

gleiche, ob wir, von 1 ausgehend, so *früh* wie möglich in 2 *eintreffen* oder so *spät* wie möglich von 1 *abreisen* wollen. Beide Probleme müssen also etwas verschieden behandelt werden. Ihre Analogie ist aber so ersichtlich, daß es hier genügen wird, uns auf das *erste* Problem, das des frühesten Eintreffens, zu beschränken. Wir gehen also aus von einer „extremalen Fahrt“ auf einer Bahnkurve $\mathfrak{A}$, welche zwischen den Zeitpunkten $t = t_1$ und $t = t_2$ überall den Differentialgleichungen (1) und (12) des Problems genügt, und vergleichen sie mit einer anderen „steuerbaren“ Fahrt auf einer „Vergleichskurve“ $\mathfrak{E}$, welche *gleichzeitig* mit der ersten zur Zeit $t = t_1$ in 1 startet und dann (früher oder später) das gleiche Ziel 2 erreicht.

Zu diesem Zwecke verbinden wir *jeden* Punkt 3 der Vergleichsbahn $\mathfrak{E}$ mit dem Zielpunkte 2 durch eine mit $\mathfrak{A}_3$ bezeichnete Extremale oder „Brachistochrone“ 3 2, welche *gleichzeitig* zur Zeit t_3 mit dem Vergleichsfahrzeug von 3 ausgeht und dann zur Zeit τ in 2 eintrifft. Indem wir nun einem Fahrgast von 1 zur Zeit t_1 auf dem Fahrzeug $\mathfrak{E}$ bis zum Punkte 3 mitnehmen und dann auf die Extremalbahn 3 2 *umsteigen* lassen, gewinnen wir einen *kontinuierlichen Übergang* zwischen den beiden zu vergleichenden Fahrten auf $\mathfrak{A}$ und $\mathfrak{E}$. Sollte sich nun dabei herausstellen, daß für zwei aufeinanderfolgende Umsteigestationen 3, 3′ mit $t_3 < t_3'$ die Fahrzeit über 1 3′ 2 immer größer ausfällt als die über 1 3 2, so würde daraus a fortiori folgen, daß auch die Gesamtfahrzeit auf $\mathfrak{E}$ zwischen 1 und 2 größer ist als die auf der Extremalen $\mathfrak{A}$. Zunächst handelt es sich aber um die *Ausführbarkeit* der geforderten Konstruktion.

Da die Integration der Differentialgleichung⟦en⟧ (1) und (12) auf drei Integrationskonstanten $\alpha_1, \alpha_2, \alpha_3$ führt,

$$x = f(t; \alpha_1, \alpha_2, \alpha_3)\ , \quad y = g(t; \alpha_1, \alpha_2, \alpha_3)\ , \quad \varphi = h(t; \alpha_1, \alpha_2, \alpha_3)\ ,$$

so wird man „im allgemeinen“ auch drei Bedingungen erfüllen, oder durch Einführung einer neuen Unbekannten τ die folgenden 4 Gleichungen befriedigen können

$$x_3 = f(t_3, \alpha_i)\ , \quad y_3 = g(t_3, \alpha_i)\ ; \qquad x_2 = f(\tau, \alpha_i)\ , \quad y_2 = g(\tau, \alpha_i)\ ,$$

die für $t_3 = t_1$, $\tau = t_2$ durch die Grundextremale $\mathfrak{A}$ schon erfüllt werden. Bei der kontinuierlichen Änderung von x, y, t auf $\mathfrak{E}$ würde diese Erfüllbarkeit nach

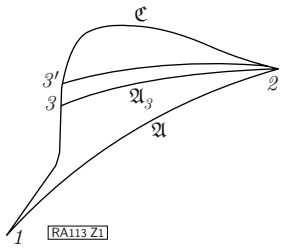


Fig. 1.

want to *arrive* at 2 as *early* as possible or whether we want to *leave* from 1 as *late* as possible. Thus, the two problems must be dealt with somewhat differently. But the analogy between the two is so obvious that we may restrict ourselves here to the *first* problem, the problem of the earliest arrival. Hence, we first posit an "extremal journey" on a path $\mathfrak{A}$ satisfying the differential equations (1) and (12) of the problem everywhere between the points in time $t = t_1$ and $t = t_2$, and compare it with another "steerable" journey on a "comparison curve" $\mathfrak{E}$ that starts from 1 *simultaneously* with the former at time $t = t_1$ and then arrives at the same destination 2 (*earlier or later* than the former).

To this end, we join *each* point 3 of the comparison path $\mathfrak{E}$ with the terminal point 2 by an extremal or "brachistochrone curve" 3 2, denoted by $\mathfrak{A}_3$, which starts from 3 *simultaneously* with the comparison vehicle at time t_3 and arrives at 2 at time τ. By picking up a passenger at 1 at time t_1, giving him a ride in vehicle $\mathfrak{E}$ to point 3, and letting him *transfer* to the extremal path 3 2, we obtain a *continual transition* between the two journeys on $\mathfrak{A}$ and $\mathfrak{E}$, which are to be compared. Now, if it turned out that for two successive transfer stations 3, 3′ with $t_3 < t_3'$ the travel time via 1 3′ 2 is always greater than the travel time via 1 3 2, then it would follow a fortiori that also the total travel time along $\mathfrak{E}$ between 1 and 2 is greater than that along the extremal $\mathfrak{A}$. But first we need to consider the *feasibility* of the required construction.

Since the integration of the differential equation⟦s⟧ (1) and (12) leads to three integration constants $\alpha_1, \alpha_2, \alpha_3$,

$$x = f(t; \alpha_1, \alpha_2, \alpha_3)\ , \quad y = g(t; \alpha_1, \alpha_2, \alpha_3)\ , \quad \varphi = h(t; \alpha_1, \alpha_2, \alpha_3)\ ,$$

it will "in general" be possible to also satisfy three conditions, or, by introduction of a new unknown τ, the following 4 equations

$$x_3 = f(t_3, \alpha_i)\ , \quad y_3 = g(t_3, \alpha_i)\ ; \qquad x_2 = f(\tau, \alpha_i)\ , \quad y_2 = g(\tau, \alpha_i)\ ,$$

which for $t_3 = t_1$, $\tau = t_2$ are already satisfied by the basic extremal $\mathfrak{A}$. According to the laws of implicit functions, the continual change of x, y, t

den Gesetzen der impliziten Funktionen nicht aufhören, solange die „Funktionaldeterminante“ J der 4 rechten Seiten nach den Variablen $\tau, \alpha_1, \alpha_2, \alpha_3$ von Null verschieden bleibt. Ihr Verschwinden für einen Punkt 3 auf $\mathfrak{C}$ würde bedeuten, daß hier zwei Wurzelsysteme der Gleichungen und damit zwei Zwischenextremalen 3 2 „zusammenfallen“. Die Frage nach der Konstruierbarkeit des „Feldes“ hängt also auf das engste zusammen mit den „Doppelelementen“ unseres Extremalensystems, und ihre Beantwortung liefert uns die unserem
118 Problem angepaßte „*Jacobi*sche Bedingung“ | der „konjugierten Punkte“. Sie hängt immer mit der besonderen „Windverteilung“ u, v zusammen und kann nur in besonderen Fällen erschöpfend beantwortet werden. Hier soll darauf nicht näher eingegangen, sondern mit der Verwertung der vorausgesetzten Konstruktion für unser Variationsproblem fortgefahren werden. Nur folgendes werde noch bemerkt. Da auch die Funktionaldeterminante J im allgemeinen eine *stetige* Funktion der vorkommenden Elemente sein wird, so genügt es, ihr Nichtverschwinden auf der „Grundextremalen“ $\mathfrak{A}$ selbst festzustellen, um dann auch auf ihr Nichtverschwinden in einer gewissen „Umgebung“ schließen zu können.

Die „Feldextremalen“ $\mathfrak{A}_3$, die zu einer „steuerbaren Vergleichsbahn“ $\mathfrak{C}$ gehören, bilden eine einparametrige Schar

$$x = f(t, \alpha) \,, \quad y = g(t, \alpha) \,, \quad \varphi = h(t, \alpha) \,,$$

als deren Parameter die Zeit t_3 gewählt werden kann, zu der unser Fahrgast auf die Extremale umsteigt. Für jede solche Extremalenschar gilt aber wegen (3), (4), (10) die Beziehung

$$\frac{d}{dt}\Big(\lambda \frac{\partial x}{\partial \alpha} + \mu \frac{\partial y}{\partial \alpha}\Big) = 0 \tag{14}$$

oder

$$\lambda \frac{\partial x}{\partial \alpha} + \mu \frac{\partial y}{\partial \alpha} = \varrho\Big(\frac{\partial x}{\partial \alpha} \cos\varphi + \frac{\partial y}{\partial \alpha} \sin\varphi\Big) = \text{const} \tag{14'}$$

in dem Sinne, daß der Ausdruck links auf jeder einzelnen Extremalen $\mathfrak{A}_3$ in der Zeit konstant bleibt, also eine Funktion des Parameters $\alpha = t_3$ darstellt. Da nun alle diese Bahnen vom (veränderlichen) Punkte 3 ausgehen und alle (sukzessive) zur Zeit $t = \tau$ im festen Endpunkte 2 einlaufen, so gilt für jedes $\alpha = t_3$

$$\begin{aligned} &\varrho_3 \left(f_\alpha(t_3, \alpha) \cos\overline{\varphi}_3 + g_\alpha(t_3, \alpha) \sin\overline{\varphi}_3\right) \\ &\quad = \varrho_2 \left(f_\alpha(\tau, \alpha) \cos\overline{\varphi}_2 + g_\alpha(\tau, \alpha) \sin\overline{\varphi}_2\right) , \end{aligned} \tag{15}$$

wo $\overline{\varphi}_3$ und $\overline{\varphi}_2$ die Steuerrichtungen der Feldextremalen in den Punkten 3 und 2, sowie ϱ_3 und ϱ_2 die Werte der in (11) eingeführten und durch (13′) bestimmten Funktion $\varrho(t, \alpha)$ auf der Feldextremalen in 3 und 2 bedeuten, also

$$\varrho_3 = \varrho\,(t_3, \alpha) = \varrho\,(\alpha, \alpha) \,, \quad \varrho_2 = \varrho\,(\tau, \alpha) \,. \tag{16}$$

along $\mathfrak{C}$ would not put an end to this satisfiability, provided that the "functional determinant" J of the 4 right sides after the variables $\tau, \alpha_1, \alpha_2, \alpha_3$ remain different from zero. Its disappearance for a point 3 on $\mathfrak{C}$ would mean that two root systems of the equations, and hence two extremals between 3 and 2, "coincide" here. The question as to whether it is possible to construct the "field" is therefore intimately related to the "double elements" of our system of extremals, and its answer yields the "*Jacob*ian condition" of the "conjugate points" that is tailored to our problem. It is always related to the particular "wind distribution" u, v and can fully be answered only in special cases. We shall not dwell on this matter here but proceed to utilize the assumed construction for our variational problem. I shall only mention the following. Since the functional determinant J will generally be a *continuous* function of the occurring elements, it is sufficient to ascertain that it does not vanish on the "basic extremal" $\mathfrak{A}$ itself in order to be able to conclude that it does not vanish in a certain "neighborhood" either.

The "field extremals" $\mathfrak{A}_3$ belonging to a "steerable comparison path" $\mathfrak{C}$ form a one-parameter family

$$x = f(t, \alpha) \ , \quad y = g(t, \alpha) \ , \quad \varphi = h(t, \alpha) \ ,$$

for whose parameter we may choose the time t_3 at which our passenger transfers to the extremal. But, on account of (3), (4), (10), for every such family of extremals the following relation holds:

$$\frac{d}{dt}\Big(\lambda \frac{\partial x}{\partial \alpha} + \mu \frac{\partial y}{\partial \alpha}\Big) = 0 \tag{14}$$

or

$$\lambda \frac{\partial x}{\partial \alpha} + \mu \frac{\partial y}{\partial \alpha} = \varrho\Big(\frac{\partial x}{\partial \alpha} \cos\varphi + \frac{\partial y}{\partial \alpha} \sin\varphi\Big) = \text{const} \tag{14'}$$

in the sense that the expression on the left side remains constant over time on each individual extremal $\mathfrak{A}_3$, and hence represents a function of the parameter $\alpha = t_3$. Since all these paths start at the (variable) point 3 and all of them (successively) arrive at the fixed endpoint 2 at time $t = \tau$, we have for every $\alpha = t_3$

$$\varrho_3 \left(f_\alpha(t_3, \alpha) \cos\overline{\varphi}_3 + g_\alpha(t_3, \alpha) \sin\overline{\varphi}_3\right) \tag{15}$$

$$= \varrho_2 \left(f_\alpha(\tau, \alpha) \cos\overline{\varphi}_2 + g_\alpha(\tau, \alpha) \sin\overline{\varphi}_2\right) ,$$

where $\overline{\varphi}_3$ and $\overline{\varphi}_2$ denote the steering directions of the field extremals in points 3 and 2, and ϱ_3 and ϱ_2 the values of the function $\varrho(t, \alpha)$, which was introduced in (11) and is determined by (13′), on the field extremal in 3 and 2, and hence

$$\varrho_3 = \varrho\,(t_3, \alpha) = \varrho\,(\alpha, \alpha) \ , \quad \varrho_2 = \varrho\,(\tau, \alpha) \ . \tag{16}$$

Aus den Grenzbedingungen der Feldextremalen in 3 und 2

$$\left.\begin{aligned} x_3 = f(t_3,\alpha) = f(\alpha,\alpha)\ ,\ y_3 = g(t_3,\alpha) = g(\alpha,\alpha) \\ x_2 = f(\tau,\alpha)\ ,\qquad\qquad y_2 = g(\tau,\alpha) \end{aligned}\right\} \tag{17}$$

ergibt sich nun durch (totale) Differentiation nach α, wobei der Akzent die partielle Differentiation nach dem ersten, der Index α die nach dem zweiten Argument bezeichnet,

$$\left.\begin{aligned} \frac{dx_3}{d\alpha} = f_\alpha(\alpha,\alpha) + f'(\alpha,\alpha)\ ,\ \frac{dy_3}{d\alpha} = g_\alpha(\alpha,\alpha) + g'(\alpha,\alpha) \\ 0 = f_\alpha(\tau,\alpha) + f'(\tau,\alpha)\frac{d\tau}{d\alpha}\ ,\ 0 = g_\alpha(\tau,\alpha) + g'(\tau,\alpha)\frac{d\tau}{d\alpha} \end{aligned}\right\} \tag{18}$$

und wir erhalten unter Hinzuziehung der „Steuergleichungen“ (1)

$$\left.\begin{aligned} f_\alpha(t_3,\alpha) = f_\alpha(\alpha,\alpha) = \frac{dx_3}{d\alpha} - f'(\alpha,\alpha) = x'_3 - \overline{x}'_3 = \cos\varphi_3 - \cos\overline{\varphi}_3 \\ g_\alpha(t_3,\alpha) = g_\alpha(\alpha,\alpha) = \frac{dy_3}{d\alpha} - g'(\alpha,\alpha) = y'_3 - \overline{y}'_3 = \sin\varphi_3 - \sin\overline{\varphi}_3\ , \end{aligned}\right\} \tag{19}$$

wo $x'_3 = \dfrac{dx_3}{dt_3}$, $y'_3 = \dfrac{dy_3}{dt_3}$ die Geschwindigkeitskomponenten der Vergleichsbahn im (bewegten) Punkte 3 zur Zeit $t = t_3 = \alpha$, sowie $\overline{x}'_3 = f'(t_3,\alpha)$, $\overline{y}'_3 = g'(t_3,\alpha)$ die entsprechenden der Feldextremalen im nämlichen Raumzeitpunkte (x_3, y_3, α) bezeichnen. Ebenso erhalten wir im (festen) Endpunkte 2 aus (18)

$$\left.\begin{aligned} -\frac{d\alpha}{d\tau} f_\alpha(\tau,\alpha) = f'(\tau,\alpha) = \overline{x}'_2 = u(x_2,y_2,\tau) + \cos\overline{\varphi}_2 \\ = u_2(\tau) + \cos\overline{\varphi}_2 \\ -\frac{d\alpha}{d\tau} g_\alpha(\tau,\alpha) = g'(\tau,\alpha) = \overline{y}'_2 = v(x_2,y_2,\tau) + \sin\overline{\varphi}_2 \\ = v_2(\tau) + \sin\overline{\varphi}_2 \end{aligned}\right\} \tag{20}$$

119 | in der entsprechenden Bezeichnungsweise. So ergibt sich durch Einsetzung in (15)

$$\begin{aligned} \frac{d\tau}{d\alpha} &= \frac{\varrho_3}{\varrho_2} \cdot \frac{(\overline{x}'_3 - x'_3)\cos\overline{\varphi}_3 + (\overline{y}'_3 - y'_3)\sin\overline{\varphi}_3}{\overline{x}'_2\cos\overline{\varphi}_2 + \overline{y}'_2\sin\overline{\varphi}_2} \\ &= \frac{\varrho_3}{\varrho_2} \cdot \frac{1 - \cos(\overline{\varphi}_3 - \varphi_3)}{1 + u_2\cos\overline{\varphi}_2 + v_2\sin\overline{\varphi}_2}\ . \end{aligned} \tag{21}$$

Hier ist der Quotient

$$\frac{\varrho_3}{\varrho_2} = \frac{\varrho\,(t_3,\alpha)}{\varrho\,(\tau,\alpha)} = \frac{\varrho\,(\alpha,\alpha)}{\varrho\,(\tau,\alpha)}$$

From the boundary conditions for the field extremals in 3 and 2

$$\left.\begin{aligned} x_3 = f(t_3,\alpha) = f(\alpha,\alpha)\ ,\ y_3 = g(t_3,\alpha) = g(\alpha,\alpha) \\ x_2 = f(\tau,\alpha)\ ,\qquad\qquad y_2 = g(\tau,\alpha) \end{aligned}\right\} \tag{17}$$

we now obtain by (total) differentiation with respect to α, where the accent denotes the partial differentiation after the first argument, and the index α that after the second argument,

$$\left.\begin{aligned} \frac{dx_3}{d\alpha} = f_\alpha(\alpha,\alpha) + f'(\alpha,\alpha)\ ,\ \frac{dy_3}{d\alpha} = g_\alpha(\alpha,\alpha) + g'(\alpha,\alpha) \\ 0 = f_\alpha(\tau,\alpha) + f'(\tau,\alpha)\frac{d\tau}{d\alpha}\ ,\ 0 = g_\alpha(\tau,\alpha) + g'(\tau,\alpha)\frac{d\tau}{d\alpha} \end{aligned}\right\} \tag{18}$$

and, relying on the "steering equations" (1), we obtain

$$\left.\begin{aligned} f_\alpha(t_3,\alpha) = f_\alpha(\alpha,\alpha) = \frac{dx_3}{d\alpha} - f'(\alpha,\alpha) = x'_3 - \overline{x}'_3 = \cos\varphi_3 - \cos\overline{\varphi}_3 \\ g_\alpha(t_3,\alpha) = g_\alpha(\alpha,\alpha) = \frac{dy_3}{d\alpha} - g'(\alpha,\alpha) = y'_3 - \overline{y}'_3 = \sin\varphi_3 - \sin\overline{\varphi}_3\ , \end{aligned}\right\} \tag{19}$$

where $x'_3 = \dfrac{dx_3}{dt_3}$, $y'_3 = \dfrac{dy_3}{dt_3}$ denote the velocity components of the comparison path in the (moving) point 3 at time $t = t_3 = \alpha$, and $\overline{x}'_3 = f'(t_3,\alpha)$, $\overline{y}'_3 = g'(t_3,\alpha)$ denote the corresponding ones of the field extremal in the same space-time-point (x_3, y_3, α). Likewise, using the appropriate designations, we obtain from (18) at the (fixed) endpoint 2

$$\left.\begin{aligned} -\frac{d\alpha}{d\tau} f_\alpha(\tau,\alpha) = f'(\tau,\alpha) = \overline{x}'_2 = u(x_2,y_2,\tau) + \cos\overline{\varphi}_2 \\ = u_2(\tau) + \cos\overline{\varphi}_2 \\ -\frac{d\alpha}{d\tau} g_\alpha(\tau,\alpha) = g'(\tau,\alpha) = \overline{y}'_2 = v(x_2,y_2,\tau) + \sin\overline{\varphi}_2 \\ = v_2(\tau) + \sin\overline{\varphi}_2\ . \end{aligned}\right\} \tag{20}$$

We thus obtain by substitution in (15)

$$\begin{aligned} \frac{d\tau}{d\alpha} &= \frac{\varrho_3}{\varrho_2} \cdot \frac{(\overline{x}'_3 - x'_3)\cos\overline{\varphi}_3 + (\overline{y}'_3 - y'_3)\sin\overline{\varphi}_3}{\overline{x}'_2\cos\overline{\varphi}_2 + \overline{y}'_2\sin\overline{\varphi}_2} \\ &= \frac{\varrho_3}{\varrho_2} \cdot \frac{1 - \cos(\overline{\varphi}_3 - \varphi_3)}{1 + u_2\cos\overline{\varphi}_2 + v_2\sin\overline{\varphi}_2}\ . \end{aligned} \tag{21}$$

Here, the quotient

$$\frac{\varrho_3}{\varrho_2} = \frac{\varrho\,(t_3,\alpha)}{\varrho\,(\tau,\alpha)} = \frac{\varrho\,(\alpha,\alpha)}{\varrho\,(\tau,\alpha)}$$

wegen (13′) *wesentlich positiv*, ebenso auch der Zähler, der höchstens für $\varphi_3 = \overline{\varphi}_3$ verschwindet, wo $x'_3 = \overline{x}'_3$, $y'_3 = \overline{y}'_3$ wird und Feldextremale und Vergleichsbahn sich berühren. Das Vorzeichen von $\dfrac{d\tau}{d\alpha}$ wird also nur noch durch das des Nenners bestimmt und ist *positiv* jedenfalls in jedem „unbeschränkt manövrierbaren Felde“, in welchem überall die Windstärke $\sqrt{u^2+v^2} < 1$ *kleiner* ist als die Eigengeschwindigkeit des Fahrzeugs. Im anderen Falle $u^2+v^2 > 1$ gehören entsprechend den Gl. (1) zu jeder Bahnrichtung $x' : y'$ im allgemeinen zwei Steuerrichtungen $\varphi_{\mathrm{I}}, \varphi_{\mathrm{II}}$, von denen die eine einen *spitzen,* die andere einen *stumpfen* Winkel mit der Bahnrichtung einschließt, so daß der Ausdruck

$$x' \cos\varphi + y' \sin\varphi = 1 + u\cos\varphi + v\sin\varphi \tag{22}$$

im ersten Falle *positiv*, im zweiten *negativ* ausfällt. Nur im Grenzfalle, wo beide zusammenfallen, *verschwindet* der Ausdruck. Dieses Verhalten erhellt unmittelbar aus der nebenstehenden Abb. 2, wie sie für das im einfachsten Spezialfalle des ruhenden, homogenen Windfeldes durch R. v. *Mises* („Ztschr. f. Flugt. u. Motorluftschiffahrt, 1917, S. 145) zur Illustration der „Steuergleichungen“ angegeben wurde. Hier bezeichnen $\mathfrak{w}, \mathfrak{q}_{\mathrm{I}}, \mathfrak{q}_{\mathrm{II}}$ die Endpunkte der von 0 ausgehenden Vektoren, die der Windstärke u, v und den beiden in gleicher Richtung möglichen Bahngeschwindigkeiten q_1, q_2 zugehören, während die Vektoren $\mathfrak{w}\mathfrak{q}_1, \mathfrak{w}\mathfrak{q}_2$ den beiden Steuerrichtungen $\varphi_{\mathrm{I}}, \varphi_{\mathrm{II}}$ entsprechen. Im Grenzfalle, wo der Bahnvektor $\mathfrak{q}$ den Einheitskreis *berührt,* verschwindet das innere Produkt $x' \cos\varphi + y' \sin\varphi$ aus Bahn- und Steuervektor, und solche Bahnelemente (x, y, x', y') wollen wir als „singulär“ bezeichnen, während die übrigen „regulär“ heißen sollen. Dann sind also die „singulären“ Bahnelemente die einzigen, in denen der oben (21) gegebene Ausdruck für $\dfrac{d\tau}{d\alpha}$ sein Vorzeichen wechseln kann, und in einem (räumlich und zeitlich begrenzten oder unbegrenzten) „regulären“ Felde von Extremalen hat der Ausdruck ein *definites* Vorzeichen und ist überall *positiv* oder *negativ*, je nachdem die erzeugenden Extremalen dem „spitzwinkligen“ oder dem „stumpfwinkligen“ Typus zuge-

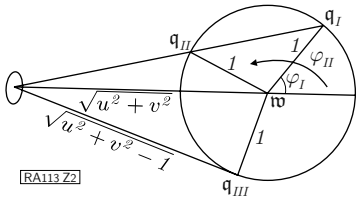


Abb. 2.

hören, d. h. je nach dem im Sinne oder entgegen der Bahnrichtung gesteuert wird. Nun ergibt aber bei unserer Konstruktion des Extremalenfeldes für jede

is *essentially positive* on account of (13′), as is the numerator, which, at most, vanishes for $\varphi_3 = \overline{\varphi}_3$, where $x'_3 = \overline{x}'_3$, $y'_3 = \overline{y}'_3$ and the field extremal and the comparison path make contact with one another. Hence, the sign of $\dfrac{d\tau}{d\alpha}$ is still determined only by that of the denominator and is *positive* at least in every "field maneuverable without limits" in which the wind strength $\sqrt{u^2+v^2} < 1$ is everywhere *smaller* than the airspeed of the vehicle. In the other case, where $u^2+v^2 > 1$, to every path direction $x' : y'$ generally belong two steering directions $\varphi_{\mathrm{I}}, \varphi_{\mathrm{II}}$ in accordance with the eq⟦uations⟧ (1) one of which encloses an *acute,* and the other of which an *obtuse* angle with the path direction, so that the expression

$$x' \cos\varphi + y' \sin\varphi = 1 + u\cos\varphi + v\sin\varphi \tag{22}$$

is *positive* in the first case and *negative* in the second case. The expression *vanishes* only in the borderline case where both coincide. This behavior is immediately suggested by the adjoining[1] figure 2, which was used by R. v. *Mises* (*1917*) for the simplest case of the homogeneous wind field at rest to illustrate the "steering equations". Here, $\mathfrak{w}, \mathfrak{q}_{\mathrm{I}}, \mathfrak{q}_{\mathrm{II}}$ denote the endpoints of the vectors that start from 0 and belong to the wind strength u, v and to the two path velocities q_1, q_2 possible in the same direction, while the vectors $\mathfrak{w}\mathfrak{q}_1, \mathfrak{w}\mathfrak{q}_2$ correspond to the two steering directions $\varphi_{\mathrm{I}}, \varphi_{\mathrm{II}}$. In the borderline case, where the path vector $\mathfrak{q}$ makes *contact* with the unit circle, the inner product $x' \cos\varphi + y' \sin\varphi$ of the path and steering vector⟦s⟧ vanishes, and such path elements (x, y, x', y') shall be called "singular", while the others are called "regular". Thus the "singular" path elements are the only ones in which the expression specified above (21) for $\dfrac{d\tau}{d\alpha}$ can change its sign, and, in a (spatially and temporally limited or unlimited) "regular" field of extremals, the expression has a *definite* sign and is everywhere *positive* or *negative*, depending on whether the generating extremals belong to the "acute-angled"

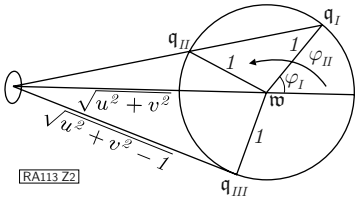


Fig. 2.

or "obtuse-angled" type, i. e., depending on whether one steers in the path direction or in the opposite direction. But the integration over the variable

[1] ⟦In contrast to the original paper, the figure follows below.⟧

innerhalb desselben verlaufende „Vergleichsbahn“ $\mathfrak{C}$ die Integration über die Variable t_3, die wir jetzt einfach durch t ersetzen, zwischen den Grenzen t_1 und t_2 die Beziehung

$$\begin{aligned}\tau_2 - t_2 = \tau(t_2) - \tau(t_1) &= \int_{t_1}^{t_2} \frac{d\tau(t)}{dt}\, dt \\ &= \int_{t_1}^{t_2} \frac{\varrho(t,t)}{\varrho(\tau,t)} \cdot \frac{1 - \cos(\overline{\varphi} - \varphi)}{1 + u_2 \cos\overline{\varphi}_2 + v_2 \sin\overline{\varphi}_2}\, dt\end{aligned} \tag{23}$$

und diese Differenz der beiden zwischen den Endpunkten 1 und 2 auf $\mathfrak{C}$ und $\mathfrak{A}$ gesteuerten Fahrtzeiten ist im „spitzwinkligen“ Falle immer positiv, im „stumpfwinkligen“ immer *negativ*. Im ersteren haben wir also ein *Minimum*, im letzteren ein *Maximum* der Fahrtdauer der extremalen Bahn gegenüber jeder ganz im Felde verlaufenden Vergleichsbahn. Der Fall des *Maximums*, also der Fall der „Megistochrone“, kann freilich nur eintreten innerhalb eines nur „beschränkt manövrierbaren“ Windfeldes, in dem die Windstärke überall größer ist als die Eigenbewegung. Innerhalb eines „unbeschränkt manövrierbaren“ Windfeldes dagegen liefert *jede* „extremale Bahn“ zwischen den durch „konjugierte Elemente“ gegebenen Grenzen ein *Minimum* der Fahrtdauer gegenüber allen im Felde verlaufenden Vergleichsbahnen.

§ 3. Das Problem im dreidimensionalen Raume.

Unsere bisherigen Betrachtungen waren auf die *Ebene* beschränkt, also bei der Luftfahrt auf eine Navigation in gleichbleibender Höhe. Nun variiert aber bekanntlich die Windverteilung außerordentlich stark mit der Höhenlage, ja noch stärker als mit den Ortskoordinaten, und dieser Tatsache pflegt auch die praktische Navigation durch geeignete Verwendung des „Höhensteuers“ Rechnung zu tragen. Es dürfte daher auch eine theoretische Behandlung des dreidimensionalen Problems nicht als überflüssig erscheinen, wenn auch bei der praktischen Aufwärts- und Abwärtsbewegung des Fahrzeuges, sofern
120 sie schnell erfolgt, | die vertikale Relativgeschwindigkeit nicht ohne weiteres der horizontalen gleichgesetzt werden kann. Jedenfalls soll im folgenden das mathematisch genau entsprechende Problem für den Raum, wie im vorausgehenden für die Ebene, behandelt werden. Dabei sind natürlich rein theoretisch alle Koordinaten gleichberechtigt; bei allen Hinweisen auf die Praxis soll aber immer die z-Richtung als die der „Vertikalen“ bevorzugt und dementsprechend jede Änderung des Winkels ϑ der „Steuerrichtung“ gegen die Vertikale als „Höhensteuerung“ bezeichnet werden im Gegensatze zu der durch φ bestimmten „Seitensteuerung“. Die *Methode* der mathematischen Behandlung wird im wesentlichen die gleiche sein wie im zweidimensionalen Falle. Doch werden wir bei der Herleitung der Differentialgleichungen ein wenig anders

t_3, which we simply replace by t, yields in our construction of the field of extremals for every "comparison path" $\mathfrak{C}$ lying within it between the boundaries t_1 and t_2 the relation

$$\begin{aligned}\tau_2 - t_2 = \tau(t_2) - \tau(t_1) &= \int_{t_1}^{t_2} \frac{d\tau(t)}{dt}\, dt \\ &= \int_{t_1}^{t_2} \frac{\varrho(t,t)}{\varrho(\tau,t)} \cdot \frac{1 - \cos(\overline{\varphi} - \varphi)}{1 + u_2 \cos\overline{\varphi}_2 + v_2 \sin\overline{\varphi}_2}\, dt\end{aligned} \tag{23}$$

and this difference of the two travel durations steered between the endpoints 1 and 2 along $\mathfrak{C}$ and $\mathfrak{A}$ is always positive in the "acute-angled" case, and always negative in the "obtuse-angled" case. Thus, in the first case we have a *minimum*, and in the latter case a *maximum*, of the travel duration of the extremal path with respect to every comparison path lying entirely within the field. The case of the *maximum*, that is, the case of the "megistochrone" curve, can of course only occur within a wind field that is only "maneuverable to a limited extent", in which the wind strength is everywhere greater than the airspeed. Within a wind field that is "maneuverable without limits", by contrast, *every* "extremal path" between the boundaries given by "conjugate elements" yields a *minimum* of the travel duration with respect to all comparison paths lying within the field.

§ 3. The problem in three-dimensional space.

Our previous considerations were limited to the *plane*, and hence, in aviation, to navigation at constant height. But, as is well-known, the wind distribution varies extraordinarily with altitude, and even more so than with space coordinates, and practical navigation allows for this fact by suitable use of the "elevator". Hence, a theoretical treatment of the three-dimensional problem ought not to seem superfluous, even if the vertical relative velocity cannot be equated with the horizontal relative velocity without further ado for the real upward-and-downward motion of the vehicle, provided it is a rapid motion. However, in what follows, we shall consider the problem in space that, mathematically speaking, precisely corresponds to the problem in the plane considered earlier. Of course, in this case, all coordinates are on a par from a purely theoretical point of view; but in all references to the practice, we shall always favor the z-direction as that of the "vertical" and, accordingly, refer to any change of the angle ϑ of the "steering direction" against the vertical as "altitude control" in contrast with the "side control" determined by φ. The *method* of the mathematical treatment will essentially be the same as in the two-dimensional case. But we shall proceed somewhat differently for the deduction of the differential equations in order to make the entire development

verfahren, um hier, wo die Formeln natürlich etwas länger werden, die ganze Entwicklung durchsichtiger zu machen. Es handelt sich dabei um die Einführung einer Funktionaldeterminante und einer „definiten“ oder „semi-definiten“ quadratischen Form, die sich hier als notwendig herausstellt, aber auch schon im ebenen Falle möglich und vielleicht auch zweckmäßig ist.

Durch Einführung der Windkomponenten u, v, w längs der Koordinatenachsen und der die „Steuerrichtung“ (oder die Relativbewegung) charakterisierenden Richtungswinkel φ, ϑ erhalten wir zunächst für jede „steuerbare“ Bewegung des Fahrzeuges die „Steuergleichungen“

$$\left.\begin{aligned} \frac{dx}{dt} &= u + \cos\vartheta\cos\varphi \\ \frac{dy}{dt} &= v + \cos\vartheta\sin\varphi \\ \frac{dz}{dt} &= w + \sin\vartheta\ , \end{aligned}\right\} \tag{24}$$

wo u, v, w gegebene differentiierbare Funktionen von x, y, z, t sein sollen.

Hieraus gewinnen wir durch „Variation“, d. h. durch partielle Differentiation δ (bei konstanter Zeit) nach einem beliebigen Parameter α einer Schar „steuerbarer Bahnen“ die „Variationsgleichungen“

$$\left.\begin{aligned} \frac{d\delta x}{dt} &= u_x\delta x + u_y\delta y + u_z\delta z - \cos\vartheta\sin\varphi\delta\varphi - \sin\vartheta\cos\varphi\delta\vartheta \\ \frac{d\delta y}{dt} &= v_x\delta x + v_y\delta y + v_z\delta z + \cos\vartheta\cos\varphi\delta\varphi - \sin\vartheta\sin\varphi\delta\vartheta \\ \frac{d\delta z}{dt} &= w_x\delta x + w_y\delta y + w_z\delta z + \cos\vartheta\delta\vartheta\ . \end{aligned}\right\} \tag{25}$$

Diese Differentialgleichungen bestimmen, wenn eine „steuerbare Grundbahn“ durch $x, y, z, \varphi, \vartheta$ als Funktionen von t gegeben ist, die Variationen $\delta x, \delta y, \delta z$ *eindeutig* durch die willkürlich vorgeschriebenen Variationen $\delta\varphi, \delta\vartheta$ und durch ihre „Anfangswerte“ bei $t = t_0$, die wir in den nächstfolgenden Betrachtungen als *verschwindend* ansehen wollen: $\delta x_0 = \delta y_0 = \delta z_0 = 0$. Diese Bestimmung erfolgt nach der Theorie der linearen Differentialgleichungen durch Ausdrücke der Form

$$\left.\begin{aligned} \delta x &= \int_{t_0}^{t} (X_1(t)\delta\varphi + X_2(t)\delta\vartheta)dt \\ \delta y &= \int_{t_0}^{t} (Y_1(t)\delta\varphi + Y_2(t)\delta\vartheta)dt \\ \delta z &= \int_{t_0}^{t} (Z_1(t)\delta\varphi + Z_2(t)\delta\vartheta)dt\ , \end{aligned}\right\} \tag{26}$$

more transparent here, where the formulas are a bit longer. This is a matter of introducing a functional determinant and of a "definite" or "semi-definite" quadratic form which turns out to be necessary in this case but is already possible, and perhaps even advantageous, in the planar case.

By introducing the wind components u, v, w alongside the coordinate axes and the azimuths φ, ϑ characterizing the "steering direction" (or the relative motion), we first obtain for every "steerable" motion of the vehicle the "steering equations"

$$\left.\begin{aligned} \frac{dx}{dt} &= u + \cos\vartheta\cos\varphi \\ \frac{dy}{dt} &= v + \cos\vartheta\sin\varphi \\ \frac{dz}{dt} &= w + \sin\vartheta\ , \end{aligned}\right\} \tag{24}$$

where u, v, w are supposed to be given differentiable functions of x, y, z, t.

From this we obtain by "variation", i. e., by partial differentiation δ (for constant time) with respect to some parameter α of a family of "steerable paths" the "variation equations"

$$\left.\begin{aligned} \frac{d\delta x}{dt} &= u_x\delta x + u_y\delta y + u_z\delta z - \cos\vartheta\sin\varphi\delta\varphi - \sin\vartheta\cos\varphi\delta\vartheta \\ \frac{d\delta y}{dt} &= v_x\delta x + v_y\delta y + v_z\delta z + \cos\vartheta\cos\varphi\delta\varphi - \sin\vartheta\sin\varphi\delta\vartheta \\ \frac{d\delta z}{dt} &= w_x\delta x + w_y\delta y + w_z\delta z + \cos\vartheta\delta\vartheta\ . \end{aligned}\right\} \tag{25}$$

If a "steerable basic path" is given by $x, y, z, \varphi, \vartheta$ as functions of t, these differential equations *uniquely* determine the variations $\delta x, \delta y, \delta z$ by means of the arbitrarily prescribed variations $\delta\varphi, \delta\vartheta$ and by means of their "initial values" for $t = t_0$, which we shall assume to be *vanishing* in the next considerations: $\delta x_0 = \delta y_0 = \delta z_0 = 0$. According to the theory of linear differential equations this determination is effected by expressions of the form

$$\left.\begin{aligned} \delta x &= \int_{t_0}^{t} (X_1(t)\delta\varphi + X_2(t)\delta\vartheta)dt \\ \delta y &= \int_{t_0}^{t} (Y_1(t)\delta\varphi + Y_2(t)\delta\vartheta)dt \\ \delta z &= \int_{t_0}^{t} (Z_1(t)\delta\varphi + Z_2(t)\delta\vartheta)dt\ , \end{aligned}\right\} \tag{26}$$

wobei die „unabhängigen Variationen“ $\delta\varphi, \delta\vartheta$, als Faktoren gegebener Zeitfunktionen unter den Integralzeichen erscheinen. Wie im ebenen Falle führen wir auch hier „Multiplikatoren“ λ, μ, ν ein, die den entsprechenden (linearen *und* homogenen) Differentialgleichungen genügen sollen

$$\left.\begin{aligned} \frac{d\lambda}{dt} + u_x\lambda + v_x\mu + w_x\nu = 0 \\ \frac{d\mu}{dt} + u_y\lambda + v_y\mu + w_y\nu = 0 \\ \frac{d\nu}{dt} + u_z\lambda + v_z\mu + w_z\nu = 0 \, , \end{aligned}\right\} \tag{27}$$

121 | durch die sie als Funktionen der Zeit bestimmt werden in der Form

$$\begin{aligned} \lambda &= c_1\lambda_1 + c_2\lambda_2 + c_3\lambda_3 \, , \\ \mu &= c_1\mu_1 + c_2\mu_2 + c_3\mu_3 \, , \\ \nu &= c_1\nu_1 + c_2\nu_2 + c_3\nu_3 \end{aligned} \tag{28}$$

mit den „willkürlichen Konstanten“ c_1, c_2, c_3. Für jedes solches Lösungssystem λ, μ, ν ergibt sich dann aus (25) und (27)

$$\left.\begin{aligned} \frac{d}{dt}(\lambda\delta x + \mu\delta y + \nu\delta z) = (-\lambda\cos\vartheta\sin\varphi + \mu\cos\vartheta\cos\varphi)\delta\varphi \\ + (-\lambda\sin\vartheta\cos\varphi - \mu\sin\vartheta\sin\varphi + \nu\cos\vartheta)\delta\vartheta \, , \end{aligned}\right\} \tag{29}$$

und es soll nun gezeigt werden, daß auf den gesuchten „Brachistochronen“ die beiden mit $\delta\varphi$ bzw. $\delta\vartheta$ multiplizierten Ausdrücke (bei geeigneter Wahl der Konstanten c_1, c_2, c_3) für variable $t > t_0$ *einzeln* verschwinden müssen. Zu diesem Zwecke wählen wir die „willkürlichen Variationen“ $\delta\varphi, \delta\vartheta$, die auch keiner „Grenzbedingung“ zu genügen brauchen, so, daß sie den mit ihnen multiplizierten Faktoren gleich werden:

$$\left.\begin{aligned} \delta\varphi = \Phi = (-\lambda\sin\varphi + \mu\cos\varphi)\cos\vartheta \\ \delta\vartheta = \Theta = -\lambda\cos\varphi\sin\vartheta - \mu\sin\varphi\sin\vartheta + \nu\cos\vartheta \, , \end{aligned}\right\} \tag{30}$$

und erhalten dabei durch Benutzung von (28) Ausdrücke der Form

$$\left.\begin{aligned} \delta\varphi = c_1\Phi_1 + c_2\Phi_2 + c_2\Phi_3 \\ \delta\vartheta = c_1\Theta_1 + c_2\Theta_2 + c_2\Theta_3 \, , \end{aligned}\right\} \tag{31}$$

wo sich die Paare Φ_i, Θ_i immer auf die entsprechenden Tripel λ_i, μ_i, ν_i in (28) beziehen. Ihnen entsprechen auf der linken Seite von (29) wegen (26) wieder Ausdrücke der Form

$$\begin{aligned} \delta x = \xi &= c_1\xi_1 + c_2\xi_2 + c_3\xi_3 \, , \\ \delta y = \eta &= c_1\eta_1 + c_2\eta_2 + c_3\eta_3 \, , \\ \delta z = \zeta &= c_1\zeta_1 + c_2\zeta_2 + c_3\zeta_3 \, , \end{aligned} \tag{32}$$

where the "independent variations" $\delta\varphi, \delta\vartheta$ appear as factors of given functions of time under the integral sign. Here, as in the planar case, we introduce "multipliers" λ, μ, ν that are supposed to satisfy the corresponding (linear *and* homogeneous) differential equations

$$\left.\begin{aligned} \frac{d\lambda}{dt} + u_x\lambda + v_x\mu + w_x\nu = 0 \\ \frac{d\mu}{dt} + u_y\lambda + v_y\mu + w_y\nu = 0 \\ \frac{d\nu}{dt} + u_z\lambda + v_z\mu + w_z\nu = 0\ , \end{aligned}\right\} \tag{27}$$

by means of which they are determined as functions of time in the form

$$\begin{aligned} \lambda &= c_1\lambda_1 + c_2\lambda_2 + c_3\lambda_3\ , \\ \mu &= c_1\mu_1 + c_2\mu_2 + c_3\mu_3\ , \\ \nu &= c_1\nu_1 + c_2\nu_2 + c_3\nu_3 \end{aligned} \tag{28}$$

with the "arbitrary constants" c_1, c_2, c_3. For every such system of solutions λ, μ, ν we then obtain from (25) and (27)

$$\left.\begin{aligned} \frac{d}{dt}(\lambda\delta x + \mu\delta y + \nu\delta z) = (-\lambda\cos\vartheta\sin\varphi + \mu\cos\vartheta\cos\varphi)\delta\varphi \\ + (-\lambda\sin\vartheta\cos\varphi - \mu\sin\vartheta\sin\varphi + \nu\cos\vartheta)\delta\vartheta\ , \end{aligned}\right\} \tag{29}$$

and it shall now be shown that on the sought "brachistochrone curves" the two expressions multiplied by $\delta\varphi$ and $\delta\vartheta$ resp. (for suitably chosen constants c_1, c_2, c_3) must vanish *individually* for variable $t > t_0$. To this end, we choose the "arbitrary variations" $\delta\varphi, \delta\vartheta$, which do not have to satisfy any "boundary condition" so that they become equal to the factors multiplied by them

$$\left.\begin{aligned} \delta\varphi = \Phi = (-\lambda\sin\varphi + \mu\cos\varphi)\cos\vartheta \\ \delta\vartheta = \Theta = -\lambda\cos\varphi\sin\vartheta - \mu\sin\varphi\sin\vartheta + \nu\cos\vartheta\ , \end{aligned}\right\} \tag{30}$$

and obtain by use of (28) expressions of the form

$$\left.\begin{aligned} \delta\varphi = c_1\Phi_1 + c_2\Phi_2 + c_2\Phi_3 \\ \delta\vartheta = c_1\Theta_1 + c_2\Theta_2 + c_2\Theta_3\ , \end{aligned}\right\} \tag{31}$$

where the pairs Φ_i, Θ_i always refer to the corresponding triples λ_i, μ_i, ν_i in (28). To these there in turn correspond on the left side of (29) on account of (26) expressions of the form

$$\begin{aligned} \delta x = \xi &= c_1\xi_1 + c_2\xi_2 + c_3\xi_3\ , \\ \delta y = \eta &= c_1\eta_1 + c_2\eta_2 + c_3\eta_3\ , \\ \delta z = \zeta &= c_1\zeta_1 + c_2\zeta_2 + c_3\zeta_3\ , \end{aligned} \tag{32}$$

wo die Tripel ξ_i, η_i, ζ_i als durch (26) gegebene Funktionen von t anzusehen sind. Durch Integration nach der Zeit erhalten wir dann aus (29) unter Verwendung von (28) und (32)

$$\begin{aligned}\lambda\xi + \mu\eta + \nu\zeta &= \int_{t_0}^{t} dt \left[(c_1\Phi_1 + c_2\Phi_2 + c_3\Phi_3)^2 + (c_1\Theta_1 + c_2\Theta_2 + c_3\Theta_3)^2\right] \\ &= Q(c_1, c_2, c_3; t) \,,\end{aligned} \tag{33}$$

als eine *homogene quadratische Form* der c_1, c_2, c_3, deren Koeffizienten gegebene Funktionen von t sind. Diese Form Q ist, wie aus der rechten Seite ersichtlich, *wesentlich positiv*, und zwar „semidefinit“ oder „definit“, je nachdem sie durch Wahl der Konstanten auf Null reduziert werden kann oder nicht.

Jetzt aber sei

$$x = \overline{x}(t) \,, \quad y = \overline{y}(t) \,, \quad z = \overline{z}(t) \tag{34}$$

eine *Brachistochrone* in dem hier gemeinten Sinne und

$$x = f(t; \alpha, \beta, \gamma) \,, \quad y = g(t; \alpha, \beta, \gamma) \,, \quad z = h(t; \alpha, \beta, \gamma) \tag{35}$$

eine dreiparametrige Schar benachbarter „steuerfähiger“ Bahnen, welche alle für $t = t_0$ von demselben Punkte x_0, y_0, z_0 ausgehen sollen und sich für $\alpha = \overline{\alpha}$, $\beta = \overline{\beta}$, $\gamma = \overline{\gamma}$ auf die „Grundbahn“ (34) reduzieren. Dann muß die aus den partiellen Ableitungen der Funktionen (35) nach den Parametern α, β, γ gebildete „Funktionaldeterminante“

$$D(t; \alpha, \beta, \gamma) = \begin{vmatrix} f_\alpha & f_\beta & f_\gamma \\ g_\alpha & g_\beta & g_\gamma \\ h_\alpha & h_\beta & h_\gamma \end{vmatrix} \tag{36}$$

auf der Brachistochrone $\alpha = \overline{\alpha}, \beta = \overline{\beta}, \gamma = \overline{\gamma}$ selbst überall, d. h. für variable t im ganzen betrachteten Intervall verschwinden:

$$\overline{D}(t) = D(t; \overline{\alpha}, \overline{\beta}, \overline{\gamma}) = 0 \,. \tag{37}$$

Denn wäre etwa $\overline{D}(t^*) \neq 0$ für $t^* > t_0$, so könnte man nach dem Satze der „impliziten Funktionen“ in der Umgebung der Stelle $t^*, \overline{x}(t^*), \overline{y}(t^*), \overline{z}(t^*)$ die Gl. (35) nach α, β, γ auflösen, wobei auch Wertsysteme $x = \overline{x}(t^*)$, $y = \overline{y}(t^*)$, $z = \overline{z}(t^*)$, $t \gtrless t^*$ in hinreichender Nähe *beliebig* vorgeschrieben werden könnten — und unsere „Grundbahn“ wäre keine Extremale oder Brachistochrone.
122 | Diese Eigenschaft (37) gilt insbesondere auch für die dreiparametrige Bahnenschar (35), die sich an unsere Brachistochrone (34) anschließt und den Bedingungen genügt

$$\overline{f}_\alpha = \xi_1(t) \,, \ \overline{g}_\alpha = \eta_1(t) \,, \ \overline{h}_\alpha = \zeta_1(t) \,, \ \overline{f}_\beta = \xi_2(t) \,, \ \overline{g}_\beta = \eta_2(t) \text{ usw.} \,, \tag{38}$$

where the triples ξ_i, η_i, ζ_i are to be considered functions of t given by (26). By integration with respect to time we then obtain from (29) by use of (28) and (32)

$$\begin{aligned}\lambda\xi + \mu\eta + \nu\zeta &= \int_{t_0}^{t} dt \left[(c_1\Phi_1 + c_2\Phi_2 + c_3\Phi_3)^2 + (c_1\Theta_1 + c_2\Theta_2 + c_3\Theta_3)^2\right] \\ &= Q(c_1, c_2, c_3; t)\end{aligned} \tag{33}$$

as a *homogeneous quadratic form* of the c_1, c_2, c_3 whose coefficients are given functions of t. This form Q is, as is evident from the right side, *essentially positive*, and in particular "semi-definite" or "definite", depending on whether or not they can be reduced to zero by choice of the constants.

But now let

$$x = \overline{x}(t)\ ,\quad y = \overline{y}(t)\ ,\quad z = \overline{z}(t) \tag{34}$$

be a *brachistochrone curve* in the sense intended here, and

$$x = f(t; \alpha, \beta, \gamma)\ ,\quad y = g(t; \alpha, \beta, \gamma)\ ,\quad z = h(t; \alpha, \beta, \gamma) \tag{35}$$

a three-parameter family of neighboring "steerable" paths all of which are supposed to start from the same point x_0, y_0, z_0 for $t = t_0$ and which reduce to the "basic path" (34) for $\alpha = \overline{\alpha}$, $\beta = \overline{\beta}$, $\gamma = \overline{\gamma}$. Then the "functional determinant" formed from the partial derivatives of the functions (35) with respect to the parameters α, β, γ

$$D(t; \alpha, \beta, \gamma) = \begin{vmatrix} f_\alpha & f_\beta & f_\gamma \\ g_\alpha & g_\beta & g_\gamma \\ h_\alpha & h_\beta & h_\gamma \end{vmatrix} \tag{36}$$

must vanish everywhere on the brachistrochrone curve $\alpha = \overline{\alpha}$, $\beta = \overline{\beta}$, $\gamma = \overline{\gamma}$ itself, i. e., for variable t in the entire interval under consideration:

$$\overline{D}(t) = D(t; \overline{\alpha}, \overline{\beta}, \overline{\gamma}) = 0\ . \tag{37}$$

For if we had, say, $\overline{D}(t^*) \neq 0$ for $t^* > t_0$, then, according to the theorem of "implicit functions", we could solve the eq⟦uation⟧ (35) with respect to α, β, γ in the neighborhood of the point $t^*, \overline{x}(t^*), \overline{y}(t^*), \overline{z}(t^*)$, where the system of values $x = \overline{x}(t^*)$, $y = \overline{y}(t^*)$, $z = \overline{z}(t^*)$, $t \gtrless t^*$, too, could be *arbitrarily* prescribed in sufficient proximity—and our "basic path" would not be an extremal or brachistochrone curve.

This property (37) holds in particular also for the three-parameter family of paths (35) that is attached to our brachistochrone curve (34) and satisfies the conditions

$$\overline{f}_\alpha = \xi_1(t)\ ,\ \overline{g}_\alpha = \eta_1(t)\ ,\ \overline{h}_\alpha = \zeta_1(t)\ ,\ \overline{f}_\beta = \xi_2(t)\ ,\ \overline{g}_\beta = \eta_2(t)\ \text{etc.}\ , \tag{38}$$

deren Existenz in der Theorie der partiellen Differentialgleichungen bewiesen werden kann. Dann folgt aber aus (37) für eine willkürliche Stelle $t = t^* > t_0$ auch die Existenz eines nicht verschwindenden Wertesystems c_1^*, c_2^*, c_3^* von der Eigenschaft

$$\left.\begin{aligned} c_1^*\xi_1(t^*) + c_2^*\xi_2(t^*) + c_3^*\xi_3(t^*) &= 0 \\ c_1^*\eta_1(t^*) + c_2^*\eta_2(t^*) + c_3^*\eta_3(t^*) &= 0 \\ c_1^*\zeta_1(t^*) + c_2^*\zeta_2(t^*) + c_3^*\zeta_3(t^*) &= 0\,, \end{aligned}\right\} \tag{39}$$

d. h. für diese Werte $c_1 = c_1^*$, $c_2 = c_2^*$, $c_3 = c_3^*$ verschwinden die „erlaubten Variationen“ ξ, η, ζ gleichzeitig an der Stelle $t = t^*$, und es ist daher hier wegen (33) weiter

$$\lambda^*\xi(t^*) + \mu^*\eta(t^*) + \nu^*\zeta(t^*) = Q(c_1^*, c_2^*, c_3^*; t^*) = 0\,, \tag{40}$$

d. h. unsere (positive) quadratische Form Q ist im ganzen Intervall der Brachistochrone *semi-definit*. Dann müssen aber auch in (33) für diese Werte c_1^*, c_2^*, c_3^* die beiden Quadrate unter dem Integralzeichen einzeln verschwinden, und wir erhalten für die entsprechenden Multiplikatoren $\lambda = \lambda^*$, $\mu = \mu^*$, $\nu = \nu^*$ gemäß (30) die Gleichungen

$$\left.\begin{aligned} \Phi &\equiv (-\lambda\sin\varphi + \mu\cos\varphi)\cos\vartheta = 0 \\ \Theta &\equiv -\lambda\sin\vartheta\cos\varphi - \mu\sin\vartheta\sin\varphi + \nu\cos\vartheta = 0 \end{aligned}\right\} \tag{41}$$

zusammen mit (24) als Bedingungsgleichungen der Brachistochrone zur Bestimmung von $x, y, z, \varphi, \vartheta, \lambda, \mu, \nu$ als Funktionen von t. Hier können wir aber die Multiplikatoren λ, μ, ν wieder eliminieren durch den Ansatz

$$\lambda = \varrho\cos\vartheta\cos\varphi\,,\ \mu = \varrho\cos\vartheta\sin\varphi\,,\ \nu = \varrho\sin\vartheta\,,\ \varrho = \sqrt{\lambda^2+\mu^2+\nu^2}\,, \tag{42}$$

der augenscheinlich die Gleichungen (41) identisch erfüllt. Es ergibt sich nämlich

$$\begin{aligned} -\varrho\frac{d\varrho}{dt} = -\lambda\frac{d\lambda}{dt} - \mu\frac{d\mu}{dt} - \nu\frac{d\nu}{dt} &= u_x\lambda^2 + (v_x+u_y)\lambda\mu + v_y\mu^2 \\ &+ (u_z+w_x)\lambda\nu + (v_z+w_y)\mu\nu + w_z\nu^2 \end{aligned}$$

oder wegen (42)

$$\left.\begin{aligned} -\frac{1}{\varrho}\frac{d\varrho}{dt} &= (u_z\cos^2\varphi + (u_y+v_x)\cos\varphi\sin\varphi + v_y\sin^2\varphi)\cos^2\vartheta \\ &+\big[(u_z+w_x)\cos\varphi + (v_z+w_y)\sin\varphi\big]\cos\vartheta\sin\vartheta \\ &+w_z\sin^2\vartheta \equiv R(t) \end{aligned}\right\} \tag{43}$$

und hieraus weiter

$$\varrho = Ce^{-\int_{t_0}^{t} R(t)dt} = \sqrt{\lambda^2+\mu^2+\nu^2} \tag{43'}$$

whose existence can be proved in the theory of partial differential equations. But then also follows from (37) for any point $t = t^* > t_0$ the existence of a non-vanishing system of values c_1^*, c_2^*, c_3^* with the property

$$\left.\begin{aligned} c_1^*\xi_1(t^*) + c_2^*\xi_2(t^*) + c_3^*\xi_3(t^*) &= 0 \\ c_1^*\eta_1(t^*) + c_2^*\eta_2(t^*) + c_3^*\eta_3(t^*) &= 0 \\ c_1^*\zeta_1(t^*) + c_2^*\zeta_2(t^*) + c_3^*\zeta_3(t^*) &= 0\,, \end{aligned}\right\} \tag{39}$$

i. e., for these values $c_1 = c_1^*$, $c_2 = c_2^*$, $c_3 = c_3^*$ the "permissible variations" ξ, η, ζ simultaneously vanish at the point $t = t^*$, and, on account of (33), furthermore

$$\lambda^*\xi(t^*) + \mu^*\eta(t^*) + \nu^*\zeta(t^*) = Q(c_1^*, c_2^*, c_3^*; t^*) = 0\,, \tag{40}$$

i. e., our (positive) quadratic form Q is *semi-definite* in the entire interval of the brachistochrone curve. But then the two squares under the integral sign must vanish individually also in (33) for these values c_1^*, c_2^*, c_3^*, and we obtain for the corresponding multipliers $\lambda = \lambda^*$, $\mu = \mu^*$, $\nu = \nu^*$ in accordance with (30) the equations

$$\left.\begin{aligned} \Phi &\equiv (-\lambda\sin\varphi + \mu\cos\varphi)\cos\vartheta = 0 \\ \Theta &\equiv -\lambda\sin\vartheta\cos\varphi - \mu\sin\vartheta\sin\varphi + \nu\cos\vartheta = 0 \end{aligned}\right\} \tag{41}$$

along with (24) as constraint equations of the brachistochrone curve for the determination of $x, y, z, \varphi, \vartheta, \lambda, \mu, \nu$ as functions of t. But here we can again eliminate the multipliers λ, μ, ν by means of the ansatz

$$\lambda = \varrho\cos\vartheta\cos\varphi\,,\ \mu = \varrho\cos\vartheta\sin\varphi\,,\ \nu = \varrho\sin\vartheta\,,\ \varrho = \sqrt{\lambda^2+\mu^2+\nu^2}\,, \tag{42}$$

which obviously identically satisfies the equations (41). For we have

$$\begin{aligned} -\varrho\frac{d\varrho}{dt} = -\lambda\frac{d\lambda}{dt} - \mu\frac{d\mu}{dt} - \nu\frac{d\nu}{dt} &= u_x\lambda^2 + (v_x+u_y)\lambda\mu + v_y\mu^2 \\ &+ (u_z+w_x)\lambda\nu + (v_z+w_y)\mu\nu + w_z\nu^2 \end{aligned}$$

or, on account of (42),

$$\left.\begin{aligned} -\frac{1}{\varrho}\frac{d\varrho}{dt} = (u_z\cos^2\varphi + (u_y+v_x)\cos\varphi\sin\varphi + v_y\sin^2\varphi)\cos^2\vartheta \\ +\big[(u_z+w_x)\cos\varphi + (v_z+w_y)\sin\varphi\big]\cos\vartheta\sin\vartheta \\ +w_z\sin^2\vartheta \equiv R(t) \end{aligned}\right\} \tag{43}$$

and from this furthermore

$$\varrho = Ce^{-\int_{t_0}^{t} R(t)dt} = \sqrt{\lambda^2+\mu^2+\nu^2} \tag{43'}$$

als eine Funktion, die in einem regulären Intervall (wo alle Ableitungen $u_x, u_y \ldots.$ endlich bleiben) weder verschwinden noch ihr Zeichen wechseln kann.

Die Einsetzung der Werte (42) in die Differentialgleichungen (27) ergibt dann vermöge einer einfachen Rechnung die beiden „Navigationsgleichungen“

$$\left.\begin{aligned} \frac{d\varphi}{d\vartheta} &= -u_y \cos^2\varphi + (u_x - v_y)\cos\varphi\sin\varphi + v_x \sin^2\varphi \\ &\quad +(w_x \sin\varphi - w_y\cos\varphi)\operatorname{tg}\vartheta \\ \frac{d\tau}{dt} &= -(u_z\cos\varphi + v_z\sin\varphi)\cos^2\vartheta \\ &\quad + \left[u_x\cos^2\varphi + (v_x + u_y)\cos\varphi\sin\varphi + v_y\sin^2\varphi - w_z\right]\cdot \\ &\quad \cdot\cos\vartheta\sin\vartheta + (w_x\cos\varphi + w_y\sin\varphi)\sin^2\vartheta\ , \end{aligned}\right\} \tag{44}$$

welche zusammen mit den „Steuergleichungen“ (24) die 5 Variablen $x, y, z, \varphi, \vartheta$ bestimmen als Funktionen der Zeit und der 5 „Anfangswerte“ $x_0, y_0, z_0, \varphi_0, \vartheta_0$ zur Zeit $t = t_0$.

In dem praktisch vielleicht wichtigen Sonderfalle, wo die Windkomponenten ausschließlich mit der Höhe z und mit der Zeit variieren, während die Vertikalkomponente verschwindet, ergibt sich

$$\frac{d\varphi}{dt} = 0\ , \quad \frac{d\vartheta}{dt} = -(u_z\cos\varphi + v_z\sin\varphi)\cos^2\vartheta$$

oder

$$\varphi = \text{const}\ , \quad \operatorname{tg}\vartheta_0 - \operatorname{tg}\vartheta = \cos\varphi\int_{t_0}^{t} u_z dt + \sin\varphi\int_{t_0}^{t} v_z dt = F(z,t)\ , \tag{45}$$

123 | d. h. die Horizontalkomponente der Steuerrichtung bleibt sich selbst parallel, während ihre Vertikalkomponente durch die jeweilige Höhenlage bestimmt wird.

Um nun auch hier *hinreichende* Bedingungen des Extremums zu finden, machen wir die *Weierstraßsche* Feldkonstruktion ganz analog wie im Falle des ebenen Problems, nur daß wir jetzt zur Abwechslung das Terminproblem der *spätesten Abreise* unserer Betrachtung zugrunde legen. Dann legen wir durch den Ausgangspunkt 1 ein dreiparametriges Bündel von Extremalbahnen, die alle zu verschiedenen Zeiten $t = \tau$ von 1 ausgehen, so daß jeder „Vergleichsbahn“ $\mathfrak{C}$ ein einparametrisches Büschel von Extremalen $\mathfrak{A}_\alpha$ entspricht, die von 1 nach dem variablen Punkte 3 auf $\mathfrak{C}$ laufen und *gleichzeitig* mit ihm zur Zeit $t_3 = \alpha$ dort eintreffen. Auch hier benutzen wir auf Grund von (29) und (41) die Haupteigenschaft der Extremalen

$$\lambda\frac{\partial x}{\partial\alpha} + \mu\frac{\partial y}{\partial\alpha} + \nu\frac{\partial z}{\partial\alpha} = \text{const} \tag{46}$$

as a function that can neither vanish nor change its sign in a regular interval (where all derivatives $u_x, u_y \ldots$. remain finite).

The substitution of the values (42) in the differential equations (27) then yields by virtue of a simple calculation the two "navigation equations"

$$\left.\begin{aligned} \frac{d\varphi}{d\vartheta} &= -u_y \cos^2\varphi + (u_x - v_y)\cos\varphi\sin\varphi + v_x \sin^2\varphi \\ &\quad +(w_x \sin\varphi - w_y \cos\varphi)\,\mathrm{tg}\,\vartheta \\ \frac{d\tau}{dt} &= -(u_z \cos\varphi + v_z \sin\varphi)\cos^2\vartheta \\ &\quad + \left[u_x \cos^2\varphi + (v_x + u_y)\cos\varphi\sin\varphi + v_y \sin^2\varphi - w_z\right]\cdot \\ &\quad \cdot \cos\vartheta \sin\vartheta + (w_x \cos\varphi + w_y \sin\varphi)\sin^2\vartheta\ , \end{aligned}\right\} \tag{44}$$

which together with the "steering equations" (24) determine the 5 variables $x, y, z, \varphi, \vartheta$ as functions of time and the 5 "initial values" $x_0, y_0, z_0, \varphi_0, \vartheta_0$ at time $t = t_0$.

In the case where the wind components vary solely with height z and with time, while the vertical component vanishes, which may be of practical significance, we have

$$\frac{d\varphi}{dt} = 0\ , \quad \frac{d\vartheta}{dt} = -(u_z \cos\varphi + v_z \sin\varphi)\cos^2\vartheta$$

or

$$\varphi = \text{const}\ , \quad \mathrm{tg}\,\vartheta_0 - \mathrm{tg}\,\vartheta = \cos\varphi \int_{t_0}^{t} u_z dt + \sin\varphi \int_{t_0}^{t} v_z dt = F(z,t)\ , \tag{45}$$

i. e., the horizontal component of the steering direction remains parallel to itself, while its vertical component is determined by the respective altitude.

In order to find *sufficient* conditions for the extremum also in this case we proceed with the *Weierstrass* field construction in a fashion entirely analogous to the case of the planar problem, except that, in this case, we base our considerations on the schedule problem of the *latest departure* for variety's sake. We then lay a three-parameter cluster of extremal paths through the starting point 1 so that all paths start at different times $t = \tau$ from 1 and so that to each "comparison curve" $\mathfrak{C}$ there corresponds a one-parameter cluster of extremals $\mathfrak{A}_\alpha$ that run from 1 to the variable point 3 along $\mathfrak{C}$ and *simultaneously* with it arrive there at time $t_3 = \alpha$. Here, too, on account of (29) and (41), we use the main property of extremals

$$\lambda \frac{\partial x}{\partial \alpha} + \mu \frac{\partial y}{\partial \alpha} + \nu \frac{\partial z}{\partial \alpha} = \text{const}\ , \tag{46}$$

und wenn wir für die Feldextremalen setzen

$$\left.\begin{aligned} &x = f(t,\alpha)\ , \quad y = g(t,\alpha)\ , \quad z = h(t,\alpha) \\ \text{und entsprechend}\quad &\lambda = \lambda(t,\alpha)\ , \quad \mu = \mu(t,\alpha)\ , \quad \nu = \nu(t,\alpha)\ , \end{aligned}\right\} \tag{47}$$

so erhalten wir für den Start in 1

$$x_1 = f(\tau,\alpha)\ , \quad y_1 = g(\tau,\alpha)\ , \quad z_1 = h(\tau,\alpha)\ , \tag{48}$$

also

$$\begin{aligned} 0 &= f_\alpha(\tau,\alpha) + f'(\tau,\alpha)\frac{d\tau}{d\alpha} \\ &= g_\alpha(\tau,\alpha) + g'(\tau,\alpha)\frac{d\tau}{d\alpha} + h_\alpha(\tau,\alpha) + h'(\tau,\alpha)\frac{d\tau}{d\alpha}\ ; \end{aligned} \tag{49}$$

und entsprechend für die Umsteigestation 3

$$x_3 = f(t_3,\alpha) = f(\alpha,\alpha)\ ,\ \ y_3 = g(\alpha,\alpha)\ ,\ \ z_3 = h(\alpha,\alpha)\ , \tag{50}$$

also

$$\begin{aligned} \frac{dx_3}{d\alpha} &= f_\alpha(\alpha,\alpha) + f'(\alpha,\alpha)\ , \\ \frac{dy_3}{d\alpha} &= g_\alpha(\alpha,\alpha) + g'(\alpha,\alpha)\ , \\ \frac{dz_3}{d\alpha} &= h_\alpha(\alpha,\alpha) + h'(\alpha,\alpha) \end{aligned} \tag{51}$$

oder, wenn wir wieder für die Geschwindigkeitskomponenten auf der Extremalen $\overline{x}', \overline{y}', \overline{z}'$ und für die auf der Vergleichsbahn $\mathfrak{C}$ einfach x', y', z' einführen, so haben wir

$$f_\alpha(\tau,\alpha) = -\overline{x}'_1\frac{d\tau}{d\alpha}\ ,\ \ g_\alpha(\tau,\alpha) = -\overline{y}'_1\frac{d\tau}{d\alpha}\ ,\ \ h_\alpha(\tau,\alpha) = -\overline{z}'_1\frac{d\tau}{d\alpha}\ , \tag{52}$$

sowie

$$f_\alpha(\alpha,\alpha) = x'_3 - \overline{x}'_3\ ,\ \ g_\alpha(\alpha,\alpha) = y'_3 - \overline{y}'_3\ ,\ \ h_\alpha(\alpha,\alpha) = z'_3 - \overline{z}'_3\ . \tag{53}$$

Dann ergibt sich aus (46):

$$\left.\begin{aligned} &\lambda(\tau,\alpha)f_\alpha(\tau,\alpha) + \mu(\tau,\alpha)g_\alpha(\tau,\alpha) + \nu(\tau,\alpha)h_\alpha(\tau,\alpha) \\ &= \lambda(\alpha,\alpha)f_\alpha(\alpha,\alpha) + \mu(\alpha,\alpha)g_\alpha(\alpha,\alpha) + \nu(\alpha,\alpha)h_\alpha(\alpha,\alpha) \end{aligned}\right\} \tag{54}$$

und durch Einsetzung der $f_\alpha, g_\alpha, h_\alpha$ aus (52) und (53):

$$\begin{aligned} &[-\lambda(\tau,\alpha)\overline{x}'_1 - \mu(\tau,\alpha)\overline{y}'_1 - \nu(\tau,\alpha)\overline{z}'_1]\,\frac{d\tau}{d\alpha} \\ &= \lambda(\alpha,\alpha)(x'_3 - \overline{x}'_3) + \mu(\alpha,\alpha)y'_3(y'_3 - \overline{y}'_3) + \nu(\alpha,\alpha)(z'_3 - \overline{z}'_3)\ , \end{aligned} \tag{55}$$

and if we set for the field extremals

$$\left.\begin{aligned} &x = f(t,\alpha)\ , \quad y = g(t,\alpha)\ , \quad z = h(t,\alpha) \\ \text{and, correspondingly,}\quad &\\ &\lambda = \lambda(t,\alpha)\ , \quad \mu = \mu(t,\alpha)\ , \quad \nu = \nu(t,\alpha)\ , \end{aligned}\right\} \tag{47}$$

we then obtain for the start at 1

$$x_1 = f(\tau,\alpha)\ , \quad y_1 = g(\tau,\alpha)\ , \quad z_1 = h(\tau,\alpha)\ , \tag{48}$$

and hence

$$\begin{aligned} 0 &= f_\alpha(\tau,\alpha) + f'(\tau,\alpha)\frac{d\tau}{d\alpha} \\ &= g_\alpha(\tau,\alpha) + g'(\tau,\alpha)\frac{d\tau}{d\alpha} + h_\alpha(\tau,\alpha) + h'(\tau,\alpha)\frac{d\tau}{d\alpha}\ ; \end{aligned} \tag{49}$$

and, correspondingly, for the transfer station 3

$$x_3 = f(t_3,\alpha) = f(\alpha,\alpha)\ ,\ \ y_3 = g(\alpha,\alpha)\ ,\ \ z_3 = h(\alpha,\alpha)\ , \tag{50}$$

and hence

$$\begin{aligned} \frac{dx_3}{d\alpha} &= f_\alpha(\alpha,\alpha) + f'(\alpha,\alpha)\ , \\ \frac{dy_3}{d\alpha} &= g_\alpha(\alpha,\alpha) + g'(\alpha,\alpha)\ , \\ \frac{dz_3}{d\alpha} &= h_\alpha(\alpha,\alpha) + h'(\alpha,\alpha)\ , \end{aligned} \tag{51}$$

or, if, once again, we simply introduce x', y', z' for the velocity components on the extremal $\overline{x}', \overline{y}', \overline{z}'$ and for those on the comparison curve $\mathfrak{C}$, then we obtain

$$f_\alpha(\tau,\alpha) = -\overline{x}'_1\frac{d\tau}{d\alpha}\ ,\ \ g_\alpha(\tau,\alpha) = -\overline{y}'_1\frac{d\tau}{d\alpha}\ ,\ \ h_\alpha(\tau,\alpha) = -\overline{z}'_1\frac{d\tau}{d\alpha}\ , \tag{52}$$

and also

$$f_\alpha(\alpha,\alpha) = x'_3 - \overline{x}'_3\ ,\ \ g_\alpha(\alpha,\alpha) = y'_3 - \overline{y}'_3\ ,\ \ h_\alpha(\alpha,\alpha) = z'_3 - \overline{z}'_3\ . \tag{53}$$

From (46) we then obtain

$$\left.\begin{aligned} &\lambda(\tau,\alpha)f_\alpha(\tau,\alpha) + \mu(\tau,\alpha)g_\alpha(\tau,\alpha) + \nu(\tau,\alpha)h_\alpha(\tau,\alpha) \\ &= \lambda(\alpha,\alpha)f_\alpha(\alpha,\alpha) + \mu(\alpha,\alpha)g_\alpha(\alpha,\alpha) + \nu(\alpha,\alpha)h_\alpha(\alpha,\alpha) \end{aligned}\right\} \tag{54}$$

And, by substitution of the $f_\alpha, g_\alpha, h_\alpha$ from (52) and (53),

$$\begin{aligned} &[-\lambda(\tau,\alpha)\overline{x}'_1 - \mu(\tau,\alpha)\overline{y}'_1 - \nu(\tau,\alpha)\overline{z}'_1]\,\frac{d\tau}{d\alpha} \\ &= \lambda(\alpha,\alpha)(x'_3 - \overline{x}'_3) + \mu(\alpha,\alpha)y'_3(y'_3 - \overline{y}'_3) + \nu(\alpha,\alpha)(z'_3 - \overline{z}'_3)\ , \end{aligned} \tag{55}$$

oder, wenn man gemäß (42) die λ, μ, ν durch $\varrho, \varphi, \vartheta$ ersetzt,

$$\left.\begin{aligned} &\frac{\varrho(\tau,\alpha)}{\varrho(\alpha,\alpha)}(\overline{x}'_1\cos\overline{\vartheta}_1\cos\overline{\varphi}_1+\overline{y}'_1\cos\overline{\vartheta}_1\sin\overline{\varphi}_1+\overline{z}'_1\sin\overline{\vartheta}_1)\frac{d\tau}{d\alpha}\\ &=(\overline{x}'_3-x'_3)\cos\overline{\vartheta}_3\cos\overline{\varphi}_3+(\overline{y}'_3-y'_3)\cos\overline{\vartheta}_3\sin\overline{\varphi}_3\\ &\quad+(\overline{z}'_3-z'_3)\sin\overline{\vartheta}_3\,. \end{aligned}\right\} \tag{56}$$

Durch Anwendung der „Steuergleichungen“ (24) auf die Extremalen in 1 und auf Extremale *und* Vergleichsbahn in 3 ergibt sich dann weiter $\overline{x}'_3 - x'_3 = \cos\overline{\vartheta}_3\cos\overline{\varphi}_3 - \cos\vartheta_3\cos\varphi_3$ usw., also

$$\left.\begin{aligned} &\tfrac{d\tau}{d\alpha}=\tfrac{\varrho(\alpha,\alpha)}{\varrho(\tau,\alpha)}\cdot\tfrac{1-\cos\vartheta_3\cos\varphi_3\cos\overline{\vartheta}_3\cos\overline{\varphi}_3-\cos\vartheta_3\sin\varphi_3\cos\overline{\vartheta}_3\sin\overline{\varphi}_3-\sin\vartheta_3\sin\overline{\vartheta}_3}{1+u_1(\tau)\cos\overline{\vartheta}_1\cos\overline{\varphi}_1+v_1(\tau)\cos\overline{\vartheta}_1\sin\overline{\varphi}_1+w_1(\tau)\sin\overline{\vartheta}_1}\\ &=\frac{\overline{\varrho}_3}{\overline{\varrho}_1}\cdot\frac{1-\cos\psi}{1+w_1^*} \end{aligned}\right\} \tag{57}$$

124 | wo ψ den Winkel zwischen den beiden „Steuerrichtungen“ in 3 (*nicht* den zwischen den Bahnrichtungen!) und w_1^* die Windkomponente in der Steuerrichtung der Feldextremalen in 1 zur Zeit τ bedeutet. Hier ist wieder wegen (43′) der Faktor $\frac{\overline{\varrho}_3}{\overline{\varrho}_1}$ wesentlich positiv, ebenso der Zähler, der nur verschwindet, wenn in 3 beide Steuerrichtungen und damit auch beide Bahnrichtungen zusammenfallen. Das Vorzeichen des Nenners aber ist gleichfalls positiv, solange die Windstärke in 1 kleiner ist als die Eigengeschwindigkeit, und hängt sonst nur von der Auswahl unter den beiden einer Bahnrichtung zugehörigen Steuerrichtungen und entsprechenden Feldextremalen ab, genau wie im ebenen Falle am Schlusse von § 2. Es ist immer positiv im „spitzwinkligen“, negativ im „stumpfwinkligen“ Falle, und durch Integration der Gl. (57) über die ganze Vergleichsbahn $\mathfrak{C}$ zwischen 1 und 2 erhält man, wenn man $\alpha = t_3$ durch t ersetzt,

$$\tau_2 - t_1 = \int_{t_1}^{t_2} \frac{\overline{\varrho}_3}{\overline{\varrho}_1}\cdot\frac{1-\cos\psi}{1+w_1^*}\,dt \quad \begin{matrix} >0 \\ <0 \end{matrix} \quad \begin{matrix} \text{(im spitzwinkligen bzw.} \\ \text{stumpfwinkligen Falle)} \end{matrix} \tag{58}$$

für die ersparte Zeit, wenn man, anstatt auf der Vergleichsbahn $\mathfrak{C}$ vielmehr auf der letzten zum Endpunkt 2 gehörenden Feldextremalen fährt, um pünktlich zur Zeit t_2 in 2 einzutreffen. Auch hier ergibt sich also wie im zweidimensionalen Falle je nach der Wahl der „spitzwinkligen“ oder „stumpfwinkligen“ Extremalen ein Minimum oder ein Maximum der Fahrzeit innerhalb des ganzen durch diese Extremalen einfach überdeckten Feldes.

or, replacing λ, μ, ν by $\varrho, \varphi, \vartheta$ in accordance with (42),

$$\left.\begin{aligned} &\frac{\varrho(\tau,\alpha)}{\varrho(\alpha,\alpha)}(\overline{x}_1' \cos\overline{\vartheta}_1 \cos\overline{\varphi}_1 + \overline{y}_1' \cos\overline{\vartheta}_1 \sin\overline{\varphi}_1 + \overline{z}_1' \sin\overline{\vartheta}_1)\frac{d\tau}{d\alpha} \\ &= (\overline{x}_3' - x_3') \cos\overline{\vartheta}_3 \cos\overline{\varphi}_3 + (\overline{y}_3' - y_3') \cos\overline{\vartheta}_3 \sin\overline{\varphi}_3 \\ &\quad + (\overline{z}_3' - z_3') \sin\overline{\vartheta}_3 . \end{aligned}\right\} \tag{56}$$

By application of the "steering equations" (24) to the extremals at 1 and to the extremal *and* comparison path at 3 we then obtain also $\overline{x}_3' - x_3' = \cos\overline{\vartheta}_3 \cos\overline{\varphi}_3 - \cos\vartheta_3 \cos\varphi_3$ etc., and hence

$$\left.\begin{aligned} \frac{d\tau}{d\alpha} &= \frac{\varrho(\alpha,\alpha)}{\varrho(\tau,\alpha)} \cdot \frac{1-\cos\vartheta_3\cos\varphi_3\cos\overline{\vartheta}_3\cos\overline{\varphi}_3 - \cos\vartheta_3\sin\varphi_3\cos\overline{\vartheta}_3\sin\overline{\varphi}_3 - \sin\vartheta_3\sin\overline{\vartheta}_3}{1+u_1(\tau)\cos\overline{\vartheta}_1\cos\overline{\varphi}_1 + v_1(\tau)\cos\overline{\vartheta}_1\sin\overline{\varphi}_1 + w_1(\tau)\sin\overline{\vartheta}_1} \\ &= \frac{\overline{\varrho}_3}{\overline{\varrho}_1} \cdot \frac{1-\cos\psi}{1+w_1^*} \end{aligned}\right\} \tag{57}$$

where ψ denotes the angle between the two "steering directions" at 3 (*not* the one between the path directions!) and w_1^* denotes the wind component in the steering direction of the field extremal at 1 at time τ. Here, too, the factor $\frac{\overline{\varrho}_3}{\overline{\varrho}_1}$ is essentially positive on account of (43′), as is the numerator, which only vanishes if the two steering directions, and hence also the two path directions, coincide at 3. But the sign of the denominator is also positive as long as the wind strength in 1 is smaller than the airspeed and otherwise depends only on the choice among the two steering directions belonging to a path direction and the corresponding field extremals, just as in the planar case at the end of § 2. It is always positive in the "acute-angled" and always negative in the "obtuse-angled" case. By integration of the eq⟦uation⟧ (57) over the entire comparison path $\mathfrak{C}$ between 1 and 2 we obtain, replacing $\alpha = t_3$ by t,

$$\tau_2 - t_1 = \int_{t_1}^{t_2} \frac{\overline{\varrho}_3}{\overline{\varrho}_1} \cdot \frac{1-\cos\psi}{1+w_1^*}\, dt \quad \begin{matrix} > 0 \\ < 0 \end{matrix} \quad \begin{matrix} \text{(in the acute-angled and} \\ \text{in the obtuse-angled case)} \end{matrix} \tag{58}$$

for the time saved, if one travels along the last field extremal belonging to the endpoint 2 instead of taking the comparison curve $\mathfrak{C}$ in order to punctually arrive at 2 at time t_2. Hence here, as in the two-dimensional case, we obtain a minimum or maximum of the travel time within the entire field simply covered by these extremals depending on the choice of the "acute-angled" or "obtuse-angled" extremals.

Introductory note to 1933a

Heinz-Dieter Ebbinghaus

Zermelo's *1933a* is a piece of elementary mathematics. It is included in this second volume as being "applied" for two *ad hoc* reasons: It gives a mathematical answer to a non-mathematical question, and it was submitted to the *Zeitschrift für angewandte Mathematik und Mechanik* (*ZAMM*). The question asks how a piece of sugar, considered as a rectangular prism, splits if one tries to break it by gripping it at two opposite corners. Zermelo first analyzes the physical conditions and then provides a mathematically simple solution together with generalizations resulting from replacing the cross-section rectangle by ellipses and parallelograms.

Because of its elementary character, *1933a* does not need commentary on the mathematical methods Zermelo uses. With its inclusion at the end of this volume, we on the other hand take this opportunity to consider two aspects of Zermelo's involvements with mathematics: (1) his fondness for applications and (2) his interest in communicating mathematics.

As to (1), the present volume shows that Zermelo's work in mathematics of a more applied character is concentrated at the beginning of his academic life and in a period of five years around 1930. The first phase ended when Zermelo, under the influence and the guidance of David Hilbert, turned to questions in the foundations of mathematics. At the end of the second period he became absorbed by discussions about the future of mathematical logic, positioning himself against the finitary approaches of Kurt Gödel, Thoralf Skolem, and others. However, as at other times Zermelo never neglected applied aspects.[1] When announcing a talk about navigation in the air for the annual meeting of the Deutsche Mathematiker-Vereinigung 1929 in Prague[2] he wrote to the chairman of the association:[3] "I still clearly remember our conversation in Breslau[4] where you urged me to take up the calculus of variations again instead of the fashionable foundational research. Well, you see that I am not a hopeless case where 'classical' mathematics is concerned, and that my old, even though hitherto mostly unhappy love for the 'applications' has secretly kept glowing."[5]

[1] One should remark here that the first draft of *Zermelo 1928* goes back to 1919.

[2] The extended abstract of Zermelo's talk is published as *Zermelo 1930c*.

[3] Letter to Adolf Kneser of 7 September 1929; Zermelo *Nachlass*, Universitätsarchiv Freiburg, sign. C 129/63.

[4] Now Wrocław, Poland.

[5] „Auch erinnere ich mich noch deutlich unseres Gesprächs in Breslau, wo sie mir eine Wiederaufnahme der Variationsrechnung anstelle der modischen Grundlagen-Forschung ans Herz zu legen suchten. Sie sehen also, daß bei mir noch nicht

H.-D. Ebbinghaus, A. Kanamori (Eds.), *Ernst Zermelo - Collected Works/ Gesammelte Werke II*, DOI 10.1007/978-3-540-70856-8_15, Schriften der Mathematisch-naturwissenschaftlichen Klasse der Heidelberger Akademie der Wissenschaften 23,

Zermelo's use of "applications" is to the point as the majority of his work of a more applied character, as is the case with *1933a*, starts from specific non-mathematical questions and provides mathematical models and solutions. One may refer here, for example, to *1902a* which arises from trying to calculate the behaviour of areas of low pressure in the atmosphere, to *1928* which asks for a method for evaluating the result of chess tournaments, or to *1930c* which concerns the optimal navigation of airships.

Zermelo submitted *1933a* to *ZAMM* in response to an invitation to contribute a paper to a volume commemorating the 50th anniversary of Richard von Mises, the founder of *ZAMM*. The letter of invitation from the editorial board[6] addresses Zermelo as "a 'pure' mathematician who, as a friend and colleague, was personally and scientifically close to von Mises."[7] Zermelo was undoubtedly inclined to accept. However, as mentioned above, at this time he was intensively involved in foundational questions. So he may not have felt disposed to work on a substantial paper of an applied character. Zermelo very likely had an early draft which he developed into *1933a*,[8] seasoning it with nice side remarks about a double coincidence with the purpose of the paper and its starting point: about the "wise" physicist and philosopher Gustav Theodor Fechner who, under the pseudonym "Dr. Mises", had published a collection of essays (his *Kleine Schriften*, *Fechner 1865*), one of which treated a likewise everyday question ("Why are sausages cut at an angle?") in a humorous way. Thanks to his broad education and interest in literature and philosophy, Zermelo surely knew Fechner's writings well.

As to (2), Zermelo's interest in communicating mathematics has several aspects. During his academic life he paid careful attention to the preparation of his lecture courses. Because of his nervous character he may not have been an ideal teacher for beginners; however, there are several assessments, in par-

---

Hopfen und Malz der ‚klassischen' Mathematik verloren ist und daß meine alte, wenn auch bisher meist unglückliche Liebe zu den ‚Anwendungen' im Stillen weiter geglommen ist."—Very probably the conversation with Kneser took part in 1913 when Zermelo had unsuccessfully applied for a full professorship of mathematics at the Technical University of Breslau. Astonishingly, one of the reasons for not taking him into consideration was his apparent lack of interest in technical questions—a clear misjudgment.

[6] Letter from Hans Reissner of 23 July 1932; Zermelo *Nachlass*, Universitätsarchiv Freiburg, sign. C 129/93.

[7] Zermelo and von Mises knew each other well. They had exchanged letters about technical questions concerning, for instance, monorail trains and steam and gas turbines. Von Mises judged Zermelo's *1930c* as an important paper as it "could not be without interest for practical aeronautics"; in *von Mises 1931* he shows that Zermelo's use of the calculus of variations may be avoided in the two-dimensional case. In turn, Zermelo submitted his extension of *1930c* to von Mises's *ZAMM* where it was published as *1931a*.

[8] Zermelo's *Nachlass* contains various notes on scattered mathematical problems which may stem from the period in question.

ticular by David Hilbert, which praise the quality of his courses for more advanced students. Furthermore Zermelo took part in the edition or translation of textbooks; examples are *Glazebrook 1897* and *Serret 1899, 1904*. At the end of his scientific life, around 1940, he made up his mind to bring mathematics to a larger public, namely by essentially anticipating—on a somewhat higher level—mathematical books of today (such as *Crilly 2008*) that aim at promoting interest in mathematics by presenting a variety of mathematical notions and results in an accessible way.[9]

In October 1940 he wrote down a list of sixteen topics which he planned to present in his *Mathematical miniatures—A collection of entertaining exercises for mathematicians and friends of mathematics*.[10] The majority of the miniatures come from the calculus of variations and applied mathematics;

[9] Perhaps, Zermelo was inspired by Hans Rademacher's and Otto Toeplitz's *Von Zahlen und Figuren* (*1930*). As is obvious from his *1934*, he knew this book well.

[10] *Mathematische Miniaturen. Eine Sammlung unterhaltender Aufgaben für Mathematiker und Freunde der Mathematik*; Zermelo *Nachlass*, Universitätsarchiv Freiburg, sign. C 129/278.

Über die Bruchlinien zentrierter Ovale. Wie zerbricht ein Stück Zucker?

1933a

In den „Vermischten Schriften des Dr. *Mises*“ aus dem Jahre 1875 hat G. Th. *Fechner* in einem humoristischen Aufsatze die Frage behandelt: warum man die Wurst schief anschneidet? Die Namensgleichheit mit dem Pseudonym jenes geistvollen Physikers und das Ausgehen von einer Frage des praktischen Lebens mögen es rechtfertigen, wenn ich in dieser Zeitschrift zu Ehren ihres Herausgebers, R. v. *Mises*, eine mathematische Verallgemeinerung der Frage nach der Bruchlinie eines Zuckerstücks behandle.

1. Wenn man ein prismatisches, quadratisches oder rechteckiges Zuckerstück mit der bloßen Hand zu halbieren versucht, so hält man es zweckmäßig an zwei diagonal entgegengesetzten Eckpunkten zwischen Daumen und Zeigefinger je einer Hand fest und probiert mit wechselndem und beständig verstärktem Druck zwischen beiden Händen, bis es endlich in der Mitte auseinanderbricht. Dabei zeigt sich, daß die geradlinige Bruchlinie zwar regelmäßig durch den Mittelpunkt O des Rechteckes, aber niemals längs der Diagonale, auch niemals parallel einer Rechteckseite verläuft, sondern in der Regel längs eines Durchmessers CD, der die größere Rechteckseite in einem Punkte C
169 schneidet, wo $AC = AO$, also AOC ebenso wie BOD ein | gleichschenkliges

among them are several that are drawn from papers included in this volume such as a proof of Poincaré's recurrence theorem ("The eternal recurrence according to Nietzsche and Poincaré: A theorem of dynamics"; from *1896a*), problems with inequalities in the calculus of variations ("Building roads in the mountains: Shortest lines of bounded steepness"; from *1902d*), the evaluation of a chess tournament ("Chess tournament: Evaluation of its participants—A maximality problem in the theory of probability"; from *1928*), and the solution of the navigation problem for airships ("Navigation in the air: Making use of winds in aeronautics"; from *1930c*).[11] Miniature 8 ("How does a piece of sugar break up? Breaking lines of a rectangle") rests on *1933a*. Zermelo sketched some parts, but never completed the book. If well crafted, it would have formed a collection of mathematical topics enriched with interesting observations reflecting a classical education and peppered with witty, ironical, and ingenious remarks—a collection reflecting a good part of Zermelo's personality.

[11] For the full list (in German) see *Ebbinghaus 2007*, 303.

On the lines of fracture of central ovals. How does a piece of sugar break up?

1933a

In a humorous paper included in the 1875 publication "Vermischte Schriften des Dr. *Mises*",[1] G. Th. *Fechner* addresses the following question: Why is a sausage cut diagonally? The homonymy with the pseudonym of that brilliant physicist and the origin from a question in everyday life may justify that in this journal in honor of its editor, R. v. *Mises*, I treat a mathematical generalization of the question about the break line of a piece of sugar.

1. If one attempts to divide a prismatic, square or rectangular piece of sugar into two halves using one's bare hands, it will be convenient to grip it between the thumb and the index finger of each hand at two diagonally opposite corners and to alternately apply constantly increasing pressure with both hands until it finally breaks apart in the middle. It turns out that while the rectilinear break line regularly passes through the center 0 of the rectangle, it never runs along the diagonal nor ever parallel to an edge of the rectangle but usually along a diameter CD intersecting the greater edge of the rectangle in a point C, where $AC = AO$, and hence AOC, as well as BOD,

[1] ⟦*Fechner 1875*.⟧

Dreieck ist. Die Erklärung dieser zunächst auffallenden Tatsache beruht auf dem Umstande, daß bei dem eingeschlagenen Verfahren der zu überwindende Widerstand der Bruchlänge $CD = 2l$, das wirksame Drehmoment aber der Länge des Lotes $AE = h$ proportional ist, und daß der Bruch bei dem kleinstmöglichen in A und B ausgeübten Kräftepaar p erfolgt. Es muß also, wenn k eine Konstante bedeutet, für ein minimales p

$$ph = 2kl$$

sein oder m. a. W., es muß der Quotient

$$\frac{l}{h} = \frac{OC}{AE}$$

durch Wahl von C möglichst klein gemacht werden. Es handelt sich demnach um die Bestimmung eines Dreieckes AOC, in welchem eine Seite $AO = d$ und ein anliegender Winkel $OAC = \alpha$ gegeben ist und in welchem das Verhältnis der gegenüberliegenden Seite OC zur Höhe AE möglichst klein sein soll (Abb. 2). Diese elementare Minimumsaufgabe führt, auf die gewöhnliche Weise behandelt, zu dem Ergebnis, daß die Winkel β und γ bei O und C einander gleich, das Dreieck also gleichschenklig sein muß. Es ist nämlich

$$OC = l = d\frac{\sin\alpha}{\sin\gamma}\,, \quad AE = h = d\sin\beta\,, \quad \text{also}$$

$$\frac{h}{l} = \frac{\sin\beta\sin\gamma}{\sin\alpha} = \frac{\cos(\beta-\gamma)-\cos(\beta+\gamma)}{2\sin\alpha} = \frac{\cos(\beta-\gamma)}{2\sin\alpha} + \frac{\cos\alpha}{2\sin\alpha}$$

$$\leqq \frac{1+\cos\alpha}{2\sin\alpha} = \frac{1}{2}\cot\frac{\alpha}{2}\,,$$

wo das =-Zeichen nur im Falle $\beta = \gamma$ gültig ist.

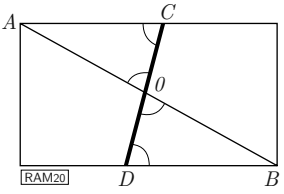


Abb. 1.

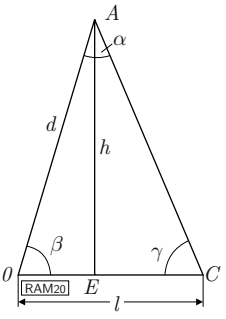


Abb. 2.

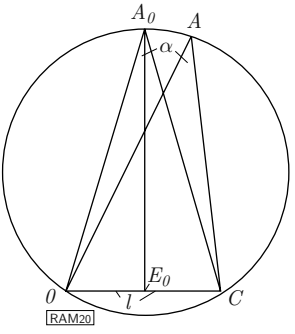


Abb. 3.

Fig. 4

is an isosceles triangle. The explanation of this fact, which is remarkable at first, rests on the fact that, given the method used here, the resistance to be overcome is proportional to the length of the break line $CD = 2l$, while the active torque is proportional to the length of the perpendicular $AE = h$, and that the break occurs for the smallest possible force pair p acting in A and B. Hence, for a minimal p, assuming that k denotes a constant, we must have

$$ph = 2kl$$

or, in other words, the quotient

$$\frac{l}{h} = \frac{OC}{AE}$$

must be made as small as possible through the choice of C. Thus, it is a matter of determining a triangle AOC in which one edge $AO = d$ and an adjacent angle $OAC = \alpha$ are given and in which the ratio of the opposite edge OC to the height AE is supposed to be as small as possible (Fig. 2). This elementary minimum problem, when dealt with in the usual way, leads to the result that the angles β and γ at O and C must be equal, and hence the triangle must be isosceles. For we have

$$OC = l = d\frac{\sin\alpha}{\sin\gamma}\,, \quad AE = h = d\sin\beta\,, \quad \text{and hence}$$

$$\frac{h}{l} = \frac{\sin\beta\sin\gamma}{\sin\alpha} = \frac{\cos(\beta-\gamma)-\cos(\beta+\gamma)}{2\sin\alpha} = \frac{\cos(\beta-\gamma)}{2\sin\alpha} + \frac{\cos\alpha}{2\sin\alpha}$$

$$\leqq \frac{1+\cos\alpha}{2\sin\alpha} = \frac{1}{2}\cot\frac{\alpha}{2}\,,$$

where the $=$-sign is only valid in the case $\beta = \gamma$.

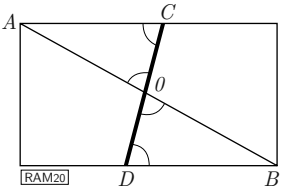


Fig. 1.

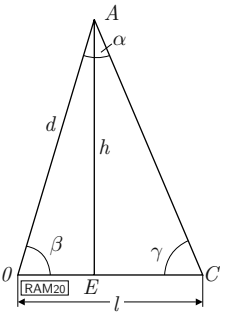


Fig. 2.

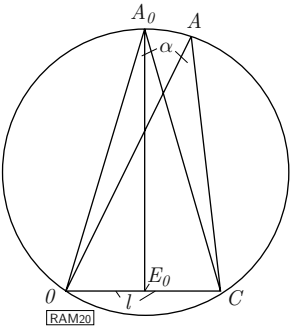


Fig. 3.

Fig. 4

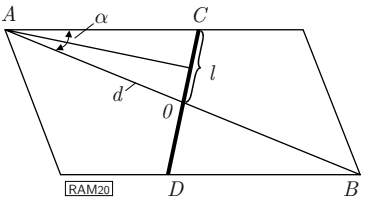


Abb. 4.

Zu dem gleichen Ergebnis gelangt man aber auch auf elementar-geometrischem Wege folgendermaßen (Abb. 3):

Die Spitzen A aller Dreiecke AOC mit *gemeinsamer* Basis $OC = l$ und dem gleichen Winkel $OAC = \alpha$ liegen auf einem über OC beschriebenen Kreisbogen, dessen *maximalen* Abstand $AE = h$ von der Basis man erhält, indem man über OC das Mittellot $E_0A_0 = h_0$ errichtet, das dem *gleichschenkligen* Falle entspricht und augenscheinlich auch das *Maximum* des Quotienten $\frac{h}{l}$ liefert.

Genau die gleichen Schlußfolgerungen ergeben sich, wenn an die Stelle des Rechteckes ein beliebiges *Parallelogramm* gesetzt wird. Auch hier zeigt es sich, daß das entsprechende $\triangle\ OAC$ über OC gleichschenklig mit A als Spitze sein muß (Abb. 4).

2. Jetzt betrachten wir eine beliebige *konvexe Figur*, mit Mittelpunkt dargestellt als Tangentengebilde in der Form

$$x\cos t + y\sin t = h(t) > 0 \quad (h(t) = h(t+\pi)) \tag{1}$$

wo t den Winkel der Normalen mit der x-Achse und t den Abstand der Tangente vom Mittelpunkt bezeichnet. Die Einhüllende der Tangentenschar ist dann bestimmt durch

$$-x\sin t + y\cos t = h'(t)\,, \tag{2}$$

also wird

$$\left.\begin{array}{l} x = h\cos t - h'\sin t \\ y = h\sin t + h'\cos t \end{array}\right\} \tag{3}$$

die Parameter-Darstellung der konvexen Kurve.

170 | Ein Durchmesser $AB = 2s$, der zum Parameter t gehört und dessen Lot mit der x-Achse den Winkel τ einschließt, wird dann bestimmt durch die Gleichung

$$\begin{gathered} x\cos\tau + y\sin\tau = 0\,, \quad \text{also} \\ h\cos(\tau - t) + h'\sin(\tau - t) = 0\,, \quad \text{oder} \\ \operatorname{tg}(\tau - t) = -\frac{h(t)}{h'(t)}\,, \end{gathered} \tag{4}$$

$$\text{d. h.} \qquad \frac{h'(t)}{h(t)} = -\cot(\tau - t) = \operatorname{tg}\left(\frac{\pi}{2} + \tau - t\right)\,.$$

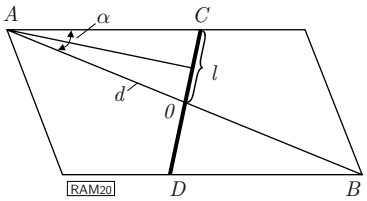


Fig. 4.

But the same result is also obtained in an elementary geometric way as follows (Fig. 3):

The vertices A of all triangles AOC with *common* basis $OC = l$ and the same angle $OAC = \alpha$ lie on an arc drawn through OC whose *maximal* distance $AE = h$ from the basis is obtained by constructing the middle perpendicular $E_0A_0 = h_0$ over OC that corresponds to the *isosceles* case and evidently also yields the *maximum* of the quotient $\dfrac{h}{l}$.

Precisely the same conclusions are obtained when the rectangle is replaced by any *parallelogram*. Here, too, it turns out that the corresponding $\triangle\, OAC$ through OC must be isosceles with A as vertex (Fig. 4).

2. We now consider any *convex figure* whose center is represented as a tangential structure in the form

$$x\cos t + y\sin t = h(t) > 0 \qquad (h(t) = h(t+\pi)) \tag{1}$$

where t denotes the angle of the normal with the x-axis, and $h(t)$[2] the distance of the tangent from the center. The envelope of the family of tangents is then determined by

$$-x\sin t + y\cos t = h'(t)\,, \tag{2}$$

and hence we have

$$\left.\begin{array}{l} x = h\cos t - h'\sin t \\ y = h\sin t + h'\cos t \end{array}\right\} \tag{3}$$

as the parametric representation of the convex curve.

A diameter $AB = 2s$ which belongs to the parameter t and whose perpendicular forms the angle τ with the x-axis is then determined by the equation

$$\begin{aligned} x\cos\tau + y\sin\tau = 0\,, &\quad \text{and hence} \\ h\cos(\tau - t) + h'\sin(\tau - t) = 0\,, &\quad \text{or} \\ \operatorname{tg}(\tau - t) = -\frac{h(t)}{h'(t)}\,, & \qquad (4) \\ \text{i. e.} \qquad \frac{h'(t)}{h(t)} = -\cot(\tau - t) = \operatorname{tg}\left(\frac{\pi}{2} + \tau - t\right)\,. & \end{aligned}$$

[2] [[Zermelo erroneously writes "t" instead of "$h(t)$".]]

Hier ist aber $\tau - t$ der Winkel, den die Tangente in A mit dem Durchmesser AB bildet.

Soll nun AB eine *Bruchlinie* unserer konvexen Figur darstellen, so muß das Kräftepaar in den Punkten P und P' eingesetzt werden, die von AB einen maximalen Abstand $PQ = P'Q'$ besitzen und daher, da in ihnen die Tangente (bzw. „Stützgerade“) der Richtung AB parallel ist, zum Parameter τ gehören. Außerdem muß für dieses Wertepaar t, τ der Quotient

$$u = \frac{OA}{PQ} = \frac{s(t)}{h(\tau)} = \frac{\sqrt{h(t)^2 + h'(t)^2}}{h(\tau)} \tag{5}$$

unter gleichzeitiger Gültigkeit von (4) einen *minimalen* Wert annehmen. Diese elementare Minimumsaufgabe ergibt nach dem gewöhnlichen Verfahren durch eine einfache Rechnung als notwendige Bedingung:

$$\frac{du}{dt} = \frac{(h(t) + h''(t))(h'(t)h(\tau) - h(t)h'(\tau))}{\sqrt{h(t)^2 + h'(t)^2 \cdot h(\tau)^2}} = 0\,, \tag{6}$$

also entweder

$$h(t) + h''(t) = 0 \tag{6a}$$

oder aber

$$\frac{h'(t)}{h(t)} = \frac{h'(\tau)}{h(\tau)}\,. \tag{6b}$$

Der erste Fall (6a) entspricht einem Minimum von $s(t) = \sqrt{h(t)^2 + h'(t)^2}$ des Durchmessers selbst und wird z. B. realisiert durch eine Ellipse oder eine beliebige *symmetrische* konvexe Figur mit Mittelpunkt, in welcher der größte und kleinste Durchmesser aufeinander senkrecht stehen und der Tangentenabstand vom Mittelpunkt immer kleiner ist als die große Halbachse (Abb. 6).

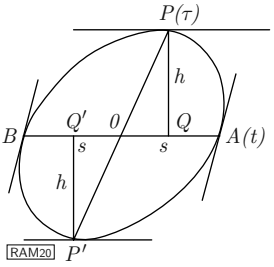


Abb. 5.

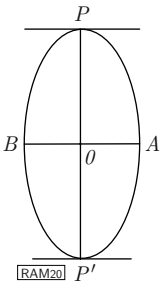


Abb. 6.

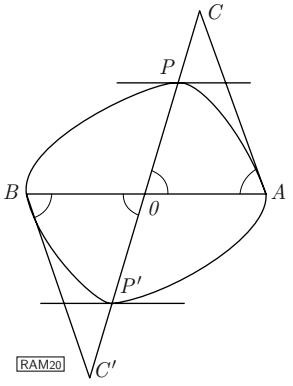


Abb. 7.

But here $\tau - t$ is the angle formed by the tangent in A with the diameter AB.

Now, if AB is to represent a *break line* of our convex figure, we must substitute the force pair in the points P and P' that have maximal distance $PQ = P'Q'$ from AB, and hence belong to the parameter τ since in P and P' the tangent (or "support line") is parallel to the direction AB. Moreover, for this pair of values t, τ, the quotient

$$u = \frac{OA}{PQ} = \frac{s(t)}{h(\tau)} = \frac{\sqrt{h(t)^2 + h'(t)^2}}{h(\tau)} \tag{5}$$

must assume a *minimal* value, given the simultaneous validity of (4). By an easy calculation, we obtain for this elementary minimum problem, following the usual method, as a necessary condition

$$\frac{du}{dt} = \frac{(h(t) + h''(t))(h'(t)h(\tau) - h(t)h'(\tau))}{\sqrt{h(t)^2 + h'(t)^2 \cdot h(\tau)^2}} = 0\,, \tag{6}$$

and hence either

$$h(t) + h''(t) = 0 \tag{6a}$$

or

$$\frac{h'(t)}{h(t)} = \frac{h'(\tau)}{h(\tau)}\,. \tag{6b}$$

The first case (6a) corresponds to a minimum of $s(t) = \sqrt{h(t)^2 + h'(t)^2}$ of the diameter itself and is realized by, e. g., an ellipse or any *symmetric* convex figure with center in which the greatest and smallest diameters are perpendicular to one another and the tangent distance from the center is always smaller than the major semi-axis (Fig. 6).

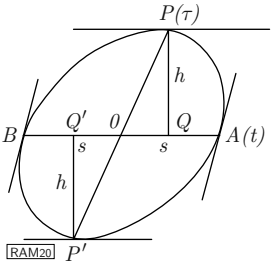


Fig. 5.

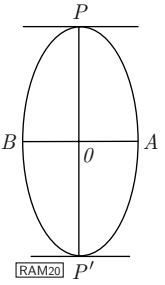


Fig. 6.

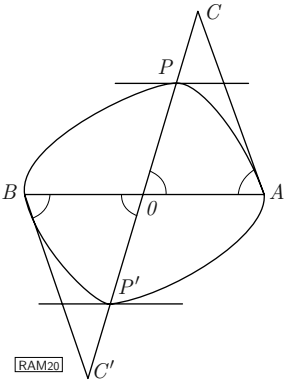


Fig. 7.

Der zweite Fall (6b) aber bedeutet nach (4), daß die Winkel in A und O einander gleich sind und so ein *gleichschenkliges* Dreieck OAC mit OA als Basis entsteht, ebenso aber auch entsprechend über OB ein gleichschenkliges Dreieck OBC', wie in Abb. 7 ersichtlich. Dieser Fall steht also in überraschender Analogie zu dem in Nr. 1 behandelten Spezialfalle des Parallelogrammes. Nur ist eben hier, wo es sich um kontinuierlich veränderliche Tangentenrichtungen handelt, das Minimum nicht mehr *eindeutig*, und es ist bei jeder Lösung besonders zu untersuchen, ob wirklich ein Minimum von $\frac{s}{h}$, eine mögliche Bruchlinie vorliegt.

However, according to (4), the second case (6b) means that the angles in A and O are equal so that we have an *isosceles* triangle OAC with OA as basis, but also, correspondingly, an isosceles triangle OBC' through OB, as is evident in Fig. 7. This case thus shows a surprising analogy to the special case of the parallelogram considered in no. 1. But in this case, which is concerned with continually changing tangent directions, the minimum is no longer *unique*, and it is necessary to separately examine for each solution whether there really obtains a minimum of $\frac{s}{h}$, a possible break line.

Bibliography

Abel, Emil

1906 Zur Theorie der elektromotorischen Kräfte in mehrphasigen und nicht-wässerigen einphasigen Systemen, *Zeitschrift für physikalische Chemie, Stöchiometrie und Verwandtschaftslehre* **56**, 612–623.

Alexandrow, Waldemar

1915 *Elementare Grundlagen für die Theorie des Maßes*, Ph. D. Thesis, University of Zurich.

Aref, Hassan, Nicholas Rott, and Hans Thomann

1992 Gröbli's solution of the three-vortex problem, *Annual Review of Fluid Mechanics* **24**, 1–20.

Aref, Hassan and Dmitri Vainchtein

1998 Point vortices exhibit asymmetric equilibria, *Nature* **392**, 769–770.

Armanini, Egidio

1900 Sulla superficie di minima resistenza, *Annali di matematica pura ed applicata*, 3rd series, **4**, 131–149.

Arnold, Vladimir I.

1989 *Mathematical methods of classical mechanics*, Springer-Verlag: New York.

Arnold, Vladimir I. and Boris A. Khesin

1988 *Topological methods in hydrodynamics,* Springer-Verlag: New York.

Baltzer, Richard

1883 *Die Elemente der Mathematik*, vol. 2: *Planimetrie, Stereometrie, Trigonometrie*, 6th edition, Hirzel: Leipzig.

Bao, David, Colleen Robles, and Zhongmin Shen

2004 Zermelo navigation on Riemannian manifolds, *Journal of Differential Geometry* **66**, 377-435.

Bartel, Hans-Georg and Rudolf P. Huebener

2007 *Walther Nernst—Pioneer of Physics and of Chemistry*, World Scientific: Hackensack.

H.-D. Ebbinghaus, A. Kanamori (Eds.), *Ernst Zermelo - Collected Works/ Gesammelte Werke II*, DOI 10.1007/978-3-540-70856-8_BM1, Schriften der Mathematisch-naturwissenschaftlichen Klasse der Heidelberger Akademie der Wissenschaften 23,

Batterman, Robert W.

1990 Irreversibility and statistical mechanics: A new approach? *Philosophy of Science* **57**, 395–419.

Ben-Asher, Joseph Z.

2010 *Optimal control with aerospace applications*, The American Institute of Aeronautics and Astronautics: Reston, VA.

Bliss, Gilbert Ames

1902a The second variation of a definite integral when one end-point is variable, *Transactions of the American Mathematical Society* **3**, 132–141.

1902b The geodesic lines on the anchor ring, *Annals of Mathematics*, 2nd series, **4**, 1–21.

1904 Jacobi's criterion when both end-points are variable, *Mathematische Annalen* **58**, 70–80.

1920 Some recent developments in the calculus of variations, *Bulletin of the American Mathematical Society* **26**, 343–361.

1925 *Calculus of variations*, Carus Mathematical Monograph no. 1, The Open Court Publishing Company: Chicago.

1946 *Lectures on the calculus of variations*, The University of Chicago Press: Chicago.

Boatto, Stefanella and Jair Koiller *see* Koiller, Jair and Stefanella Boatto

Bôcher, Maxime

1900 Randwertaufgaben bei gewöhnlichen Differentialgleichungen, in *Encyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen*, vol. 2: *Analysis*, part 1.1, B. G. Teubner: Leipzig 1899–1916, 437–463.

Boerner, Hermann

1936 Über die Extremalen und geodätischen Felder in der Variationsrechnung der mehrfachen Integrale, *Mathematische Annalen* **112**, 187–213.

Bogomolov, V. A.

1977 The dynamics of vorticity on a sphere, *Fluid Dynamics* **16**, 863–870.

1979 Two-dimensional hydrodynamics on a sphere, *Izvestiya Atmospheric and Oceanic Physics* **15**, 18–22.

1985 On the motion of a vortex on a rotating sphere, *Izvestiya Atmospheric and Oceanic Physics* **21**, 298–302.

Bolsinov, Aleksey V. and Anatoly T. Fomenko

1999 *Integrable Hamiltonian systems. Geometry, topology, classification*, vols. 1, 2, Izdatel'skii Dom "Udmurtskii Universitet": Izhevsk (Russian); English translation: Chapman & Hall/CRC: Boca Raton 2004.

Boltzmann, Ludwig

1868 Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten, *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, Abtheilung IIa*, **58**, 517–560; reprinted in *Boltzmann 1909*, vol. 1, 49–96.

1871a Einige allgemeine Sätze über Wärmegleichgewicht, *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, Abtheilung IIa*, **63**, 679–711; reprinted in *Boltzmann 1909*, vol. 1, 259–287.

1871b Analytischer Beweis des zweiten Hauptsatzes der mechanischen Wärmetheorie aus den Sätzen über das Gleichgewicht der lebendigen Kraft, *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, Abtheilung IIa*, **63**, 712–732; reprinted in *Boltzmann 1909*, vol. 1, 288–308.

1872 Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen, *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, Abtheilung IIa*, **66**, 275–370; reprinted in *Boltzmann 1909*, vol. 1, 316–402.

1877a Bemerkungen über einige Probleme der mechanischen Wärmetheorie, *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, Abtheilung IIa*, **75**, 62–100; reprinted in *Boltzmann 1909*, vol. 2, 112–148.

1877b Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewicht, *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, Abtheilung IIa*, **76**, 373–435; reprinted in *Boltzmann 1909*, vol. 2, 164–223.

1878 Über die Beziehung der Diffusionsphänomene zum zweiten Hauptsatze der mechanischen Wärmetheorie, *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, Abtheilung IIa*, **78**, 733–763; reprinted in *Boltzmann 1909*, vol. 2, 289–317.

1880 Zur Theorie der Gasreibung I, *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der*

Wissenschaften, Abtheilung IIa, **81**, 117–158; reprinted in *Boltzmann 1909*, vol. 2, 388–430.

1881a Zur Theorie der Gasreibung II, *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, Abtheilung IIa*, **84**, 40–135; reprinted in *Boltzmann 1909*, vol. 2, 431–522.

1881b Zur Theorie der Gasreibung III, *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, Abtheilung IIa*, **84**, 1230–1263; reprinted in *Boltzmann 1909*, vol. 2, 523–556.

1886 Der zweite Hauptsatz der mechanischen Wärmetheorie, *Almanach der Kaiserlichen Akademie der Wissenschaften* **36**, 225–259; reprinted in *Boltzmann 1905a*, 25–50.

1892 III. Teil der Studien über Gleichgewicht der lebendigen Kraft, *Sitzungsberichte der königlich bayerischen Akademie der Wissenschaften. Mathematisch-physikalische Klasse* **22**, 329–358; reprinted in *Boltzmann 1909*, vol. 3, 428–453.

1894 Über den Beweis des Maxwellschen Geschwindigkeitsvertheilungsgesetzes unter Gasmolecülen, *Sitzungsberichte der königlich bayerischen Akademie der Wissenschaften. Mathematisch-physikalische Klasse* **24**, 207–210; reprinted in *Annalen der Physik und Chemie,* new series, **53**, 955–958, and in *Boltzmann 1909*, vol. 3, 528–531.

1895a Nochmals das Maxwellsche Vertheilungsgesetz der Geschwindigkeiten, *Sitzungsberichte der königlich bayerischen Akademie der Wissenschaften. Mathematisch-physikalische Klasse* **25**, 25–26; reprinted in *Annalen der Physik und Chemie,* new series, **55**, 223–224, and in *Boltzmann 1909*, vol. 3, 532–534.

1895b On certain questions of the theory of gases, *Nature* **51**, 413–415; reprinted in *Boltzmann 1909*, vol. 3, 535–544.

1895c On the minimum theorem in the theory of gases, *Nature* **52**, 221; reprinted in *Boltzmann 1909*, vol. 3, 546.

1896 Entgegnung auf die wärmetheoretischen Betrachtungen des Hrn. E. Zermelo, *Annalen der Physik und Chemie*, new series, **57**, 773–784; reprinted in *Boltzmann 1909*, vol. 3, 567–578.

1896a *Vorlesungen über Gastheorie*, vol. 1: *Theorie der Gase mit einatomigen Molekülen, deren Dimensionen gegen die mittlere Weglänge verschwinden*, J. A. Barth: Leipzig; English translation in *Boltzmann 1964*.

1897 Zu Hrn. Zermelo's Abhandlung 'Ueber die mechanische Erklärung irreversibler Vorgänge', *Annalen der Physik und Chemie*, new series, **60**, 392–398; reprinted in *Boltzmann 1909*, vol. 3, 579–586.

1897a Über die Unentbehrlichkeit der Atomistik in der Naturwissenschaft, *Annalen der Physik und Chemie,* new series, **60**, 231–247; reprinted in *Boltzmann 1905a*, 141–157.

1897b Über einen mechanischen Satz Poincaré's, *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, Abtheilung IIa*, **106**, 12–20; reprinted in *Boltzmann 1909*, vol. 3, 587–595.

1897c Über die Frage nach der objektiven Existenz der Vorgänge in der unbelebten Natur, *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, Abtheilung IIa*, **106**, 83–109; reprinted in *Boltzmann 1905a*, 162–187.

1898 *Vorlesungen über Gastheorie*, vol. 2: *Theorie van der Waals*'; *Gase mit zusammengesetzten Molekülen*; *Gasdissociation*; *Schlussbemerkungen*, J. A. Barth: Leipzig; English translation in *Boltzmann 1964*.

1905a *Populäre Schriften*, J. A. Barth: Leipzig.

1905b Reise eines deutschen Professors ins Eldorado, in *Boltzmann 1905a*, 403–435.

1909 *Wissenschaftliche Abhandlungen*, vols. 1–3, edited by Friedrich Hasenöhrl, J. A. Barth: Leipzig.

1964 *Lectures on gas theory*, English translation of *Boltzmann 1896a* and *Boltzmann 1898* by Stephen G. Brush, University of California Press: Berkeley.

Boltzmann, Ludwig and Joseph Nabl

1907 Kinetische Theorie der Materie, in *Encyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen*, vol. 5: *Physik*, part 1, article finished in October 1905, B. G. Teubner: Leipzig 1903–1921, 493–557.

Bolza, Oskar

1901a Some instructive examples in the calculus of variations, *Bulletin of the American Mathematical Society*, 2nd series, **9**, 1–10.

1901b New proof of a theorem of Osgood's in the calculus of variations, *Transactions of the American Mathematical Society* **2**, 422–427.

1902a Proof of the sufficiency of Jacobi's condition for a permanent sign of the second variation in the so-called isoperimetric problem, *Transactions of the American Mathematical Society* **3**, 305–311.

1902b Proof of the sufficiency of Jacobi's condition for a permanent sign of the second variation in the so-called isoperimetric problem, *The Decennial Publications of the University of Chicago* **9**, 7 pp.; reprint of *Bolza 1902a*.

1902c Concerning the isoperimetric problem on a given surface, *The Decennial Publications of the University of Chicago* **9**, 13–21.

1903a Zur zweiten Variation bei isoperimetrischen Problemen, *Mathematische Annalen* **57**, 44–47.

1903b Über das isoperimetrische Problem auf einer gegebenen Fläche, *Mathematische Annalen* **57**, 48–52.

1904 *Lectures on the calculus of variations*, *The Decennial Publications of the University of Chicago*, 2nd series, vol. XIV, The University of Chicago Press: Chicago.

1909 *Vorlesungen über Variationsrechnung*, B. G. Teubner: Leipzig; extended translation of *Bolza 1904*.

Borisov, Alexey V., Aleksander A. Kilin, and Ivan S. Mamaev

2005a Absolute and relative choreographies in the problem of the motion of point vortices on a plane, *Doklady Mathematics* **71**, 139–144.

2005b New periodic solutions for three or four identical vortices on a plane and a sphere, *Discrete and Continuous Dynamical Systems*, supplement, 110–120.

2006 Transition to chaos in the dynamics of four point vortices on a plane, *Doklady Physics* **51**, 262–267.

2008 A new integrable problem of motion of point vortices on the sphere, in Alexey V. Borisov et al. (eds.), *IUTAM Symposium on Hamiltonian dynamics, vortex structures, turbulence (Moscow, 25–30 August, 2006)*, Springer: Dordrecht, 39–53.

Borisov, Alexey V. and Vladimir G. Lebedev

1998a Dynamics of three vortices on a plane and a sphere. II. General compact case, *Regular and Chaotic Dynamics* **3**, no. 2, 99–114.

1998b Dynamics of three vortices on a plane and a sphere. III. Non-compact case. Problems of collapse and scattering, *Regular and Chaotic Dynamics* **3**, no. 3, 76–90.

Borisov, Alexey V. and Ivan S. Mamaev

1999 *Poisson structures and Lie algebras in Hamiltonian mechanics*, RCD/Institute of Computer Science: Ìoscow, Izhevsk (Russian).

2005a *Rigid body dynamics. Hamiltonian methods, integrability, chaos*, RCD/Institute of Computer Science: Ìoscow, Izhevsk (Russian).

2005b *Mathematical methods in the dynamics of vortex structures*, RCD/Institute of Computer Science: Ìoscow, Izhevsk (Russian).

Borisov, Alexey V., Ivan S. Mamaev, and Sergei M. Ramodanov

2008 Algebraic reduction of systems on two- and three-dimensional spheres, *Russian Journal of Nonlinear Dynamics* **4**, 407–416 (Russian).

2010 Coupled motion of a rigid body and point vortices on a two-dimensional spherical surface, *Regular and Chaotic Dynamics* **5**, 319–343.

Borisov, Alexey V. and Alexander E. Pavlov

1998 Dynamics and statics of vortices on a plane and a sphere. I, *Regular and Chaotic Dynamics* **3**, no. 1, 28–39.

Bradley, Ralph A. and Milton E. Terry

1952 The rank analysis of incomplete block designs. I. The method of paired comparisons, *Biometrika* **39**, 324–345.

Brown, Harvey R., Wayne Myrvold, and Jos Uffink

2009 Boltzmann's H-theorem, its discontents, and the birth of statistical mechanics, *Studies in History and Philosophy of Modern Physics* **40**, 174–191.

Brush, Stephen G.

1999 Gadflies and geniuses in the history of gas theory, *Synthèse* **119**, 11–43; in *Brush 2003*, 421–450.

2003 *The kinetic theory of gases*, Imperial College Press: London.

Bryan, George Hartley

1894 Letter to the editor, *Nature* **51**, 175.

Bryson, Jr., Arthur E. and Yu-Chi Ho

1975 *Applied optimal control*, Hemisphere Publishing Company: Washington, New York, and London.

Bühlmann, Hans and Peter J. Huber

1963 Pairwise comparison and ranking in tournaments, *Annals of Statistics* **34**, 501–510.

Burbury, Samuel H.

1894a Boltzmann's minimum function, *Nature* **51**, 78.

1894b The kinetic theory of gases, *Nature* **51**, 175–176.

Burkhardt, Heinrich and W. Franz Meyer

1900 Potentialtheorie (Theorie der Laplace-Poissonschen Differentialgleichung), in *Encyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen*, vol. 2: *Analysis*, part 1.1, B. G. Teubner: Leipzig 1899–1916, 464–503.

Busemann, Herbert

1943 On spaces in which two points determine a geodesic, *Transactions of the American Mathematical Society* **54**, 171–184.

Cantor, Georg

1932 *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts. Mit erläuternden Anmerkungen sowie mit Ergänzungen*

aus dem Briefwechsel Cantor–Dedekind herausgegeben von Ernst Zermelo, Springer-Verlag: Berlin; reprinted with biographical supplement 1980 and 1990.

Carathéodory, Constantin

1904 *Über die diskontinuierlichen Lösungen in der Variationsrechnung*, Ph. D. Thesis, University of Göttingen; reprinted in *Carathéodory 1954/57*, vol. I, 3–79.

1910 Reviews of *Bolza 1909* and *Hadamard 1910*, *Archiv der Mathematik und Physik* **16**, 221–224; reprinted in *Carathéodory 1954/57*, vol. V, 304–308.

1927 *Vorlesungen über reelle Funktionen*, 2nd edition, B. G. Teubner: Leipzig and Berlin.

1928 Review of *Weierstraß 1927*, *Deutsche Literaturzeitung für Kritik der Internationalen Wissenschaft* **49**, columns 57–58; reprinted in *Carathéodory 1954/57*, vol. V, 343–344.

1928/29 Reviews of *Weierstraß 1927* and *Forsyth 1927*, *Mathematical Gazette* **16**, 310–313; reprinted in *Carathéodory 1954/57*, vol. V, 345–349.

1929 Über die Variationsrechnung bei mehrfachen Integralen, *Acta scientiarum mathematicarum* (*Szeged*) **4**, 193–216; reprinted in *Carathéodory 1954/57*, vol. I, 401–426.

1935 *Variationsrechnung und partielle Differentialgleichungen erster Ordnung*, B. G. Teubner: Leipzig; 2nd edition 1956.

1954/57 *Gesammelte Mathematische Schriften*, vol. I–V, C. H. Beck: Munich.

1967 *Calculus of variations and partial differential equations of the first order*, part I: *Partial differential equations of the first order*, part II: *Calculus of variations*, English translation of the 2nd edition of *1935* by Robert B. Dean, Holden-Day: San Francisco.

1994 *Variationsrechnung und partielle Differentialgleichungen erster Ordnung: Variationsrechnung*, reprint of part 2 of *Carathéodory 1935*, edited by Rolf Klötzler, Teubner-Archiv zur Mathematik, vol. 18, B. G. Teubner: Stuttgart and Leipzig 1994.

Cercignani, Carlo

1998 *Ludwig Boltzmann. The man who trusted atoms*, Oxford University Press: Oxford.

Chernov, Nikolai I. and Yakov G. Sinai *see* Sinai, Yakov G. and Nikolai I. Chernov

Clebsch, Rudolf Friedrich Alfred

1858a Ueber die Reduktion der zweiten Variation auf ihre einfachste Form, *Journal für die reine und angewandte Mathematik* **55**, 254–273.

1858b Ueber diejenigen Probleme der Variationsrechnung, welche nur eine unabhängige Variable enthalten, *Journal für die reine und angewandte Mathematik* **55**, 335-355.

1859 Ueber die zweite Variation vielfacher Integrale, *Journal für die reine und angewandte Mathematik* **56**, 122–148.

Cosby, Byron

1937 Fields for multiple integrals in the calculus of variations, in Gilbert Ames Bliss and Lawrence Murray Graves (eds.), *Contributions to the calculus of variations, 1933–37*, University Press: Chicago, 53–83.

Courant, Richard

1926 Über direkte Methoden bei Variations- und Randwertproblemen, *Jahresbericht der Deutschen Mathematiker-Vereinigung* **34**, 90–117.

Courant, Richard and David Hilbert

1924 *Methoden der Mathematischen Physik*, vol. 1, Springer-Verlag: Berlin.

Crilly, Tony

2008 *50 mathematical ideas you really need to know*, Quercus: London.

Crowdy, Darren

2006 Point vortex motion on the surface of a sphere with impenetrable boundaries, *Physics of Fluids* **18**, 036602.

Culverwell, Edward P.

1894 Dr. Watson's proof of Boltzmann's theorem on permanence of distributions, *Nature* **50**, 617.

Curd, Martin

1982 Popper on the direction of time, in Roman Ulrich Sexl and John T. Blackmore (eds.), *Ludwig Boltzmann. Internationale Tagung anlässlich des 75. Jahrestages seines Todes, 5.-8. September 1981: Ausgewählte Abhandlungen*, 263–303, Akademische Druck- und Verlagsanstalt: Graz.

Darboux, Jean Gaston

1889 *Leçons sur la théorie générale des surfaces et sur les applications géométriques du calcul infinitesimal*, vol. 2, Gauthier-Villars: Paris.

1894 *Leçons sur la théorie générale des surfaces et sur les applications géométriques du calcul infinitesimal*, vol. 3, Gauthier-Villars: Paris.

David, Herbert A.

1988 *The method of paired comparisons,* 2nd edition, Oxford University Press: New York.

de Regt, Henk W.

2005 Scientific realism in action: Molecular models and Boltzmann's Bildtheorie, *Erkenntnis* **63**, 205–230.

Dias, Penha Maria Cardoso

1994 Will someone say exactly what the H-theorem proves? A study of Burbury's Condition A and Maxwell's Proposition II. *Archive for History of Exact Sciences* **46**, 341–366.

Dresden, Arnold

1926 Some recent work in the calculus of variations, *Bulletin of the American Mathematical Society* **32**, 475–521.

du Bois-Reymond, Paul

1879a Erläuterung zu den Anfangsgründen der Variationsrechnung, *Mathematische Annalen* **15**, 283–314.

1879b Fortsetzung der Erläuterungen zu den Anfangsgründen der Variationsrechnung, *Mathematische Annalen* **15**, 564–576.

Dykstra, Otto

1956 A note on the rank analysis of incomplete block designs – applications beyond the scope of existing tables, *Biometrics* **12**, 301–306.

Earman, John

2006 The "Past Hypothesis": not even false, *Studies in History and Philosophy of Modern Physics* **37**, 399–430.

Ebbinghaus, Heinz-Dieter

2007 (in cooperation with Volker Peckhaus) *Ernst Zermelo. An approach to his life and work*, Springer-Verlag: Berlin et al.

Ehrenfest, Paul and Tatiana Ehrenfest-Afanassjewa

1912 Begriffliche Grundlagen der statistischen Auffassung der Mechanik, in *Encyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen*, vol. 4: *Mechanik*, part 4. D. 32, Teubner: Leipzig; English translation *The conceptual foundations of the statistical approach in mechanics*, Cornell University Press: Ithaca, 1959.

Elo, Arpad E.

1978 *The rating of chessplayers, past and present,* Arco Publishing: New York.

Erdmann, Georg

1877 Ueber unstetige Lösungen in der Variationsrechnung, *Journal für die reine und ungewandte Mathematik* **82**, 21–30.

Escherich, Gustav von *see* von Escherich, Gustav

Euler, Leonhard

1744 *Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes*, Bousquet: Lausanne and Geneva; reprinted as *Euler 1952*.

1952 *Opera omnia*, ser. 1, vol. 24, edited by Constantin Carathéodory, Orell Füssli: Zurich.

Fechner, Gustav Theodor

1875 *Kleine Schriften*, Breitkopf und Härtel: Leipzig.

Fick, Adolf

1855 Ueber Diffusion, *Annalen der Physik und Chemie*, 4th series, **94**, 59–86.

Finsler, Paul

1918 *Über Kurven und Flächen*, Ph. D. Thesis, University of Göttingen; reprint Birkhäuser: Basel 1951.

Fomenko, Anatoly T. and Aleksey V. Bolsinov *see* Bolsinov, Aleksey V. and Anatoly T. Fomenko

Ford, Lester R. Jr.

1957 Solution of a ranking problem from binary comparisons, *The American Mathematical Monthly* **64(8)**, 28–33.

Forsyth, Andrew Russell

1927 *Calculus of variations*, Cambridge University Press: Cambridge.

Fox, Charles

1950 *An introduction to the calculus of variations*, Oxford University Press: Oxford.

Frank, Philipp

1933 Die schnellste Flugverbindung zwischen zwei Punkten, *Zeitschrift für angewandte Mathematik und Mechanik* **13**, 88–91.

Fraser, Craig G.

2003 The calculus of variations: A historical survey, in Hans Niels Jahnke (ed.), *A history of analysis*, American Mathematical Society and London Mathematical Society: Providence, 355–384.

Frigg, Roman

2008 A field guide to recent work on the foundations of statistical mechanics, in Dean Rickles (ed.), *The Ashgate companion to contemporary philosophy of physics*, Ashgate: London, 99–196.

Funk, Paul

1929 Über Geometrien, bei denen die Geraden die Kürzesten sind, *Mathematische Annalen* **101**, 226–237.

1962 *Variationsrechnung und ihre Andwendung in Physik und Technik*, Springer-Verlag: Berlin et al.

Gauß, Carl Friedrich

1828 Disquisitiones generales circa superficies curvas, *Commentationes societatis regiae scientiarum Gottingensis recentiores* (*classis mathematicae*) **6**, 99–146; reprinted in *Gauß 1873*, 217–258.

1829 Über ein neues allgemeines Grundgesetz der Mechanik, *Journal für die reine und angewandte Mathematik* **4**, 232–235; reprinted in *Gauß 1877*, 23–28.

1873 *Werke*, vol. 4, Königliche Gesellschaft der Wissenschaften zu Göttingen: Göttingen.

1877 *Werke*, vol. 5, Königliche Gesellschaft der Wissenschaften zu Göttingen: Göttingen.

Gernet, Nadeschda

1902 *Untersuchungen zur Variationsrechnung. Über eine neue Methode in der Variationsrechnung*, Ph. D. Thesis, University of Göttingen.

Giaquinta, Mariano and Stefan Hildebrandt

1996 *Calculus of variations*, 2 vols., Springer-Verlag: Berlin et al.

Gibbs, Josiah Willard

1875 On the equilibrium of heterogeneous substances, *Transactions of the Connecticut Academy of Arts and Sciences* **3**, 108–248.

1892 *Thermodynamische Studien. Unter Mitwirkung des Verfassers aus dem Englischen übersetzt von Wilhelm Ostwald*, Engelmann: Leipzig.

1902 *Elementary principles in statistical mechanics, developed with especial reference to the rational foundation of thermodynamics*, Scribner: New York.

1905 *Elementare Grundlagen der statistischen Mechanik entwickelt besonders im Hinblick auf eine rationelle Begründung der Thermodynamik*, German edition of *Gibbs 1902* by Ernst Zermelo, J. A. Barth: Leipzig.

Glass, Kathryn

1997 Equilibrium configurations for a system of N particles in the plane, *Physics Letters A* **235**, 591–596.

Glazebrook, Richard Tedley

1894 *Light, an elementary textbook theoretical and practical*, Cambridge University Press: Cambridge.

1897 *Das Licht.Grundriß der Optik, für Studierende und Schüler*, German edition of *Glazebrook 1894* by Ernst Zermelo, S. Calvary & Co.: Berlin.

Glickman, Mark E.

1999 Parameter estimation in large dynamic paired comparison experiments, *Applied Statistics* **48**, 377–394.

Goldstine, Herman H.

1980 *A history of the calculus of variations from the 17th through the 19th century*, Springer-Verlag: New York.

Good, Irving J.

1955 On the marking of chess-players, *Mathematical Gazette* **39**, 292–296.

Goursat, Édouard Jean-Baptiste

1904 A simple proof of a theorem in the calculus of variations, *Transactions of the American Mathematical Society* **5**, 110–112.

1905 *Cours d'analyse mathématique*, vol. 2, Gauthier-Villars: Paris; Chapter 23, pp. 590–632, is devoted to the calculus of variations.

Grelling, Kurt

1910 *Die Axiome der Arithmetik, mit besonderer Berücksichtigung der Beziehungen zur Mengenlehre*, Ph. D. Thesis, University of Göttingen.

Gröbli, Walter

1877 *Specielle Probleme über die Bewegung geradliniger paralleler Wirbelfäden*, Züricher und Furrer: Zurich; also published in *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich* **22**, 37–81, 129–165.

Gromeka, Ippolit Stepanovich

1881 *Some cases of the motion of an incompressible fluid*, Ph. D. Thesis, University of Kazan; reprinted in *Gromeka 1952*, 76–148.

1885 On vortex motions of a fluid on a sphere, *Uchenye zapiski Kasanskogo un-ta*, Session 45, April 13, 1885; in: *Gromeka 1952*, 184–205.

1952 *Collected papers*, Akademiya Nauk SSSR: Moscow.

Guldberg, Alf

1902 Ueber die Maxima und Minima der Integrale, die eine continuirliche Gruppe gestatten, *Skrifter udgivne af Videnskabsselskabet i Christiania, I.* Mathematisk-naturvidenskabelig Klasse **1902**, no. 7.

Hadamard, Jaques

1902 Sur une question de calcul des variations, *Bulletin de la société mathématique de France* **30**, 253–256.

1910 *Leçons sur le calcul des variations*, A. Hermann et Fils: Paris.

Hahn, Hans

1903a *Zur Theorie der zweiten Variation einfacher Integrale*, Ph. D. Thesis, University of Vienna, *Monatshefte für Mathematik und Physik* **14**, 3–57; reprinted in *Hahn 1996*, 19–74.

1903b Über die Lagrangesche Multiplikationsmethode in der Variationsrechnung, *Monatshefte für Mathematik und Physik* **14**, 325–342; reprinted in *Hahn 1996*, 75–92.

1904 Bemerkungen zur Variationsrechnung, *Mathematische Annalen* **58**, 148–168; reprinted in *Hahn 1996*, 93–113.

1906 Über das allgemeine Problem der Variationsrechnung, *Monatshefte für Mathematik und Physik* **17**, 295–304; reprinted in *Hahn 1996*, 147–156.

1927 Variationsrechnung, in Ernesto Pascal (ed.), *Repertorium der höheren Mathematik*, vol. 1/2, 626–684, B. G. Teubner: Leipzig.

1996 *Gesammelte Abhandlungen / Collected Works*, Bd. 2 / vol. 2, edited by Leopold Schmetterer and Karl Sigmund, Springer-Verlag: Wien.

Hahn, Hans and Ernst Zermelo

1904 Weiterentwicklung der Variationsrechnung in den letzten Jahren, in *Enzyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen*, vol. 2: *Analysis*, part 1.1, B. G. Teubner: Leipzig 1899–1916, 626–641.

Hamel, Georg

1901 *Über die Geometrieen, in denen die Geraden die Kürzesten sind*, Ph. D. Thesis, University of Göttingen.

1903 Über die Geometrieen, in denen die Geraden die Kürzesten sind, *Mathematische Annalen* **57**, 231–264.

Hausdorff, Felix

1914 *Grundzüge der Mengenlehre*, Verlag von Veit: Leipzig; reprinted in Felix Hausdorff, *Gesammelte Werke*, vol. II: *Grundzüge der Mengenlehre*, Springer-Verlag: Berlin et al. 2002, 91–576.

Hedrick, Earle Raymond

1901a Supplementary note on the calculus of variations, *Bulletin of the American Mathematical Society*, 2nd series, **9**, 245–247.

1901b *Über den analytischen Charakter der Lösungen von Differentialgleichungen*, Ph. D. Thesis, University of Göttingen.

1903 On the characteristics of differential equations, *Annals of Mathematics*, 2nd series, **4**, 121–159.

Helmholtz, Hermann von *see* von Helmholtz, Hermann

Hertz, Heinrich

1894 *Die Principien der Mechanik, in einem neuen Zusammenhange dargestellt*, Wilhelm Engelmann: Leipzig.

Hertz, Wilhelm Arthur

1903 *Über partielle Differentialgleichungen, die in der Variationsrechnung vorkommen*, Ph. D. Thesis, University of Kiel.

Hesse, Ludwig Otto

1857 Ueber die Criterien des Maximums und Minimums der einfachen Integrale, *Journal für die reine und angewandte Mathematik* **54**, 227–273; reprinted in *Hesse 1897*, 413–467.

1897 *Gesammelte Werke*, Königlich Bayerische Akademie der Wissenschaften: Munich; reprint Chelsea Publishing Company: New York 1972.

Hestenes, Magnus R.

1937 A navigation problem in the calculus of variations, *American Journal of Mathematics* **59**, 327–334.

Hilbert, David

1895 Ueber die gerade Linie als kürzeste Verbindung zweier Punkte, *Mathematische Annalen* **46**, 91–96.

1899 *Grundlagen der Geometrie*, B. G. Teubner: Leipzig; English translation *Hilbert 1902c*.

1900a Mathematische Probleme. Vortrag, gehalten auf dem internationalen Mathematiker-Kongreß zu Paris 1900, *Nachrichten von der königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-physikalische Klasse. Aus dem Jahre 1900*, 253–297; reprinted in *Archiv der Mathematik und Physik*, 3rd series, **1**, 44–63 and 213–237, and in *Hilbert 1935*, 290–329; English translation *Hilbert 1902a*, French translation *Hilbert 1902b*.

1900b Über das Dirichlet'sche Princip, *Jahresbericht der Deutschen Mathematiker-Vereinigung* **8**, 184–188; reprinted in *Hilbert 1935*, 10–14.

1901a Mathematische Probleme, *Archiv der Mathematik und Physik*, 3rd series, **1**, 44–63 and 213–237; reprinted in *Hilbert 1935*, 290–329.

1901b Über das *Dirichlet*'sche Princip, in *Festschrift zur Feier des hundertfünfzigjährigen Bestehens der königlichen Gesellschaft der Wissenschaften zu Göttingen, Abhandlungen der mathematisch-physikalischen Klasse*, Weidmannsche Buchhandlung: Berlin, 3rd contribution; reprinted as *Hilbert 1904*.

1902a Mathematical problems. Lecture delivered before the International Congress of Mathematicians at Paris 1900, *Bulletin of the American Mathematical Society* **8**, 437–479; reprinted in vol. **37** (2000), 407–436.

1902b Sur les problèmes futurs des mathématiques, in Ernest Duporcq (ed.), *Compte rendu du deuxième congrès international des mathématiciens, tenu a Paris du 6 au 12 Août 1900*, Gauthier-Villars: Paris, 58–114.

1902c *The foundations of geometry*, The Open Court Publishing Company: Chicago.

1904 Über das Dirichletsche Princip, *Mathematische Annalen* **59**, 161–186.

1905 Zur Variationsrechnung, *Nachrichten von der königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-physikalische Klasse. Aus dem Jahre 1905*, 159–180; reprinted in *Mathematische Annalen* **62** (1906), 351–370, and in *Hilbert 1935*, 38–55.

1935 *Gesammelte Abhandlungen*, vol. 3: *Analysis, Grundlagen der Mathematik, Physik, Verschiedenes, Lebensgeschichte*, Springer-Verlag: Berlin and Heidelberg.

Hilbert, David and Richard Courant *see* Courant, Richard and David Hilbert

Hildebrandt, Stefan and Mariano Giaquinta *see* Giaquinta, Mariano and Stefan Hildebrandt

Hirsch, Arthur

1897 Ueber eine charakteristische Eigenschaft der Differentialgleichungen der Variationsrechnung, *Mathematische Annalen* **49**, 49–72.

Ho, Yu-Chi and Arthur E. Bryson, Jr., *see* Bryson, Jr., Arthur E. and Yu-Chi Ho

Höflechner, Walter

1994 *Ludwig Boltzmann. Leben und Briefe*, Akademische Druck- und Verlagsanstalt: Graz.

Hsu, Jay C. and Andrew U. Meyer

1968 *Modern control principles and applications*, McGraw-Hill Book Company: New York et al.

Huber, Peter J. and Hans Bühlmann *see* Bühlmann, Hans and Peter J. Huber

Huebener, Rudolf P. and Hans-Georg Bartel *see* Bartel, Hans-Georg and Rudolf P. Huebener

Husserl, Edmund

1882 *Beiträge zur Theorie der Variationsrechnung*, Ph. D. Thesis, University of Vienna; French translation *Contributions à la théorie du calcul des variations* by Mlle Devouard, edited by Jacques Vauthier, Queen's Papers in Pure and Applied Mathematics, vol. 65, Queen's University: Kingston, Ontario, 1983.

Ioffe, Aleksandr Davidovich and Vladimir Mikhailovich Tichomirov

1974 *Theory of extremal problems*, Nauka: Moscow (Russian); English translation *Theory of extremal problems, Studies in mathematics and its applications*, vol. 6, North-Holland Publishing Company: Amsterdam, New York, and Oxford, 1979.

Jacobi, Carl Gustav Jacob

1837 Zur Theorie der Variations-Rechnung und der Differential-Gleichungen, *Journal für die reine und angewandte Mathematik* **17**, 68–82.

1838 Sur le calcul des variations et sur la théorie des équations différentielles, *Journal de Mathématiques Pures et Appliquées* **3**, 44–59; French translation of *Jacobi 1837*.

1866 *Vorlesungen über Dynamik nebst fünf hinterlassenen Abhandlungen desselben*, edited by Alfred Clebsch, Georg Reimer: Berlin.

Khesin, Boris A. and Vladimir I. Arnold *see* Arnold, Vladimir I. and Boris A. Khesin

Kidambi, Rangachari and Paul Newton

1998 Motion of three point vortices on a sphere, *Physica D* **116**, 143–175.

1999 Collapse of three vortices on a sphere, *Il Nuovo Cimento C* **22**, 779–791.

2000a Point vortex motion on a sphere with solid boundaries, *Physics of Fluids* **12**, 581–588.

2000b Streamline topologies for integrable vortex motion on a sphere, *Physica D* **140**, 95–125.

Kilin, Aleksander A., Alexey V. Borisov, et al. *see* Borisov, Alexey V., Aleksander A. Kilin, and Ivan S. Mamaev

Kirchhoff, Gustav Robert

1894 *Vorlesungen über die Theorie der Wärme*, *Vorlesungen über mathematische Physik*, vol. 4, edited by Max Planck, B. G. Teubner: Leipzig.

Kirsten, Christa and Hans-Günther Körber

1975 *Physiker über Physiker. Wahlvorschläge zur Aufnahme von Physikern in die Berliner Akademie 1870 bis 1929 von Hermann v. Helmholtz bis Erwin Schrödinger*, Akademie-Verlag: Berlin, 169–171.

Klein, Felix

1884 *Vorlesungen über das Ikosaeder und die Auflösung der Gleichungen vom fünften Grade*, B. G. Teubner: Leipzig.

Klein, Martin J.

1973 The development of Boltzmann's statistical ideas, in Ezechiel Godert David Cohen and Walter Thirring (eds.), *The Boltzmann equation*, Springer-Verlag: Wien, 53–106.

Klötzler, Rolf

1970 *Mehrdimensionale Variationsrechnung*, Birkhäuser: Basel et al.

1994 Kommentare und Erweiterungen, in *Carathéodory 1994*, 307–359.

Kneser, Adolf

1898 Zur Variationsrechnung, *Mathematische Annalen* **50**, 27–50.

1900 *Lehrbuch der Variationsrechnung*, Vieweg: Braunschweig; extended edition 1925.

1902a Beiträge zur Theorie und Anwendung der Variationsrechnung, *Mathematische Annalen* **55**, 86–107.

1902b Ein Beitrag zur Frage nach der zweckmäßigsten Gestalt der Geschossspitzen, *Archiv der Mathematik und Physik*, 3rd series, **2**, 267–278.

1902c Die Jacobische Bedingung des Extremums bei einem allgemeinen Typus von Aufgaben der Variationsrechnung, *Communications of the Kharkov Mathematical Society*, 2nd series, **7**, 253–267.

1903a Beiträge zur Theorie und Anwendung der Variationsrechnung (Zweiter Aufsatz), *Mathematische Annalen* **56**, 169–232.

1903b Die Stabilität des Gleichgewichts hängender schwerer Fäden, *Journal für die reine und angewandte Mathematik* **125**, 189–206.

1904 Variationsrechnung, in *Encyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen*, vol. 2: *Analysis*, part 1.1, B. G. Teubner: Leipzig 1899–1916, 571–625.

Koiller, Jair and Stefanella Boatto

2008a Vortices on closed surfaces, arXiv: 0802.43132.

2008b Vortex pairs on surfaces, in Fernando Etayo, Mario Fioravanti, and Rafael Santamaría (eds.), *Geometry and physics: XVII international fall workshop on geometry and physics, AIP Conference Proceedings* **1130**, 77–88.

Körber, Hans-Günther and Christa Kirsten *see* Kirsten, Christa & Hans-Günther Körber

Korn, Arthur

1902 Ueber den einfachsten semidefiniten Fall in der eigentlichen Variationsrechnung, *Sitzungsberichte der königlich bayerischen Akademie der Wissenschaften. Mathematisch-physikalische Klasse* **32**, 75–90.

Krylov, Nikolai Sergeevich

1979 *Works on the foundations of statistical physics*, Princeton University Press: Princeton.

Kuhn, Thomas S.

1978 *Black-body theory and the quantum discontinuity, 1894–1912*, Oxford University Press: Oxford.

Kurakin, Leonid G. and Victor I. Yudovich

2002 The stability of stationary rotation of a regular vortex polygon, *Chaos* **12**, 574–595.

Kürschák, Josef

1903 Ueber die Transformation der partiellen Differentialgleichungen der Variationsrechnung, *Mathematische Annalen* **56**, 155–164.

Lagrange, Joseph Louis de

1788 *Méchanique analytique*, Chez La Veuve Desaint: Paris; reprinted as *Lagrange 1888* and *Lagrange 1889*; English translation *Lagrange 1997*.

1806 *Leçons sur le calcul des fonctions*, 3rd edition, Chez Courcier: Paris; reprinted in *Lagrange 1884*, 7–451.

1869 *Œuvres*, vol. 4, Gauthier-Villars: Paris.

1884 *Œuvres*, vol. 10, Gauthier-Villars: Paris.

1888 *Œuvres*, vol. 11: *Mécanique analytique*, part 1: *La statique*, edited by Joseph Alfred Serret and Gaston Darboux, Gauthier-Villars: Paris.

1889 *Œuvres*, vol. 12: *Mécanique analytique*, part 2: *La dynamique*, edited by Joseph Alfred Serret and Gaston Darboux, Gauthier-Villars: Paris.

1997 *Analytical mechanics*, English edition of *Lagrange 1788* by Auguste Boissonnade and Victor N. Vagliente, Kluwer Academic Publishers: Dordrecht and Boston.

Lamb, Horace

1895 *Hydrodynamics*, Cambridge University Press: Cambridge.

Landau, Edmund

1914 Über Preisverteilung bei Spielturnieren, *Zeitschrift für Mathematik und Physik* **63**, 192–202.

Lanford, Oscar E.

1975 Time evolution of large classical systems, in Jürgen Moser (ed.), *Dynamical systems. Theory and applications*, Lecture Notes in Physics, vol. 38, Springer-Verlag: Berlin et al., 1–111.

Laura, Ernesto

1902 Sul moto parallelo ad un piano di un fluido in cui vi sono N vortici elementari, *Atti della Accademia delle Scienze di Torino. Classe di Scienze Fisiche, Matematiche e Naturali* **37**, 469–476.

1905 Sulle equazioni differenziali canoniche del moto di un sistema di vortici elementari, rettilinei e paralleli, in un fluido incompressibile indefinito, *Atti della Accademia delle Scienze di Torino. Classe di Scienze Fisiche, Matematiche e Naturali* **40**, 296–312.

Lebedev, Vladimir G. and Alexey V. Borisov *see* Borisov, Alexey V. and Vladimir G. Lebedev

Lebowitz, Joel L.

1999 Statistical mechanics: A selective review of two central issues, *Reviews of Modern Physics* **71**, 346–357.

Lecat, Maurice

1913 Calcul des variations, extended French translation of *Hahn and Zermelo 1904*, *Encyclopédie des sciences mathématiques*, 2/6/1, Gauthier-Villars: Paris, 1–288.

Legendre, Adrien Marie

1786 Sur la manière de distinguer les maxima des minima dans le calcul des variations, *Histoire de l'Académie Royale des Sciences, année M. DCCLXXXVI. Avec les Mémoirs de Mathématique & de Physique, pour la même année*, 7–37; supplement, **1787**, 340–354.

Leonard, Thomas

1977 An alternative Bayesian approach to the Bradley-Terry model for paired comparisons, *Biometrics* **33**, 121–132.

Lepage, Théophile

1941 Sur les champs géodésiques des intégrales multiples, *Bulletin de l'Académie Royale Belgique* **27**, 27–46.

Levi-Civita, Tullio

1931 Über Zermelos Luftfahrtproblem, *Zeitschrift für angewandte Mathematik und Mechanik* **11**, 314–322.

Lewis, Frank L.

1986 *Optimal control*, John Wiley & Sons: New York et al.

Lim, Chjan C., James Montaldi, and Mark Roberts

2001 Relative equilibria of point vortices on the sphere, *Physica D* **148**, 97–135.

Lindelöf, Ernst Leonard and François Napoléon Marie Moigno

1861 *Leçons de calcul différentiel et de calcul intégral*, vol. 4: *Calcul des variations*, Mallet-Bachelier: Paris.

Lord Kelvin *see* Thomson, William

Loschmidt, Josef

1876a Ueber den Zustand des Wärmegleichgewichtes eines Systems von Körpern mit Rücksicht auf die Schwerkraft I, *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, II. Abtheilung* **73**, 128–142.

1876b Ueber den Zustand des Wärmegleichgewichtes eines Systems von Körpern mit Rücksicht auf die Schwerkraft II, *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, II. Abtheilung* **73**, 366–372.

1877a Ueber den Zustand des Wärmegleichgewichtes eines Systems von Körpern mit Rücksicht auf die Schwerkraft III, *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, II. Abtheilung* **75**, 287–298.

1877b Ueber den Zustand des Wärmegleichgewichtes eines Systems von Körpern mit Rücksicht auf die Schwerkraft IV, *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, II. Abtheilung* **76**, 209–225.

Lyusternik, Lazar Aronovich

1934 On the conditional extrema of functionals, *Mathematicheskij Sbornik* **41**, no. 3, 390–401 (Russian).

Mamaev, Ivan S. and Alexey V. Borisov *see* Borisov, Alexey V. and Ivan S. Mamaev

Mamaev, Ivan S., Alexey V. Borisov, et al. *see* Borisov, Alexey V., Aleksander A. Kilin, and Ivan S. Mamaev and *see* Borisov, Alexey V., Ivan S. Mamaev, and Sergei M. Ramodanov

Manià, Basilio
1937 Sopra un problema di navigazione di Zermelo, *Mathematische Annalen* **113**, 584–599.

Marcus, David J.
2001 New table-tennis rating system, *The Statistician* **50**, 191–208.

Mason, Charles Max
1903 *Randwertaufgaben bei gewöhnlichen Differentialgleichungen*, Ph. D. Thesis, University of Göttingen.

Maxwell, James Clerk
1860 Illustrations of the dynamical theory of gases, *Philosophical Magazine* **19**, 19–32, and **20**, 21–37.
1867 On the dynamical theory of gases, *Philosophical Transactions of the Royal Society of London* **157**, 49–88.
1873 *A treatise on electricity and magnetism*, 2 vols., Clarendon Press: Oxford.
1879 On Boltzmann's theorem on the average distribution of energy in a system of material points, *Transactions of the Cambridge Philosophical Society* **12**, 547–570.
1890 *The scientific papers*, vol. 2, edited by William D. Niven, Cambridge University Press: Cambridge.

Mayer, Adolph
1866 *Beiträge zur Theorie der Maxima und Minima der einfachen Integrale*, *Habilitation* thesis, University of Leipzig.
1868 Ueber die Kriterien des Maximums und Minimums der einfachen Integrale, *Journal für die reine und angewandte Mathematik* **69**, 238–263.
1885 Begründung der Lagrange'schen Multiplicatorenmethode in der Variationsrechnung, *Berichte über die Verhandlungen der Königlich-Sächsischen Gesellschaft der Wissenschaften zu Leipzig, mathematisch-physikalische Classe* **37**, 7–14; reprinted and slightly extended as *Mayer 1886*.
1886 Begründung der Lagrange'schen Multiplicatorenmethode in der Variationsrechnung, *Mathematische Annalen* **26**, 74–82.
1889 Zur Theorie des gewöhnlichen Maximums und Minimums, *Berichte über die Verhandlungen der Königlich-Sächsischen Gesellschaft der Wissenschaften zu Leipzig, mathematisch-physikalische Classe* **41**, 122–144.
1895 Die Lagrange'sche Multiplicatorenmethode und das allgemeinste Problem der Variationsrechung bei einer unabhängigen Variablen, *Berichte über die Verhandlungen der Königlich-Sächsischen Gesellschaft der Wissenschaften zu Leipzig, mathematisch-physikalische Classe* **47**, 129–144.

1899 Über die Aufstellung der Differentialgleichungen für reibungslose Punktsysteme, die Bedingungsgleichungen unterworfen sind, *Berichte über die Verhandlungen der Königlich-Sächsischen Gesellschaft der Wissenschaften zu Leipzig, mathematisch-physikalische Classe* **51**, 224–244.

1903 Über den Hilbertschen Unabhängigkeitssatz in der Theorie des Maximums und Minimums der einfachen Integrale, *Berichte über die Verhandlungen der Königlich-Sächsischen Gesellschaft der Wissenschaften zu Leipzig, mathematisch-physikalische Classe* **55**, 131–145; reprinted as *Mayer 1904*.

1904 Über den Hilbertschen Unabhängigkeitssatz in der Theorie des Maximums und Minimums der einfachen Integrale, *Mathematische Annalen* **58**, 235–248.

1905 Über den Hilbertschen Unabhängigkeitssatz in der Theorie des Maximums und Minimums der einfachen Integrale. II. Mitteilung, *Berichte über die Verhandlungen der Königlich-Sächsischen Gesellschaft der Wissenschaften zu Leipzig, mathematisch-physikalische Classe* **57**, 49–67, 313–314; reprinted as *Mayer 1906*.

1906 Über den Hilbertschen Unabhängigkeitssatz in der Theorie des Maximums und Minimums der einfachen Integrale. II. Mitteilung, *Mathematische Annalen* **62**, 335–350.

McCullagh, Peter and John A. Nelder

1989 *Generalized linear models,* Chapman and Hall: London.

Meyer, Andrew U. and Jay C. Hsu see Hsu, Jay C. and Andrew U. Meyer

Meyer, W. Franz and Heinrich Burkhardt *see* Burkhardt, Heinrich and W. Franz Meyer

Minkowski, Hermann

1900 Über die Begriffe Länge, Oberfläche und Volumen, *Jahresbericht der Deutschen Mathematiker-Vereinigung* **9**, 115–121.

1903 Volumen und Oberfläche, *Mathematische Annalen* **57**, 447–495.

Mises, Richard von *see* von Mises, Richard

Moigno, François Napoléon Marie and Lindelöf, Ernst Leonard *see* Lindelöf, Ernst Leonard and François Napoléon Marie Moigno

Montaldi, James, Chjan C. Lim, et al. *see* Lim, Chjan C., James Montaldi, and Mark Roberts

Mosteller, Frederick

1951a Remarks on the method of paired comparisons: I. The least squares solution assuming equal standard deviations and equal correlations, *Psychometrika* **16**, 3–9.

1951b Remarks on the method of paired comparisons: II. The effect of an aberrant standard deviation when equal standard deviations and equal correlations are assumed, *Psychometrika* **16**, 203–206.

1951c Remarks on the method of paired comparisons: III. A test of significance for paired comparisons when equal standard deviations and equal correlations are assumed, *Psychometrika* **16**, 207–218.

Müller, Johann Oswald

1902 Ueber die Minimaleigenschaft der Kugel, *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Aus dem Jahre 1902*, 176–181.

1903 *Über die Minimaleigenschaft der Kugel*, Ph. D. Thesis, University of Göttingen.

Myrvold, Wayne, Harvey R. Brown, and Jos Uffink *see* Brown, Harvey R., Wayne Myrvold, and Jos Uffink

Nabl, Joseph and Ludwig Boltzmann *see* Boltzmann, Ludwig and Joseph Nabl

Nelder, John A. and Peter McCullagh *see* McCullagh, Peter and John A. Nelder

Nernst, Walther

1890 Ueber die Verteilung eines Stoffes zwischen zwei Lösungsmitteln, *Nachrichten von der Königlichen Gesellschaft der Wissenschaften und der Georg-August-Universität zu Göttingen. Aus dem Jahre 1890*, 401–416.

1891 Verteilung eines Stoffes zwischen zwei Lösungsmitteln und zwischen Lösungsmittel und Dampfraum, *Zeitschrift für physikalische Chemie, Stöchiometrie und Verwandtschaftslehre* **8**, 110–139.

1897 Ueber das chemische Gleichgewicht, elektromotorische Wirksamkeit und elektrolytische Abscheidung von Metallgemischen, *Zeitschrift für physikalische Chemie, Stöchiometrie und Verwandtschaftslehre* **22**, 539–542.

1899 Zur Theorie der elektrischen Reizung, *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Aus dem Jahre 1899*, 104–108.

1908 Zur Theorie des elektrischen Reizes, *Pflügers Archiv für die gesamte Physiologie des Menschen und der Tiere* **122**, 275–314.

Nernst, Walther and Ernst Hermann Riesenfeld

1901 Ueber elektrolytische Erscheinungen an der Grenzfläche zweier Lösungsmittel, *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Aus dem Jahre 1901*, 54–61.

1902 Ueber elektrolytische Erscheinungen an der Grenzfläche zweier Lösungsmittel, *Annalen der Physik*, 4th series, **313**, 600–608.

Nernst, Walther and Arthur Schoenflies

1895 *Einführung in die Mathematische Behandlung der Naturwissenschaften — Kurzgefasstes Lehrbuch der Differential- und Integralrechnung mit besonderer Berücksichtigung der Chemie*, E. Wolff: Munich; further editions 1898, 1901; R. Oldenbourg: Munich, Berlin, 1904, 1907, 1910, 1913, 1918, 1919, 1923, and (revised by Walther Nernst and Wilhelm Orthmann) 1931; English edition: Young, Jacob William Albert and Charles Elijah Linebarger, *The elements of the differential and integral calculus, based on Kurzgefasstes Lehrbuch der Differential- und Integralrechnung von W. Nernst und A. Schönflies*, Hirschfeld: London 1900.

Newton, Paul

2001 *The N-vortex problem. Analytical techniques*, Springer-Verlag: New York.

Newton, Paul and Rangachari Kidambi *see* Kidambi, Rangachari and Paul Newton

Noble, Charles Albert

1901 *Eine neue Methode in der Variationsrechnung*, Ph. D. Thesis, University of Göttingen.

Noyes, Arthur Amos and Willis R. Whitney

1897 Ueber die Auflösungsgeschwindigkeit von festen Stoffen in ihren eigenen Lösungen, *Zeitschrift für physikalische Chemie, Stöchiometrie und Verwandtschaftslehre* **23**, 689–692.

Olsen, Elwood T.

1993 Classical mechanics and entropy, *Foundations of Physics Letters* **6**, 327–337.

O'Neil, Kevin

2007 Relative equilibrium and collapse configurations of four point vortices, *Regular and Chaotic Dynamics* **12**, 117–126.

2008a Equilibrium configurations of point vortices on a sphere, *Regular and Chaotic Dynamics* **13**, 1–8.

2008b Relative equilibria of point vortices that lie on a great circle of a sphere, *Nonlinearity* **21**, 2043–2051.

Osgood, William Fogg

1900/01 Sufficient conditions in the calculus of variations, *Annals of Mathematics*, 2nd series, **2**, 105–129.

1901a On the existence of a minimum of the integral $\int_{x_0}^{x_1} F(x, y, y')dx$ when x_0 and x_1 are conjugate points, and the geodesics on an ellipsoid of revolution: a revision of a theorem of Kneser's, *Transactions of the American Mathematical Society* **2**, 166–182.

1901b On a fundamental property of a minimum in the calculus of variations and the proof of a theorem of Weierstrass's, *Transactions of the American Mathematical Society* **2**, 273–295.

Pars, Leopold Alexander

1962 *An introduction to the calculus of variations*, Heinemann Educational Books Ltd.: London.

Pavlov, Alexander E. and Alexey V. Borisov *see* Borisov, Alexey V. and Alexander E. Pavlov

Picard, Emile

1890 Mémoire sur la théorie des équations aux dérivées partielles et la méthode des approximations successives, *Journal de Mathématiques Pures et Appliquées*, 4th series, **6**, 145–210.

Planck, Max

1887 Über die molekulare Konstitution verdünnter Lösungen, *Zeitschrift für Physikalische Chemie* **1**, 577–582.

1891 Allgemeines zur neueren Entwicklung der Wärmetheorie, *Zeitschrift für Physikalische Chemie* **8**, 647–656.

1894a Ueber den Beweis des Maxwellschen Geschwindigkeitsvertheilungsgesetzes unter Gasmolecülen, *Sitzungsberichte der königlich bayerischen Akademie der Wissenschaften. Mathematisch-physikalische Klasse* **24**, 391–394; reprinted as *Planck 1895*.

1894b Antrittsrede, *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin* **2**, 641–644.

1895 Ueber den Beweis des Maxwellschen Geschwindigkeitsvertheilungsgesetzes unter Gasmolecülen, *Annalen der Physik und Chemie*, new series, **55**, 220–222.

1909 Die Einheit des physikalischen Weltbildes, *Physikalische Zeitschrift* **10**, 62–75.

1946 Persönliche Erinnerungen aus alten Zeiten, *Die Naturwissenschaften* **8**, 230–235.

Poincaré, Henri

1889 Sur les tentatives d'explication mécanique des principes de la termodynamique, *Comptes rendus de l'Académie des sciences* **108**, 550–553; reprinted in *Poincaré 1954b*, 231–233.

1890 Sur le probléme des trois corps et les équations de la dynamique, *Acta mathematica* **13**, 1–270; reprinted in *Poincaré 1952*, 262–479; English translation of pages 8, 10, and 67–73 in *Brush 2003*, 368–376.

1893a *Théorie des tourbillons*, George Carré: Paris.
1893b *Thermodynamik – Vorlesungen gehalten von H. Poincaré*, authorized German edition by Wilhelm Jaeger and Ernst Gumlich, Springer-Verlag: Berlin.
1893c Le mécanisme et l'expérience, *Revue de Metaphysique et de Morale* **1**, 534–537; English translation in *Brush 2003*, 377–381.
1899 *Les méthodes nouvelles de la mécanique céleste*, vol. 3, Gauthier-Villars: Paris.
1906 Réflexions sur la théorie cinétique des gaz, *Journal de Physique* **5**, 369–403; reprinted in *Poincaré 1954a*, 587–619.
1952 *Œuvres de Henri Poincaré*, vol. VII, Gauthier-Villars: Paris.
1954a *Œuvres de Henri Poincaré*, vol. IX, Gauthier-Villars: Paris.
1954b *Œuvres de Henri Poincaré*, vol. X, Gauthier-Villars: Paris.

Poisson, Siméon Denis

1823 Mémoire sur le calcul numérique des intégrales définies, in *Mémoires de l'Académie Royale des Sciences de l'Institut de France* **6** (1823), 571–602.

Rademacher, Hans and Otto Toeplitz

1930 *Von Zahlen und Figuren*, Springer-Verlag: Berlin.

Ramodanov, Sergei M., Alexey V. Borisov, et al. *see* Borisov, Alexey V., Ivan S. Mamaev, and Sergei M. Ramodanov

Regt, Henk W. de *see* de Regt, Henk W.

Reinhold, Bernhard

1910 *Abhängigkeit der Überführungszahlen von der Konzentration und der Wahl des Lösungsmittels*, Ph. D. Thesis, University of Freiburg.

Reinhold, Bernhard and Ernst Hermann Riesenfeld *see* Riesenfeld, Ernst Hermann and Bernhard Reinhold

Riecke, Eduard

1891 Das thermische Potential für verdünnte Lösungen, *Zeitschrift für physikalische Chemie, Stöchiometrie und Verwandtschaftslehre* **7**, 97–114.

Riemann, Bernhard and Heinrich Weber

1900 *Die partiellen Differential-Gleichungen der mathematischen Physik*, vol. 1, Vieweg: Braunschweig.
1901 *Die partiellen Differential-Gleichungen der mathematischen Physik*, vol. 2, Vieweg: Braunschweig.

Riesenfeld, Ernst Hermann

1901 Über elektrolytische Erscheinungen an der Grenzfläche zweier Lösungsmittel, *Zeitschrift für Elektrochemie* **7**, 645–648.

1902a *Ueber elektrolytische Erscheinungen und elektromotorische Kräfte an der Grenzfläche zweier Lösungsmittel*, Ph. D. Thesis, University of Göttingen.

1902b Bestimmung der Ueberführungszahl einiger Salze in Phenol, *Annalen der Physik*, 4th series, **313**, 609–615.

1902c Concentrationsketten mit nichtmischbaren Lösungsmitteln, *Annalen der Physik*, 4th series, **313**, 616–624.

Riesenfeld, Ernst Hermann and Walther Nernst *see* Nernst, Walther and Ernst Hermann Riesenfeld

Riesenfeld, Ernst Hermann and Bernhard Reinhold

1909 Berechnung der Ionenhydratation aus der Überführungszahl und der Ionenbeweglichkeit, *Zeitschrift für physikalische Chemie, Stöchiometrie und Verwandtschaftslehre* **66**, 672–686.

1910a Die Überführungszahl der Salzsäure, *Zeitschrift für physikalische Chemie, Stöchiometrie und Verwandtschaftslehre* **68**, 440–458.

1910b Bestimmung von Überführungszahlen aus elektromotorischen Kräften in Lösungsmitteln, welche mit Wasser nur beschränkt mischbar sind, *Zeitschrift für physikalische Chemie, Stöchiometrie und Verwandtschaftslehre* **68**, 459–470.

Riesenfeld, Ernst Heinrich and Ernst Zermelo

1909 Die Einstellung der Grenzkonzentrationen an der Trennungsfläche zweier Lösungsmittel, *Physikalische Zeitschrift* **10**, 958–961.

Ritz, Walter

1908 Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik, *Journal für die reine und angewandte Mathematik* **135**, 1–61.

Roberts, Mark, Chjan C. Lim, et al. *see* Lim, Chjan C., James Montaldi, and Mark Roberts

Robles, Colleen, David Bao, and Zhongmin Shen *see* Bao, David, Colleen Robles, and Zhongmin Shen

Rott, Nicholas, Hassan Aref, et al. *see* Aref, Hassan, Nicholas Rott, and Hans Thomann

Routh, Edward John

1877 *A Treatise on the stability of a given state of motion, particularly steady motion,* MacMillan & Co: London.

Sagan, Hans

1969 *Introduction to the calculus of variations*, McGraw Hill Book Company: New York.

Scheeffer, Ludwig

1885 Die Maxima und Minima der einfachen Integrale zwischen festen Grenzen, *Mathematische Annalen* **25**, 522–593; supplement "Bemerkungen zu dem vorstehenden Aufsatze", ibid., 594–595.

1886 Ueber die Bedeutung der Begriffe "Maximum und Minimum" in der Variationsrechnung, *Mathematische Annalen* **26**, 197–208.

Schoenflies, Arthur and Walther Nernst *see* Nernst, Walther and Arthur Schoenflies

Schwarz, Hermann Amandus

1884 Beweis des Satzes, dass die Kugel kleinere Oberfläche besitzt, als jeder andere Körper gleichen Volumens, *Nachrichten von der Königlichen Gesellschaft der Wissenschaften und der Georg-August-Universität zu Göttingen* **1**, 1–13; reprinted in *Schwarz 1890b*, 327–340.

1885 Ueber ein die Flächen kleinsten Flächeninhalts betreffendes Problem der Variationsrechnung, Festschrift zum siebzigsten Geburtstage des Herrn Karl Weierstrass, *Acta societatis scientiarum Fennicae* **15**, 315-362; reprinted in *Schwarz 1890a*, 223–269.

1890a *Gesammelte mathematische Abhandlungen*, vol. 1, Julius Springer: Berlin.

1890b *Gesammelte mathematische Abhandlungen*, vol. 2, Julius Springer: Berlin.

Serret, Joseph Alfred

1899 *Lehrbuch der Differential- und Integralrechnung*, German edition by Axel Harnack, 2nd revised edition, vol. 2: *Integralrechnung*, edited by Georg Bohlmann with assistance of Heinrich Liebmann and Ernst Zermelo, B. G. Teubner: Leipzig.

1904 *Lehrbuch der Differential- und Integralrechnung*, German edition by Axel Harnack, 2nd revised edition, vol. 3: *Differentialgleichungen und Variationsrechnung*, edited by Georg Bohlmann and Ernst Zermelo, B. G. Teubner: Leipzig.

Shen, Zhongmin, David Bao, and Colleen Robles *see* Bao, David, Colleen Robles, and Zhongmin Shen

Simányi, Nándor and Domokos Szász

1999 Hard ball systems are completely hyperbolic, *Annals of Mathematics* **149**, 35–96.

Sinai, Yakov G. and Nikolai I. Chernov

1987 Ergodic properties of certain systems of two-dimensional discs and three-dimensional balls, *Russian Mathematical Surveys* **42**, no. 3, 181–207.

Sommerfeld, Arnold

1900a Bemerkungen zur Variationsrechnung, *Jahresbericht der Deutschen Mathematiker-Vereinigung* **8**, 188–193.

1900b Randwertaufgaben in der Theorie der partiellen Differentialgleichungen, in *Enzyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen*, vol. 2: *Analysis*, part 1.1, B. G. Teubner: Leipzig 1899–1916, 504–570.

Stokes, George Gabriel

1880 *Mathematical and physical papers*, vol. 1, Cambridge University Press: Cambridge.

Synge, John

1949 On the motion of three vortices, *Canadian Journal of Mathematics* **1**, 257–270.

Szász, Domokos

1996 Boltzmann's ergodic hypothesis, a conjecture for centuries?, *Studia scientiarum mathematicarum Hungarica* **31**, 299–322.

Szász, Domokos and Nándor Simányi *see* Simányi, Nándor and Domokos Szász

Terry, Milton E. and Ralph A. Bradley *see* Bradley, Ralph A. and Milton E. Terry

Thiele, Rüdiger

2007 *Von der Bernoullischen Brachistochrone zum Kalibrator-Konzept. Ein historischer Abriß zur Entstehung der Feldtheorie in der Variationsrechnung (hinreichende Bedingungen in der Variationsrechnung)*, De diversis artibus, vol. 80, Brepols: Turnhout.

2011 *Felix Klein in Leipzig*, Edition am Gutenbergplatz: Leipzig.

Thomann, Hans, Hassan Aref, et al. *see* Aref, Hassan, Nicholas Rott, and Hans Thomann

Thomson, William (Lord Kelvin)

1878 Floating magnets (illustrating vortex-systems), *Nature (London)* **18**, 13–14; reprinted in Joseph Larmor (ed.), *Mathematical and physical papers by Sir William Thomson, Baron Kelvin*, vol. 4, Cambridge University Press: Cambridge, 1910, 135–140.

Thurstone, Louis L.

1927a A law of comparative judgment, *Psychological Review* **34**, 273–286.

1927b Psychophysical analysis, *American Journal of Psychology* **38**, 368–389.

1927c The method of paired comparisons for social values, *Journal of Abnormal and Social Psychology* **21**, 384–400.

Tichomirov, Vladimir Mikhailovich and Aleksandr Davidovich Ioffe *see* Ioffe, Aleksandr Davidovich und Vladimir Mikhailovich Tichomirov

Todhunter, Isaac

1861 *A history of the progress of the calculus of variations during the nineteenth century*, Macmillan: Cambridge and London; reprint *A history of the calculus of variations during the nineteenth century*, Chelsea Publishing Company: New York 1962.

1871 *Researches in the calculus of variations, principally on the theory of discontinuous solutions*, Macmillan: Cambridge and London.

Toeplitz, Otto and Hans Rademacher *see* Rademacher, Hans and Otto Toeplitz

Tonelli, Leonida

1921/23 *Fondamenti di calcolo delle variazioni*, 2 vols., Zanichelli: Bologna.

1926 Sur la méthode d'adjonction dans le calcul des variations, *Comptes rendus hebdomadaires des séances de l'Acadèmie des sciences* **182**, 678–679.

Uffink, Jos, Harvey R. Brown, and Wayne Myrvold *see* Brown, Harvey R., Wayne Myrvold, and Jos Uffink

Vainchtein, Dmitri and Hassan Aref *see* Aref, Hassan and Dmitri Vainchtein

Valentine, Frederick A.

1937 The problem of Lagrange with differential inequalities as added side conditions, in Gilbert Ames Bliss (ed.), *Contributions to the calculus of variations, 1933–1937*, The University of Chicago Press: Chicago, 403–447.

van der Waals, Johannes Diderik

1873 *Over de Continuiteit van den Gas- en Vloesistoftoestand*, Ph. D. Thesis, University of Leiden, A. W. Sijthoff: Leiden.

Vasilyev, N. S.

1913 *On the motion of infinitely thin vortices*, *Zapiski Matematicheskogo otdeleniya Novorossiyskogo obschestva Estesvoznaniya*, vol. 22: Odessa (Russian).

Vessiot, Ernest

1900 Gewöhnliche Differentialgleichungen; elementare Integrationsmethoden, in *Enzyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen*, vol. 2: *Analysis*, part 1.1, B. G. Teubner: Leipzig 1899–1916, 230–293.

von Escherich, Gustav

1898a Die zweite Variation der einfachen Integrale (I. Mittheilung), *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, II. Abtheilung* **107**, 1191–1252.

1898b Die zweite Variation der einfachen Integrale (II. Mittheilung), *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, II. Abtheilung* **107**, 1267–1326.

1898c Die zweite Variation der einfachen Integrale (III. Mittheilung), *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, II. Abtheilung* **107**, 1381–1428.

1899 Die zweite Variation der einfachen Integrale (IV. Mittheilung), *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, II. Abtheilung* **108**, 1269–1340.

1901 Die zweite Variation der einfachen Integrale (V. Mittheilung), *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, II. Abtheilung* **110**, 1355–1421.

1902 Ueber eine hinreichende Bedingung für das Maximum und Minimum einfacher Integrale, *Mathematische Annalen* **55**, 108–118.

von Helmholtz, Hermann

1858 Über Integrale der hydrodynamischen Gleichungen, welche den Wirbelbewegungen entsprechen, *Journal für die reine und angewandte Mathematik* **55**, 25-55; reprinted in Hermann von Helmholtz, *Wissenschaftliche Abhandlungen*, vol. I, J. A. Barth: Leipzig 1882, 101–134.

1882 Die Thermodynamik chemischer Vorgänge, *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, **1882**, 22–39.

1884 Studien zur Statik monocyklischer Systeme, *Sitzungsberichte der königlich preussischen Akademie der Wissenschaften zu Berlin*, **1884**, 159–177.

von Mises, Richard

1917 Zur Theorie des Tragflächenauftriebs, *Zeitschrift für Flugtechnik und Motorluftschiffahrt* **8**, 157–163.

1931 Zum Navigationsproblem der Luftfahrt, *Zeitschrift für angewandte Mathematik und Mechanik* **11**, 373–381.

Voss, Aurel

1901 Die Prinzipien der rationellen Mechanik, in *Enzyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen*, vol. 4: *Mechanik*, part 1.1, B. G. Teubner: Leipzig 1901–1908, 3–121.

Watson, Henry William

1893 *A treatise on the kinetic theory of gases*, Clarendon Press: Oxford.

Weber, Heinrich

1869 Ueber die Integration der partiellen Differentialgleichung: $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + k^2 u = 0$, *Mathematische Annalen* **1**, 1–36.

1895 *Lehrbuch der Algebra*, vol. I, Vieweg: Braunschweig.

Weber, Heinrich and Bernhard Riemann *see* Riemann, Bernhard and Heinrich Weber

Weierstraß, Karl

1927 *Mathematische Werke*, vol. 7: *Vorlesungen über Variationsrechnung*, edited by Rudolf Rothe, Akademische Verlagsgesellschaft: Leipzig.

Whitney, Willis R. and Arthur Amos Noyes *see* Noyes, Arthur Amos and Willis R. Whitney

Whittemore, James K.

1901a Lagrange's equation in the calculus of variations and the extension of a theorem of Erdmann, *Annals of Mathematics*, 2nd series, **2**, 130–136.

1901b The isoperimetrical problem on any surface, *Annals of Mathematics*, 2nd series, **2**, 175–178.

Wirtinger, Wilhelm

1901 Algebraische Funktionen und ihre Integrale, in *Enzyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen*, vol. 2: *Analysis*, part 2, B. G. Teubner: Leipzig 1899–1916, 115–175.

Yodovich, Victor I. and Leonid G. Kurakin *see* Kurakin, Leonid G. and Victor I. Yudovich

Yoshiye, Takuji

1903 Anwendungen der Variationsrechnung auf partielle Differentialgleichungen mit zwei unabhängigen Variabeln, *Mathematische Annalen* **57**, 185–194.

Young, Laurence Chisholm

1969 *Lectures on the calculus of variations and optimal control theory*, W. B. Saunders Company: Philadelphia, London, and Toronto.

Zermelo, Ernst

1894 *Untersuchungen zur Variations-Rechnung*, Ph. D. Thesis, University of Berlin.

1896a Ueber einen Satz der Dynamik und die mechanische Wärmetheorie, *Annalen der Physik und Chemie,* new series, **57**, 485–494.

1896b Ueber mechanische Erklärungen irreversibler Vorgänge. Eine Antwort auf Hrn. Boltzmann's "Entgegnung", *Annalen der Physik und Chemie,* new series, **59**, 793–801.

1899a Ueber die Bewegung eines Punktsystems bei Bedingungsungleichungen, *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Aus dem Jahre 1899,* 306–310.

s1899b Wie bewegt sich ein unausdehnbarer materieller Faden unter dem Einfluß von Kräften mit dem Potentiale $W(x, y, z)$?, manuscript, 2 pp., Niedersächsische Staats- und Universitätsbibliothek Göttingen, Cod. Ms. D. Hilbert 722.

1900 Über die Anwendung der Wahrscheinlichkeitsrechnung auf dynamische Systeme, *Physikalische Zeitschrift* **1**, 317–320.

1901 Ueber die Addition transfiniter Cardinalzahlen, *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Aus dem Jahre 1901*, 34–38; reprinted with English translation in *Zermelo 2010*, 70–79.

1902a Hydrodynamische Untersuchungen über die Wirbelbewegungen in einer Kugelfläche (Erste Mitteilung), *Zeitschrift für Mathematik und Physik* **47**, 201–237.

s1902b Hydrodynamische Untersuchungen über die Wirbelbewegungen in einer Kugelfläche (Zweite Mitteilung), manuscript with pages 61–117, Universitätsarchiv Freiburg, Zermelo *Nachlass*, C 129/225.

s1902c §5. Die absolute Bewegung, manuscript with pages 119–131, Universitätsarchiv Freiburg, Zermelo *Nachlass*, C 129/253.

1902d Zur Theorie der kürzesten Linien, *Jahresbericht der Deutschen Mathematiker-Vereinigung* **11**, 184–187.

1904 Beweis, daß jede Menge wohlgeordnet werden kann, *Mathematische Annalen* **59**, 514–516; reprinted with English translation in *Zermelo 2010*, 114–119.

1904a Über die Herleitung der Differentialgleichung bei Variationsproblemen, *Mathematische Annalen* **58**, 558–564.

1906 Review of *Gibbs 1902* and *Gibbs 1905*, *Jahresbericht der Deutschen Mathematiker-Vereinigung* **15**, 232–242.

1908a Neuer Beweis für die Möglichkeit einer Wohlordnung, *Mathematische Annalen* **65**, 107–128; reprinted with English translation in *Zermelo 2010*, 120–159.

1908b Untersuchungen über die Grundlagen der Mengenlehre. I, *Mathematische Annalen* **65**, 261–281; reprinted with English translation in *Zermelo 2010*, 188–229.

1909a Sur les ensembles finis et le principe de l'induction complète, *Acta mathematica* **32**, 185–193; reprinted with English translation in *Zermelo 2010*, 236–253.

1909b Ueber die Grundlagen der Arithmetik, in G. Castelnuovo (ed.), *Atti del IV Congresso Internazionale dei Matematici (Roma, 6–11 Aprile 1908)*, vol. 2: *Comunicazioni delle sezioni I e II*, Tipografia della R. Accademia dei Lincei: Roma, 8–11; reprinted with English translation in *Zermelo 2010*, 252–259.

1913 Über eine Anwendung der Mengenlehre auf die Theorie des Schachspiels, in E. W. Hobson and A. E. H. Love (eds.), *Proceedings of the 5th international congress of mathematicians, Cambridge 1912*, vol. 2, Cambridge University Press: Cambridge, 501–504; reprinted with English translation in *Zermelo 2010*, 266–273.

1914 Über ganze transzendente Zahlen, *Mathematische Annalen* **75**, 434–442; reprinted with English translation in *Zermelo 2010*, 278–295.

s1921 Thesen über das Unendliche in der Mathematik, typescript, 1 p., Universitätsarchiv Freiburg, Zermelo *Nachlass*, part of C 129/278; reprinted with English translation in *Zermelo 2010*, 306–307.

1927 Über das Maß und die Diskrepanz von Punktmengen, *Journal für die reine und angewandte Mathematik* **158**, 154–167; reprinted with English translation in *Zermelo 2010*, 312–347.

1928 Die Berechnung der Turnier-Ergebnisse als ein Maximumproblem der Wahrscheinlichkeitsrechnung, *Mathematische Zeitschrift* **29**, 436–460.

1929 = *Zermelo 1928*; see fn.* on page 616.

1929a Über den Begriff der Definitheit in der Axiomatik, *Fundamenta mathematicae* **14**, 339–344; reprinted with English translation in *Zermelo 2010*, 358–367.

s1929b Vortrags-Themata für Warschau 1929, typescript, 9 pp., Universitätsarchiv Freiburg, Zermelo *Nachlass*, C 129/288; reprinted with English translation in *Zermelo 2010*, 374–389.

1930a Über Grenzzahlen und Mengenbereiche. Neue Untersuchungen über die Grundlagen der Mengenlehre, *Fundamenta mathematicae* **16**, 29–47; reprinted with English translation in *Zermelo 2010*, 400–431.

1930b Über die logische Form der mathematischen Theorien, *Annales de la Société Polonaise de Mathématique* **9**, 187; reprinted with English translation in *Zermelo 2010*, 430–431.

1930c Über die Navigation in der Luft als Problem der Variationsrechnung, *Jahresbericht der Deutschen Mathematiker-Vereinigung* **39**, 2nd section, 44–48.

s1930d Bericht an die Notgemeinschaft der Deutschen Wissenschaft über meine Forschungen betreffend die *Grundlagen der Mathematik,* typescript, 5 pp. with 2 pp. appendices, dated 3 December 1930, Universitätsarchiv Freiburg, Zermelo *Nachlass*, part of C 129/140; reprinted with English translation in *Zermelo 2010*, 434–443.

s1930e Über das mengentheoretische Modell, typescript, 5 pp., Universitätsarchiv Freiburg, Zermelo *Nachlass*, part of C 129/292. Copy of first 4 pages, dated September 1930, in C 129/212; reprinted with English translation in *Zermelo 2010*, 446–453.

1930f Aus Homers Odyssee, *Wiener Blätter für die Freunde der Antike* **6**, 93–94; reprinted with English translation in *Zermelo 2010*, 462–467.

1931a Über das Navigationsproblem bei ruhender oder veränderlicher Windverteilung. *Zeitschrift für angewandte Mathematik und Mechanik* **11**, 114–124.

s1931b Letter to Kurt Gödel of 21 September 1931, Gödel *Nachlass*, Institute for Advanced Study, 013270. Draft in Universitätsarchiv Freiburg, Zermelo *Nachlass*, part of C 129/36; reprinted with English translation in *Gödel 2003*, 420–423, and in *Zermelo 2010*, 486–489.

s1931c Letter to Reinhold Baer of 7 October 1931, in the possession of Ulrich Felgner and given to him by Reinhold Baer; reprinted with English translation in *Zermelo 2010*, 490–493.

s1931d Letter to Kurt Gödel of 29 October 1931, Gödel *Nachlass*, Institute for Advanced Study, 013271. Carbon copy in Universitätsarchiv Freiburg, Zermelo *Nachlass*, part of C 129/36; reprinted with English translation in *Gödel 2003*, 430–431, and in *Zermelo 2010*, 500–501.

s1931e Seven notes on ordinal numbers and large cardinals, typescript, 8 pp., undated, ca. 1931, Universitätsarchiv Freiburg, Zermelo *Nachlass*, C 129/236; reprinted with English translation in *Zermelo 2010*, 504–515.

s1931f Sätze über *geschlossene Bereiche*, typescript, 2 pp., undated, ca. 1931, Universitätsarchiv Freiburg, Zermelo *Nachlass*, C 129/236; reprinted with English translation in *Zermelo 2010*, 520–523.

s1931g Allgemeine Theorie der mathematischen Systeme, manuscript, dated July 1931, Universitätsarchiv Freiburg, Zermelo *Nachlass*,

C 129/294; reprinted with English translation in *Zermelo 2010*, 528–531.

1932a Über Stufen der Quantifikation und die Logik des Unendlichen, *Jahresbericht der Deutschen Mathematiker-Vereinigung* **41**, 2nd section, 85–88; reprinted with English translation in *Zermelo 2010*, 542–549.

1932b Über mathematische Systeme und die Logik des Unendlichen, *Forschungen und Fortschritte* **8**, 6–7; reprinted with English translation in *Zermelo 2010*, 550–555.

1932c Vorwort. Preface to *Cantor 1932*, in *Cantor 1932,* iii–v; reprinted with English translation in *Zermelo 2010*, 560–565.

s1932d Mengenlehre 1932, typescript, 5 pp., Universitätsarchiv Freiburg, Zermelo *Nachlass*, part of C 129/290; reprinted with English translation in *Zermelo 2010*, 564–571.

1933a Über die Bruchlinien zentrierter Ovale. Wie zerbricht ein Stück Zucker?, *Zeitschrift für angewandte Mathematik und Mechanik* **13**, 168–170.

s1933b Die unbegrenzte Zahlenreihe und die exorbitanten Zahlen, typescript, 1 p., November 1933, Universitätsarchiv Freiburg, Zermelo *Nachlass*, part of C 129/239; reprinted with English translation in *Zermelo 2010*, 572–573.

1934 Elementare Betrachtungen zur Theorie der Primzahlen, *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Fachgruppe I; Mathematik*, new series, **1**, 43–46; reprinted with English translation in *Zermelo 2010*, 576–581.

1935 Grundlagen einer allgemeinen Theorie der mathematischen Satzsysteme (Erste Mitteilung), *Fundamenta mathematicae* **25**, 136–146; reprinted with English translation in *Zermelo 2010*, 582–599.

s1937 Der Relativismus in der Mengenlehre und der sogenannte Skolemsche Satz, manuscript, dated 4 October 1937, 3 pp., Universitätsarchiv Freiburg, Zermelo *Nachlass*, part of C 129/217; reprinted with English translation in *Zermelo 2010*, 602–605.

s1941 Letter to Paul Bernays of 1 October 1941, ETH-Bibliothek Zürich, *Nachlass* Bernays, Hs. 975.5259; reprinted with English translation in *Zermelo 2010*, 608–609.

2010 *Collected Works / Gesammelte Werke*, vol. I / Bd. I: *Set theory, miscellanea / Mengenlehre, Varia*, edited by / herausgegeben von Heinz-Dieter Ebbinghaus and Akihiro Kanamori, Schriften der Mathematisch-naturwissenschaftlichen Klasse der Heidelberger Akademie der Wissenschaften, vol. 21, Springer-Verlag: Berlin et al.

Zermelo, Ernst and Hans Hahn *see* Hahn, Hans and Ernst Zermelo

Zermelo, Ernst and Ernst Heinrich Riesenfeld *see* Riesenfeld, Ernst Heinrich and Ernst Zermelo

Index