1

Berlin *1871–1897*

1.1 Family and Youth

Ernst Friedrich Ferdinand Zermelo was born in Berlin on 27 July 1871.[1] It was the time when Georg Cantor wrote his papers on trigonometric series.[2] As Zermelo expressed it about 60 years later, these papers led with inherent necessity to the conception of transfinite ordinal numbers; they hence could be considered the birthplace of Cantorian set theory, i. e., the theory whose transformation into an axiomatic form would be an essential part of Zermelo's scientific achievements.

There exist several hypotheses about the origin of the name "Zermelo." They all seem to go back to Zermelo himself. According to the most well-known (*Reid 1970*, 98), Zermelo is said to have replied to newcomers in Göttingen when asked about his strange name, that it resulted from the word "Walzermelodie" ("waltz melody") by deleting the first and the last syllable. In conversations with his wife Gertrud, Zermelo suggested the more seriously meant possibility that it was derived from "Termeulen," a form of the name "Zurmühlen" ("at the mill") current in the lower Rhine and Friesian areas. The name most probably originates from the variant "Tormählen"[3] which evolved in northeastern Germany, the region the Zermelo family comes from, during the 16th century; the Low German "T" might have been transformed into the High German "Z" in the course of one of the frequently observed hypercorrections.[4] With the death of Zermelo's widow Gertrud in December 2003 the name "Zermelo" disappeared from the German telephone listings. She seems to have been its last bearer.

[1] Details about Zermelo's family are taken from documents held in MLF.
[2] *Cantor 1870a, 1870b, 1871a, 1871b, 1872.*
[3] The "ä" is pronounced like the "ea" in "bear."
[4] *Kunze 1998*, 159, and oral communication with Konrad Kunze.

H.-D. Ebbinghaus, *Ernst Zermelo*,
DOI 10.1007/978-3-662-47997-1_1

Zermelo had five sisters, the older Anna (*1870*–1917)[5] and the younger ones Elisabeth (1875–1952), Margarete (1876–1959), Lena (*1877*–1906) and Marie (*1878*–1908). His parents had married in Delitzsch, Saxony, on 18 May 1869. His mother, Maria Auguste née Zieger, was born there on 11 February 1847 as the only child of the surgeon Dr. Ottomar Hugo Zieger and his wife Auguste née Meißner. Both her parents died early, her mother just one week after giving birth to her, her father one year later from tuberculosis. She then grew up in the wider family, which included the natural historian Christian Gottfried Ehrenberg (1795–1876) who was to become known for his work on microorganisms and who accompanied Alexander von Humboldt on his Russian expedition of 1829.[6] Like her parents, Maria Auguste suffered from poor health. Exhausted from the strain of multiple pregnancies, she died on 3 June 1878, soon after the birth of Marie. After her death a maid, Olga Pahlke, took care of the household.

Not only Lena and Marie, who died at a young age, but also the other children inherited their mother's delicate health; they all fell ill with tuberculosis. Neither of the three surviving sisters married. The letters which Zermelo wrote to them attest to a close relationship between the siblings, and care and concern on Zermelo's side.

Zermelo's father, Theodor Zermelo, was born on 9 June 1834 in the East Prussian town of Tilsit (today Sovetsk, Russia) on the river Memel (Neman). Theodor's father Ferdinand Zermelo, a bookbinder who had graduated from the college of arts and crafts in Königsberg (today Kaliningrad, Russia), took over a shop for books, arts, and stationery in Tilsit in 1839. Documentation about him and his wife Bertha née Haberland is sparse: a carefully written diary about a trip to Paris and London, which Ferdinand undertook in the summer of 1851, and newspaper reports and obituary notices on the occasion of his death showing that he was an integral part and a supporter of cultural life in Tilsit and engaged in the town's social institutions. He served as a councillor and was a co-founder of the Tilsit Schiller Society and the Tilsit Art Society.

After finishing his secondary school education in Tilsit, Theodor Zermelo studied mainly history and geography, first in Königsberg and then in Berlin. Having attained his doctorate in 1856 with a dissertation in history, he completed his studies with the exams *pro facultate docendi,* i. e., his state exams for teaching at secondary schools, which allowed him to teach history and geography as main subjects and Greek and Latin as subsidiary ones; a later supplementary examination also permitted him to teach French. He refrained from a further examination in the natural sciences, but completed an additional diploma in mathematics. In its evaluation of 8 March 1857, the

[5] The birth dates and death dates concerning Zermelo's sisters which cannot be traced back to reliable documents are rendered in italics.

[6] In a *curriculum vitae* written after 1933 (UAF, C 129/273) Zermelo counts not only Ehrenberg, but also the physicist Arnold Sommerfeld among his relatives.

Königlich-Wissenschaftliche Prüfungskommission (Royal Academic Examination Board) found him to have "an adequate knowledge of the elements of geometry and algebra for teaching mathematics in the lower grades." It further observes: "As concerns his philosophical knowledge, he has a solid grasp of the basic notions of logic and general grammar (allgemeine Grammatik) and shows an assured competence in their application." Theodor Zermelo became a teacher at the Friedrichswerder Gewerbeschule in Berlin, a secondary school that was vocationally oriented and did not have Latin and Greek as subjects like the traditional *Gymnasium.* In 1886 he was appointed *Gymnasialprofessor.*

Zermelo's parents Maria Auguste and Theodor Zermelo

Altogether, Ernst Zermelo grew up in a family with an academic background, but not one with a significant orientation towards mathematics and the sciences. His sisters Anna and Elisabeth became painters, his sister Margarete a teacher of history and modern languages. Too little information is available about Zermelo's childhood and youth to come to reliable conclusions concerning his education. As his mother died when he was seven years old, her influence was limited to his early childhood. Zermelo described her to his wife Gertrud as a delicate and sensitive woman. She confided many of her thoughts to a diary.

The influence of his father is more obvious. As a teacher, Theodor Zermelo appears to have been popular and highly regarded. In a letter which former pupils wrote to congratulate him on his nomination as a professor it says:

> We all are delighted about this event, since each of us has had the good fortune to have been taught by you just a few years ago. We sincerely regret that we now have to do without. We all owe you a great part of our education, and therefore we wish you with all our heart a long and happy life.

Left: Zermelo as a baby with his mother
Right: Zermelo with his parents and his sisters Anna (right) and Elisabeth

Zermelo also respected his father and held him in high esteem, as is illustrated by a letter in which he thanks him for the presents he had received on his 17th birthday:[7]

> Dear father,
> Your affectionate congratulations on my birthday yesterday and your presents, which are as generous as ever, caused me great pleasure. Above all I would like to express my heartfelt thanks to you. Your very kind offer of a trip which I certainly do not particularly deserve was a complete surprise to me. I can only hope that your plan will indeed be realised, not only in my own interest, but especially in yours, as it is to depend precisely on your state of health. It goes without saying that I will gladly obey your "conditions." [...]
> Your obedient son
> Ernst. (OV 1.01)

Theodor Zermelo's state of health got worse and led to his death half a year later, on 24 January 1889. Zermelo and his sisters had become orphans. Thanks to the family's assets, their livelihood was secured for the time being.

[7]Letter of 28 July 1888; MLF. A letter which Zermelo wrote to his father in April (copy in MLF) shows that Theodor Zermelo was not living together with his children at that time; perhaps he was staying in hospital because of a serious illness. Earlier letters (MLF) document that the Zermelo siblings often spent their holidays without their father.

Zermelo had a kind heart, as is attested by some of his contemporaries.[8] However, he was also prone to sharp and polemical reactions and did not refrain from trenchant irony when he was convinced of his opinions, both in scientific and in political matters. He did so even if he had to suffer from the consequences, as with his criticism of the Nazis. Whereas the first trait might be attributable to the influence of his mother, his determination to speak his mind may be due to the example of his father. In 1875, Theodor Zermelo had published a widely acknowledged[9] paper (*1875*) on the historian and philologist August Ludwig von Schlözer (1735–1809) of Göttingen, the most influential anti-absolutist German publicist of the early age of enlightenment.[10] It was his aim to raise awareness of those values expressed by Schlözer's plea for free speech, sincerity, tolerance, and humanity on which, in his opinion, the newly founded German Reich should be built. There are indications that the views of his father as reflected in this treatise and the example of von Schlözer had their effect on Zermelo. In the copy of the treatise, which he owned, many passages are underlined, among them those that concern the determined defence of one's own convictions. In a 1935 letter to his sisters, which describes his attitude towards the Nazis (4.11.3), he argues in the spirit of his father. Moreover, he used to talk about von Schlözer with his wife Gertrud, thereby arousing her interest in von Schlözer's daughter Dorothea (1770–1825), known as "the loveliest incarnation of enlightenment." In 1787, at the age of 17, she was the first woman to receive a Ph. D. at the University of Göttingen.[11]

Zermelo shared with his father an interest in poetry. As a young teacher, Theodor Zermelo had compiled a large selection of his own translations of poems from England, the United States, France, and Italy under the title "Aus der Fremde" ("From Foreign Lands"). Zermelo was later to translate parts of the Homeric epics, and he enjoyed commenting on daily events in the form of poems. In 1885, at the age of thirteen, he made "a metrical translation from the first book of Virgil's Aeneid" as a birthday present for his father (MLF).

1.2 School and University Studies

In 1880 Zermelo entered the Luisenstädtisches Gymnasium in Berlin, situated in Brandenburgstraße (today Lobeckstraße).[12] Its name refers to the neighbourhood Luisenstadt honouring the Prussian Queen Luise Auguste Wilhelmine Amalie (1776–1810), the wife of King Friedrich Wilhelm III (1770–1840). The Luisenstädtisches Gymnasium had been opened on 10 October 1864, after the Berlin authorities had decided to establish new secondary

[8] So Wilhelm Süss in a letter to Zermelo of 27 July 1951; MLF.
[9] Cf. *Brockhaus' Konversationslexikon*, 14th edition, vol. 14 (1895), 525.
[10] Cf., e. g., *Peters 2003*.
[11] Cf. *Kern and Kern 1988* or *Küssner 1976*.
[12] The certificates relating to Zermelo's education at school and university are held in MLF.

Zermelo as a schoolboy with his sisters Anna (right) and Elisabeth

schools and to extend existing ones in a reaction to the large increase in the population of Berlin.[13] Teaching started in the autumn of 1864 with 86 students. When Zermelo entered the school it had 587 students with 50 students in each entry class.[14] After the turn of the century the number of students in Luisenstadt decreased steadily. The school was eventually closed at its first site and transferred further to the west. The building was then destroyed in World War II.[15]

Shortly after the death of his father, Zermelo finished school. His final school certificate (*Zeugnis der Reife*), dated 1 March 1889, certifies good results in religious education, German language, Latin, Greek, history, geography, mathematics, and physics. His performance in French was adequate. The

[13] *Erster Jahresbericht über das Luisenstädtische Gymnasium in Berlin*, Berlin 1865, 33–34.

[14] *Sechzehnter Jahresbericht über das Luisenstädtische Gymnasium in Berlin*, Berlin 1881, 7.

[15] In 1928 the school was renamed after Heinrich Schliemann; since 1993 its name is "Heinrich-Schliemann-Oberschule (Gymnasium)".

comment in mathematics says that he followed the instructions with good understanding and that he has reliable knowledge and the ability to use it for skilfully solving problems. In physics it is testified that he is well-acquainted with a number of phenomena and laws. Under the heading "Behaviour and Diligence" it is remarked that he followed the lessons with reflection, but that he occasionally showed a certain passivity as a result of physical fatigue, an indication that Zermelo suffered from poor health already during his school days. The certificate furthermore informs us that he was exempted from oral examinations and that he was now going to take up university studies in mathematics and physics.

In the summer semester 1889 Zermelo matriculated at the Friedrich Wilhelm University (now Humboldt University) in Berlin. During his course of studies he took one semester each at the University of Halle-Wittenberg (winter semester 1890/91) and at the University of Freiburg (summer semester 1891). After his final examinations in 1987 he continued his studies at Göttingen.

When Zermelo's father died in 1889, he left behind six children who were still minors, now under the guardianship of Amtsgerichtsrat (judge of a county court) Muellner. On 24 December 1891 Muellner testified that the father's estate was needed to provide for the younger siblings, thereby helping Zermelo get a grant to finish his university studies. Zermelo was awarded grants from the Moses Mendelssohn Foundation (probably 1891/92) and the Magnus Foundation (1892/93).[16] Both foundations belonged to the Friedrich Wilhelm University, set up to support gifted students. The first was named after Moses Mendelssohn (1729–1786), the Berlin philosopher of German enlightenment, the second after the physicist Heinrich Gustav Magnus (1802–1870) who is regarded as the progenitor of Berlin physics. The Magnus Foundation was the most important one in the second half of the 19th century. Magnus' widow had donated 60000 Marks to the University. The interest allowed the support of two excellent students of mathematics or the sciences with an annual grant of 1200 Marks, a significant sum at that time. Karl Weierstraß himself had outlined and defended its statutes.[17] As stipulated, Zermelo had to pass special examinations in order to prove that he had the level of knowledge expected at that stage of his studies. Two testimonials in mathematics, one for the Mendelssohn Foundation signed by Lazarus Fuchs and one for the Magnus Foundation signed by Hermann Amandus Schwarz, can be found among his papers.

In Berlin Zermelo enrolled for philosophy and took courses in mathematics, above all with Johannes Knoblauch, a student of Karl Weierstraß, and with Fuchs. Furthermore, he attended courses on experimental physics and heard a course on experimental psychology by Hermann Ebbinghaus. In Halle he enrolled for mathematics and physics, attending Georg Cantor's courses on

[16] Cf. *Biermann 1988*, 126.
[17] Ibid., 126.

elliptical functions and number theory, Albert Wangerin's courses on differential equations and on spherical astronomy, but also a course given by Edmund Husserl on the philosophy of mathematics at the time when Husserl's *Philosophie der Arithmetik* (*1891a*) was about to be published. Zermelo also took part in the course on logic given by Benno Erdmann, then one of the leading philosophical (psychologistic) logicians. In Freiburg he studied mathematical physics with Emil Warburg, analytical geometry and the method of least squares with Jakob Lüroth, experimental psychology with Hugo Münsterberg, and history of philosophy with Alois Riehl. Furthermore he attended a seminar on Heinrich von Kleist.

Having returned to Berlin, Zermelo attended several courses by Max Planck on theoretical physics, among them a course on the theory of heat in the winter semester 1893/94. He also attended a course on the principle of the conservation of energy by Wilhelm Wien (summer semester 1893). In mathematics he took part in courses on differential equations by Fuchs, algebraic geometry by Ferdinand Georg Frobenius, and non-Euclidean geometry by Knoblauch. The calculus of variations, the central topic of Zermelo's later work, was taught by Hermann Amandus Schwarz in the summer semester 1892. In philosophy he attended "Philosophische Übungen" by the neo-Kantian Friedrich Paulsen and Wilhelm Dilthey's course on the history of philosophy. He also took a course on psychology, again by Hermann Ebbinghaus (winter semester 1893/94).

Zermelo's Ph. D. thesis *Untersuchungen zur Variations-Rechnung* (*Investigations In the Calculus of Variations*) (cf. 1.3) was suggested and guided by Hermann Amandus Schwarz. Schwarz, who had started his studies in chemistry, was brought over to mathematics by Weierstraß and was to become his most eminent student. In particular, he aimed at "keeping a firm hold on and handing down the state of mathematical exactness which Weierstraß had reached" (*Hamel 1923*, 9). In 1892, then a professor in Göttingen, he was appointed Weierstraß' successor in Berlin. Zermelo became his first Ph. D. student there.

On 23 March 1894 Zermelo, then 22 years old, applied to begin the Ph. D. procedure.[18] Schwarz and the second referee Fuchs delivered their reports in July 1894. The oral examination took place on 6 October 1894. It included the defence of three theses that could be proposed by the candidate. Zermelo had made the following choice (*1894*, 98):

I. In the calculus of variations one has to place greater emphasis than before on a precise definition of the maximum or minimum.[19]

[18] Details follow the files of Zermelo's Ph. D. procedure in the archives of the Humboldt University in Berlin and are quoted from *Thiele 2006*, 298–303.

[19] "In der Variations-Rechnung ist auf eine genaue Definition des Maximums oder Minimums grösserer Wert als bisher zu legen."

II. It is not justified to charge physics with the task of reducing all phenomena in nature to the mechanics of atoms.[20]

III. Measurement is to be conceived of as a universally applicable expedient to distinguish and to compare continuously varying qualities.[21]

The latter two theses, in particular the second one, are clearly aimed against early atomism in physics and the mechanical explanations coming with it. They will gain in substance within the next two years and lead to a serious debate with Ludwig Boltzmann about the scope of statistical mechanics (cf. 1.4).

After having received his doctorate, Zermelo got a position as an assistant to Max Planck at the Berlin Institute for Theoretical Physics from December 1894 to September 1897.[22] During this time he edited a German translation (*Glazebrook 1897*) of Richard Tetley Glazebrook's elementary textbook on light (*Glazebrook 1894*). He also prepared for his exams *pro facultate docendi*, which he passed successfully on 2 February 1897. In philosophy he wrote an essay entitled "Welche Bedeutung hat das Prinzip der Erhaltung der Energie für die Frage nach dem Verhältnis von Leib und Seele?" ("What is the Significance of the Principle of the Conservation of Energy for the Question of the Relation Between Body and Mind?").[23] According to the reports Zermelo was well-informed about the history of philosophy and systematical subjects. He was examined in religious education (showing excellent knowledge of the Holy Bible and church history) and in German language and literature. As to his exams in mathematics the report says that although he was not always aware of the methods in each domain, it was nevertheless evident that he had acquired an excellent mathematical education. The physics part was rated excellent. The exams in geography showed that he was excellent in respect to theoretical explanations, but that he was not that affected by "studying reality." Expressing its conviction that he would be able to close these gaps in the future, the committee allowed him to teach geography in higher classes. Finally it was certified that he had the knowledge for teaching mineralogy in higher classes and chemistry in the middle classes.

Already in the summer of 1896, while still an assistant to Plank, Zermelo applied for a position as assistant at the Deutsche Seewarte in Hamburg, the central institution for maritime meteorology of the German Reich. The application was supported by Schwarz who reported on 20 July 1896 that he knew Zermelo from mathematics courses at the University of Berlin, from a private course ("Privatissimum"), and as supervisor of his dissertation. He expressed his conviction that Zermelo had excellent skills for investigations

[20] "Mit Unrecht wird der Physik die Aufgabe gestellt, alle Naturerscheinungen auf Mechanik der Atome zurückzuführen."

[21] "Die Messung ist aufzufassen als das überall anwendbare Hilfsmittel, stetig veränderliche Qualitäten zu unterscheiden und zu vergleichen."

[22] Assessment from Max Planck of 3 November 1897; MLF.

[23] Draft in shorthand in UAF, C 129/270.

in theoretical mathematics. Planck confirmed on 7 July 1896 (MLF) that he was extraordinarily satisfied with Zermelo's achievements where he "made use of his special mathematical talent in an utmost conscentious manner" (OV 1.02).

For unknown reasons, however, Zermelo ultimately decided to pursue the aim of obtaining an academic position, taking up further studies at the University of Göttingen. He enrolled for mathematics on 4 November 1897. In the winter semester 1897/98 he heard David Hilbert's course on irrational numbers and attended the mathematical seminar on differential equations in mechanics directed by Hilbert and Felix Klein. He took exercises in physics with Eduard Riecke and heard thermodynamics with Oskar Emil Meyer. During the next semester he took part in a course on set theory given by Arthur Schoenflies and in Felix Klein's mathematical seminar.

Looking back to Zermelo's university studies, one may emphasize the following points:

He acquired a solid and broad knowledge in both mathematics and physics, his main subjects. His advanced studies directed him to his early specialities of research, to mathematical physics and thermodynamics, and to the calculus of variations, the topic of his dissertation.

He attended courses by Georg Cantor, but no courses on set theory, neither in Halle[24] nor at other places, at least until the summer semester 1898 when Schoenflies gave his course in Göttingen. Furthermore, he had no instruction in mathematical logic at a time, however, when German universities offered this subject only in Jena with Gottlob Frege and in Karlsruhe with Ernst Schröder. He got his logical training from Benno Erdmann, who defined logic as "the general, formal and normative science of the methodological preconditions of scientific reasoning" (*1892*, 25).

Given his teachers in philosophy, Riehl, Paulsen, Dilthey, and Erdmann, it can be assumed that he had a broad overall knowledge of philosophical theories. In Halle he even got acquainted with Husserl's phenomenological philosophy of mathematics *in statu nascendi*. His interest in experimental psychology as taught by Hermann Ebbinghaus and Hugo Münsterberg is quite evident.

1.3 Ph. D. Thesis and the Calculus of Variations

We recall that Zermelo's Ph. D. thesis *Untersuchungen zur Variations-Rechnung* (*Investigations In the Calculus of Variations*), *1894*, was guided by

[24]Cantor lectured on set theory only in the summer semester 1885 in a course entitled "Zahlentheorie, als Einleitung in die Theorie der Ordnungstypen" ("Number Theory, as an Introduction to the Theory of Order Types") (information by Rüdiger Thiele; cf. also *Grattan-Guinness 1970*, 81, or *Purkert and Ilgauds 1987*, 104).

Hermann Amandus Schwarz. It may have been inspired by Schwarz's lectures in the summer semester 1892. Schwarz proposed to generalize methods and results in the calculus of variations, which Weierstraß had obtained for derivatives of first order, to higher derivatives.[25]

The calculus of variations treats problems of the following kind: Given a functional J from the set M into the set of reals, what are the elements of M for which J has an extremal value? The transition from J to $-J$ shows that it suffices to treat either minima or maxima. In a classical example, a special form of the so-called *isoperimetric problem,* M is the set of closed curves in the plane of a fixed given perimeter, and J maps a curve in M to the area it encloses. The problem asks for the curves enclosing an area of maximal size. In Zermelo's Ph. D. thesis M is a set of curves and J the formation of integrals along curves in M for a given integrand which may now be of a more general kind. In Zermelo's own words (*1894*, 24):

> [Given] an integral
>
> $$J = \int_{t_1}^{t_2} F\,dt,$$
>
> [...] our task is as follows:
> Suppose that
>
> $$F\Big(x^{(\mu)}, y^{(\mu)}\Big) = F\Big(x, x', .. x^{(n)};\, y, y', .. y^{(n)}\Big)$$
>
> 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
>
> $$x = \varphi(t), \ \ y = \psi(t)$$
>
> satisfying certain prescribed conditions, we seek to determine a special curve a furnishing a value of the integral
>
> $$J = \int_{t_1}^{t_2} F\Big(x^{(\mu)}, y^{(\mu)}\Big) \quad \Big(x^{(\mu)} = \frac{d^\mu x}{dt^\mu},\ y^{(\mu)} = \frac{d^\mu y}{dt^\mu}\Big)$$
>
> taken along the curve between certain limits that is *smaller* than that for all *neighbouring* curves $\bar{a}$ of the same totality A.

The integrability conditions to which Zermelo refers ensure that the value of the integral does not depend on the parametrization of the curve in question.

Weierstraß had treated his solution for $n = 1$ in his courses, in particular in the course given in the summer semester 1879.[26] This course had been

[25] Also here we follow the files of Zermelo's Ph. D. procedure in the archives of the Humboldt University in Berlin as quoted in *Thiele 2006*, 298–303.

[26] A comprehensive version of Weierstraß' lectures has been edited as *Weierstraß 1927*; cf. also *Thiele 2006*, 183–243, or *Goldstine 1980*, Ch. 5.

written up under the auspices of the Berlin Mathematical Society. Edmund Husserl had taken part in this project.[27]

Zermelo studied the lecture notes in the summer of 1892, when he attended Schwarz's course on the calculus of variations (*Zermelo 1894*, 1). Less than two years later he succeeded in solving the task Schwarz had set. In his report of 5 July 1894 Schwarz describes the subject as "very difficult;" he is convinced that Zermelo provided the very best solution and predicts a lasting influence of the methods Zermelo had developed and of the results he had obtained:

> According to my judgement the author succeeds in generalizing the main investigations of Mr. Weierstraß [...] in the correct manner. In my opinion he thus obtained a valuable completion of our present knowledge in this part of the calculus of variations. Unless I am very much mistaken, all future researchers in this difficult area will have to take up the results of this work and the way they are deduced. (OV 1.03)

He marked the thesis with the highest degree possible, *diligentia et acuminis specimen egregium.* Co-referee Fuchs shared his evaluation.

The dissertation starts by exhibiting the integrability conditions mentioned above. Taking them as "a problem of independent interest" (*1894*, 14), Zermelo develops them in a more general framework than needed for the later applications. In the second part he provides a careful definition of minimum (ibid., 25-29), quite in accordance with the first thesis he had chosen for the oral examination. The notion of minimum results from the conditions which he imposes on the totalities A of (parametrizations of) curves and on the relation "$\overline{a}$ is a neighbouring curve of a."[28] In the final parts he carries out the Weierstraßian programme in the framework thus created.[29]

The dissertation ends (p. 96) with "the first clear formulation and proof of the important envelope theorem,"[30] according to Gilbert A. Bliss (*1946*, 24) "one of the most interesting and most beautiful theorems in the domain

[27] Husserl (1859–1938) studied mathematics, first in Leipzig (1876–1878) and then in Berlin (1878–1880) mainly with Weierstraß. His Ph. D. thesis *Beiträge zur Theorie der Variationsrechnung* (*Contributions to the Theory of the Calculus of Variations*) (*1882*) is written in the spirit of Weierstraß; cf. *Thiele 2006*, 293–295.

[28] Partly, these conditions generalize those of Weierstraß for the case "n = 1" in a natural way, demanding, for instance, that for curves in A the nth derivative exists and is continuous. Over and above that, the curves in A have to have not only the same boundary values, but also the same so-called "boundary osculation invariants," among them as a simple example the torsion.

Analogously, the definition of neighbourhood does not only refer to a natural distance between a and $\overline{a}$, but also to quantities corresponding to the osculation invariants, among them again the torsion.

[29] For technical details see *Thiele 2006*, 298–303.

[30] So Hilbert in his assessment of Zermelo of 16 January 1910, written for the University of Zurich; SUB, Cod. Ms. D. Hilbert 492, fols. 4/1-2. – For a sketch of the proof see *Goldstine 1980*, 340–341.

of geometrical analysis." The theorem generalizes the following proposition[31] about geodetic lines to one-parameter families of extremals, likewise providing a criterion for the non-existence of a minimum of the corresponding variational problem among the extremals under consideration:[32]

> If $\mathcal{F}$ is a one-parameter family of geodetic lines issuing from a point P of a surface and E an envelope of $\mathcal{F}$, and if, moreover, $G_1 \in \mathcal{F}$ and $G_2 \in \mathcal{F}$ are tangent to E in P_1 and P_2, respectively, and P_1 precedes P_2 on E, then
>
> $$\text{length of } PP_2 = \text{length of } PP_1 + \text{length of } P_1P_2,$$
>
> where, for example, PP_1 denotes the arc of G_1 between P and P_1, and P_1P_2 denotes the arc on E between P_1 and P_2.

In the "triangle" formed by P, P_1, and P_2, the arc P_1P_2 can be replaced by a shorter curve (the envelope is not geodetic). Hence, there is a curve connecting P with P_2 that is shorter than PP_2, i.e., the (geodetic) arc PP_2 does not provide a curve of minimal length connecting P with P_2.

Later, Adolf Kneser provided generalizations of Zermelo's theorem, discussing also the existence of an envelope – a desideratum Zermelo had left open.[33] Zermelo is, however, more general than Kneser in respect to the variational integrand, allowing derivatives of arbitrary order. Goldstine speaks of the Zermelo-Kneser results as of the "Zermelo-Kneser envelope theorem," giving credit to Zermelo for the "most elegant argument" (*Goldstine 1980*, 340).

Various voices confirm the significance of Zermelo's dissertation for the development of the Weierstraßian direction in the calculus of variations.

David Hilbert writes in an assessment on the occasion of Zermelo's *Habilitation* in 1899, that "the dissertation is of scientific value and enjoys a high estimation among experts."[34]

Adolf Kneser's 1900 monograph on the calculus of variations recommends the dissertation as giving valuable information about the Weierstraßian methods and the case of higher derivatives. In the preface Kneser says (*1900*, IV):

> The relationship between my work and the investigations of Weierstraß requires a special remark. By his creative activity Weierstraß opened new ways in the calculus of variations. As is well-known, his research is not available in a systematic representation; as the most productive sources I

[31] Attributed to Jean Gaston Darboux; cf. *Bolza 1904*, 166, or *Goldstine 1980*, 339.

[32] Namely, by providing a proof of the necessity of the so-called Jacobi condition; cf., for example, *Bliss 1946*, Sect. 10.

[33] *Kneser 1898*, 27; *Kneser 1900*, §15. Cf. *Goldstine 1980*, Sect. 7.5, or *Bolza 1909*, §43, for details. For a further generalization of the envelope theorem obtained by Carathéodory in 1923, cf. *Carathéodory 1935*, 292–293 and 398.

[34] OV 2.03, second paragraph.

> used the Ph. D. thesis of Zermelo and a paper of Kobb[35] [...]. In a modified and generalized form they contain all essential ideas of Weierstraß as far as they are related to our topic.

Oskar Bolza (1857–1942), one of the most influential proponents of the calculus of variations, also gives due respect to Zermelo: his epochal monograph (*1909*) contains numerous quotations of Zermelo's work.[36]

Constantin Carathéodory, in his similarly influential monograph on the subject, comments that "in the beginning Weierstraß' method was known to only a few people and opened to a larger public by Zermelo's dissertation" (*1935*, 388).

The quality of the dissertation played a major role when Zermelo was considered for university positions after his *Habilitation.* In 1903 the search committee for an extraordinary professorship[37] of mathematics at the University of Breslau (now Wrocław, Poland) remarks that "[Zermelo's] doctoral dissertation, when it appeared, was taken note of considerably more than usually happens to such writings."[38] In 1909 Zermelo was shortlisted for the succession of Eduard von Weber at the University of Würzburg. The report of the Philosophical Faculty states that "in the calculus of variations his work had a really epoch making influence."[39] In 1913, when Zermelo was chosen for the first place in a list of three for a full professorship in mathematics at the Technical University of Breslau, the report to the minister characterizes his thesis as an investigation "basic (grundlegend) to the development of the calculus of variations."[40]

One hundred years later, in a wider historical perspective, Craig Fraser writes (*2013a*, 22–23):

> 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. [...]
>
> 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 [...], it virtually disappeared from the textbook literature in the twentieth century. Instead one developed the theory for n

[35] The two parts of the paper are *Kobb 1892a* and *Kobb 1892b*.

[36] The monograph is a substantial extension of Bolza's *Lectures in the Calculus of Variations* (*Bolza 1904*). The latter is rooted in a series of eight talks about the history and recent developments of the subject which he had given at the 1901 meeting of the American Mathematical Society in Ithaca, N. Y.

[37] Corresponding to the position of an associate professor.

[38] UAW, signature F73, p. 114.

[39] Archives of the University of Würzburg, UWü ZV PA Emil Hilb (No. 88).

[40] UAW, signature TH 156, p. 14.

> dependent variables with variational integrands that contain only the first derivatives of the variables. [...]
>
> 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 [...], Osgood [...] and Bolza as well as the other researchers of the period.

Zermelo's interest in the calculus of variations never weakened. He gave lecture courses (for example, in the summer semester 1910 at the University of Zurich and in the winter semester 1928/29 at the University of Freiburg) and continued publishing papers in the field.

In the first of these papers (*1902d*), he provides an intuitive description of some extensions of the problem of shortest lines on a surface, namely for the case of bounded steepness with or without bounded torsion, illustrating them with railroads and roads, respectively, in the mountains.[41] About 40 years later he will choose this topic as one of the chapters of a planned book *Mathematische Miniaturen* (*Mathematical Miniatures*) under the title "Straßenbau im Gebirge" ("Building Roads In the Mountains").

In the next paper (*1904b*) Zermelo gives two simplified proofs of a result of Paul du Bois-Reymond (*1879a, 1879b*) which says that, given n and an analytical F, any function y for which $y^{(n)}$ exists and is continuous and which yields an extremum of $\int_a^b F(x, y, \ldots, y^{(n)})dx$, possesses derivatives of arbitrarily high order;[42] the theorem shows that the Lagrangian method for solving the related variational problem which uses the existence of $y^{(2n)}$ and its continuity, does not exclude solutions for which $y^{(n)}$ exists and is continuous. Constantin Carathéodory characterizes Zermelo's proofs as "unsurpassable with respect to simplicity, shortness, and classical elegance."[43]

Roughly, Zermelo proceeds as follows: Using partial integration according to Lagrange together with a suitable transformation of the first variation according to du Bois-Reymond, he obtains the first variation of the given problem in the form

$$\int_a^b \Lambda(x)\eta^{(n)}(x)dx,$$

where the variation η is supposed to possess a continuous n-th derivative and to satisfy

$$\eta^{(i)}(a) = \eta^{(i)}(b) = 0 \text{ for } 0 \le i \le n.$$

---

[41] Technically, the paper is concerned with so-called discontinuous solutions; cf. *Carathéodory 1935*, 400.

[42] Zermelo adds a variant of the second proof which he attributes to Erhard Schmidt.

[43] *Carathéodory 1910*, 222; *Carathéodory 1957*, 305.

The main simplification of du Bois-Reymond's proof consists in showing that Λ is a polynomial in x of a degree $\leq n-1$.[44] In particular, Λ is arbitrarily often differentiable. It is easy to show that this property is transferred to the solutions y.

Together with Hans Hahn, one of the initiators of linear functional analysis, Zermelo writes a continuation of Adolf Kneser's contribution (*Kneser 1904*) on the calculus of variations for the *Encyklopädie der Mathematischen Wissenschaften,* giving a clear exposition of the Weierstraßian method (*Hahn and Zermelo 1904*).

Finally, he formulates and solves "Zermelo's navigation problem" concerning optimal routes of airships under changing winds.[45]

1.4 The Boltzmann Controversy

We recall the second thesis that Zermelo defended in the oral examination of his Ph. D. procedure: "It is not justified to charge physics with the task of reducing all phenomena in nature to the mechanics of atoms." Probably he had met the questions framing his thesis in Max Planck's course on the theory of heat in the winter semester 1893/94 and had come to acknowledge Planck's critical attitude against early atomism and the mechanical explanations of natural phenomena accompanying it. During his time as an assistant to Planck (1894–1897) his criticism consolidated, now clearly aiming also against statistical mechanics, in particular against Ludwig Boltzmann's statistical explanation of the second law of thermodynamics. Having a strong mathematical argument and backed by Planck, he published it as aiming directly against Boltzmann (*1896a*), thus provoking a serious controversy which took place in two rounds in the *Annalen der Physik und Chemie* in 1896/97. Despite being interspersed with personal sharpness, it led to a re-consideration of foundational questions in physics at a time when probability was emerging, competing with the paradigm of causality.[46]

[44]For $n = 1$ this result had already been obtained by du Bois-Reymond himself in 1879 (*1879a*) and, with a different proof, by Hilbert in 1899 (cf. *Bolza 1904*, §6). – It is here that Zermelo provides two arguments.

[45]*Zermelo 1930c, 1931a*; cf. 4.2.2.

[46]For a description of the debates on the kinetic theory of gas, also in the context of 19th century cultural development, cf., for example, *Brush 1966* (containing reprints or English translations of the most important contributions) and *Brush 1978*. There is an abundance of analyses of the controversy; besides the items by Brush cf., e. g., *Batterman 1990, Behrend 1969, Cercignani 1998*, 96–119, *Curd 1982, Earman 2006, Frigg 2008, Jungnickel and MacCormmach 1986*, 213–217, *Kaiser 1988*, 125–127, *Lebowitz 1999, Uffink 2007*, 983–992, and, in particular, *Uffink 2013a*.

1.4.1 The Situation

Around 1895 the atomistic point of view, while widely accepted in chemistry, was still under debate in physics, Max Planck being among the sceptics. Looking back to this situation, Planck remembers (*1914*, 87):

> To many a cautious researcher the huge jump from the visible, directly controllable region into the invisible area, from macrocosm into microcosm, seemed to be too daring.

Discussions mainly crystallized in the theory of heat. The so-called energeticists such as Ernst Mach and Wilhelm Ostwald regarded energy as the most fundamental physical entity and the basic physical principles of heat theory as autonomous phenomenological laws that were not in need of further explanation. Among these principles we find the first law of thermodynamics[47] stating the conservation of energy and the second law of thermodynamics[48] concerning the spontaneous transition of physical systems into some state of equilibrium. The latter may be illustrated by a system A_0 consisting of a container filled with some kind of gas and being *closed*, i. e., entirely isolated from its surroundings. As long as temperature or pressure vary inside A_0, there is a balancing out towards homogeneity. If the universe is considered as a closed system, the equilibrium state or *heat death* it is approaching can be characterized by a total absence of processes that come with some amount of differentiation. The parameter measuring the homogeneity, the so-called *entropy*, can be thought to represent the amount of heat energy that is no longer freely available.

From the early atomistic point of view and the mechanical conceptions accompanying it, physical systems consist of microscopic particles called *atoms* which obey the laws of mechanics. For an "atomist," the principles governing heat theory may no longer be viewed as unquestionable phenomena, but are reducible to the mechanical behaviour of the particles that constitute the system under consideration. In particular, the heat content of the system is identified with the kinetic energy of its particles. Determining its behaviour means determining the behaviour of all its constituents. However, as a rule, the number of atoms definitely precludes the possibility of calculating exactly how each of them will behave. To overcome this dilemma, atomists used statistical methods to describe the expected behaviour at least approximately. For justification they argued that the phenomena of thermodynamics as observed in nature result from the global and "statistical" view, the only view we have at our disposal of the behaviour of the invisible particles which lies at their root. Josiah Willard Gibbs, who shaped the mathematical theory of statistical mechanics, describes the new approach in the preface of his classical book (*1902*, vii–viii):

---

[47] Henceforth also called "First Law."
[48] Henceforth also called "Second Law."

> The usual point of view in the study of mechanics is that where the attention is mainly directed to the changes which take place in the course of time in a given system. The principal problem is the determination of the condition of the system with respect to configuration and velocities at any required time, when its condition in these respects has been given for some one time, and the fundamental equations are those which express the changes continually taking place in the system. [...]
>
> The laws of thermodynamics, as empirically determined, express the approximate and probable behavior of systems of a great number of particles, or, more precisely, they express the laws of mechanics for such systems as they appear to beings who have not the fineness of perception to enable them to appreciate quantities of the order of magnitude of those which relate to single particles, and who cannot repeat their experiments often enough to obtain any but the most probable results.

According to the last part of this quotation, the statistical view includes a change of paradigm: It replaces causality by probability. In Boltzmann's presentation the probability $W(s)$ of a system A to be in state s is measured by the relative number (with respect to some suitable measure) of the configurations of A that macroscopically represent s. According to this interpretation the Second Law now says that physical systems tend toward states of maximal probability.

Coming back to the system A_0 defined above, let us assume that the container is divided into two parts P_1 and P_2 of equal size with no wall in between. Furthermore, let s be a state where all atoms are in part P_1, and let s' be a state where the gas is homogeneously distributed over both parts. Then the probability $W(s')$ is maximal and $W(s)$ is low compared to $W(s')$. Therefore, if A_0 is in a state macroscopically represented by s, it will tend toward a state macroscopically represented by s'; in other words, the atoms will tend to distribute homogeneously over the whole container. The entropy $S(s)$ of a state s may now be interpreted as to measure the probability $W(s)$. In fact, it is identified with a suitable multiple of its logarithm:

$$S = k \log W,$$

where k is the so-called Boltzmann constant.

By providing systems with a tendency towards states of increasing entropy or probability, the Second Law imposes a direction on the physical processes concerned. Similarly, the distribution of the velocities of their constituents will converge to a final equilibrium distribution, the so-called Maxwell distribution.

1.4.2 The First Round

It was, in particular, Ludwig Boltzmann (1844–1906) who developed the kinetic theory of gas according to the principles of statistical mechanics, thereby

Ludwig Boltzmann in 1898
Courtesy of Österreichische Nationalbibliothek

providing the statistical interpretation of the Second Law sketched above.[49] In the mid 1890s, attacked already by energeticists and other non-atomists, he found himself confronted with a serious mathematical counterargument written up by Zermelo in December 1895 in the paper entitled "Ueber einen Satz der Dynamik und die mechanische Wärmetheorie" ("On a Theorem of Dynamics and the Mechanical Heat Theory") (*1896a*), the *Wiederkehreinwand* or *recurrence objection.* For an account we return to the system A_0 described above. Its microscopic description will only depend on the spatial coordinates of its atoms together with their velocities. Under reasonable assumptions which Boltzmann had taken for granted, the function describing these data in dependence of time falls under the so-called *recurrence theorem* proved by Henri Poincaré in 1890:[50] System A_0 will infinitely often approach its initial state. Hence, one may argue, also the entropy, depending only on states, will infinitely often approach the initial entropy. It thus cannot steadily

[49]The sketch does not take into consideration the development of the respective notions in Boltzmann's work itself and the way they differ from those in *Gibbs 1902.* For respective details cf. *Klein 1973, Uffink 2007, Uffink 2013a, Uffink 2013b*, or *Kac 1959*, Ch. III.

[50]*Poincaré 1890*, esp. Section 8, "Usage des invariants intégraux," 67–72.

increase. Moreover, the velocities of the atoms of A_0 will not tend to a final distribution. To be fully exact, one has to assume that the initial state of A_0 is different from some exceptional states for which Poincaré's argument does not work. These states, however, are "surrounded" by non-exceptional ones.[51] So there really seems to be a contradiction.

Zermelo starts his paper with a lucid proof of the recurrence theorem[52] and then presents the preceding argument. In order to avoid the inconsistency from the statistical point of view, he continues (*1896a*, 492), one would have to assume that nature always realizes one of the exceptional initial states that do not lead to recurrence. However, as each exceptional state has non-exceptional ones arbitrarily near, such an assumption would contradict

> the spirit of the mechanical view of nature itself, which will always compel us, at least within certain boundaries, to consider all *conceivable* mechanical initial states as physically *possible*.

He, therefore, pleads for a modification of the mechanical model underlying the theory and *expressis verbis* claims that Boltzmann's probabilistic deduction of the Maxwell distribution[53] could no longer be maintained. He confidently concludes (ibid., 494):

> Faced with the difficulties of the subject matter, I have for now refrained from a detailed examination of the previous attempts at such a derivation, in particular the one undertaken by *Boltzmann* and *Lorentz* [...]. Rather, I chose to present here as clearly as possible what appears to me as being rigorously provable and essentially important in order to contribute to a renewed discussion of the question at hand and to its eventual resolution.

Where does Zermelo's self-confidence come from?

There is a first point: As he opens his paper with a clear proof of Poincaré's theorem, one may guess that the mathematical facts strengthened his conviction. According to Pierre Brémaud (*1998*, 133) "he held a strong position" with his "striking and seemingly inescapable argument." But there was a second, likewise important point. As mentioned above, supporters of atomism such as Boltzmann together with the majority of chemists, were opposed by influential chemists and physicists such as Ostwald and Mach. Among the opponents we also find Planck, however with a different reason. Whereas Ostwald and Mach were not willing to accept a principal difference between the irreversible process of the transition of heat from a hot to a cold body and the reversible process of a body swinging from a higher level to a lower one, Planck

[51] In precise terms the exceptional states form a set of measure zero.

[52] Poincaré and Zermelo did not have measure theory at their disposal and performed the measure-theoretic arguments, which are required for the proof, in a naive way. The first rigorous proof of the recurrence theorem was given about 25 years later by Carathéodory (*1919*).

[53] Apparently Zermelo does not refer to one specific deduction, but to various ones Boltzmann performed, for example, in *Boltzmann 1868, 1872, 1877*.

was convinced of the fundamental difference. He clearly distinguished between the character of the First Law stating the conservation of energy and that of the Second Law involving the existence of irreversible processes. So he should have supported Boltzmann in the debate with Ostwald. But he didn't. In his "Personal Reminiscences" he makes this disagreement clear (*1949b*, 12–13):

> For Boltzmann knew very well that my point of view was in fact rather different from his. In particular he got angry that I was not only indifferent towards the atomistic theory which was the basis of all of his research, but even a little bit negative. The reason was that at this time I attributed the same validity without exception 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.

Zermelo being his assistant, Planck of course got acquainted with the recurrence argument and – as we shall see in a moment – supported (or maybe even initiated) its publication.[54] Hence, Zermelo should not only be seen in the role of a capable mathematician who forwarded a strong mathematical argument, but also in the role of Planck's mouthpiece against Boltzmann's atomistic theories.

Boltzmann knew about this situation. Therefore he took Zermelo's paper seriously and felt himself compelled to give an immediate answer (*1896*).[55] In his reply he defends his theory by arguing that the Second Law as well as the Maxwell law about the distribution of velocities would in itself be subject to probability (an argument that, as just described, asked for Planck's criticism). The increase of entropy was not absolutely certain and a decrease or a return to the neighbourhood of the initial state was possible, but so improbable that one was fully justified in excluding it. Zermelo's paper would demonstrate that this point had not been understood. Boltzmann's résumé is definite (ibid., 773): "While the theorem by Poincaré that Zermelo discusses in the beginning of his paper, is of course correct, its application to the theory of heat is not." The contradiction Zermelo thought he had found could only apply if he had been able to prove that a return to the initial state (or its neighbourhood) had to happen within an observable time. He compares Zermelo with a dice player who calculates that the probability of throwing one thousand consecutive ones

[54]For the debate between Boltzmann and Planck himself cf. *Hollinger and Zenzen 1985*, 19–28.

[55]Three years earlier Poincaré had already formulated the recurrence objection as one of the major difficulties when bringing together experimental results and mechanism (*1893a*, 537). However, there had not been a reaction by physicists. Also Zermelo is not aware of Poincaré's argument; in his paper he says (*1896a*, 485) that "apparently Poincaré did not notice the applicability [of the recurrence theorem] to systems of molecules or atoms and, hence, to the mechanical theory of heat." More than a decade later the encyclopedic contributions *Boltzmann and Nabl 1907* and *Ehrenfest and Ehrenfest 1911* still attribute the recurrence objection only to Zermelo.

is different from zero and, faced with the fact that he never met such an event, concludes that his dice is loaded. Boltzmann summarizes (ibid., 779-780):

> If one considers heat as a motion of molecules that takes place in accordance with the general equations of mechanics, and assumes that the complex of bodies that we perceive is currently in a highly improbable state, then a theorem follows that agrees with the Second Law for all phenomena so far observed.

Full of self-confidence like Zermelo, but more stridently, he ends (ibid., 781–782):

> All objections raised against the mechanical approach to nature are therefore unfounded and based on errors. Those however, who cannot overcome the difficulties attendant on a clear comprehension of the gas-theoretic theorems really ought to follow Mr. *Zermelo*'s advice and decide to abandon the theory altogether.

In his "Personal Reminiscences" Planck describes the impression he got from Boltzmann's reply (*1949b*, 13):

> In any case he answered the young Zermelo with scathing sharpness a part of which also hit me because actually Zermelo's paper had appeared with my permission.

As a result of this sharpness the exchange of arguments which followed saw a more polemical tone also on Zermelo's side.

1.4.3 The Second Round

As explained above, Planck had rejected Boltzmann's theory because of the resulting probabilistic character of the Second Law. So he did not agree with his arguments either. As Zermelo shared Planck's opinion, it is not surprising to see him publishing a reply to Boltzmann's reply (*1896b*). To begin with, he considers Boltzmann's comments rather as a confirmation of his view than as a refutation. For Boltzmann had acknowledged that in a strict mathematical sense the behaviour of a closed system of gas molecules was essentially periodic and, hence, not irreversible. He thus had accepted just the point which he, Zermelo, had had in mind when writing his paper. Boltzmann's claim to be in correspondence with the Second Law if the system under consideration was in an improbable state suffered from not providing a *mechanical* explanation for the fact that the systems one observes are always in a state of growing entropy (ibid., 795):

> 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.

Among further objections, he puts forward as "*a priori* clear" (ibid., 799) that the notion of probability did not refer to time and, hence, could not be used to impose a direction on physical processes as is the case with the Second Law. He thereby touches Josef Loschmidt's *Umkehreinwand* or *reversibility objection* (*Loschmidt 1876*, 139): As the mechanical laws are invariant under the inversion of time, a deduction based on these laws and yielding an increase of entropy in time t should also yield an increase in time $-t$, i. e., a *decrease* in time t.[56]

Faced with the Second Law as a well-established empirical fact and the insufficiencies of the mechanical theory of gas he had exhibited, Zermelo condenses the essence of his criticism in the following methodological conclusion (*1896b*, 793–794):

> 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.

Already on 16 December 1896 Boltzmann completes a reply (*1897a*) where he systematically answers the questions Zermelo had raised, among them the request to explain *mechanically* why the systems observed should be in an improbable state. Boltzmann's explanation is based on what he calls "Assumption A" (ibid., 392): Viewed as a mechanical system, the universe (or at least an extended part of it surrounding us) has started in a highly improbable state and still is in such a state. Therefore, any system which is separated from the rest of the world, as a rule will be in a highly improbable state as well, and as long as it is closed, will tend to more probable states. Hence, if the mechanical theory of heat is applied not to *arbitrary* systems, but to systems which are part of our actual world, it will be in accordance with our experience. Concerning a justification of "the of course unprovable" Assumption A, he basically repeats an argument from his first answer (ibid., 392–393):

> Assumption A is the physical explanation, comprehensible in accordance with the laws of mechanics, of the peculiarity of the initial states [...]. 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

[56] For time reversal cf., for example, *Greene 2004*, 146–149 and 506–507.

> in each special case; for nobody will demand that the ultimate principle of explanation itself be explained as well.

At the end he varies his model, thereby taking care of the reversibility objection as well: The universe may already be in a state of equilibrium, of heat death; however, there might be probabilistic fluctuations, i. e, relatively small areas, possibly as large as our "space of stars" – *single worlds* (*Einzelwelten*) in his terminology – which are in an improbable state, among them worlds which presently tend to more probable states and worlds which tend to less probable ones. In this sense the universe might realize both directions of time and not prefer one of them; in other words (*Boltzmann and Nabl 1907*, 522):

> The statistical method shows that the irreversibility of processes in a relatively small part of the world or, under special initial conditions also in the whole world, certainly is compatible with the symmetry of the mechanical equations with respect to the positive and the negative direction of time.

Faced with the possibility of successful mechanical explanations of a series of phenomena, certain discrepancies between the phenomenological view and the mechanical method (for example, with respect to "uncontrollable questions" such as that for the development of the world for a very long time) should not be taken to justify Zermelo's demand for a methodological change. For just such discrepancies would support the view that (*Boltzmann 1897a*, 397)

> the full versatility of our mental pictures will be increased by also examining the implications of the mechanical[57] picture alongside those of the Second Law in the *Carnot-Clausius* version.

There was no further reply by Zermelo.

1.4.4 After the Debate

For ten more years, Zermelo maintained interest in statistical mechanics. In his *Habilitation* address (*1900*), given 1899 in Göttingen, he offered a solution of a problem which he had mentioned at the end of his rejoinder to Boltzmann (*1896b*, 800):

> 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, 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.

[57] The translation in *Zermelo 2013* erroneously uses the term "mental" instead of "mechanical".

In order to obtain corresponding results which are "mathematically reliable, at least by the standards of probability theory" (*1900*, 318), Zermelo pleads for a notion of probability that takes this problem into account,[58] and ends by treating mean values and promising a deduction of the Maxwell distribution of velocities, which would be "based solely on the definition of probability given here without any further hypothesis" (ibid., 320). The promise remained unfulfilled, perhaps for the reason that soon statistical mechanics would find a representation along similar lines in Gibbs' monograph from which we have quoted above. As described below, Zermelo appreciated it to such an extent that he provided a German translation.

Some of the members of the Göttingen *Habilitation* committee comment also on Zermelo's anti-Boltzmann papers.[59] Hilbert states that Zermelo's recurrence objection is "very remarkable and found the attention it deserves."[60] Felix Klein is more critical: In connection with "a certain tendency to one-sided criticism" which he sees in Zermelo, he emphasizes

> that the two notes directed against Boltzmann are very astute and, in particular, that one has to acknowledge the courage by which the young author opposes a man like Boltzmann, but that, finally, it was Boltzmann who was right; it must be said, however, that Boltzmann's deep and original way of looking at these things was difficult to understand. (OV 1.04)

Woldemar Voigt, professor of mathematical physics, who follows Klein's assessment, adds that

> already the intensive preoccupation with the highly interesting papers of Boltzmann, which are not really taken into account by the majority of physicists, represents some kind of achievement, and that Dr. Zermelo's criticism undoubtedly attributed to a clarification of the difficult subject. (OV 1.05)

[58] Let S be a system of n particles in, say, 3-dimensional space whose possible states are given by the coordinates $x = (x_i)_{i \leq 3n}$ and by the components $m = (m_i)_{i \leq 3n}$ of the momenta of its particles and whose behaviour can be described by Hamiltonian equations. The totality of states $s = (x, m)$ fills a body Q of finite contents $|Q|$. Let s be a state of S and q a cuboid in Q of contents $|q|$. Zermelo considers the probability $p(s, q) := |q|/|Q|$ for s belonging to q. This notion of probability is independent of physical developments in the following sense: If $t_0 < t_1$ and s is considered as an initial state at time t_0 which develops into state $s(t_1)$ at time t_1 and if, moreover, q is considered as a set of initial states at time t_0 which develops into $q(t_1)$ at time t_1, then, by Liouville's theorem, $|q| = |q(t_1)|$ and, hence, $p(s, q) = |q|/|Q| = |q(t_1)|/|Q| = p(s(t_1), q(t_1))$. – Similar relations between probability and phase space volume were well-known at that time, originating, e. g., in *Boltzmann 1868*. However, so Massimiliano Badino (written communication), Zermelo's paper "is terser and much clearer than Boltzmann's" and "expresses an important critique of the use of probability in physics."

[59] UAG, Philosophische Fakultät 184a, 149seq.

[60] OV 2.03, last but one paragraph.

Six years later Zermelo published his German translation (*Gibb 1905*) of Gibbs' *Elementary Principles* (*Gibbs 1902*), "a major contribution in making Gibbs' work accessible to the German-speaking community" (*Uffink 2013b*, 570). In the preface he appreciates Gibbs' book as being (*Gibbs 1905*, iii)

> the first attempt to develop strictly and on a secure mathematical basis the statistical and the probability theoretical considerations in mechanics, as they are indispensable in various areas of physics, in particular in the kinetic theory of gas, independent of their applications.

However, despite the success of the statistical method, his reservations concerning the range of statistical mechanics are not settled (ibid., iii):

> I intend to work out soon and elsewhere the objections I have, in particular, against the considerations in the twelfth chapter as far as they attribute to the mechanical systems a tendency towards a state of statistical balance and, based on this, a full analogy to thermodynamical systems in the sense of the Second Law.

The elaboration appeared in his review (*1906a*) of Gibbs' book, where he repeats his objections from 1896, in particular pointing to the irreversibility problem (ibid., 241):

> Neither *Boltzmann's* nor *Gibbs'* 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 one decides to base it not on Hamilton's equations of motion, but on differential equations already containing the principle of irreversibility in themselves.

Contrary to Zermelo's scepticism, Boltzmann's views gained general acknowledgement. Already at the end of 1900 even Planck *expressis verbis* used Boltzmann's probabilistic approach (without instantly accepting the atomistic hypothesis). He deduced an improved radiation formula by working with a probabilistic interpretation of the distribution of oscillator frequencies, i.e. "by introducing probability considerations into the electromagnetic theory of radiation, whose importance for the second law of thermodynamics was first discovered by Mr. L. *Boltzmann*" (*1900*, 238). Finally Planck fully accepted Boltzmann's atomism (*1914*; *1949a*, 87):

> It was L. Boltzmann who reduced the contents of the Second Law and, hence, the totality of irreversible processes, whose properties had caused unsurmountable difficulties for a general dynamic explanation, to their true root by introducing the atomistic point of view.

As a matter of course, it was Boltzmann who was asked by Felix Klein to contribute an article on the kinetic theory in the *Encyklopädie der Mathematischen Wissenschaften.* Boltzmann remembers (*1905b*, 406):

> When Klein asked me for an encyclopedic article, I refused for a long time. Finally he wrote: "If you do not do it, I will give it to Zermelo." Zermelo represents a view that is just diametrically opposed to mine. That view should not become the dominant one in the encyclopedia. Hence, I answered immediately: "Rather than this is done by that Pestalutz, I'll do it." [61]

The finished article (*Boltzmann and Nabl 1907*) discusses also the Boltzmann-Zermelo controversy, repeating essential parts of Boltzmann's arguments (ibid., 519–522). Boltzmann himself did not see it appear; mentally weakened he committed suicide in 1906.[62]

Already in his first answer to Zermelo, Boltzmann had illustrated by an example that the recurrence time in Poincaré's theorem should be far greater than the age of the universe even for systems comprising only a small number of particles. Paul and Tatiana Ehrenfest were able to support his arguments and substantially contribute to a more precise elaboration (*Ehrenfest and Ehrenfest 1907*). They conceived a simple model consisting of two adjacent containers connected by an opening and filled with gas atoms, per time unit allowing one atom, chosen by chance, to change its container. By easy calculations they obtained the somewhat paradoxical facts that this model meets Boltzmann's assertions, at the same time being invariant under a reversion of time. They thus showed (ibid., 314) that neither the recurrence objection nor the reversibility objection are sufficient or suitable to refute Boltzmann's theory, in other words, that "it is always possible to reconcile both time reversibility and recurrence with 'observable' irreversible behaviour" (*Kac 1959*, 73).[63] So in 1908 Planck could write (*1908*, 41):

> Nature simply prefers more probable states to less probable ones by performing only transitions in the direction of larger probability. [...] By this point of view all at once the second law of the theory of heat is moved out of its isolated position, what is mysterious with the preference of nature vanishes, and the principle of entropy as a well-founded theorem of probability theory gets connected with the introduction of atomism into the physical conception of the world.

Nevertheless, problems remained, among them, for example, Zermelo's question for a physical explanation why, say, our universe might have been in a state of low entropy when it came to exist, or the question about how Boltzmann's *Einzelwelten* would fit the cosmological picture. They initiated

[61] Boltzmann refers to a minor figure in Friedrich Schiller's drama *Wallenstein's Tod* (*The Death of Wallenstein*). Pestalutz does not appear on the stage. When three conspirators are haggling over who is to assassinate Wallenstein, their ringleader persuades the other ones to perform the assassination by threatening to turn to Pestalutz in case they would not.

[62] Cf., e. g., *Lindley 2001* for Boltzmann's life.

[63] For the recurrence problem see also *Ehrenfest and Ehrenfest 1911*, 22–33, and *Brush 1966*, 15–18; for an extended treatment of the irreversibility problem cf. *Hollinger and Zenzen 1985* and *Zeh 1989*.

a development leading to difficult conceptual challenges and ever more sophisticated technical resources (*Sklar 1993*, xiii), finally reaching the limits of scientific research at the roots of the cosmological phenomena, the big bang.[64]

Looking back to the Boltzmann controversy, Jos Uffink (*2013a*, 214–215) concludes:

> 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 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.

And Zermelo? After 1906, having contributed to the growing effort to clarify the basis of the theory of heat, he left the lively scene. His scientific energy focused on the foundations of mathematics. Already in 1904 he had formulated the axiom of choice and used it to prove the well-ordering theorem, thereby opening a new era in set theory and at the same time one of its most serious debates.

[64]For a detailed discussion cf., e. g., *Greene 2004*, in particular Chs. 6 and 11.

http://www.springer.com/978-3-662-47996-4

Ernst Zermelo
An Approach to His Life and Work
Ebbinghaus, H.-D.; Peckhaus, V.
2015, XVI, 384 p. 43 illus., Hardcover
ISBN: 978-3-662-47996-4