Chapter 2

The Foundation of Physics:

The Lectures (1916–1917)

Introduction

Hilbert gave a two-semester course on the "Foundations of Physics" in the summer semester 1916 and the following winter semester 1916/17. In this Chapter we present the *Ausarbeitungen* of this two-part course.

While Hilbert was preparing his "First Communication" on the "Foundations of Physics" for publication (*Hilbert 1915*, this Volume pp. 28–46), classes had resumed in Göttingen for the winter term of 1915/16. But Hilbert did not lecture on relativity during this semester. Instead, he lectured on differential equations (*Hilbert 1915**). Thus, in contrast to his lectures and publications on radiation theory and his second note, the publication of his first note is not an example of Hilbert's teaching on a topic prior to or simultaneous with his current research concerns. But in the following summer term of 1916, Hilbert started to lecture on relativity theory. The lecture was announced as "Einleitung in die Prinzipien der Physik" and took place on Thursdays, 9–11 a.m.[1] As is clear from the handwritten equations, Hilbert's summer lectures were worked out by the Swiss mathematician and physicist Richard Bär, who had come to work as Hilbert's assistant some time around Easter 1916.[2]

While the title of the lecture course in summer 1916 bears the same title as Hilbert's "First Communication" of the previous winter, it does not present the same material. Hilbert's intention was to bring his students up to speed with his own research work in this field. In order to do so, Hilbert goes back to the foundations of special relativity and first presents an axiomatic discussion of special relativistic kinematics. This work may be regarded as the starting point of a subsequent tradition of attempts to axiomatize relativity theory, an endeavour that is currently associated with Hans Reichenbach's writings. It is unclear, though, whether and to what extent Reichenbach was aware of Hilbert's investigations when he was working out his own axiomatic analysis of relativity theory.

In the second part of the summer course, Hilbert also discusses special relativistic dynamics and, in particular, dynamic concepts of the electron. In this context, Hilbert first presents the rigid electron model developed by the former Göttingen physicist Max Abraham and later refined by Hilbert's own students Erich Hecke and Wilhelm Behrens. This model of electron dynamics is then contrasted with Gustav Mie's field theoretic concept of the electron as a static and spherically symmetric solution of non-linear generalizations of Maxwell's equations. Particularly interesting is Hilbert's lucid discussion of Mie's explicit solution of one such non-linear set of Maxwell equations in which the electromagnetic potential enters the field equations to fifth power. This solution had already been discussed by Mie in his own original paper

[1] Hilbert also lectured on "Partielle Differentialgleichungen" (Mondays, 9-11 a.m., *Hilbert 1916b**); see Hilbert's Lecture Courses, 1886–1934, this Volume, p. 719.

[2] See the discussion in the Introduction to this Volume, sec. 4.

on the subject. It is taken up here by Hilbert and discussed in its simple mathematical essence. This explicit solution became a paradigmatic example for later field theoretic programs that attempted to overcome the duality of matter and fields by conceiving material particles as non-singular, static, and spherically symmetric solutions to generalized field equations of the gravitational and electromagnetic fields.

It is only in the last lecture of the 1916 summer course that Hilbert addresses the question of generalizing special relativity and Mie's non-linear electrodynamics into a generally covariant theory. Mie had proposed a theory of gravitation within the Lorentz covariant framework of his non-linear electrodynamics. But it had been one of Hilbert's central insights of the fall of 1915 that Mie's electromagnetic theory can be combined with Einstein's theory of gravitation. His main idea had been to generalize Mie's electrodynamics to a generally covariant theory where, following Einstein, a metric tensor field would represent the inertio-gravitational field. Mie's equations would be recovered from the generally covariant theory in the special relativistic limit. A didactic exposition of Einstein's theory of general relativity is hence announced for the winter term 1916/17 at the end of the first part of Hilbert's summer course.

During the years of World War I, Felix Klein studied this *Ausarbeitung* as he prepared the second volume of his lectures on the development of mathematics in the nineteenth century (*Klein 1927*). Excerpt notes by Klein are extant in his papers,[3] and some of his comments have been pointed out in the annotation.

The second part of Hilbert's course on the "Foundations of Physics" was given in the winter term 1916/17. Again, the lecture was not explicitly announced under this title. In fact, it was not announced at all.[4] The *Ausarbeitung* of this course, which was given for four hours per week, was again prepared by Richard Bär as is indicated explicitly on its title page. It is here presented in its entirety.

This second part of the course covers the subject matter of Hilbert's first note. But again, it starts with a discussion that will be the foundation for what is to follow. Hilbert first gives an introduction to differential geometry, proceeding step-by-step from standard two-dimensional surface theory to the case of three and four dimensions, and from a positive-definite to an indefinite line element. Hilbert spends a great deal of time carefully explaining the basic concepts of differential geometry for the simple case of two-dimensional surface theory. He introduces what he calls Riemannian coordinates, i.e., geodesic normal coordinates, which he then uses to motivate and derive the Riemann

[3]SUB Göttingen, *Cod. Ms. F. Klein* 22A, f. 31–32.
[4]See Hilbert's Lecture Courses, 1886–1934, this Volume, p. 719.

curvature tensor. He explicitly illustrates the concepts and equations by discussing the example of the sphere, and then moves on to introduce Gaussian coordinates showing that the Riemann curvature in two dimensions is proportional to the Gaussian curvature which is given geometrically in terms of principal radii of curvature. Hilbert also discusses the case of non-Euclidean geometries of constant curvature and what he calls the pseudosphere. Before moving on to the three-dimensional case, Hilbert devotes a few paragraphs to the case of an indefinite metric. Here we find him using the term "pseudo-Euclidean geometry," perhaps a local Göttingen expression. In his treatment of the three-dimensional case, he proceeds along the same lines by emphasizing the structural generalizations that are necessitated by the additional dimension. The three-dimensional case also gives him occasion to discuss the theory of characteristics of the Hamilton-Jacobi and Monge differential equations. Of particular interest is his introduction and use of the invariant conditions that express the character of the geometry, and which develop into his so-called reality conditions in the four-dimensional case. After such long preparation, Hilbert's treatment of the physically relevant four-dimensional case can then be rather brief.

While the first part of the course presents a purely mathematical discussion, it is in § 37, entitled *Zusammenhang der Theorie mit der Wirklichkeit*, that Hilbert then begins to address the specific issues with which he is most concerned. Among the first interpretational and foundational problems to be treated are the questions of the principal measurability of the components of the metric tensor and the necessary reinterpretation of the notion of causality in a generally covariant physical theory. By this time, Hilbert's lecture has caught up with his current research interest, and parts of the material covered in his lectures are quite similar to the discussion in his "Second Communication" to the "Foundations of Physics" (*Hilbert 1917*, this Volume, pp. 47–72). Nevertheless, his lectures are more explicit and sometimes more lucid than his published account. It is in this second part of the lecture course that Hilbert brings to bear his knowledge of a rich mathematical tradition to the analysis of foundational problems raised by general relativity. Several themes are discussed in the course of his lectures. He first gives a derivation of the gravitational and electromagnetic field equations, then discusses the question of the empirical status of Minkowskian space-time. This question is debated on the level of the open question of the existence and uniqueness of the Minkowski line element as a solution of the vacuum field equation. In a second step, he looks at the case of spatial spherical symmetry, rederiving Schwarzschild's solution. He also discusses the subtle issue of singularities of the metric.

Further discussion of the implications of the field equations involve the solution of the equations of motion for a particle in a gravitational field. The conceptual framework for this problem is set up by means of provisional axioms for the motion of material particles and for the geometry of light rays in curved space-times. The axiomatic postulates here serve as a substitute for a complete derivation of the equations of motion from the field equations themselves. Given the mathematical problem of solving a differential equation for geodesic motion in a curved space-time, Hilbert skilfully gives a state-of-the-art discussion of the predictions for the three classical tests of general relativity.

In the final chapters of his second course, Hilbert deals with the field theoretic problem of the integration of electrodynamics into the gravito-inertial framework. He outlines a research program in which a full theoretical account of the real world is to be achieved by a series of successive approximations. An expansion of the metric field in terms of the coupling constant to the electromagnetic field around the vacuum solution defines a series of ever better approximations of the real world. As a first approximation, Hilbert takes Einstein's gravitational wave solution of the linearized field equations. The second approximation then would include the electron as a spherically symmetric solution of the field equations. This program is only hinted at, and we have little evidence about any progress that Hilbert would have made along these lines.[5] The final part of the lecture course discusses the problem of the concept of energy in a general relativistic field theory and the problem of formulating a generalized theorem about energy-momentum conservation. These problems had played a role already in the proofs and published version of *Hilbert 1915* (see this Volume, pp. 317–329 and pp. 28–46) and were further discussed in Göttingen until its principal solution by Felix Klein and Emmy Noether in 1918.[6]

Several other copies of the *Ausarbeitung* of the second course of this lecture are known to the Editors. The papers of Erich Hückel in the Berlin Staatsbibliothek contain an unbound copy[7] that is identical to the Göttingen copy and shows the same page breaks throughout the whole text. The copy has the same equations and the same occasional corrections. These were added in R. Bär's hand for pages 1 to 155. Following page 156, however, all handwritten additions are in a different hand. This copy contains page 107, which is missing in the Göttingen copy. The papers of Max Born in the Berlin Staatsbibliothek contain a complete bound copy[8] that is identical in text but has different page breaks. Its title page says that the *Ausarbeitung* was prepared by Paul Scherrer. A comparison of this copy with the Göttingen copy and the copy in the Hückel papers shows that the handwritten corrections

[5] See his notes on the "Foundations of Physics," this Volume, pp. 331–333.

[6] For further discussion of Hilbert's lectures on the *Foundations of Physics*, see *Corry 2004*, ch. 8.3, *Renn and Stachel 2007*, *Brading and Ryckman 2008*.

[7] Staatsbibliothek Preußischer Kulturbesitz, Berlin. Handschriftenabteilung. Nachl. Hückel 2.11.

[8] Staatsbibliothek Preußischer Kulturbesitz, Berlin. Handschriftenabteilung. Nachl. Born 1818.

of the latter versions are incorporated into the typewritten text of the former. It appears that the copy in the Born papers is merely a fresh typescript of the earlier Göttingen version, which according to its title page was prepared by Richard Bär, and otherwise is an identical copy. Finally, the Archives of the California Institute of Technology own a bound copy that was part of the papers of Paul Epstein.[9] Its title page also credits Paul Scherrer with the *Ausarbeitung*, it shows the same page breaks as the copy in the Born papers, and appears to be an identical copy of the version in the Born papers.

Tilman Sauer

[9] California Institute of Technology, Pasadena, Archives. Call No. QA401.H5.

Die Grundlagen der Physik

Inhalt

Einleitung

Die Vorlesung, die ich in diesem Semester mit „Grundlagen der Physik" angezeigt habe,[1] soll sich wesentlich mit den modernen relativistischen Ideen beschäftigen, die den Namen „Grundlagen" wohl mit ebenso grossem, wenn nicht mit grösserem Recht für sich in Anspruch nehmen dürfen, wie andere moderne Theorien und Ideenbildungen. Der Anfang soll mit dem sogenannten „kleinen Relativitätsprinzip"[2] gemacht werden, einmal aus pädagogischen Gründen, dann aber auch des grossen und selbständigen Interesses wegen, das es bietet. In einem späteren Teil werden wir dann zu der allgemeinsten Fassung des Relativitätsprinzips übergehen.[3] Unsere mathematischen Hilfsmittel werden hier die Variationsrechnung und die Invariantentheorie sein. Letztere wird von Fräulein Dr. Noether in einem Zyklus von Vorträgen im Seminar behandelt werden.[4]

§ 1. Das Problem und seine Geschichte

Es handelt sich bei den relativistischen Theorien ganz wesentlich um die fundamentalsten Begriffe der Physik, um Raum und Zeit. Die Philosophen untersuchten schon frühzeitig das gegenseitige Verhältnis von Raum und Zeit und kamen, bald instinktiv, bald aus philosophischen Gründen, zur Erkenntnis der Zusammengehörigkeit beider. Hier ist, um von älteren ganz zu schweigen, Kant zu nennen, in dessen Erkenntnistheorie (vgl. Kritik der reinen Vernunft, wo in der „transzendentalen Aesthetik"[5] die Frage: „Wie ist Mathematik als reine Wissen|schaft möglich?"[6] untersucht wird) Raum- und Zeitbegriff eine 2

[1]The lecture course was announced as "Einleitung in die Prinzipien der Physik" (*Hilbert 2004*, p. 617) and took place on Thursdays, 9–11 am (*Verzeichnis 1916*, p. 15). In the page margin, a typewritten date ("4.V.16") was added next to the first line. The summer semester 1916 started on 16 April (*Verzeichnis 1916*, p. 1).

[2]The terminology "kleines Relativitätsprinzip" is non-standard. In excerpt notes about Hilbert's course notes, dated 26 August 1916, Felix Klein remarked: "Neue Termini: "Kleines Relativitätsprinzip" "Zukunfts- bez. Vergangenheitskegel". Das Historische ist immer unzureichend! Vielfach die Termini. Wie viel kommt auf unvollkommene Ausarbeitung?" (SUB *Cod. Ms. Klein 22A*, sheet 29).

[3]General Relativity is introduced only in the last paragraph (§ 31.) of this lecture; see pp. 154ff below. It is treated *in extenso* in the second part of the lecture course on *Grundlagen der Physik* given in winter 1916/17, *Hilbert 1916/17**, this Volume, pp. 162–307.

[4]Emmy Noether (1882–1935) had come to Göttingen in summer 1915 in order to work with Klein and Hilbert. Supported by Hilbert, she had attempted twice to obtain a *venia legendi* in spite of an explicit decree against the habilitation of women at Prussian universities dating from 1908. Both attempts had failed due to the conservative majority of the Göttingen philosophical faculty (*Tollmien 1991*), and Noether had to announce her lectures under Hilbert's name. In early 1917, after receiving a call to Berlin, Hilbert negotiated with the Prussian ministry that funds be allocated to support Noether's teaching in Göttingen after the war (*Sauer 2000*, p. 191). Noether finally obtained her habilitation in 1919.

[5]*Kant 1956*, pp. 63–93.

fundamentale Rolle spielen.

Neben dieser philosophischen Entwicklung läuft eine zweite Reihe von Untersuchungen her, die sich zwar ursprünglich nur auf den Raum selbst beziehen, aber zu einer Erweiterung auf Raum und Zeit drängten. Wir meinen hier die Untersuchungen über die Grundlagen der Geometrie. Diese bezwecken eine axiomatisch-logische Analyse der Grundbegriffe unserer Raumanschauung und suchen ein System von Axiomen, durch die es möglich ist, die ganze Wissenschaft vom Raume lückenlos und widerspruchsfrei aufzubauen. Nimmt man zum Raum noch die Zeit hinzu, so kommt man von hier aus zur Kinematik.

Der wichtigste und folgenreichste Anstoss zur Weiterentwicklung ist aber von keiner der skizzierten Richtungen gekommen, sondern vom physikalischen Experiment. Hieran anschliessend vollzog sich eine der grössten Umwälzungen, die unsere Wissenschaft kennt und noch immer ist das Ende dieser Revolution nicht gekommen. Soviel können wir aber sagen: Was die Philosophen instinktiv gefühlt, was sie immer gefordert, hier wird es zur Tat. Raum und Zeit stehen nicht mehr getrennt nebeneinander, sondern gehen eine so enge Verbindung ein, dass Zeit ohne Raum, Raum ohne Zeit nicht mehr denkbar ist.[7]

Wir schliessen uns im folgenden dem historischen Weg an, denn es ist nötig, die alte Auffassung von Raum und Zeit zu kennen, um die neue zu verstehen. Der Weg, den wir einschlagen, wird der axiomatische sein.

3 § 2. Die Axiome der Geometrie

Entsprechend dem zuletzt gesagten, gehen wir aus von den Axiomen der Geometrie, wie sie aus den Untersuchungen über die Grundlagen der Geometrie bekannt sind, und lassen es dahingestellt, ob sich nicht schon hier Schwierigkeiten ergeben, d.h. ob es erlaubt ist, hier die Zeit so vom Raum abzutrennen, insbesondere, ob hier nicht vielleicht schon von dem Begriff der Gleichzeitigkeit implizite Gebrauch gemacht ist. Jedenfalls müssen wir, um mit der Erfahrung in Uebereinstimmung zu bleiben, axiomatisch fordern, dass die „Welt“ (die Gesamtheit von Raum und Zeit) eine vierdimensionale Mannigfaltigkeit ist, also durch vier Parameter völlig beschrieben werden kann, von denen drei, die „Raumkoordinaten“, zusammengehören und der vierten, der Zeitkoordinate gegenüberstehen, aber nicht so, dass sie streng von einander geschieden wären.

Wir wollen aus pädagogischen Gründen im folgenden für den Raum nie mehr Dimensionen annehmen, als wir unbedingt brauchen, also nur dann drei Dimensionen, wenn dies unbedingt notwendig ist, sonst zwei oder womöglich nur

[6] In the introduction to the second edition of the *Kritik der reinen Vernunft*, Kant asks the questions: “Wie ist reine Mathematik möglich? Wie ist reine Naturwissenschaft möglich?” *Kant 1956*, p.52*.

[7] An allusion to Minkowski, cf. note 29 below.

eine. Es würde die Anschauung eben unleugbar erschwert, wenn wir immer mit drei Dimensionen operieren wollten.

Wir legen in den Raum ein festes Koordinatensystem. Schon hier tritt uns die erste Schwierigkeit entgegen: Was heisst „fest" oder, wie wir auch sagen werden, „starr"? Die Starrheit ist natürlich ursprünglich das Resultat eines aus der Erfahrung entsprungenen physikalischen Grenzprozesses, muss aber hier axiomatisch gefasst werden. Wir kommen hierauf | noch zurück.[8] 4

Wir haben also ein Cartesisches Koordinatensystem in einer Euklidischen Ebene. Jeder Punkt ist Träger eines Zahlenpaares (x, y). Wir können uns dies so versinnbildlichen, dass wir uns in jedem Punkt einen Kilometerstein denken, auf dem die beiden Zahlen x und y stehen. Wir denken uns weiter ein physikalisches Instrument, den starren[9] „Massstab", der als „Entfernung" zweier Punkte mit den Koordinaten (x', y') u. (x'', y'') überall $\sqrt{(x'-x'')^2 + (y'-y'')^2}$ angibt. Von diesem Massstab werden wir sagen, er sei „starr". In unserer Ebene soll, wie wir schon oben sagten, die euklidische Geometrie, (also insbesondere der Satz von der Winkelsumme im Dreieck und die ganze analytische Geometrie) gelten. Die reinste Form dieser Voraussetzung ist in folgenden Forderungen enthalten:

1. Die Gleichung jeder geraden Linie ist linear.

2. Alle möglichen Benennungen der Ebene durch Cartesische Koordinaten werden erhalten durch Transformationen von der Form

$$\begin{aligned} \xi &= \alpha_{11}x + \alpha_{12}y \\ \eta &= \alpha_{21}x + \alpha_{22}y \end{aligned} \tag{I}$$

mit der Nebenbedingung:

$$\xi^2 + \eta^2 = x^2 + y^2$$

und von der Form

$$\begin{aligned} \xi &= x + \alpha_1 \\ \eta &= y + \alpha_2 \end{aligned} \tag{II}$$

und durch Zusammensetzungen von Transformationen beider Arten. Die Nebenbedingung bei I sagt aus, dass die „Entfernung" zweier | Punkte etwas vom 5
Koordinatensystem unabhängiges ist, bei den Transformationen II ist eine solche Nebenbedingung überflüssig, da sie die Entfernung von selbst invariant lassen. Da uns die Transformationen I und II stets erlauben, ein Koordinatensystem zu finden, dessen Anfangspunkt ein vorgegebener Punkt ist und dessen eine Achse durch einen zweiten gegebenen Punkt hindurch geht, so dürfen wir sagen, dass die Transformationen I und II das Anlegen des Massstabes repräsentieren.

Wir haben angenommen, dass die Geometrie die Euklidische sei, man könnte natürlich auch eine der nichteuklidischen Geometrien zu Grunde legen und

[8] See the discussion in § 6, pp. 88–90, and in § 25, pp. 136–140, below.

[9] „starren" was interlineated.

darauf weiterbauen. Die Frage, ob unsere Annahme zulässig ist, muss durch das Experiment entschieden werden, da wir nur dann hoffen dürfen, mit der Wirklichkeit in Uebereinstimmung zu bleiben, wenn unsere Annahmen der Erfahrung nicht widersprechen. Für unsere Frage ist das Experiment schon von Gauss angestellt worden; unter der Annahme, dass die Lichtstrahlen Gerade sind, ergaben sich für die Winkelsumme im Dreieck so geringe Abweichungen von 180°, dass sie völlig innerhalb der Grenzen der Beobachtungsfehler fielen, also keine Entscheidung zu Ungunsten der Euklidischen Geometrie lieferten.[10] Für den dreidimensionalen Raum verläuft alles genau so, wie wir es eben für den zweidimensionalen darlegten, wir gehen also darauf nicht weiter ein.

§ 3. Die Axiome und Definition der Zeitmessung

Zu dem Raum müssen wir nun die Zeit hinzunehmen und da ist es nötig,
6 von der Zeitmessung und den ihr | zugrunde liegenden Axiomen zu sprechen, um dann zu den Axiomen überzugehen, die Zeit und Raum in Verbindung bringen, also für den Aufbau der Mechanik und damit der ganzen Physik entscheidend sind.

In der Geometrie müssen wir von gewissen Begriffen Gebrauch machen, die wir der Anschauung entnehmen und die durch ihre für die Geometrie allein wesentlichen Eigenschaften noch nicht völlig bestimmt sind, wie die „Gerade Linie" u. s. w. Es gibt ja z. B. Kreisgebüsche, in denen die Euklidische Geometrie herrscht, wenn man statt „Kreis des Gebüsches" überall „Gerade" sagt, diese Kreise haben also alle Eigenschaften der euklidischen Geraden, sind aber für unsere Anschauung jedenfalls keine Geraden. Genau so ist es mit der Zeit und Bewegung. Was „Bewegung", was „Zeit" ist, ist einer einwandfreien Definition nicht zugänglich, alle Definitionen sind Scheindefinitionen, wir drehen uns bei ihnen immer im Kreise. Wir wollen den Begriff der Zeit also als gegeben voraussetzen. Bewegung ist dann die Abhängigkeit des Ortes (d. h. der Koordinaten) von der Zeit. Es handelt sich nur darum, die Zeit zu messen, und wir messen sie durch eine Bewegung, beziehen also alle Bewegungen auf eine Normalbewegung. Das Axiom, das wir zugrunde legen müssen, ist, dass es eine Bewegung gibt, mit der wir alle anderen Bewegungen vergleichen, auf die wir die anderen Bewegungen beziehen können, m.a.W., dass es möglich ist, eine Uhr zu konstruieren.

7 Als diese Normalbewegung wollen wir die Lichtbewe|gung nehmen und uns

[10]In *Gauss 1828*, § 28, Gauß reported measurements of the spherical excess of the angular sum of the triangle formed by the three mountain peaks of Brocken, Hohehagen, and Inselsberg in the Hanoverian kingdom. According to his contemporary and friend Sartorius von Waltershausen, Gauß took these most precise geodesic measurements of a large terrestial triangle as evidence that the embedding three-dimensional space is Euclidean to high approximation (*Sartorius von Waltershausen 1856*, p. 81). The historical accuracy of this claim has been a topic of debate in the history of science. For a critical summary of

eine „Lichtuhr" wie folgt konstruieren: Wir stellen in dem Punkte $\frac{1}{2}$ der x-Achse einen Spiegel auf,[11] lassen vom Koordinatenanfang 0 ein Lichtsignal ausgehen und sagen, es sei die Zeiteinheit verflossen, wenn der reflektierte Strahl wieder in 0 ankommt. Ebenso soll die Zeit t verflossen sein, wenn der Spiegel in der Entfernung $\frac{t}{2}$ von 0 aufgestellt war.[12] Der Grund, warum wir die Lichtbewegung als Normalbewegung nehmen, ist hier leicht ersichtlich; es ist möglich, das Licht zu zwingen, zweimal denselben Ort zu passieren. Wir sind aber mit der Definition der Lichtuhr noch nicht am Ziel, soll sie brauchbar sein, so muss zunächst noch folgendes Axiom erfüllt sein:

Die Lichtuhr ist unabhängig vom Koordinatensystem. Es ist also gleichgültig, dass wir den Spiegel gerade im Punkt $\frac{1}{2}$ der x-Achse aufstellten; jeder andere Punkt in der Entfernung $\frac{1}{2}$ von 0 hätte uns dieselbe Stellung der Uhr angezeigt. Wir können nun in jedem Punkt der Ebene eine solche Lichtuhr konstruieren (dass die Uhr im Koordinatenanfang konstruiert war, ist ja keine Einschränkung, da man durch die im vorigen Paragraphen angegebenen Transformationen II jeden Punkt zum Koordinatenanfang machen kann). Es handelt sich nun darum, diese verschiedenen Uhren, die noch völlig unabhängig voneinander sind, in Beziehung zu setzen, nach einer von ihnen ⟨die anderen⟩ zu stellen.[13] Wir tun dies wie folgt:

Man sendet von 0 beim Uhrstand $t = t_0$ ein Licht|signal aus und stelle alle 8
Uhren in der Entfernung r von 0 auf $t_0 + r$, sobald das Signal dort ankommt. In dieser Art des Stellens liegen zwei Axiome, wenn sie Sinn haben soll, nämlich:

1) Die Stellung der Uhr soll von t_0 unabhängig sein, d. h. habe ich die Uhr wie angegeben gestellt und vergleiche die Uhrstände, indem ich dasselbe zur Zeit t_1 in 0 wiederhole, so stehen die Uhren schon alle richtig, oder anders ausgedrückt: Der Gang der Uhr ist vom Ort unabhängig.

2) Die Stellung der Uhr ist vom Punkt 0 nicht wesentlich abhängig, d. h. hätte ich die Uhren alle nach der Uhr in einem Punkt $0'$ gestellt, statt nach der Uhr in 0, so wäre der Unterschied der Stellung nach 0 und nach $0'$ bei allen Uhren derselbe, also konstant, und zwar gleich der Zeit, um die die Uhr in $0'$ gegen die Uhr in 0 falsch geht.

Wir haben nun in jedem Raumpunkt ausser seinem Kilometerstein noch eine Uhr, die vierdimensionale Mannigfaltigkeit Raum und Zeit wollen wir „Welt" nennen und einen Kilometerstein mit einem bestimmten Uhrstand infolgedessen einen „Weltpunkt" oder auch „Ereignis".

---

the debate and a detailed historical discussion that supports the account of Sartorius von Walterhausen, see *Scholz 2004*.

[11] The preceding sentence was corrected from: "Wir stellen in einem der Punkte in der x-Achse (dem Koordinatenanfang) einen Spiegel auf, ...".

[12] For this definition of a light clock, cf. *Einstein 1912*, p. 366, cf. also *Einstein 1913*, p. 1254, where this kind of clock is compared with a gravitational clock.

[13] "in Beziehung zu setzen, nach einer von ihnen zu stellen" was corrected from "in Beziehung zu stellen, nach einer von ihnen".

Wir nennen nun zwei Ereignisse *gleichzeitig*, wenn der Uhrstand bei beiden derselbe ist.

§ 4. Zusammenstellung der Axiome und Definitionen der beiden ersten Axiomgruppen

Wir wollen die bisher besprochenen Axiome und Definitionen nochmals kurz hier rekapitulieren, bevor wir zur nächsten, der wichtigsten Gruppe von Axio-
9 men und Definitionen über|gehen. Wir hatten zunächst zwei Gruppen von Axiomen nebst den zugehörigen Definitionen:

I. Axiome des Raumes

II. " der Zeit.

Und zwar waren dies im einzelnen folgende:
I. Axiome des Raumes.[14]

Axiom 1. Es gibt ein Bezugssystem $x\ y$ der Ebene, so dass die Gleichung der geraden Linie linear ist.

Axiom 2. Alle möglichen solchen Bezugssysteme gehen aus ihm hervor durch die Transformation:

$$\text{I.}\quad \begin{aligned}\xi &= \alpha_{11}x + \alpha_{12}y\\ \eta &= \alpha_{21}x + \alpha_{22}y\end{aligned} \qquad \text{nebst}\quad \xi^2 + \eta^2 = x^2 + y^2$$

$$\text{II.}\quad \begin{aligned}\xi &= x + \alpha_1\\ \eta &= y + \alpha_2\end{aligned}$$

Definition: Wir nennen $\sqrt{(x' - x'')^2 + (y' - y'')^2}$ "Entfernung" der Punkte $P' = (x', y')$ und $P'' = (x'', y'')$.

II. Axiome der Zeit:

Axiom 1. Es gibt eine Normalbewegung, die Lichtbewegung, die als Vergleichsbewegung brauchbar ist, also zur Konstruktion einer Uhr dienen kann.
Definition 1.: Definition der Lichtuhr und der Zeiteinheit.
Axiom 2.: Die Lichtuhr ist unabhängig vom Koordinatensystem. [15]
Hier ist enthalten:
10 (2a: Die Lichtuhr kann in jedem Punkt der Ebene | konstruiert werden).
Definition 2.: Stellung der Uhren in der Ebene.

[14] The following two sentences were corrected from:
"Es gibt eine Bewegung xy der Ebene, so dass
Axiom 1. die Gleichung der geraden Linie linear ist
2. alle möglichen Bewegungen gehen aus ihr hervor durch die Transformation:"

[15] "vom Koordinatensystem" was corrected in pencil to "von der Richtung".

Axiom 3. Der Gang der Uhr ist vom Ort unabhängig.
4. Die Stellung der Uhr ist von 0 nicht wesentlich[16] abhängig. (Gleichberechtigungsaxiom)

Definition 3.: Ereignis.
4. Gleichzeitigkeit.

§ 5. Der Begriff der Bewegung

Unter "Bewegung" verstehen wir, wie schon oben gesagt,[17] die Abhängigkeit der Koordinaten von der Zeit. Es handelt sich für uns nun darum, dies mathematisch so zu fassen, wie wir es im folgenden brauchen.

Kontinuum nennen wir jede stetige Punktmenge des x, y, z Raumes. Das Kontinuum bewegt sich, wenn für jeden Punkt der Menge die Koordinaten x, y, z Funktionen von t sind, wir haben also, da das Kontinuum eine dreifach unendliche Punktmenge ist, die Bewegung gegeben, wenn x, y, z in der Form

$$\begin{aligned} x &= f(\lambda, \mu, \nu, t), \\ y &= g(\lambda, \mu, \nu, t), \\ z &= h(\lambda, \mu, \nu, t) \end{aligned}$$

gegeben sind. Von diesen Funktionen f, g, h müssen wir, wenn die Definition der Bewegung mit der Erfahrung übereinstimmen soll, verschiedenes voraussetzen (so z.B. mehrfache Differenzierbarkeit usw.), hier brauchen wir nur 1) die Stetigkeit, 2) dass bei konstantem t die Zahlentripel x, y, z und λ, μ, ν einander eineindeutig zugeordnet sind.[18]

Die letztere Forderung besagt: Zwei Punkte des Kontinuums, die zur Zeit t_1 11
verschieden sind, sind auch zu jeder anderen Zeit t_2 verschieden. Die Parameter λ, μ, ν kann man noch auf unendlich viele Arten wählen, man wird dies natürlich so tun, dass sie physikalisch anschauliche Bedeutung haben.

Bahnkurven der Punkte des Kontinuums nennen wir die dreiparametrige Schar der Kurven, die durch

$$\begin{aligned} x &= f(\lambda, \mu, \nu, t), \\ y &= g(\lambda, \mu, \nu, t), \\ z &= h(\lambda, \mu, \nu, t) \end{aligned}$$

in Parameterstellung mit t als Parameter gegeben sind. Die Länge des Wegs, den ein Punkt in der Zeit t_0 bis t_1 zurücklegt, ist die Länge der Bahnkurve zwischen den zu $t = t_0$ und $t = t_1$ gehörigen Kurvenpunkten.

[16] "wesentlich" was corrected from "unendlich".

[17] § 3, p. 6, above.

[18] *Boltzmann 1897*, § 3, discusses differentiability of the trajectories of material particles as an independent axiom for mechanics, with reference to the existence of functions which are continuous everywhere but nowhere differentiable.

Mehr als das Gesagte brauchen wir vom allgemeinen Bewegungsbegriff nicht, da wir uns im folgenden auf spezielle Bewegungen, die geradlinig gleichförmigen beschränken.

§ 6. Die gleichförmige gradlinige Bewegung starrer Systeme

Wir nennen die Bewegung „*geradlinig*“, wenn die Bahnkurven aller Punkte des Kontinuums gerade Linien sind, wir nennen sie „*gleichförmig*“, wenn der Weg eines Punktes in einem Zeitintervall $t_1 - t_0$ nur von der Grösse dieses Intervalls und nicht von t_0 und t_1 selbst abhängt, also, für die *geradlinige* Bewegung:

$$\sqrt{(x_{t_1} - x_{t_0})^2 + (y_{t_1} - y_{t_0})^2 + (z_{t_1} - z_{t_0})^2} = f(t_1 - t_0)$$

12 unabhängig von x, y, z, t_0, t_1 | ist. Da f, wie aus der Definition folgt, noch die Funktionalgleichung:

$$f(t_2 - t_1) + f(t_1 - t_0) = f(t_2 - t_0)$$

erfüllen muss, muss f notwendig linear sein, also

$$f(t_1 - t_0) = (t_1 - t_0)\varphi(\lambda, \mu, \nu),$$

und daraus folgt, dass bei *geradliniger gleichförmiger Bewegung ist*:

$$\begin{aligned} x &= f_1(\lambda, \mu, \nu) + \varphi_1 \cdot t, \\ y &= f_2(\lambda, \mu, \nu) + \varphi_2 \cdot t, \\ z &= f_3(\lambda, \mu, \nu) + \varphi_3 \cdot t. \end{aligned}$$

Wir sagen, das Kontinuum bewege sich als starrer Körper, wenn die Entfernungen zweier bestimmter Punkte des Kontinuums unabhängig von der Zeit sind:

$$(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 = \text{unabhängig von } t,$$

$$x_1 = f(\lambda_1, \mu_1, \nu_1, t) \qquad x_2 = f(x_2, \mu_2, \nu_2, t) \qquad \cdots$$

u. s. w.

Im Falle der geradlinigen gleichförmigen Bewegung heisst das:

$$\left.\begin{aligned} \varphi_1 &= \text{const.} = c_1 \\ \varphi_2 &= \text{const.} = c_2 \\ \varphi_3 &= \text{const.} = c_3 \end{aligned}\right\} \quad \text{unabhängig von } \lambda,\ \mu,\ \nu,$$

also

$$\begin{aligned} x &= f_1 + c_1 t, \\ y &= f_2 + c_2 t, \\ z &= f_3 + c_3 t. \end{aligned}$$

Wir müssen nun die Parameter λ, μ, ν so wählen, resp. statt ihrer andere so einführen, dass sie eine anschauliche physikalische Bedeutung erhalten.
13 Dies machen wir so, indem wir zunächst sehen, was bei der Bewegung erhalten bleibt. Wir behaupten zunächst, dass Punkte des Kontinuums, die einmal in einer Geraden liegen, für jedes t auf einer Geraden liegen. Es sei z.B. für $t = t_0$

$$(ax + by + cz)_{t=t_0} = \text{const.} = d, \quad \text{d. h. aber:}$$

$$af_1 + bf_2 + bf_3 = d - (ac_1 + bc_2 + cc_3)t_0 = \text{const.}$$

dann ist auch

$$(ax + bx + cz)_{t=t_1} = \text{const.} = d_1,$$

wie man sofort sieht. Unsere Behauptung ist also richtig.

Die zweite Eigenschaft der geradlinig-gleichförmigen Bewegung eines starren Kontinuums ist, dass Gerade, die einmal parallel waren, parallel bleiben.

Legen wir also in das Kontinuum ein Koordinatensystem ξ, η, ζ, das mit dem Kontinuum fest verbunden ist (d. h. jeder Punkt des Kontinuums hat zu aller Zeit dieselben ξ, η, ζ-Koordinaten), so entsprechen Gerade im ξ, η, ζ-System Geraden im x, y, z-System, und insbesondere Parallelen im ξ, η, ζ-System Parallelen im x, y, z-System und umgekehrt. D. h. aber, jede Figur des ξ, η, ζ-Systems ist das affine Bild einer Figur des x, y, z-Systems, beide Ebenen sind affin aufeinander bezogen. Daraus folgt aber, dass zwischen x, y, z und ξ, η, ζ eine lineare Beziehung besteht für jedes konstante t:

$$\begin{aligned}
\xi &= A_{11}x + A_{12}y + A_{13}z + A_{14},\\
\eta &= A_{21}x + A_{22}y + A_{23}z + A_{24},\\
\zeta &= A_{31}x + A_{32}y + A_{33}z + A_{34}.
\end{aligned}$$

Dabei sind die A_{ik} Funktionen von t, die wir noch bestimmen wollen. | Löst 14
man nun nach den x, y, z auf, so muss Linearität in t herauskommen (denn die Bewegung soll ja gleichförmig geradlinig sein), da die ξ, η, ζ aber von t gar nicht abhängen, so heisst dies, nach elementaren Determinantensätzen:

$$\begin{aligned}
&A_{ik} = a_{ik} = \text{const.}, && i, k = 1, 2, 3,\\
&A_{i4} = a_{i4}t + a'_{i4}, && i \;\; = 1, 2, 3,\\
&\left.\begin{matrix} a_{i4} \\ a'_{i4} \end{matrix}\right\} = \text{const.}
\end{aligned}$$

und wir haben, wenn wir noch statt ξ, η, ζ $\quad \xi - a'_{14}$, $\eta - a'_{24}$, $\zeta - a'_{34}$ einführen (also das ξ, η, ζ-System parallel zu sich selbst verschieben) und diese Grössen wieder ξ, η, ζ nennen:

$$\begin{aligned}
\xi &= a_{11}x + a_{12}y + a_{13}z + a_{14},\\
\eta &= a_{21}x + a_{22}y + a_{23}z + a_{24},\\
\zeta &= a_{31}x + a_{32}y + a_{33}z + a_{34}.
\end{aligned}$$

Nach x, y, z aufgelöst ist das genau ein Tripel von Gleichungen wie das, durch welches wir die Bewegung definierten, nur haben wir hier für λ, μ, ν ganz spezielle Parameter gewählt, ξ, η, ζ, die aber eine anschauliche Bedeutung haben, es sind ja die Koordinaten des bewegten Punktes in einem mitbewegten Koordinatensystem. Wir haben also das Resultat:
Die geradlinig-gleichförmige Bewegung eines starren Systems lässt sich beschreiben durch drei Gleichungen von der Form:

$$\begin{aligned}\xi &= a_{11}x + a_{12}y + a_{13}z + a_{14}t,\\ \eta &= a_{21}x + a_{22}y + a_{23}z + a_{24}t,\\ \zeta &= a_{31}x + a_{32}y + a_{33}z + a_{34}t.\end{aligned}$$

Dabei können ξ, η, ζ als Koordinaten der bewegten Punkte in einem mitbewegten Koordinatensystem gedeutet werden.

15 § 7. Die Axiome der gleichförmig gradlinigen Bewegung

Wir müssen nun irgendwelche Axiome einführen, die uns erlauben, die Koeffizienten $a_{11} \ldots a_{34}$ zu bestimmen.
Wir fordern zunächst, und dies ist ein neues Axiom, dass durch jede geradlinig-gleichförmige Bewegung nur eine einzige Richtung des Raumes ausgezeichnet wird, die Richtung der Bahngeraden. Dieses Axiom führt sofort zu einigen wichtigen Folgerungen, durch die die Anzahl der wesentlichen Koeffizienten in ξ, η, ζ auf zwei reduziert wird.
Zunächst können wir die Koordinatensysteme ξ, η, ζ und x, y, z so orthogonal linear transformieren, dass die neue ξ- und die neue x-Achse Bahngerade sind, also parallel sind; dann hat die Transformation schon die speziellere Gestalt:

$$\begin{aligned}\xi &= a_{11}x + \qquad + \qquad + a_{14}t,\\ \eta &= a_{21}x + a_{22}y + a_{23}z + a_{24}t,\\ \zeta &= a_{31}x + a_{32}y + a_{33}z + a_{34}t.\end{aligned}$$

Nun machen wir davon Gebrauch, dass jede affine Transformation des Raumes sich zusammensetzt aus zwei speziellen Arten affiner Transformationen: 1) einer Aehnlichkeitstransformation, 2) einer Verkürzung des Raumes im konstanten Verhältnis in einer gegebenen Richtung, bei der eine Ebene in sich übergeht. Die erste Transformation, die Aehnlichkeitstransformation, brauchen wir hier nicht zu berücksichtigen, sie kommt ja nur darauf hinaus, dass statt ξ, η, ζ $\quad$ $\rho\xi$, $\rho\eta$, $\rho\zeta$ eingesetzt werden. (Sie lässt sich übrigens auch durch drei spezielle Transformationen 2) ersetzen) die zweite ist für uns die wesentliche.
16 Bei ihr gibt es zwei ausgezeichnete Richtungen, die Richtung der | Verkürzung und die Normale der Ebene; da es bei den durch geradlinig-gleichförmige Bewegungen vermittelten affinen Transformationen aber nur eine ausgezeichnete

Richtung geben soll, müssen beide zusammenfallen, d. h. die invariante Ebene muss senkrecht auf der Verkürzungsrichtung stehen und die Verkürzungsrichtung muss die Bahngerade sein. Daraus folgt aber $a_{21} = a_{31} = 0$ und dass a_{22}, a_{23}, a_{32}, a_{33} so beschaffen sind, dass durch eine lineare orthogonale Transformation von y, z bzw. η, ζ die Beziehung zwischen x, y, z und ξ, η, ζ die Gestalt erhält:

$$\begin{aligned}
\xi &= a_{11}x + \qquad\qquad + a_{14}t,\\
\eta &= \qquad y \qquad + a_{24}t,\\
\zeta &= \qquad\qquad z + a_{34}t.
\end{aligned}$$

Wir betrachten nun den Punkt $\xi = 0$, $\eta = 0$, $\zeta = 0$. Dieser befindet sich in der Zeit $t = 0$ in $x = 0$, $y = 0$, $z = 0$. Die x-Achse ist aber Bahnkurve, also muss er *immer* die Koordinaten $y = 0$, $z = 0$ haben (nach Definition der Bahnkurve). Also muss $a_{24} = a_{34} = 0$ sein.
Also ist schliesslich:

$$\begin{aligned}
\xi &= a_{11}x \qquad\qquad + a_{14}t,\\
\eta &= \qquad y \qquad\qquad ,\\
\zeta &= \qquad\qquad z \qquad .
\end{aligned}$$

Die geradlinig-gleichförmige Bewegung hat also höchstens zwei wesentliche Konstanten, da aber in der Definition der Bewegung nur von *einer* Konstanten die Rede war, nämlich von

$$\frac{\sqrt{(x_{t_1} - x_{t_0})^2 + (y_{t_1} - y_{t_0})^2 + (z_{t_1} - z_{t_0})^2}}{t_1 - t_0} = -\frac{a_{14}}{a_{11}},$$

so werden wir uns fragen, ob nicht Axiome nötig sind, die einen Zusammen- 17
hang herstellen zwischen a_{11} und a_{14}. Die Konstante $\frac{a_{14}}{a_{11}}$ wollen wir die *Geschwindigkeit* des Systems ξ, η, ζ gegen das System x, y, z nennen.
Aus unserer Formel folgt:

$$\begin{aligned}
x &= \frac{\xi}{a_{11}} \qquad\qquad - \frac{a_{14}}{a_{11}}t,\\
y &= \qquad \eta \qquad\qquad ,\\
z &= \qquad\qquad \zeta \qquad ,
\end{aligned}$$

und hieraus wird in Verbindung mit einem weiteren Axiom der gewünschte Zusammenhang folgen.

§ 8. Das Gleichberechtigungsaxiom der gleichförmig geradlinigen Bewegung

Im Koordinatensystem ξ, η, ζ müssen wir natürlich wieder die Zeit messen, Uhren haben. Es wird also alles darauf hinauslaufen, wie man den Uhrstand im

System ξ, η, ζ definiert. Dies wird man natürlich so versuchen, dass möglichst viel von den bisher genannten Axiomen erhalten bleibt, also insbesondere die Axiome der ersten beiden Gruppen (des Raumes und der Zeit). Zunächst ist folgendes Zwischenaxiom zu machen:

Axiom (Gleichberechtigungsaxiom der gleichförmig-geradlinigen Bewegung)

Bewegt sich das System ξ, η, ζ gleichförmig-geradlinig gegen das System x, y, z, so bewegt sich auch umgekehrt x, y, z gleichförmig-geradlinig gegen ξ, η, ζ mit der entgegengesetzt gleichen Geschwindigkeit. Waren im x, y, z System die Bahnkurven die Parallelen zur x-Achse und lag die ξ-Achse so,
18 dass sie in die x-Achse fiel, so sind umgekehrt bei der Be|wegung von x, y, z gegen ξ, η, ζ die Bahnkurven Parallelen zur ξ-Achse und die x-Achse fällt in die ξ-Achse.

Der erste Teil dieses Axioms sagt, dass stets die Zeit τ im neuen System mit der im alten System linear zusammenhängt in der Form

$$\tau = a_{41}x + a_{42}y + a_{43}z + a_{44}t,$$

wobei die a_{4i} allerdings nicht alle willkürlich sind, der zweite Teil sagt, dass bei der speziellen Orientierung unserer Koordinatensysteme, die wir zuletzt vorgenommen hatten,

$$\tau = a_{41}x + a_{44}t$$

wird. Wir haben also

$$\begin{aligned}
\xi &= a_{11}x + \cdot + \cdot + a_{14}t, \\
\eta &= \qquad y \qquad , \\
\zeta &= \qquad\quad z \qquad , \\
\tau &= a_{41}x + \cdot + \cdot + a_{44}t.
\end{aligned}$$

Nach x, y, z, t aufgelöst ergibt das:

$$\begin{aligned}
x &= \frac{a_{44}\xi}{a_{11}a_{44} - a_{41}a_{14}} - \frac{a_{14}\tau}{a_{11}a_{44} - a_{41}a_{14}} = \alpha_{11}\xi + \alpha_{14}\tau, \\
y &= \qquad\qquad \eta \qquad\qquad , \\
z &= \qquad\qquad\qquad\qquad \zeta \qquad\qquad , \\
t &= -\frac{a_{41}\xi}{a_{11}a_{44} - a_{41}a_{14}} + \frac{a_{11}\tau}{a_{11}a_{44} - a_{41}a_{14}} = \alpha_{41}\xi + \alpha_{44}\tau.
\end{aligned}$$

Nach dem Teil des Axioms, der von der Geschwindigkeit handelt, muss nun noch

$$-\frac{a_{14}}{a_{11}} = +\frac{\alpha_{14}}{\alpha_{11}}$$

sein. Dies ist aber der Fall, wenn

$$\frac{a_{14}}{a_{11}} = +\frac{a_{14}}{a_{44}},$$

d. h.

$$a_{11} = a_{44}$$

ist. Für a_{14} und a_{41} ergibt sich *keine* Bedingung.

§ 9. Das Newtonsche Axiom (Axiom von der absoluten Zeit) 19

Es liegt nun nahe, und damit kommen wir zur sogenannten *Newtonschen* Mechanik, die Annahme zu machen, dass im ξ, η, ζ System die erste Axiomgruppe erhalten bleibt, in dem Sinn, dass der ξ, η, ζ Raum ein kongruentes Bild des x, y, z Raumes (nicht nur ein affines Bild) ist. Dann muss aber

1) $a_{11} = 1$

sein. Daraus folgt weiter (infolge der letzten Formel des vorigen Paragraphen)

2) $a_{44} = 1$.

Weiter muss dann aber auch umgekehrt der x, y, z Raum ein kongruentes Bild des ξ, η, ζ Raumes sein (beide Räume sind je nicht vor einander ausgezeichnet) also muss

$$\alpha_{11} = 1 \qquad \alpha_{44} = 1$$

sein, d. h. aber:

$$a_{41}a_{14} = 0,$$

und da die Geschwindigkeit $a_{14} \neq 0$ sein soll, folgt daraus

$$a_{41} = 0,$$

womit übrigens zugleich $\alpha_{41} = 0$ miterfüllt ist.

Unser Formelsystem wird also, wenn wir für die Geschwindigkeit, die jetzt gleich a_{14} ist, noch v schreiben,

$$\begin{aligned} \xi &= x + vt, \\ \eta &= y, \\ \zeta &= z, \\ \tau &= t. \end{aligned}$$

Dies ist das Formelsystem der Newtonschen Mechanik, aus dem | wir nun 20
unsere Folgerungen ziehen wollen.

Zunächst ist klar, dass die erste Axiomengruppe wieder für das ξ, η, ζ System gilt, das ist ja in unserem Axiom enthalten, wir dürfen aber, das wird sich später zeigen, nicht behaupten, dass dies die einzig mögliche Art ist, bei der die Axiome der ersten Gruppe erhalten bleiben. Mit unserem Axiom völlig gleichwertig ist die Annahme $t = \tau$; denn einmal konnten wir $t = \tau$ aus unserem Axiom folgern, dann aber folgt auch umgekehrt, wie man nachrechnen kann, aus $t = \tau$ unser Axiom.

Man nennt unser Axiom oder das mit ihm völlig gleichwertige Axiom $t = \tau$ das *Axiom von der absoluten Zeit*; den Grund hierfür wollen wir im folgenden auseinandersetzen. Er liegt darin, dass für das ξ, η, ζ System bei unserer Annahme $t = \tau$ die Axiome der zweiten Gruppe keine Gültigkeit mehr haben.

§ 10. Die Lichtgeschwindigkeit im bewegten System nach dem Newtonschen Axiom und der Michelsonsche Versuch

Wir wollen der Einfachheit halber im folgenden wieder den Raum als nur zweidimensional annehmen, haben also

$$\begin{aligned} \xi &= x + vt, \\ \eta &= y, \\ \tau &= t \end{aligned}$$

als Grundformeln.

Die erste Folgerung, die wir aus diesen Formeln ziehen, ist das sogenannte Gesetz von der Addition der Geschwindigkeiten.

21 Es bewege sich ξ, η relativ zu x, y mit der Ge|schwindigkeit v, $\mathfrak{x}$, $\mathfrak{y}$ relativ zu ξ, η mit der Geschwindigkeit $\mathfrak{v}$; dann ist

$$\begin{aligned} \xi &= x + vt, & \mathfrak{x} &= \xi + \mathfrak{v}\tau, \\ \eta &= y, & \mathfrak{y} &= \eta, \\ \tau &= t, & \mathfrak{t} &= \tau. \end{aligned}$$

Also

$$\begin{aligned} \mathfrak{x} &= x + (v + \mathfrak{v})t, \\ \mathfrak{y} &= y, \\ \mathfrak{t} &= t. \end{aligned}$$

d. h. $\mathfrak{x}$, $\mathfrak{y}$ bewegt sich gegen x, y mit der Geschwindigkeit

$$v + \mathfrak{v}.$$

Wir betrachten nun die Konstruktion der Lichtuhr für das ξ, η-System. Senden wir zur Zeit $t = \tau = 0$ von $x = \xi = 0$ $y = \eta = 0$ ein Lichtsignal aus, so ist dies zur Zeit r in allen Punkten im Abstand r angelangt:

$$x^2 + y^2 - r^2 = 0,$$

und die Koordinaten des Punktes auf der unter dem Winkel ϑ gegen die x-Achse geneigten Geraden durch $(0, 0)$, in dem das Signal zur Zeit t anlangt, sind daher

$$x = t\cos\vartheta \qquad y = t\sin\vartheta.$$

Welcher Punkt im $\xi\eta$ System ist das nun? Offenbar, nach unserer Formel der Punkt

$$\xi = t\cos\vartheta - vt = t(\cos\vartheta - v),$$
$$\eta = t\sin\vartheta.$$

Der vom Licht im ξ, η System zurückgelegte Weg ist also:

$$\sqrt{\xi^2+\eta^2} = t\sqrt{1-2v\cos\vartheta+v^2}.$$

Nun definierten wir aber oben: Geschwindigkeit $= \frac{\text{Weg}}{\text{Zeit}}$, also ist die Geschwin- 22
digkeit des Lichtes im x, y-System konstant $= 1$ im ξ, η System aber gleich

$$\gamma_\vartheta = \sqrt{1-2v\cos\vartheta+v^2},$$

also von der Richtung abhängig. Es gilt also in der ξ, η-Ebene das Axiom von der Unabhängigkeit der Uhr vom Koordinatensystem[19] nicht mehr, wenn wir uns im ξ, η System eine Lichtuhr konstruieren. Es wäre also nicht erlaubt, im ξ, η System die Zeit mit einer Lichtuhr zu messen, sondern wir müssen die Zeit im Vorüberfahren von den Uhren des x, y Systems ablesen und dies wäre dann die richtige Zeit.

Es fragt sich nun, ob dies, d. h. also die Abhängigkeit der Lichtgeschwindigkeit in ξ, η von v und ϑ, mit der Erfahrung übereinstimmt. Wir bilden dazu

$$\frac{1}{\gamma_\vartheta} + \frac{1}{\gamma_{\vartheta+\pi}}.$$

Diese Grösse ist nämlich, wie sich gezeigt hat, der experimentellen Untersuchung leicht zugänglich. Dabei ist nach der Definition:

$$\gamma_\vartheta = \sqrt{1-2v\cos\vartheta+v^2},$$
$$\gamma_{\vartheta+\pi} = \sqrt{1+2v\cos\vartheta+v^2}.$$

$\gamma_{\vartheta+\pi}$ ist also die Lichtgeschwindigkeit in einem System ξ', η', das sich mit der Geschwindigkeit $-v$ gegen x, y bewegt, die Geschwindigkeit des Systems ξ, η gegen das System ξ', η' ist also $2v$ nach dem Additionsgesetz.

Nun ist nach dem binomischen Satz, wenn wir nur die Glieder bis zur zweiten 23
Potenz von v beibehalten:

$$\frac{1}{\gamma_\vartheta} = 1 + \frac{1}{2}(2v\cos\vartheta - v^2) + \frac{3}{8}4v^2\cos^2\vartheta,$$
$$\frac{1}{\gamma_{\vartheta+\pi}} = 1 - \frac{1}{2}(2v\cos\vartheta + v^2) + \frac{3}{8}4v^2\cos^2\vartheta.$$

Also ist

$$\frac{1}{\gamma_\vartheta} + \frac{1}{\gamma_{\vartheta+\pi}} = 2 + (3\cos^2\vartheta - 1)v^2.$$

[19]“Koordinatensystem” was corrected in blue pencil to “der Richtung”.

An dieser Formel ist folgendes zu beachten: sie gilt nur für kleine v und ferner, das ist das wichtigste, sie enthält nur die *relative* Geschwindigkeit von ξ, η gegen ξ', η', da v die Hälfte dieser Relativgeschwindigkeit ist. Auch ϑ ist aus v berechenbar, denn der Ort der Punkte, an die die Lichtsignale kommen, die im x, y System auf der Geraden $\vartheta = \text{const}$ liegen, ist die Gerade

$$\xi \sin\vartheta = \eta(\cos\vartheta - v),$$

und v eingesetzt ergibt sich ϑ. Es sind also nur Grössen vorhanden, die von der Relativbewegung der beiden Systeme ξ, η und ξ', η' abhängen. Die experimentelle Prüfung schien zunächst zu stimmen, es ergab sich $\frac{1}{\gamma_\vartheta} + \frac{1}{\gamma_{\vartheta+\pi}}$ fast genau gleich 2, die Beobachtungsfehler waren aber grösser als das Glied mit v^2, der Versuch bewies also nur das Fehlen des linearen Gliedes in v, also in dieser Hinsicht eine Uebereinstimmung mit unserer Formel.

Später erdachte *Michelson* eine Versuchsanordnung, in der er die Bewegung
der Erde benutzte, und die⟨se Anordnung⟩ erlaubte, Glieder von der Grössen-
24 ordnung v^2 noch genau zu messen. *Es* | *ergab sich* da, und dies ist der Punkt,
in dem der Fortschritt unserer Erkenntnis dem Experiment zu danken ist, *dass eine Abhängigkeit von* $\frac{1}{\gamma_\vartheta} + \frac{1}{\gamma_{\vartheta+\pi}}$ *von* v^2 *nicht vorhanden war*, dass sich in den Grenzen der Beobachtungsfehler, und das mit grosser Genauigkeit,

$$\frac{1}{\gamma_\vartheta} + \frac{1}{\gamma_{\vartheta+\pi}} = 2$$

ergab.[20]

§ 11. Die Verknüpfung von Raum und Zeit bei Erhaltung der Axiome der zweiten Gruppe (Axiom von der konstanten Lichtgeschwindigkeit)

Woher kommt nun diese Nichtübereinstimmung von Theorie und Experiment? Offenbar daher, dass das Koordinatensystem x, y *ausgezeichnet* ist; ausgezeichnet insofern, als in ihm alle Axiome der zweiten Gruppe gelten, was für das System ξ, η usw. nicht der Fall ist. Die Zeitmessung hängt aber wesentlich von diesen Axiomen ab, es ist also eine Messung in einem ausgezeichneten System, also etwas *absolutes*. Da wir schon darauf vorbereitet waren, dass wir etwas absolutes nicht erkennen können, richteten wir unsere Formel zur Prüfung unserer Axiome so ein, dass sie nur Grössen der relativen Bewegung enthielten.

Wollen wir also den Fehler korrigieren, so müssen wir dafür sorgen, dass die Axiome der zweiten Gruppe auch im ξ, η-System gelten, das wesentliche an diesen Axiomen aber war, dass die Lichtgeschwindigkeit (wir haben ja jetzt

[20] For Michelson's experiment, see *Michelson 1881* and *Michelson 1887*. For a general discussion of the experimental foundations of special relativity, see, e.g., *Laue 1911*, § 2.

den Begriff der Geschwindigkeit, den wir damals noch nicht hatten) unabhängig von der Richtung und vom Ort, konstant gleich 1 war. Dies wollen wir jetzt verlangen und dann sehen, was dabei heraus kommt. Vorher wollen wir uns aber überlegen, wie es denn | mit den Axiomen der ersten Gruppe steht. Hier 25
liegt die Sache sehr einfach. Schon dadurch, dass wir ein Koordinatensystem ξ, η haben, ist gesagt, dass im ξ, η-System wieder die Euklidische Geometrie gilt, aber nur für einen Beobachter, der im System steht; für einen Beobachter im x, y-System allerdings sehen alle Figuren verkürzt aus, ist die ξ, η-Ebene ein affines Bild der x, y-Ebene, aber nicht für Beobachter im ξ, η System.

Wir gehen jetzt zu der angekündigten Ueberlegung über. Die Lichtgeschwindigkeit ist ja jetzt, wenn wir wieder, wie vorhin zur Zeit $t = 0$ von $(0, 0)$ unter dem Winkel ϑ im x, y-System ein Signal aussenden:

$$\frac{\sqrt{\xi^2 + \eta^2}}{\tau} = \sqrt{\frac{(a_{11} t \cos\vartheta + a_{14} t)^2 + t^2 \sin^2\vartheta}{(a_{41} t \cos\vartheta + a_{44} t)^2}},$$

und diese Grösse soll konstant $= 1$ sein. Dabei müssen wir uns erinnern, dass $a_{11} = a_{44}$ war. Es ergibt sich hieraus:

$$(a_{11} \cos\vartheta + a_{14})^2 + \sin^2\vartheta = (a_{41} \cos\vartheta + a_{44})^2,$$

also

$$\begin{aligned} 1 + a_{14}^2 &= a_{44}^2, \\ a_{11} a_{14} &= a_{41} a_{44}, \\ a_{11}^2 - 1 &= a_{41}^2, \end{aligned}$$

und hieraus folgt wegen $a_{11} = a_{44}$ sofort $a_{14} = a_{41}$ und alle Gleichungen sind verträglich. Wir brauchen aber $a_{14} = a_{41}$ nicht zu benutzen, denn es ergibt sich durch Multiplikation der ersten und dritten Gleichung:

$$a_{11}^2 a_{14}^2 + a_{11}^2 - a_{14}^2 - 1 = a_{41}^2 a_{44}^2,$$

d. h. wegen der zweiten Gleichung 26

$$a_{11}^2 - a_{14}^2 - 1 = 0,$$

oder:

$$\begin{aligned} a_{41}^2 &= a_{14}^2, \\ a_{41} &= \pm a_{14}. \end{aligned}$$

Wir nehmen hier das Pluszeichen, also

$$a_{41} = a_{14},$$

daraus folgt dann $a_{11} = a_{44}$.
Setzen wir wieder die Geschwindigkeit gleich v:

$$-\frac{a_{14}}{a_{11}} = v,$$

so wird wegen der dritten Gleichung und $a_{14} = a_{41}$

$$\begin{aligned} a_{11} = a_{44} &= \frac{1}{\sqrt{1-v^2}}, \\ a_{41} = a_{14} &= -a_{11}v, \end{aligned}$$

und daher ist

$$\begin{aligned} \xi &= \frac{x - vt}{\sqrt{1-v^2}}, \\ \eta &= y, \\ \tau &= \frac{-vx + t}{\sqrt{1-v^2}}. \end{aligned}$$

Auflösung nach x, y, t ergibt:

$$\begin{aligned} x &= \frac{\xi + vt}{\sqrt{1-v^2}} \\ y &= \eta \\ t &= \frac{v\xi + \tau}{\sqrt{1-v^2}}. \end{aligned}$$

27 § 12. Erste Folgerungen; das „Relativitätsprinzip“

Wir wollen nun aus diesen beiden Formelsystemen die einfachsten Schlussfolgerungen ziehen. Dass die beiden Axiomgruppen I und II erhalten bleiben, haben wir schon gesagt. Weiter folgt aus den Formeln, dass wir die Uhren in dem einen System sehr wohl wieder nach den Uhren im anderen richten können, aber nicht, indem wir nur die Uhr ablesen, sondern auch den Kilometerstein, in dem sie steht. Dann erhalten wir in dem zweiten System die Zeit für gewisse bestimmte Uhren.

Die nächste Folgerung ist die, dass nicht jede Geschwindigkeit möglich ist, denn da nur reelle Grössen physikalische Bedeutung haben, müssen ξ, η, τ reell sein, weil x, y, t reell sind. Daraus folgt aber

$$1 - v^2 > 0,$$

d. h.

$$|v| < 1.$$

Es sind also nur Geschwindigkeiten kleiner als Lichtgeschwindigkeit möglich. Bei der früheren Annahme $t = \tau$ war dies anders, wie aus dem Satz von der Addition der Geschwindigkeiten folgt. Die grösste Geschwindigkeit ist also jetzt die Geschwindigkeit 1, die Lichtgeschwindigkeit. Das stimmt auch sehr gut mit der Erfahrung, denn fast alle Geschwindigkeiten, die wir kennen, die der Planeten, Fixsterne, Elektronen, bleiben sehr weit hinter der Lichtgeschwindigkeit zurück, nur die β-Strahlen haben eine ähnlich grosse Geschwindigkeit.[21]

Wir wollen nun das Gesetz von der Superposition der | Geschwindigkeiten 28
ableiten. Es sei also

$$\xi = \frac{x - vt}{\sqrt{1-v^2}}, \quad \eta = y, \quad \tau = \frac{-vx + t}{\sqrt{1-v^2}},$$
$$\mathfrak{x} = \frac{\xi - \mathfrak{v}\tau}{\sqrt{1-\mathfrak{v}^2}}, \quad \mathfrak{y} = \eta, \quad \mathfrak{t} = \frac{-\mathfrak{v}\xi + \tau}{\sqrt{1-\mathfrak{v}^2}}.$$

Rechnet man hieraus $\mathfrak{x}$, $\mathfrak{t}$ als Funktion von x, t aus, so ergibt sich:

$$\mathfrak{x} = \frac{x - Vt}{\sqrt{1-V^2}}, \quad \mathfrak{y} = y, \quad \mathfrak{t} = \frac{-Vx + t}{\sqrt{1-V^2}}$$

wenn V durch die Gleichung

$$V = \frac{v + \mathfrak{v}}{1 + v\mathfrak{v}}$$

definiert wird.

Nun ist, da nur Geschwindigkeiten < 1 möglich sind:

$$(1 - v^2)(1 - \mathfrak{v}^2) \geqq 0,$$

da aber

$$(1 - v^2)(1 - \mathfrak{v}^2) = (1 + \mathfrak{v}v)^2 - (v + \mathfrak{v})^2$$

ist, so folgt

$$(1 - v^2)(1 - \mathfrak{v}^2) = (1 + \mathfrak{v}v)^2(1 - V^2),$$
$$V^2 \leqq 1,$$

d. h. die zusammengesetzte Geschwindigkeit ist wieder kleiner als 1 und gleich 1, wenn eine der ursprünglichen 1 war. Man sieht durch eine elementare Ueberlegung übrigens, dass

$$V \geqq \mathfrak{v}, \quad V \geqq v$$

ist, wenn $\mathfrak{v}$ und v beide positiv sind, aber

$$v \geqq V \geqq \mathfrak{v},$$

wenn

$$v > 0, \quad \mathfrak{v} < 0$$

ist, und 29

$$v \geqq V, \qquad \mathfrak{v} \geqq V,$$

wenn v und $\mathfrak{v}$ beide negativ sind.

Eine Vereinfachung ist in unseren Formeln dadurch eingetreten, dass wir überall die Lichtgeschwindigkeit als Einheit angenommen haben. Hätten wir dies nicht getan, sondern überall die Lichtgschwindigkeit gleich c gesetzt, so hätten wir das Formelsystem erhalten, das entsteht, wenn man t und τ durch ct und $c\tau$, v aber durch $\frac{v}{c}$ ersetzt, also:

$$\xi = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}}; \quad \eta = y; \quad \tau = \frac{-\frac{vx}{c^2} + t}{\sqrt{1 - \frac{v^2}{c^2}}},$$

bezw.

$$x = \frac{\xi + vt}{\sqrt{1 - \frac{v^2}{c^2}}}; \quad y = \eta; \quad t = \frac{\frac{v}{c^2}\xi + \tau}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Machen wir hierin den Grenzübergang für $c = \infty$, so kommen die Formeln der Newtonschen Mechanik heraus; dasselbe gilt auch für das Additionsgesetz der Geschwindigkeiten:

$$V = \frac{v + \mathfrak{v}}{1 + \frac{v\mathfrak{v}}{c^2}}.$$

Dass beide Systeme für unendlich grosse Lichtgeschwindigkeit ineinander übergehen müssen, folgt auch sofort daraus, dass dann ja γ_ϑ auch unendlich ist, also eine wirkliche Abhängigkeit von ϑ bei γ_ϑ dann nicht mehr vorliegt.

Man nennt die auf den Formeln

$$\xi = \frac{x - vt}{\sqrt{1 - v^2}}; \quad \tau = \frac{-vx + t}{\sqrt{1 - v^2}}$$

30 sich aufbauende Theorie der Bewegung auch „Relativitätstheorie“, da in ihr die Gleichwertigkeit aller relativ zueinander geradlinig-gleichförmig bewegten Koordinatensysteme, d. h. die Nichtexistenz von etwas absolutem, wie der absoluten Zeit, der Nichtabhängigkeit aller physikalischen Erscheinungen vom Koordinatensystem zugrunde liegt. Dieses Prinzip der Gleichwertigkeit aller relativ zueinander geradlinig-gleichförmig bewegten Koordinatensysteme nennt man (oft zusammen mit dem ganzen Komplex von Folgerungen) das „*kleine Relativitätsprinzip*“ oder „Relativitätsprinzip“ schlechthin. Der Ausdruck „kleines“ soll es von den weitergehenden Relativitätsforderungen unterscheiden, die neueren Theorien zugrunde liegen.[22]

[21] An undated manuscript in Hilbert's papers, written perhaps by a certain Weber, lists on two pages „Einige interessante Geschwindigkeiten aus der Astronomie“, see *Cod. Ms. D. Hilbert 718.*

[22] This is Hilbert's terminology; see note 2 above.

§ 13. Zusammenhang der Relativitätstheorie mit der quadratischen Form $x^2 + y^2 + z^2 - t^2$

Die Gleichungen, die den Uebergang von dem System xy zu dem System $\xi\eta$ vermitteln, sind homogen und linear, aber von einer besonderen Form, sie transformieren nämlich den Ausdruck $x^2 - t^2$ in sich. Nach unseren Formeln ist ja:

$$\xi^2 = \frac{x^2 - 2vxt + v^2t^2}{1 - v^2},$$
$$\tau^2 = \frac{v^2x^2 - 2vxt + t^2}{1 - v^2},$$

also

$$\xi^2 - \tau^2 \equiv x^2 - t^2.$$

Wir haben also den Satz: die homogenen linearen Transformationen, die einer geradlinig-gleichförmigen Bewegung entsprechen, führen die Form

$$x^2 + y^2 + z^2 - t^2$$

in sich über. 31

Wir haben in diesem Satz schon etwas mehr gesagt, als wir bisher besprochen haben. Ueben wir nämlich auf x, y, z eine *lineare orthogonale* Transformation aus, und fügen $t = \tau$ hinzu, so ist das auch eine Transformation, die $x^2 + y^2 + z^2 - t^2$ in sich überführt, sie entspricht einer blossen Veränderung des Koordinatensystems im x, y, z-Raume, also einer Umbenennung der Punkte dieses Raumes. Dazu kommt noch die Transformation, die die Bewegung repräsentiert; bei dieser war angenommen, dass die Koordinatensysteme x, y, z und ξ, η, ζ so orthogonal transformiert waren, dass die entsprechenden Achsen parallel wurden und die x-Achse die Bewegungsrichtung angab. Unsere Behauptung ist also richtig, dass jede Transformation, die einer geradlinig-gleichförmigen Bewegung entspricht, und die also, wenn das Koordinatensystem nicht so spezialisiert wird, wie wir es spezialisiert hatten, die Form hat:

$$\xi = a_{11}x + a_{12}y + a_{13}z + a_{14}t,$$
$$\eta = a_{21}x + a_{22}y + a_{23}z + a_{24}t,$$
$$\zeta = a_{31}x + a_{32}y + a_{33}z + a_{34}t,$$
$$\tau = a_{41}x + a_{42}y + a_{43}z + a_{44}t,$$

die Form

$$x^2 + y^2 + z^2 - t^2$$

überführt in

$$\xi^2 + \eta^2 + \zeta^2 - \tau^2.$$

Wir wollen die nähere Besprechung des Zusammenhanges der quadratischen Form mit der Relativitätstheorie noch etwas zurück|stellen und uns zu einem 32 wichtigen Punkt wenden, der Frage der Gleichzeitigkeit zweier Ereignisse und dem Zusammenhang dieser Frage mit dem Kausalitätsprinzip.[23]

§ 14. Lorentzkontraktion, Gleichzeitigkeit und Kausalität

Wir haben früher den Begriff der Gleichzeitigkeit *innerhalb* des x, y-Systems, und damit, wegen der Gleichwertigkeit beider Systeme, auch *innerhalb* des ξ, η-Systems definiert. Betrachten wir nun unsere Formeln, die x, y, t mit ξ, η, τ verbinden, so sehen wir, dass aus $t = \text{const.}$ *nicht* auch $\tau = \text{const.}$ folgt. Ereignisse, die im x, y-System gleichzeitig sind, sind also im ξ, η-System nicht gleichzeitig.

Daraus folgt auch, dass Ereignisse, die nicht gleichzeitig sind, unter Umständen gleichzeitig dadurch werden können, dass man sie auf ein geeignet gewähltes Koordinatensystem bezieht. Wir fragen nun, wann können Ereignisse in dem angegebenen Sinn gleichzeitig gemacht werden? Wir gehen etwas allgemeiner vor. Wir betrachten zunächst zwei Ereignisse am selben Ort im neuen System: ξ, τ_1 und ξ, τ_2. Dann ist der zugehörige Zeitunterschied im alten System

$$t_2 - t_1 = \frac{\tau_2 - \tau_1}{\sqrt{1 - v^2}},$$

d. h.

$$t_2 - t_1 > \tau_2 - \tau_1,$$

oder: *eine Uhr im bewegten System ξ, η, verglichen mit einer Uhr im festen System x, y scheint langsamer zu gehen.*[24] Krass ausgedrückt: Durch schnelles Reisen kann man sein Leben verlängern. Weiter ist $t_2 - t_1$ nur dann Null,
33 wenn $\tau_2 - \tau_1$ Null ist: Ver|schiedene Ereignisse am selben Ort können nicht auf Gleichzeitigkeit transformiert werden.

Anders ist es mit Ereignissen ξ_1, τ u. ξ_2, τ zur selben Zeit an verschiedenen Orten. Hier ergibt sich:

$$x_2 - x_1 = \frac{\xi_2 - \xi_1}{\sqrt{1 - v^2}},$$
$$t_2 - t_1 = v\frac{\xi_2 - \xi_1}{\sqrt{1 - v^2}} = v(x_2 - x_1),$$

d. h. alle Längen im bewegten System scheinen vom ruhenden aus verkürzt[25] (Lorentzkontraktion), Ereignisse, die im bewegten System gleichzeitig sind, haben im festen eine Zeitdifferenz, die proportional ihrem Abstand ist.

Aus unserer letzten Formel folgt daher:

---

[23] Added in pencil in an unknown hand: $\xi = \frac{x - vt}{\sqrt{1-v^2}}$ and $\tau = \frac{-vx + t}{\sqrt{1-v^2}}$.

[24] In the preceding sentence "scheint" was underlined twice.

[25] In the preceding sentence "scheinen" was underlined twice.

Haben zwei Ereignisse im alten System die Zeitdifferenz $t_2 - t_1$, so werden sie im neuen System gleichzeitig, wenn die Geschwindigkeit v des neuen Systems:

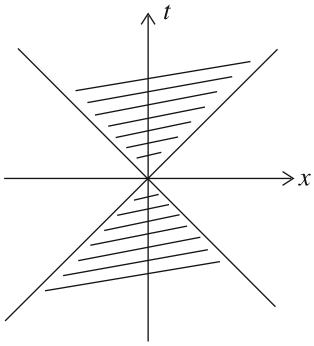
$$v = \frac{t_2 - t_1}{x_2 - x_1}$$

ist. Da aber $|v| < 1$ sein muss, so sind nur solche Ereignisse auf Gleichzeitigkeit transformierbar, für die

$$-1 \leqq \frac{t_2 - t_1}{x_2 - x_1} \leqq +1$$

ist. (Hätten wir die Lichtgeschwindigkeit nicht 1 sondern c gesetzt, so wäre[26]

$$-c > \frac{x_2 - x_1}{t_2 - t_1} > c)$$



Stellen wir uns dies in einer x,t-Ebene graphisch dar, so | heisst das ($x_1 = 0$, $t_1 = 0$ angenommen), 34
dass die Ereignisse, deren Bildpunkte in dem schraffierten Winkelraum liegen, nicht auf Gleichzeitigkeit transformiert werden können, während eine solche Transformation bei den übrigen Punkten möglich ist.

Es fragt sich nun, wie die Möglichkeit, zwei Punkte auf Gleichzeitigkeit zu transformieren, mit dem Kausalitätsprinzip verträglich ist. Soll das Kausalitätsprinzip gelten, so müssen wir verlangen, dass zwei Ereignisse nur dann auf Gleichzeitigkeit transformiert werden können, wenn sie nicht im Verhältnis von Ursache und Wirkung stehen können, d. h. wenn es nicht möglich ist, dass das eine Ereignis eine Folge des anderen ist. Soll ein Ereignis aber die Folge des anderen sein, so muss sich die Wirkung von diesem aus mit einer gewissen Geschwindigkeit fortpflanzen und diese Geschwindigkeit muss kleiner als Lichtgeschwindigkeit sein. Da wir Geschwindigkeit definiert haben als $\frac{\text{Weg}}{\text{Zeit}}$, so heisst das:

$$\left| \frac{x_2 - x_1}{t_2 - t_1} \right| < 1,$$

d. h. die Bildpunkte müssen in dem schraffierten Gebiet liegen. *Damit ist gezeigt, dass die Möglichkeit der Transformation auf Gleichzeitigkeit zum Kausalitätsprinzip nicht im Widerspruch steht: Nur solche Ereignisse können gleichzeitig gemacht | werden, die nicht im Verhältnis von Ursache und* 35
Wirkung stehen. Es ist nicht möglich, zwei Ereignisse, die auf Gleichzeitigkeit transformiert werden können, durch Signale zu verbinden.

Hätten wir statt einer Raumkoordinate x zwei ⟨Raumkoordinaten⟩ x, y genommen, so wäre an die Stelle des Winkelraumes in unserer Figur ein Rotationskegel getreten, wenn wir alle Koordinaten nehmen, ist eine geometrische

[26]In the following formula, the ">"s were corrected from "<"s.

Darstellung nicht mehr möglich, da wir dann ein Gebilde im vierdimensionalen Raum erhalten.

Die beiden Teile unseres Winkelraumes (bezw. beide Hälften des Kegels) sind übrigens nicht gleichwertig. In dem einen Teil sind nämlich die Bildpunkte aller der Ereignisse x_2, t_2, deren Uhrstand t_2 stets grösser als t_1 ist, also die für jedes Bezugssystem *zukünftigen* Ereignisse, während der andere die für jedes Bezugssystem *vergangenen* Ereignisse repräsentiert. Da diese Ereignisse allein in der Geschichte des Ereignisses x_1, t_1 eine Rolle spielen (denn nur diese Ereignisse können von ihm beeinflusst sein bezw. ihn beeinflusst haben), so wollen wir sie mit „*Zukunftskegel*" und „*Vergangenheitskegel*" des Ereignisses x_1, t_1 bezeichnen.[27]

Wir haben früher die *Bahnkurven* eines Punktes definiert, wir wollen jetzt die „*Weltlinien*" eines Punktes definieren. Die Bewegung eines Punktes war dargestellt durch drei Funktionen:

$$x = f(t),$$
$$y = g(t),$$
$$z = h(t).$$

36 Als Parameterdarstellung einer Kurve des x, y, z-Raumes gab das die Bahnkurve. Wir können diese drei Funktionen aber auch als eine Kurve im x, y, z, t-Raum (also im Raum von vier Dimensionen) auffassen, und diese Kurve nennen wir die Weltlinie des Punktes.[28]

Aus dieser Definition folgt, dass die Bahnkurve die Projektion der Weltlinie aus dem x, y, z, t-Raume in den x, y, z-Raum ist. Während sich für die Bahnkurve keine wesentlichen Einschränkungen ergaben, (jede einigermassen vernünftige Kurve kann Bahnkurve sein), ist dies für die Weltlinie anders. Jeder Punkt der Weltlinie ist ja eine Folge der Punkte mit kleinerem t, die Punkte der Weltlinie stehen also im Verhältnis von Ursache und Wirkung. Daraus folgt, dass die Neigung jeder Sekante gegen die t-Achse absolut kleiner als 1 sein muss:

$$\left| \frac{\sqrt{(x_{t_2} - x_{t_1})^2 + (y_{t_2} - y_{t_1})^2 + (z_{t_2} - z_{t_1})^2}}{t_2 - t_1} \right| < 1,$$

oder, da dies auch im Limes d. h. für die Tangente der Weltlinie gilt:

$$\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2} < 1.$$

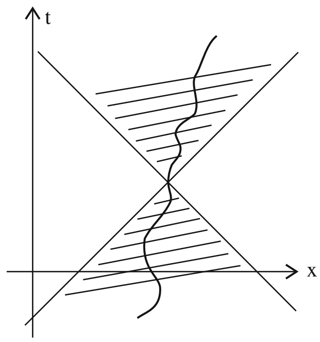
[27] In his excerpts of these lectures, Klein remarked that "Zukunftskegel" and "Vergangenheitskegel" were new terminology, introduced by Hilbert, see note 2 above. *Minkowski 1909*, p. 107, introduced the words "Nachkegel" and "Vorkegel".

[28] The term "Weltlinie" was introduced in *Minkowski 1909*, p. 104: "Wir erhalten [...] als Bild sozusagen für den ewigen Lebenslauf des substantiellen Punktes eine Kurve in der Welt, eine Weltlinie, deren Punkte sich eindeutig auf den Parameter t von $-\infty$ bis $+\infty$ beziehen lassen."

Die Quadratwurzel ist das Verhältnis des in der unendlich kleinen Zeit dt zurückgelegten Weges zu dieser Zeit, wir werden sie also als sinngemässe Verallgemeinerung des Begriffs Geschwindigkeit auffassen und Geschwindigkeit des Punktes nennen. Beschränken wir uns, um uns das graphisch besser vorstellen zu | können, auf eindimensionale Bewegungen $y = \text{const.}$, $z = \text{const.}$ 37
so muss, nach unserer Formel

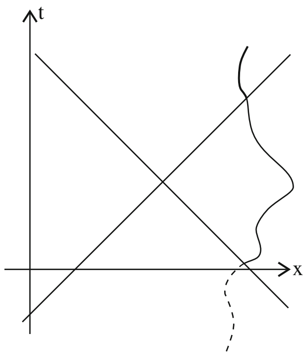
$$-1 < \frac{dx}{dt} < +1$$

sein, also die Tangente der Weltlinie steiler wie 45°. Konstruieren wir zu einem Punkte x_1, t_1 der Weltlinie Zukunfts- und Vergangenheitskegel, so zerfällt die Weltlinie in zwei Stücke, deren eines ganz im Zukunftskegel liegt, während das andere im Vergangenheitskegel verläuft.



Betrachten wir nun die Weltlinie (also die Lebensgeschichte) eines Punktes, und ein Ereignis ausserhalb der Weltlinie und fragen uns: Wann kann der Punkt unter der Einwirkung dieses Ereignisses stehen und von welchem Teil der Lebensgeschichte meines Punktes ist mein Ereignis beeinflusst? Hier brauchen wir nur zu unserem Ereignis Zukunfts- und Vergangenheitskegel zu konstruieren. Der Teil der Weltlinie, der im Zukunftskegel des Ereignisses liegt, stellt den Teil des Lebens unseres Punktes dar, der unter dem Einfluss des gegebenen Ereignisses verläuft, der Teil der im Vergangenheitskegel liegt, repräsentiert den Teil des Lebens des Punktes, der auf das Ereignis gewirkt hat.

38



§ 15. Beispiele

Wir wollen noch einige Beispiele für die im vorigen Paragraphen besprochenen Unterschiede in der Grösse von Zeit- und Raumintervallen in zwei gegen einander bewegten Systemen angeben.

1) Ein Zeitintervall von 100 Jahren auf der Sonne erscheint vom Fixsternsystem als ruhendem System aus gesehen um 6^{sec} verlängert.

2) Die Erde würde von der Sonne aus gemessen um 6 cm verkürzt erscheinen.

3) Zwei Ereignisse auf der Sonne und dem nächsten Fixstern, die von der Sonne aus betrachtet

gleichzeitig erscheinen, sind vom Fixsternsystem aus betrachtet um 2^{h} verschieden.

Die Differenzen sind also praktisch völlig bedeutungslos, die prinzipielle Wichtigkeit aber ist nicht zu verkennen. Wir verstehen jetzt das Wort Minkowskis:[29] „Raum und Zeit sinken zu blossen Schatten herab und nur eine Art Union zwischen ihnen behält Selbständigkeit.“ Die menschliche Phantasie hat wohl die nichteuklidische Geometrie aus sich heraus geschaffen, aber zu Ideenbildungen von solcher Kühnheit, wie die hier entwickelten Grundlagen der Relativitätstheorie musste der Anstoss von aussen kommen. Das Experiment gab ihn, und wenn es im Faust heisst:

„Geheimnisvoll am lichten Tag
Lässt sich Natur des Schleiers nicht berauben,
Und was sie deinem Geist nicht offenbaren mag,
Das zwingst du ihr nicht ab mit Hebeln und mit Schrauben“[30]

39 so waren hier Hebel und Schrauben stärker als die kühnste Phantasie, stärker als das reine Denken, unerbittliche Richter über eines der grössten und schönsten Systeme, die menschliches Denken zur Erklärung der Natur geschaffen.

§ 16. Die Physik als Geometrie des vierdimensionalen Raumes

Für die gleichförmig geradlinige Bewegung hatten wir das Gleichungssystem gefunden:

$$\begin{aligned}
\xi &= \alpha_{11}x + \alpha_{12}y + \alpha_{13}z + \alpha_{14}t,\\
\eta &= \alpha_{21}x + \alpha_{22}y + \alpha_{23}z + \alpha_{24}t,\\
\zeta &= \alpha_{31}x + \alpha_{32}y + \alpha_{33}z + \alpha_{34}t,\\
\tau &= \alpha_{41}x + \alpha_{42}y + \alpha_{43}z + \alpha_{44}t,
\end{aligned}$$

wo die α_{ik} der Bedingung genügten

$$\begin{vmatrix}
\alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14}\\
\alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{24}\\
\alpha_{31} & \alpha_{32} & \alpha_{33} & \alpha_{34}\\
\alpha_{41} & \alpha_{42} & \alpha_{43} & \alpha_{44}
\end{vmatrix} = 1.$$

[29] A reference to Minkowski's famous opening words of his address at the 80th annual meeting of the German Assocation of Scientists and Physicians in Cologne: “Die Anschauungen über Raum und Zeit, die ich Ihnen entwicklen möchte, sind auf experimentell-physikalischem Boden erwachsen. Darin liegt ihre Stärke. Ihre Tendenz ist eine radikale. Von Stund' an sollen Raum und Zeit für sich und Zeit für sich völlig zu Schatten herabsinken und nur noch eine Art Union der beiden soll Selbständigkeit bewahren.” *Minkowski 1909*, p. 104.

[30] From the tragedy's first part, first scene (“Nacht”), *Goethe 1993a*, lines 672–675; the same quote is used in Hilbert's Bucharest lectures, [p. 7], this Volume, p. 355.

Dann blieb der Ausdruck

$$x^2 + y^2 + z^2 - t^2$$

gegenüber dieser Transformation ungeändert.

Diese quadratische Form ist wesentlich verschieden von der entsprechenden der Geometrie

$$x^2 + y^2 + z^2.$$

Dort sind alle Vorzeichen positiv, hier dagegen sind drei positiv, ein Vorzeichen negativ, was besagt, dass Raum und Zeit nicht völlig durcheinandergehen, wenn dieses auch in quantitativer Hinsicht der Fall ist. Die Welt zerfällt in Raum und Zeit, die Vierzahl in $3 + 1$. Aeusserlich kann man das negative Zeichen leicht beseitigen, indem man statt t it einführt, und z. B.

$$it = l,$$
$$i\tau = \lambda$$

setzt. Dann geht $\xi^2 + \eta^2 + \zeta^2 - \tau^2$ über in 40

$$\xi^2 + \eta^2 + \zeta^2 + \lambda^2 = x^2 + y^2 + z^2 + l^2.$$

λ und l sind dabei rein imaginär. In diesem Sinn sind die räumlichen und zeitlichen Abmessungen völlig äquivalent. — Führen wir noch die Lichtgeschwindigkeit c ein, so lautet die quadratische Form

$$\xi^2 + \eta^2 + \zeta^2 - c^2\tau^2 = x^2 + y^2 + z^2 - c^2t^2.$$

Um auch hier überall positive[31] Vorzeichen zu bekommen, müssen wir setzen

$$ict = l,$$
$$ic\tau = \lambda.$$

Wenn wir, wie üblich, als Einheiten cm, sec. wählen und

$$c = 3 \cdot 10^{10} \text{ cm sec}^{-1}$$

setzen, so folgt hieraus

$$i \cdot 3 \cdot 10^{10} \text{ cm} = 1 \text{ sec.}$$

oder

$$i \cdot 3 \cdot 10^{5} \text{ km} = 1 \text{ sec.}$$

In dieser allgemeinen Raum- und Zeit-Auffassung spielt die Lichtgeschwindigkeit eine exceptionelle Rolle, wir kommen darauf später zurück.[32] — Durch die imaginäre Transformation

$$ict = l$$

[31] "positive" was corrected by an unknown writer from "passende".

[32] In § 30, p. 102, below, the exceptional role of light is listed as one of Hilbert's arguments against Mie's theory.

ist die genaue Analogie zur Euklidischen Geometrie des dreidimensionalen Raumes formal hergestellt. Wenn wir also jetzt die Physik aufbauen wollen, können wir uns an die Geometrie halten. Dem dreidimensionalen Raum entspricht in der Physik der vierdimensionale, die Welt; dem Punkte xyz ein
41 Punkt $xyzt$, | d. h. ein Ereignis. Es handelt sich also darum, eine Geometrie des vierdimensionalen Raumes aufzubauen, die der des dreidimensionalen Raumes formal entspricht.

Das Grundaxiom der Geometrie lautet: die Sätze der Geometrie sind von der Wahl des Bezugssystems x, y, z unabhängig, d. h. wenn wir von einem System zu einem anderen durch orthogonale Transformation[33] übergehen, erhalten wir dieselben Aussagen. Da wir nun gesehen haben, dass die Physik (Raum, Zeit) eine Verallgemeinerung der Geometrie (Raum) ist, setzen wir auch hier das Grundaxiom an die Spitze: *die Sätze der Physik sind von der Wahl des Koordinatensystems unabhängig*. Gesetze sind Aussagen. Aussagen sind Beziehungen zwischen Begriffen. Wir werden uns also mit den Beziehungen der vom Koordinatensystem[34] unabhängigen Begriffe beschäftigen. Dadurch gelangen wir in das Gebiet der Invariantentheorie. Hier speziell wird es die Vektor-Analysis sein, die uns die Aufstellung von invarianten Beziehungen vermittelt.

§ 17. Definitionen und Sätze der Dreiervektor-Analysis

Die Vektoranalysis ist eine Sprache, das Wesentliche bleibt ihr Inhalt.[35] Sie ist das Instrument, die Beziehungen der Geometrie und der Physik zu ermitteln. Im dreidimensionalen Raum mit festem Nullpunkt ist der unabhängige Begriff der Punkt selbst, er ist gegeben durch x, y, z.

(x, y, z) bezeichnet man als Radius-Vektor, drei Grössen, die sich wie (x, y, z) transformieren, als *Vektor*

$$\mathfrak{A} = (\mathfrak{A}_x, \mathfrak{A}_y, \mathfrak{A}_z).$$

42 Neben dem Vektor können wir den *Tensor* definieren.[36] Er besteht im allgemeinen Falle aus 9 Grössen:

$$\sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix},$$

[33] "zu einem anderen durch orthogonale Transformation" was corrected from "zu einem orthogonalen".

[34] „vom Koordinatensystem" was interlineated.

[35] Cf. Hilbert's exposition of three-dimensional vector calculus in his lectures on "Continuum theory" (*Hilbert 1911**, chap. II). For a historical account of vector calculus, see *Crowe 1994*. For a contemporary expositions, see, e.g. *Foeppl 1907*, chap. I.

[36] For a historical account of the notion of a tensor, see *Reich 1994*.

die sich wie die Produkte x^2, y^2, z^2, xy, $xz \cdots$ transformieren:

$$\sigma_{ik} = x_i y_k.$$

Ist speziell

$$\sigma_{ik} = \sigma_{ki} \quad \text{bez.} \quad \sigma_{ik} = -\sigma_{ki},$$

so nennt man den Tensor symmetrisch bez. schiefsymmetrisch, die 9 Tensorkomponenten reduzieren sich auf 6 bez. 3. Endlich ist

$$\sigma_{ik} = \delta_{ik} \quad \delta_{ik} = \begin{cases} 1 & \text{für} \quad i = k \\ 0 & \text{"} \quad i \neq k \end{cases}$$

ein Tensor, der bei orthogonalen Transformationen invariant bleibt.[37]
Wir stellen nun die wichtigsten Formeln der Vektorrechnung zusammen:
A) *Vektor-Algebra.*
Vektor-Addition: $\mathfrak{A} + \mathfrak{B} = (\mathfrak{A}_x + \mathfrak{B}_x, \mathfrak{A}_y + \mathfrak{B}_y, \mathfrak{A}_z + \mathfrak{B}_z)$.
Das Resultat ist wieder ein Vektor.
Bei der *Multiplikation* unterscheidet man zwei Arten.
1) *Das skalare Produkt*:

$$\mathfrak{A} \cdot \mathfrak{B} = \mathfrak{A}_x \mathfrak{B}_x + \mathfrak{A}_y \mathfrak{B}_y + \mathfrak{A}_z \mathfrak{B}_z.$$

Dies ist eine einzige Grösse, also ein Skalar. Der Radius-Vektor mit sich selbst skalar multipliziert ist $= x^2 + y^2 + z^2$ und ist daher eine invariante Grösse. Die Quadratwurzel aus dieser | Grösse bezeichnet man als Entfernung vom 43
Nullpunkt, oder allgemein als Betrag des Vektors:

$$|\mathfrak{A}| = \sqrt{\mathfrak{A}_x^2 + \mathfrak{A}_y^2 + \mathfrak{A}_z^2}.$$

2) Das *vektorielle Produkt*:

$$\mathfrak{A} \times \mathfrak{B} = (\mathfrak{A}_y \mathfrak{B}_z - \mathfrak{A}_z \mathfrak{B}_y, \mathfrak{A}_z \mathfrak{B}_x - \mathfrak{A}_x \mathfrak{B}_z, \mathfrak{A}_x \mathfrak{B}_y - \mathfrak{A}_y \mathfrak{B}_x).$$

Es stellt einen Vektor $\mathfrak{C}$ dar, der senkrecht zur Ebene der beiden Vektoren $\mathfrak{A}$ und $\mathfrak{B}$ steht, u. zwar so, dass die Vektoren $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C} = \mathfrak{A} \times \mathfrak{B}$ aufeinander folgen wie x, y, z in einem Rechtssystem; sein absoluter Betrag ist gleich dem Inhalt des von $\mathfrak{A}$ und $\mathfrak{B}$ gebildeten Parallelogramms.
Mit Hilfe dieser Grundoperationen lassen sich schon Beziehungen aufstellen. Es ist z. B.:

$$\begin{aligned} \mathfrak{A} \cdot (\mathfrak{B} \times \mathfrak{C}) &= (\mathfrak{A} \times \mathfrak{B})\mathfrak{C} \\ &= \begin{vmatrix} \mathfrak{A}_x & \mathfrak{A}_y & \mathfrak{A}_z \\ \mathfrak{B}_x & \mathfrak{B}_y & \mathfrak{B}_z \\ \mathfrak{C}_x & \mathfrak{C}_y & \mathfrak{C}_z \end{vmatrix}, \end{aligned}$$

[37]The typed words of the last sentence were corrected from "Ist endlich ... so heisst der Tensor schiefsymmetrisch." The words "bez. schiefsymmetrisch" in the preceding sentence are interlineated. Those corrections apparently were made in order to introduce the tensor δ_{ik} without retyping the page, cf. 44 below.

also $= \frac{1}{6}$ Inhalt des durch die 3 Vektoren definierten Tetraeders.
B) *Differential- und Integralrechnung der Vektoren.*
Zu diesen algebraischen Relationen treten noch die Differential- und Integraloperationen der Vektoranalysis. Auch hier handelt es sich wieder darum, dass diese Beziehungen invariant, also unabhängig von der Wahl des Bezugsystems bleiben. Ist jedem Punkt im Raum ein Vektor zugeordnet, so sind die Komponenten Funktionen von x, y, z. Man nennt nun

$$\operatorname{div}\mathfrak{A} = \frac{\partial\mathfrak{A}_x}{\partial x} + \frac{\partial\mathfrak{A}_y}{\partial y} + \frac{\partial\mathfrak{A}_z}{\partial z}.$$

44 Das ist eine einzige Grösse, ein Skalar, wie die Vektorrechnung aussagt.
Aus einem Vektor kann man wieder einen Vektor erhalten durch folgende Operation:[38] Rotation $\mathfrak{A}$ ist definiert:

$$\operatorname{rot}\mathfrak{A} = (\frac{\partial\mathfrak{A}_z}{\partial y} - \frac{\partial\mathfrak{A}_y}{\partial z}, \frac{\partial\mathfrak{A}_x}{\partial z} - \frac{\partial\mathfrak{A}_z}{\partial x}, \frac{\partial\mathfrak{A}_y}{\partial x} - \frac{\partial\mathfrak{A}_x}{\partial y}).$$

Aber auch aus einem Skalar kann man einen Vektor erhalten. Man definiert

$$\operatorname{grad} S = (-\frac{\partial S}{\partial x}, -\frac{\partial S}{\partial y}, -\frac{\partial S}{\partial z}).$$

Aus diesen Definitionen folgen nun die Identitäten:

$$\begin{aligned} \Delta S &= \frac{\partial^2 S}{\partial x^2} + \frac{\partial^2 S}{\partial y^2} + \frac{\partial^2 S}{\partial z^2} \\ &= -\operatorname{div}\operatorname{grad} S, \\ \operatorname{rot}\operatorname{rot}\mathfrak{A} &= -\operatorname{grad}\operatorname{div}\mathfrak{A} - \Delta\mathfrak{A}, \\ \operatorname{rot}(S\mathfrak{A}) &= -\mathfrak{A} \times \operatorname{grad} S + S\operatorname{rot}\mathfrak{A}. \end{aligned}$$

Dieser Differentialrechnung schliesst sich eine Integralrechnung der Vektoren an. Wichtig ist die Gauss'sche Formel. Ist K das Innere eines begrenzten Körpers, F dessen Oberfläche, und n die äussere Normalenrichtung, so besagt der Gauss'sche Satz:[39]

$$\int_K \operatorname{div}\mathfrak{A} dK = \int_F (\operatorname{grad}\mathfrak{A})_n \, dF.$$

Sei G eine samt ihren ersten Ableitungen stetige Funktion in K und auf dessen
45 Oberfläche | so ist[40]

$$\int\limits_K \operatorname{div}\mathfrak{A} dK = \int\limits_F \mathfrak{A}_n dF.$$

[38] "wieder" was corrected from "aber noch"; "durch" was corrected from ", und zwar durch".

[39] The following equation was entered in a different hand and crossed out again.

[40] The preceding half-sentence was corrected from: "wo ⟨?⟩ eine stetige Funktion von ⟨?⟩ und auf dessen Oberfläche ist, und daselbst stetige Ableitung besitzt".

§ 18. Uebertragung dieser Beziehungen auf den vierdimensionalen Raum

Um die im vorigen Paragraphen zusammengestellten Beziehungen in der Physik gebrauchen zu können, müssen wir sie erst auf den vierdimensionalen Raum übertragen. Es handelt sich um das Bezugssystem x, y, z, t, die Welt. Wir werden ein System von Formeln bekommen, wo an Stelle der Dreizahl die Vierzahl tritt. Bei manchen ⟨Formeln⟩ wird sich ganz Analoges ergeben, jedoch werden in einigen Fällen wesentliche Aenderungen auftreten. — In der Dreier-Analysis werden Vektoren mit deutschen Buchstaben, Skalare mit lateinischen Buchstaben bezeichnet. Statt rot schrieb man früher auch curl.[41]
Neben x, y, z ist dort auch die Bezeichnung x_1, x_2, x_3 gebräuchlich, die Vektor-Komponenten bezeichnet man dann mit $\mathfrak{A}_1$, $\mathfrak{A}_2$, $\mathfrak{A}_3$.
In der Vierer-Analysis ist die Bezeichnung nicht mehr so einheitlich, z. B. werden Vektoren nicht mehr durchgehend mit deutschen Buchstaben bezeichnet.[42]
Im vierdimensionalen Raum ist der Weltpunkt x, y, z, t oder in imaginären Koordinaten: x, y, z, l die einfachste Invariante. (x, y, z, l) ist auch hier der Radius-Vektor; $(\mathfrak{A}_x, \mathfrak{A}_y, \mathfrak{A}_z, \mathfrak{A}_l)$ ein Vektor. Einen solchen Vierervektor wollen wir mit kleinen lateinischen Buchstaben bezeichnen:

$$a = (x, y, z, l)$$

oder auch

$$a = (x_1, x_2, x_3, x_4).$$

Daneben können wir gleich den 16er Tensor stellen. 46

$$\sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44} \end{pmatrix} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} & \sigma_{xl} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} & \sigma_{yl} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} & \sigma_{zl} \\ \sigma_{lx} & \sigma_{ly} & \sigma_{lz} & \sigma_{ll} \end{pmatrix},$$

wo sich die 16 Grössen σ_{ik} transformieren wie die Produkte:

$$\sigma_{ik} = x_i y_k.$$

Wenn sich 16 Grössen so transformieren, sagt man, sie bilden einen 16er Tensor.[43]
Diese 16 Grössen können nun besondere Eigenschaften haben, die invariant sind. Drei der wichtigsten dieser Eigenschaften sind⟨:⟩[44]

[41] For a history of vector calculus, see *Crowe 1994*.

[42] For historical accounts of four-dimensional vector calculus, see *Norton 1992*, pp. 302–310, *Reich 1994*, sec. 5.2.2., *Walter 1999*.

[43] Added: "$\sigma_{11} + \sigma_{22} + \sigma_{33} + \sigma_{44}$ ist eine skalare Invariante."

[44] A number of corrections on this page indicate that the case $\sigma_{ik} = \delta_{ik}$ was added later, see note 37 above.

Die Symmetrie $\sigma_{ik} = \sigma_{ki}$; ferner $\sigma_{ik} = -\sigma_{ki}$, d. h. die Antisymmetrie oder Schiefsymmetrie; und schließlich die Eigenschaft $\sigma_{ik} = \delta_{ik}$, wo $\delta_{ik} = 0$ für $i \neq k$, $\delta_{ik} = 1$ für $i = k$.
Diese Eigenschaften haben also die Besonderheit, invariant zu sein, d. h. es gilt auch für die transformierten Grössen σ'_{ik}

$$\sigma'_{ik} = \sigma'_{ki}, \quad \text{bez.} \quad \sigma'_{ik} = -\sigma'_{ki}, \quad \text{bez.} \quad \sigma'_{ik} = \delta_{ik}.$$

Die Eigenschaft der Symmetrie reduziert die 16 Grössen auf 10, die der Schiefsymmetrie auf 6 und die letzte sogar auf 1, so dass der symmetrische 16er Tensor nur wesentlich 10, der schiefsymmetrische 16er Tensor nur 6 Komponenten hat, während der Tensor δ_{ik} 4 Komponenten vom Betrage eins besitzt. Man spricht dann von 10er Tensoren bezw. 6er Tensoren oder - was im letzten Falle gebräuchlicher ist - von 6er Vektoren.[45] Man wählt zur Bezeichnung eines solchen Sechservektors grosse lateinische Buchstaben:

$$F = (F_{yz}, F_{zx}, F_{xy}, F_{xl}, F_{yl}, F_{zl}),$$

47 oder als Tensor geschrieben:

$$F = \begin{pmatrix} 0 & F_{xy} & F_{xz} & F_{xl} \\ -F_{xy} & 0 & F_{yz} & F_{yl} \\ -F_{xz} & -F_{yz} & 0 & F_{zl} \\ -F_{xl} & -F_{yl} & -F_{zl} & 0 \end{pmatrix},$$

dabei ist für F_{yx} etc. $-F_{xy}$ gesetzt.
Zu F gehört ein *dualer Vektor* F^*

$$F^* = (F_{xl}, F_{yl}, F_{zl}, F_{yz}, F_{zx}, F_{xy}).$$

In dieser Reihenfolge transformieren sich die Grössen genau so wie die ursprünglichen, das liegt an der Orthogonalität der Transformation. Wenden wir auf den dualen Vektor F^* noch einmal den Prozess * an, so erhalten wir wieder F, also

$$F^{**} = F.$$

Der 6er Vektor ist von besonderer Wichtigkeit. Die entsprechende Bildung in der 3er Analysis liefert nichts besonderes, sondern nur wieder einen 3er Vektor.
Das *Rechnen mit Vektoren* wird einfach aus der 3er Analysis übertragen.
Vektor Algebra.

[45] Antisymmetric tensors in four dimensions were introduced in *Minkowski 1908*, §. 5, where they are called "Raum-Zeit-Vektoren II. Art." The term "Sechservektor" was introduced in *Sommerfeld 1910a*, cf. also *Laue 1911*, § 10; see also *Reich 1994* and *Walter 1999*.

Sind $f = (f_1, f_2, f_3, f_4) = (f_x, f_y, f_z, f_l)$ und $g = (g_1, g_2, g_3, g_4) = (g_x, g_y, g_z, g_l)$ 2 Vektoren, so ist ihre Summe definiert durch

$$f + g = (f_x + g_x, f_y + g_y, f_z + g_z, f_l + g_l)$$

Ebenso definiert man das skalare Produkt durch:

$$\begin{aligned} f \cdot g &= f_x g_x + f_y g_y + f_z g_z + f_l g_l \qquad \text{(Skalar)} \\ &= \sum_{i=1}^{4} f_i g_i \end{aligned}$$

Insbesondere ist 48

$$f \cdot f = \sum_{i=i}^{4} f_i^2$$

und

$$|f| = \sqrt{f \cdot f}.$$

Auch das vektorielle Produkt definiert man wie in der 3er Analysis, also durch die zweireihigen Unterdeterminanten. Hier erhält man wieder einen Vektor, aber wegen der Vierzahl unseres Systems einen Sechservektor

$$f \times g = F,$$

wo F_{jk} definiert ist durch[46]

$$F_{jk} = f_j g_k - f_k g_j = -F_{kj}.$$

Ein Vierervektor f skalar mit einem Sechzehnertensor σ multipliziert gibt einen Vierervektor g

$$f \cdot \sigma = g,$$

wo

$$g_j = f_x \sigma_{xj} + f_y \sigma_{yj} + f_z \sigma_{zj} + f_l \sigma_{lj}$$

bedeutet.

Speziell ist

$$f \cdot F = g,$$

wo

$$g_j = f_x F_{xj} + f_y F_{yj} + f_z F_{zj} + f_l F_{lj} = \sum_{i=1}^{4} f_i F_{ij}$$

ist.

Zwei Sechzehnertensoren σ und τ skalar multipliziert liefern

$$\sigma \cdot \tau = \sum_{jk} \sigma_{jk} \tau_{jk},$$

also einen Skalar. In der Physik kommt nur das skalare Produkt zweier spezi- 49
eller Tensoren, nämlich zweier Sechservektoren F u. G vor:[47]

$$F \cdot G = F_{yz}G_{yz} + F_{zx}G_{zx} + F_{xy}G_{xy} + F_{xl}G_{xl} + \cdots + F_{lz}G_{lz} \quad \text{(12 Glieder)},$$

$$F \cdot F = F^2 = \sum_{jk} (F_{jk})^2.$$

Wichtig ist das vektorielle Produkt zweier Tensoren[48] σ und τ: $\rho = \sigma \times \tau$, wo

$$\begin{aligned} S_{jk} &= \sigma_{jx}\tau_{xk} + \sigma_{jy}\tau_{yk} + \sigma_{jz}\tau_{zk} + \sigma_{jl}\tau_{lk} \\ &= \sum_s \sigma_{js}\tau_{sk} \end{aligned}$$

ist. Hieraus folgt für die vektorielle Mulitplikation zweier Sechservektoren F und G:

$$\begin{aligned} F \times G &= \sigma, \\ \sigma_{jk} &= \sum_s F_{js}G_{sk}. \end{aligned}$$

Alle diese Ausdrücke sind Invarianten, d. h. sie gehen nach Ausführung einer orthogonalen linearen Transformation

$$\begin{aligned} x'_1 &= \alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{13}x_3 + \alpha_{14}x_4, \\ x'_2 &= \alpha_{21}x_1 + \alpha_{22}x_2 + \alpha_{23}x_3 + \alpha_{24}x_4, \\ x'_3 &= \alpha_{31}x_1 + \alpha_{32}x_2 + \alpha_{33}x_3 + \alpha_{34}x_4, \\ x'_4 &= \alpha_{41}x_1 + \alpha_{42}x_2 + \alpha_{43}x_3 + \alpha_{44}x_4, \end{aligned}$$

wo $\sum_s (\alpha_{js})^2 = 1$ ist, über in die entsprechenden gestrichenen Grössen; z.B. geht

$$F \cdot F^* = 2(F_{yz}F_{xl} + F_{zx}F_{yl} + F_{xy}F_{zl})$$

50 über in: $2(F_{y'z'}F_{x'l'} + F_{z'x'}F_{y'l'} + F_{x'y'}F_{z'l'})$. Auch die folgende Invariante sei hier erwähnt:
Es war

$$\begin{aligned} f \times g &= F, \\ F_{jk} &= f_j g_k - f_k g_j. \end{aligned}$$

Multiplizieren wir den zu F gehörigen dualen Vektor F^* skalar mit h, so erhalten wir einen Viervektor, und durch nochmalige skalare Multiplikation

[46] In the following equation, "$= -F_{kj}$" was added in violet pencil.

[47] The preceding sentence was handwritten by the composer of the *Ausarbeitung*.

[48] The preceding sentence was corrected from: "Das skalare Produkt zweier Tensoren kommt in der Physik nicht vor. Wichtig hingegen ist ihr vektorielles Produkt".

des letzteren mit k einen Skalar, der gleich der negativen Determinante:

$$\begin{vmatrix} k_1 & k_2 & k_3 & k_4 \\ h_1 & h_2 & h_3 & h_4 \\ f_1 & f_2 & f_3 & f_4 \\ g_1 & g_2 & g_3 & g_4 \end{vmatrix}$$

ist.

Andererseits ist diese Determinante $= KF^* = K^*F$, wo $K = k \times h$ gesetzt, d. h. also

$$-k(h(f \times g)^*) = -k(hF^*) = KF^* = K^*F.$$

Setzen wir hierin $k = h$, so sind zwei Zeilen der Determinante einander gleich, d. h. also

$$k(k \cdot F^*) = 0 \quad \text{oder} \quad k(k \cdot G) = 0,$$

wo $G = F^*$ ein beliebiger Sechserverktor ist.[49]

Differentialrechnung der Vektoren.

Von grösster Wichtigkeit für die Physik ist es nun, festzustellen, was der Differentialrechnung der Dreieranalysis in der Viereranalysis entspricht.

Wie dort betrachten wir auch hier die Vektoren und ihre Komponenten als 51
Funktionen der Koordinaten x_1, x_2, x_3, x_4 bezw. x, y, z, l, d. h. wir ordnen jedem Punkte des x_1, x_2, x_3, x_4 Raumes, der Welt, einen Vektor zu. Dadurch gelangen wir wieder zu einem Vektorfelde, jetzt also zu einem Weltvektorfelde, und auf dieses übertragen wir die früheren Prozesse, die wir aber nach Sommerfeld mit grossen Anfangsbuchstaben bezeichnen.[50] Dann ist

$$\operatorname{Grad} S = (-\frac{\partial S}{\partial x}, -\frac{\partial S}{\partial y}, -\frac{\partial S}{\partial z}, -\frac{\partial S}{\partial l}).$$

Diese vier Ableitungen transformieren sich, wie man beweisen kann, wie die Komponenten eines Vierervektors. Ferner definieren wir den Ausdruck:

$$\operatorname{Div} f = \frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} + \frac{\partial f_z}{\partial z} + \frac{\partial f_l}{\partial l},$$

der, wie man leicht sieht, ein Skalar ist. Weiter wird

$$\operatorname{Div} \sigma = g, \ \text{wo}$$
$$g_j = \frac{\partial \sigma_{xj}}{\partial x} + \frac{\partial \sigma_{yj}}{\partial y} + \frac{\partial \sigma_{zj}}{\partial z} + \frac{\partial \sigma_{lj}}{\partial l};$$

[49] For the preceding paragraph, the typewritten text reads: "Andererseits ist diese Determinante wobei gesetzt ist d. h. also". The typed words "wobei gesetzt ist" were then deleted, and the remaining text was added by hand.

[50] Cf. *Sommerfeld 1910b*.

die Divergenz eines Sechzehnertensors ist also ein Vierervektor. Speziell ergibt sich hieraus:

$$\operatorname{Div} F = g,$$
$$g_j = \frac{\partial F_{xj}}{\partial x} + \frac{\partial F_{yj}}{\partial y} + \frac{\partial F_{zj}}{\partial z} + \frac{\partial F_{lj}}{\partial l}.$$

Endlich ist

$$\operatorname{Rot} f = F, \text{ wo}$$
$$F_{jk} = \frac{\partial f_j}{\partial x_k} - \frac{\partial f_k}{\partial x_j}$$

52 ist, | d. h. die Rotation eines Vierervektors ist ein Sechservektor.

Dies sind die wichtigsten Definitionen der Viereranalysis, aus ihnen ergeben sich leicht folgende *Identitäten*:[51]

$$\operatorname{Div}\operatorname{Div} F = 0, \qquad \text{(1 Gleichung)}, \tag{1}$$

denn aus

$$(\operatorname{Div} F)_j = g_j = \sum_k \frac{\partial F_{kj}}{\partial x_k}$$

folgt:

$$\begin{aligned}\operatorname{Div}\operatorname{Div} F &= \sum_j \frac{\partial}{\partial x_j} \sum_k \frac{\partial F_{kj}}{\partial x_k} \\ &= \sum_{jk} \frac{\partial^2 F_{kj}}{\partial x_k \partial x_j} \equiv 0 \quad \text{wegen } F_{kj} = -F_{jk},\end{aligned}$$

$$\operatorname{Rot}\operatorname{Grad} S = 0 \qquad \text{(6 Gleichungen)}, \tag{2}$$
$$\operatorname{Div}\operatorname{Grad} S = -\Box S \qquad \text{(1 Gleichung)}, \tag{3}$$

wo $\Box = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + \frac{\partial^2}{\partial l^2}$ bedeutet.

$$\operatorname{Div}(\operatorname{Rot} f)^* = 0 \qquad \text{(4 Gleichungen)}, \tag{4}$$
$$\operatorname{Div}\operatorname{Rot} f = -\operatorname{Div}\operatorname{Grad} f + \operatorname{Grad}\operatorname{Div} f. \tag{5}$$

Setzen wir speziell

$$\sigma = -\frac{1}{2} F \cdot F \delta_{sk} + \sum_m F_{sm} F_{km}, \quad \delta_{sk} = \begin{matrix} 0 \text{ für } s \neq k \\ 1 \text{ für } s = k, \end{matrix}$$

so gilt die Identität

$$\operatorname*{Div}_k \left\{ -\frac{1}{2} F \cdot F \delta_{sk} + \sum_m F_{sm} F_{km} \right\} \qquad \text{(4 Gleichungen)} \tag{6}$$

[51] The equation numbers on this page were added in the left page margin.

$$= \sum_k \left\{ F_{ks} \operatorname*{Div}_m F_{km} - F^*_{ks} \operatorname*{Div}_m F^*_{km} \right\}$$
$$= \sum_k \left\{ F^*_{ks} \cdot (\mathrm{Div}\, F^*)_k - F_{ks} \cdot (\mathrm{Div}\, F)_k \right\}.$$ [52]

Schließlich hatten wir schon

$$f(f \cdot F) = 0. \tag{7}$$

Die Identität: 53

$$\mathrm{Div}(\mathrm{Rot}\, f)^* = 0$$

sagt aus: Wenn wir $F = \mathrm{Rot}\, f$ bilden und zu F^* übergehen, so ist die

$$\mathrm{Div}\, F^* = 0.$$

Ist umgekehrt F ein solcher Sechservektor, dass

$$\mathrm{Div}\, F^* = 0$$

ist, so gibt es stets einen Vierervektor f, so dass

$$F = \mathrm{Rot}\, f$$

ist. Hieraus wollen wir den folgenden Satz ableiten: Gibt es einen Sechservektor F, so dass

$$\mathrm{Div}\, F^* = 0$$

ist, so gibt es auch einen Vierervektor f, so dass gleichzeitig

$$F = \mathrm{Rot}\, f$$
$$\text{und} \quad \mathrm{Div}\, F = -\Box f$$

ist.

Beweis.

Nach dem vorhergehenden Satz ist die erste Behauptung

$$F = \mathrm{Rot}\, f \quad \text{stets erfüllt.}$$

Wir können den Vierervektor aber modifizieren, ohne seine Rotation zu ändern, denn

$$\mathrm{Rot}(f + \mathrm{Grad}\, S) = \mathrm{Rot}\, f.$$

Nun fügen wir $\mathrm{Grad}\, S$ hinzu, dass

$$\mathrm{Div}(f + \mathrm{Grad}\, S) = 0 \quad \text{ist.}$$

Setzen wir $f + \mathrm{Grad}\, S = g$, so genügt g, in die obige Gleichung für f eingesetzt, unter Berücksichtigung der Iden|tität (5) 54

[52] The last line was interlineated.

$$\operatorname{Div}\operatorname{Rot} f = \operatorname{Div}\operatorname{Grad} f - \operatorname{Grad}\operatorname{Div} f$$

den beiden Anforderungen

$$\begin{aligned} F &= \operatorname{Rot} g, \\ \operatorname{Div} F &= -\Box g. \end{aligned}$$

Bezüglich der *Realitätsverhältnisse* ist noch folgendes zu bemerken: Wir haben einen Vierervektor definiert als ein System von 4 Grössen x, y, z, l, die sich wie x_1, x_2, x_3, x_4 transformieren. Um nun bei der Anwendung wieder auf reelle Grössen zurückzukommen, müssen wir für

$$x_4 = l = it$$

einführen. Nun geschieht es oft, dass man für

$$(f_1, f_2, f_3, f_4) = \overline{(f_1, f_2, f_3, if_4)}$$

schreibt, wobei f_4 reell zu nehmen ist, und man einfach das i fortlässt. Man betrachtet dann also an Stelle des Vektors

$$(f_1, f_2, f_3, f_4) \quad \text{den Vektor} \quad \overline{(f_1, f_2, f_3, f_4)}$$

dessen Komponenten alle reell sind, und sich wie

$$x, y, z, t$$

transformieren.
Wenden wir diese Regel auf die obigen Differentialoperationen an, so ergibt sich zunächst:

$$\begin{aligned} \overline{\operatorname{Div} f} &= \frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} + \frac{\partial f_3}{\partial x_3} + \frac{\partial i f_4}{\partial x_4} = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z} + \frac{\partial f_4}{\partial it} \\ &= \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z} + \frac{\partial f_4}{\partial t}. \end{aligned}$$

55 Wir erhalten also rechts wieder lauter positive Vorzeichen. Aus

$$\begin{aligned} \operatorname{Grad} S &= (-\frac{\partial S}{\partial x}, -\frac{\partial S}{\partial y}, -\frac{\partial S}{\partial z}, -\frac{\partial S}{\partial it}) \\ &= (-\frac{\partial S}{\partial x}, -\frac{\partial S}{\partial y}, -\frac{\partial S}{\partial z}, \mp i\frac{\partial S}{\partial t}) \end{aligned}$$

folgt reell geschrieben:

$$\overline{\operatorname{Grad} S} = (-\frac{\partial S}{\partial x}, -\frac{\partial S}{\partial y}, -\frac{\partial S}{\partial z}, \mp i\frac{\partial S}{\partial t}).$$

Ferner ist, wegen

$$\begin{aligned} \Box S &= -\operatorname{Div}\operatorname{Grad} S, \\ \overline{\Box S} &= \frac{\partial^2 S}{\partial x^2} + \frac{\partial^2 S}{\partial y^2} + \frac{\partial^2 S}{\partial z^2} - \frac{\partial^2 S}{\partial t^2}. \end{aligned}$$

§ 19. Analytische Formulierung des Begriffes der Weltlinie; die Eigenzeit

Es wird im folgenden unsere Aufgabe sein, die Physik unabhängig vom Koordinatensystem aufzubauen. Dass die geometrischen Gebilde vom Koordinatensystem unabhängig sind, entspricht ganz unseren anerzogenen Anschauungen. Das Bezugssystem erscheint hier als etwas künstlich Hereingebrachtes, von dem das Wesen[53] der Sache unabhängig ist. In der Physik ist es anders. Wir sind gewohnt, die physikalischen Ereignisse als an ein Bezugssystem gebunden zu betrachten. So sehr uns also die Fortschaffung des Koordinatensystems das Verständnis der Geometrie erleichtert, erschwert sie das der Physik. Wir verfahren nun am besten, wenn wir uns eine vierdimensionale Welt denken. Das einfachste Gebilde derselben ist der Weltpunkt x, y, z, t, das Ereignis. Das Lebensschicksal eines Punktes, eine Kette | von Ereignissen, bildet eine 56
Kurve, eine *Weltlinie*, die wir uns auch durch Bewegung eines Punktes erzeugt denken können. Die Bewegung eines Punktes im dreidimensionalen Raum ist gegeben durch die Bahnkurve. Kennen wir noch die Zeit t als Funktion der Bogenlänge s, so können wir angeben, wo sich der Punkt P zu jeder Zeit befindet. Im vierdimensionalen Raum stellen auch x, y, z als Funktionen von t eine Kurve dar, die die Bewegung des Punktes angibt. Wir brauchen aber hier nicht t als Funktion von s hinzuzufügen, sondern die Kurve gibt schon an sich das ganze Lebensschicksal des Punktes an. Die Projektion dieser Weltlinie auf den dreidimensionalen Raum liefert die Bahnkurve. Die Bahnkurve ist also nichts Invariantes, denn sie ist bei andern Koordinaten x', y', z' eine andere; invariant hingegen ist die Weltlinie.

Nunmehr gehen wir dazu über, den Begriff der Weltlinie analytisch zu formulieren. Eine Kurve im dreidimensionalen Raum wird in symmetrischer Form dargestellt durch die 3 Funktionen von t:

$$x = x(t),$$
$$y = y(t),$$
$$z = z(t).$$

Diese drei Gleichungen geben aber auch die Weltlinie, doch ist dann t als vierte Koordinate gegenüber den drei anderen ausgezeichnet. In der gewöhnlichen Mechanik ist diese Darstellung aber durchaus naturgemäss, denn die Zeit ist hier in der Tat | etwas Absolutes, und die Bewegung ist gegeben durch den 57
Vektor

$$\mathfrak{R} = (x(t), y(t), z(t)),$$

wo die Zeit t die Rolle eines Parameters spielt. Die Darstellung bleibt aber nicht invariant.

[53]"von dem das Wesen" was corrected from "das vom Wesen".

Um auch die Weltlinie in symmetrischer Gestalt darzustellen, geben wir x, y, z, t gleichzeitig als Funktionen eines Parameters p:

$$\begin{aligned} x &= x(p), \\ y &= y(p), \\ z &= z(p), \\ t &= t(p), \end{aligned}$$

wo der Parameter p seiner Natur nach eine skalare Grösse ist. Es handelt sich nun um die Normierung von p, wobei ja eine gewisse Willkür besteht. In der Kurventheorie wählt man, wenn eine Kurve gegeben ist durch die Gleichungen

$$\begin{aligned} x &= x(p), \\ y &= y(p), \\ z &= z(p) \end{aligned}$$

oft für p die Bogenlänge s, die definiert ist durch die Invariante

$$\left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2 + \left(\frac{dz}{ds}\right)^2 = 1,$$

woraus folgt

$$s = \int \sqrt{dx^2 + dy^2 + dz^2} = \int \sqrt{\left(\frac{dx}{dp}\right)^2 + \left(\frac{dy}{dp}\right)^2 + \left(\frac{dz}{dp}\right)^2}\, dp.$$

58 Wir wollen jetzt eine analoge Ueberlegung auch in der Physik | ausführen. Um auch hier die Willkür bezüglich p zu beseitigen, führen wir statt p einen bestimmten Parameter τ ein durch die Gleichung:

$$\tau = \int \sqrt{\left(\frac{dx_1}{dp}\right)^2 + \left(\frac{dx_2}{dp}\right)^2 + \left(\frac{dx_3}{dp}\right)^2 + \left(\frac{dx_4}{dp}\right)^2}\, dp,$$

wofür wir auch schreiben können:

$$\left(\frac{dx_1}{dp}\right)^2 + \left(\frac{dx_2}{dp}\right)^2 + \left(\frac{dx_3}{dp}\right)^2 + \left(\frac{dx_4}{dp}\right)^2 = \pm 1.$$

Das Vorzeichen bestimmt sich aus folgender Ueberlegung:
Setzen wir für

$$\begin{aligned} x_1 &= x(p), \\ x_2 &= y(p), \\ x_3 &= z(p), \\ x_4 &= it \end{aligned}$$

und für $p = t$, so erhalten wir

$$\tau = \int \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2 - 1}\, dt.$$

Hier geben nun die drei Quadrate unter der Wurzel das Quadrat der gewöhnlichen Geschwindigkeit, das ja $\leqq 1$ sein soll. Damit also die Quadratwurzel reell ausfällt, müssen wir das Vorzeichen umkehren, also setzen

$$\tau = \int \sqrt{1 - \left(\frac{dx}{dt}\right)^2 - \left(\frac{dy}{dt}\right)^2 - \left(\frac{dz}{dt}\right)^2}\, dt$$

oder

$$\tau = \int \sqrt{1 - \mathfrak{v}^2}\, dt,$$

wo $\mathfrak{v}$ die alte Bedeutung hat:

$$\mathfrak{v}^2 = \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2,$$
$$\mathfrak{v} = \frac{ds}{dt}.$$

Die Gleichung 59

$$\tau = \int \sqrt{1 - \mathfrak{v}^2}\, dt$$

lässt sich noch etwas umformen:
Es gilt

$$\left(\frac{d\tau}{dt}\right)^2 = 1 - \mathfrak{v}^2$$

oder

$$\begin{aligned} 1 &= \left(\frac{dt}{d\tau}\right)^2 - \mathfrak{v}^2 \left(\frac{dt}{d\tau}\right)^2 \\ &= \left(\frac{dt}{d\tau}\right)^2 - \left(\frac{dx}{d\tau}\right)^2 - \left(\frac{dy}{d\tau}\right)^2 - \left(\frac{dz}{d\tau}\right)^2 \end{aligned}$$

oder

$$-1 = \left(\frac{dx}{d\tau}\right)^2 + \left(\frac{dy}{d\tau}\right)^2 + \left(\frac{dz}{d\tau}\right)^2 - \left(\frac{dt}{d\tau}\right)^2.$$

Gehen wir nun zur allgemeinen Darstellung über, so erhalten wir

$$\left(\frac{dx_1}{d\tau}\right)^2 + \left(\frac{dx_2}{d\tau}\right)^2 + \left(\frac{dx_3}{d\tau}\right)^2 + \left(\frac{dx_4}{d\tau}\right)^2 = -1,$$

d. h. in unserer oben erwähnten Gleichung gilt das negative Vorzeichen.

τ ist dann für die Weltlinie der naturgemässe Parameter, der hier auch stets reell ausfallen wird. Er entspricht genau der Bogenlänge in der Geometrie und ist, wie diese dort, hier eine invariante und skalare Grösse. Wir werden uns im folgenden stets seiner bedienen, sobald es sich darum handeln wird, die Weltlinie in Parametergestalt darzustellen. Für die Weltlinie ergibt sich also die allgemeine imaginäre Darstellung

$$\begin{aligned} x_1 &= x_1(\tau), \\ x_2 &= x_2(\tau), \\ x_3 &= x_3(\tau), \\ x_4 &= x_4(\tau), \end{aligned}$$

60 oder wenn wir die letzte Gleichung reell schreiben:

$$x = x(\tau), \quad y = y(\tau), \quad z = z(\tau), \quad t = t(\tau).$$

Der der Weltlinie zugehörige Parameter τ wird die *Eigenzeit des Weltpunktes* genannt.

Um den Sinn dieser Bezeichnung zu verstehen, erinnern wir uns daran, wie wir in der Geometrie die Bogenlänge definieren. Diese ist gleich der oberen Grenze aller eingeschriebenen Polygone der Kurve. Wir betrachten dort also zunächst die Sehnen, lassen diese kleiner und kleiner werden und gehen zur Grenze über. Analog verhält es sich mit der Eigenzeit. Man gelangt zu ihr, indem man die Weltlinie als aus lauter geraden Stücken bestehend betrachtet, (d. h. indem man die Bewegung des Punktes längs der Weltlinie so auffasst, als wenn sie in jedem einzelnen Teilintervall gleichförmig geradlinig wäre[G]), und zur Grenze übergeht. Wir denken uns also einen gleichförmig bewegten Punkt, der, von einem zweiten System aus betrachtet, sich in Ruhe befindet, dann lauten die Transformationsformeln bekanntlich, wenn wir wie oben der Einfachheit halber die Bewegung als in der x, t Ebene verlaufend betrachten:

$$\begin{aligned} x &= \frac{1}{\sqrt{1-\mathfrak{v}^2}}(\xi + \mathfrak{v}\tau), \\ t &= \frac{1}{\sqrt{1-\mathfrak{v}^2}}(|\mathfrak{v}|\xi + \tau). \end{aligned}$$

61 Dann ergibt sich für ein Zeitintervall $\tau_2 - \tau_1 = \sqrt{1-\mathfrak{v}^2}(t_2 - t_1)$ | oder, wenn wir zur Grenze übergehen

$$d\tau = \sqrt{1-\mathfrak{v}^2}dt.$$

Hieraus erkennt man, dass die Eigenzeit $d\tau$ der Zeitzuwachs ist, den die Uhr anzeigt, wenn sie dies Element durchläuft. Für ein endliches Stück der

[G] Nur dann ist nämlich eine Transformation auf Ruhe möglich.

Weltlinie gelangen wir zu ihr durch Integration über elementare geradlinige Stücke.[54] Wir können also sagen: die Eigenzeit ist diejenige Zeit, die in einem mit dem Punkte mitbewegten Koordinatensystem gemessen wird.

§ 20. Geschwindigkeit und Beschleunigung

Nachdem wir τ eingeführt haben, ist es leicht, die wichtigsten Begriffe der Mechanik so zu modifizieren, dass sie auf die View-Sprache passen.

Geschwindigkeit: In der gewöhnlichen Mechanik ist die Geschwindigkeit $\mathfrak{v}$ ein Dreiervektor, den wir erhalten, indem wir den Radiusvektor

$$w = (x(t), y(t), z(t))$$

nach t differenzieren:

$$\mathfrak{v} = \frac{dw}{dt} = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right) = (\mathfrak{v}_x, \mathfrak{v}_y, \mathfrak{v}_z),$$

$$|\mathfrak{v}|^2 = \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2.$$

Analog definieren wir nun als Geschwindigkeit den Vierervektor v

$$v = \frac{dr}{d\tau} = \left(\frac{dx_1}{d\tau}, \frac{dx_2}{d\tau}, \frac{dx_3}{d\tau}, \frac{dx_4}{d\tau}\right)$$

oder in reeller Sprache 62

$$\frac{dx}{d\tau}, \frac{dy}{d\tau}, \frac{dz}{d\tau}, \frac{dt}{d\tau}.$$

v nennen wir kurz die *Vierergeschwindigkeit*. Ihre Komponenten sind also:

$$v_1 = \frac{dx_1}{d\tau},$$
$$v_2 = \frac{dx_2}{d\tau},$$
$$v_3 = \frac{dx_3}{d\tau},$$
$$v_4 = \frac{dx_4}{d\tau}$$

[54]The preceding two sentences were corrected from: "Hieraus erkennt man, dass die Eigenzeit $d\tau$ gerade das ist, was der   , das wir oben bei der geradlinigen Bewegung betrachtet haben, entsprechen würde. Nun haben wir zwar bei unserer allgemeinen Weltlinie keine geradlinige gleichförmige Bewegung, doch gelangen wir gerade zu dieser Formel durch Integration über elementare geradlinige Stücke. Diese Eigenzeit ist gerade das, was auf unsere alten Ueberlegungen passt."

oder

$$\begin{aligned} v_x &= \frac{dx}{d\tau} = \frac{\mathfrak{v}_x}{\sqrt{1-\mathfrak{v}^2}}, \\ v_y &= \frac{dy}{d\tau} = \frac{\mathfrak{v}_y}{\sqrt{1-\mathfrak{v}^2}}, \\ v_z &= \frac{dz}{d\tau} = \frac{\mathfrak{v}_z}{\sqrt{1-\mathfrak{v}^2}}, \\ v_t &= \frac{dt}{d\tau} = \frac{i}{\sqrt{1-\mathfrak{v}^2}} \end{aligned}$$

oder, wenn man reell schreibt und das i fortlässt

$$v_t = \frac{1}{\sqrt{1-\mathfrak{v}^2}}.$$

Daraus folgt umgekehrt:

$$\begin{aligned} \mathfrak{v}_x &= \frac{v_x}{v_t}, \\ \mathfrak{v}_y &= \frac{v_y}{v_t}, \\ \mathfrak{v}_z &= \frac{v_z}{v_t}. \end{aligned}$$

Die Vierergeschwindigkeit v ist kein beliebiger Vektor, denn aus der Definitionsgleichung für τ,

$$\left(\frac{dx_1}{d\tau}\right) + \left(\frac{dx_2}{d\tau}\right)^2 + \left(\frac{dx_3}{d\tau}\right)^2 + \left(\frac{dx_4}{d\tau}\right)^2 = -1,$$

63 folgt

$$v_1^2 + v_2^2 + v_3^2 + v_4^2 = -1$$

oder

$$v \cdot v = -1.$$

Man erkennt leicht, dass die einzelnen Komponenten von v beliebig gross sein dürfen und nicht der früher erwähnten Bedingung für die Geschwindigkeit $|v| < 1$ unterliegen.[55]

Beschleunigung. In der gewöhnlichen Mechanik bekommen wir die Beschleunigung durch zweimalige Differentiation des Radiusvektors w nach t:

$$\mathfrak{b} = \frac{d^2 w}{dt^2} = \left(\frac{d^2 x}{dt^2}, \frac{d^2 y}{dt^2}, \frac{d^2 z}{dt^2}\right).$$

[55] Cf. the discussion in § 14 above.

Analog definieren wir auch hier den Beschleunigungsvektor b,

$$\begin{aligned} b = \frac{d^2r}{d\tau^2} &= \left(\frac{d^2x_1}{d\tau^2}, \frac{d^2x_2}{d\tau^2}, \frac{d^2x_3}{d\tau^2}, \frac{d^2x_4}{d\tau^2}\right) \\ &= \frac{dv}{d\tau}, \end{aligned}$$

und nennen ihn die *Viererbeschleunigung*.
Durch Differentiation ⟨der⟩ Gleichung[56]

$$v \cdot v = -1$$

nach τ erhalten wir:

$$v\frac{dv}{d\tau} = 0$$

oder

$$vb = 0,$$

d. h. Der *Geschwindigkeitsvektor v steht senkrecht auf dem Beschleunigungsvektor b.*
Damit umgekehrt ein Vektor b Beschleunigungsvektor | sein kann, muss er der 64
Bedingung genügen:

$$v \cdot b = v_1\frac{dv_1}{d\tau} + v_2\frac{dv_2}{d\tau} + \frac{dv_3}{d\tau} + \frac{dv_4}{d\tau} = 0.$$

Ferner gilt die Beziehung:

$$b \cdot b \geqq 0.$$

Um sie zu beweisen, machen wir folgende Ueberlegung: Ist E ein Punkt einer Weltlinie, so können wir ein Bezugssystem wählen, von dem aus betrachtet E für den Augenblick sich in Ruhe befindet, so dass gilt

$$v_x = v_y = v_z = 0.$$

Geometrisch heisst dies: die neue t-Achse, also die t'-Achse, ist Tangente der Weltlinie. Der Punkt ist im ursprünglichen System in Ruhe, wenn die t'⟨-Achse⟩ parallel der t-Achse ist. Jeder Punkt E der Weltlinie ist also von einem geeignet gewählten Bezugssystem aus betrachtet für einen Augenblick in Ruhe, und es folgt dann aus der Beziehung

$$v \cdot b = v_x b_x + v_y b_y + v_z b_z + v_t b_t = 0$$

$$v_t b_t = 0,$$

und da v_t sicher $\neq 0$ ist, vielmehr $v_t = \frac{1}{\sqrt{1-\mathfrak{v}^2}} = 1$, so wird

$$b_t = 0.$$

[56]"Gleichung" was corrected from "der Definitionsgleichung der Eigenzeit".

Daraus ergibt sich:

$$b \cdot b = b_x^2 + b_y^2 + b_z^2 \geqq 0, \qquad \text{q. e. d.}$$

Wir wollen jetzt zwei Arten von Vektoren definieren. Man nennt einen Vierervektor f *zeitartig*, wenn das Produkt

$$f \cdot f < 0$$

65 ⟨ist,⟩ *raumartig*, wenn ⟨das Produkt⟩

$$f \cdot f > 0$$

ist. Ein Beispiel für den zeitartigen Vektor ist die Vierergeschwindigkeit v; die Beschleunigung b hingegen stellt, wie eben bewiesen, einen raumartigen Vektor dar. $f \cdot f = 0$ sind die Lichtstrahlvektoren.[57]

Unsere bisherigen Ueberlegungen beziehen sich auf die Bewegung eines Punktes, eine Weltlinie. Was im Dreidimensionalem für die Bahnkurve gilt, lässt sich auf die Weltlinie übertragen, nämlich auch[58] hier wird die Geschwindigkeit v durch die Tangente, die Beschleunigung b durch die Krümmung der Kurve bestimmt.

§ 21. Definition von Volumen, Masse und Dichte

Volumen. Zunächst soll jetzt der *Begriff des Volumens* in die Vierersprache eingeführt werden. In der gewöhnlichen Mechanik ist das Volumen eine skalare Grösse, also eine Invariante; denn es ist definiert durch

$$\iiint dxdydy.$$

Im vierdimensionalen Raum ist dies aber keine Invariante, vielmehr gilt dieses erst von dem Weltstück I, das definiert ist durch

$$I = \iiiint dxdydzdt.$$

Bilden wir also

$$\begin{aligned} V = \frac{\partial I}{\partial t} &= \frac{\partial I}{\partial \tau}\frac{\partial \tau}{\partial t} = \frac{\frac{\partial I}{\partial \tau}}{\frac{\partial t}{\partial \tau}}, \\ &= \sqrt{1 - \mathfrak{v}^2}\frac{\partial I}{\partial \tau}, \end{aligned}$$

66 so sehen wir, dass zwar nicht V selbst, wohl aber

[57]The preceding sentence was interlineated.

[58]The first part of this sentence was corrected from “Wir sind somit zum Begriff der Kurve im vierdimensionalen Raum gelangt. Auch”.

$$\frac{v}{\frac{\partial\tau}{\partial t}} = \frac{V}{\sqrt{1-\mathfrak{v}^2}}$$

eine *Invariante* ist.

Masse und Dichte. Um nun zum *Begriff der Dichte* zu gelangen, müssen wir von dem der Masse ausgehen. Wir treffen die Festsetzung, dass die Masse m eine Invariante sein soll. Dann ist in der gewöhnlichen Mechanik die Dichte ρ der Grenzwert des Quotienten $\frac{\text{Masse}}{\text{Volumen}}$. Wir bilden

$$\rho = \lim \frac{m}{V} = \frac{dt}{d\tau} \lim \frac{m}{\frac{\partial I}{\partial\tau}} = \frac{dt}{d\tau}\tilde{\tilde{\rho}},$$

wo $\tilde{\tilde{\rho}} = \lim \frac{m}{\frac{\partial I}{\partial\tau}}$ gesetzt ist. Dieser der gewöhnlichen Dichte analog gebildete Ausdruck,

$$\frac{dt}{d\tau}\tilde{\tilde{\rho}} = v_t\tilde{\tilde{\rho}},$$

hat also den Charakter der vierten Komponente eines Vierervektors, ist folglich keine Invariante. Daher liegt es nahe, nicht diese Komponente, sondern direkt eine *Viererdichte* r einzuführen durch

$$r = v\tilde{\tilde{\rho}} = (v_x\tilde{\tilde{\rho}}, v_y\tilde{\tilde{\rho}}, v_z\tilde{\tilde{\rho}}v_t\tilde{\tilde{\rho}}).$$

Befindet sich der Körper in Ruhe, so ist

$$v_x = v_y = v_z = 0,$$
$$v_t = 1,$$

r geht also über in $\tilde{\tilde{\rho}}$. Dies hat Anlass dazu gegeben, | $\tilde{\tilde{\rho}} = \lim \frac{m}{\frac{\partial I}{\partial\tau}}$ als die 67
Ruhdichte zu bezeichnen. [59] $\tilde{\tilde{\rho}}$ *und* τ *sind Invarianten.*

Damit eine Funktion Viererdichte sein kann, muss sie einer Differentialgleichung genügen. In der gewöhnlichen Mechanik ist diese Gleichung, die die Dichte ρ erfüllen muss, die Kontinuitätsgleichung:

$$\frac{\partial\rho}{\partial t} + \operatorname{div}(\rho\mathfrak{v}) = 0.$$

Wie lautet die analoge Gleichung in unserer Mechanik?

Um die Viererdichte in die Kontinuitätsgleichung einzuführen, berücksichtigen wir die ⟨in⟩ § 20 abgeleiteten Beziehungen:

$$v_x = \frac{\mathfrak{v}_x}{\sqrt{1-\mathfrak{v}^2}},$$
$$v_y = \frac{\mathfrak{v}_y}{\sqrt{1-\mathfrak{v}^2}},$$
$$v_z = \frac{\mathfrak{v}_z}{\sqrt{1-\mathfrak{v}^2}},$$
$$v_t = \frac{1}{\sqrt{1-\mathfrak{v}^2}},$$

[59] Added in an unknown hand: "In der Tat wird für $t = \tau$ $\rho = \tilde{\rho}$".

woraus folgt:

$$\begin{aligned} r_x &= \frac{\overline{\rho}\mathfrak{v}_x}{\sqrt{1-\mathfrak{v}^2}} = \rho\mathfrak{v}_x,\\ r_y &= \frac{\overline{\rho}\mathfrak{v}_y}{\sqrt{1-\mathfrak{v}^2}} = \rho\mathfrak{v}_y,\\ r_z &= \frac{\overline{\rho}\mathfrak{v}_z}{\sqrt{1-\mathfrak{v}^2}} = \rho\mathfrak{v}_z,\\ r_t &= \frac{\overline{\rho}}{\sqrt{1-\mathfrak{v}^2}} = \rho. \end{aligned}$$

68 Setzen wir dieses ein, so erhalten wir:

$$\frac{\partial r_x}{\partial x} + \frac{\partial r_y}{\partial y} + \frac{\partial r_z}{\partial z} + \frac{\partial r_t}{\partial t} = \operatorname{Div} r = 0.$$

Die Kontinuitätsgleichung sagt also aus: *die Weltdivergenz der Viererdichte r ist gleich* 0.

§ 22. Kraft und Energie

Kraft. Jetzt können wir auch den *Begriff der Kraft* einführen. Sie ist definiert durch

$$\begin{array}{ccccc} \text{Kraft} & = & \text{Masse} & \cdot & \text{Beschleunigung},\\ k & = & m & \cdot & b. \end{array}$$

Da m ein Skalar ist, gilt

$$v \cdot k = 0, \quad \text{da ja} \quad v \cdot b = 0 \quad \text{ist},$$

d. h. *auch die Kraft steht auf der Geschwindigkeit senkrecht*. Ausführlich geschrieben lauten ihre Komponenten:

$$\begin{aligned} k_x &= mb_x = m\frac{d^2x}{d\tau^2},\\ k_y &= mb_y = m\frac{d^2y}{d\tau^2},\\ k_z &= mb_z = m\frac{d^2z}{d\tau^2},\\ k_t &= mb_t = m\frac{d^2t}{d\tau^2}. \end{aligned}$$

Energie. An die letzte Gleichung,

$$k_t = mb_t = m\frac{d^2t}{d\tau^2},$$

wollen wir noch eine Betrachtung knüpfen. Die Gleichung

$$v \cdot k = 0$$

oder:

$$v_x k_x + v_y k_y + v_z k_z + v_l k_l = 0$$

oder reell geschrieben: 69

$$v_x k_x + v_y k_y + v_z k_z - v_t k_t = 0$$

lässt sich, wenn wir die gewöhnliche Geschwindigkeit einführen, schreiben:

$$\mathfrak{v}_x k_x + \mathfrak{v}_y k_y + \mathfrak{v}_z k_z = k_t = m \frac{dv_t}{d\tau}.$$

Fassen wir links die drei Viererkomponenten von k als Dreierkraft $\mathfrak{k}$ auf, und multiplizieren die Gleichung mit $d\tau$, so stellt die linke Seite im Sinne der alten Mechanik die während der Zeit $d\tau$ geleistete Arbeit dar:

$$d\tau\{\mathfrak{v}_x \mathfrak{k}_x + \mathfrak{v}_y \mathfrak{k}_y + \mathfrak{v}_z \mathfrak{k}_z\} = mdV_t.$$

Hieraus folgt: Die linke Seite hat die Dimension einer Arbeit[60] In diesem Sinne dürfen wir

$$mv_t - m = k$$

als die *kinetische Energie* des Massenpunktes ansprechen. Die Integrationskonstante m ist so normiert, daß für $\mathfrak{v}_t = 0$ auch $k = 0$ wird.[61] Unsere obige Gleichung lautet dann

$$dk = d\tau\{\mathfrak{v}_x \mathfrak{k}_x + \mathfrak{v}_y \mathfrak{k}_y + \mathfrak{v}_z \mathfrak{k}_z\},$$

d. h. die geleistete Arbeit ist gleich der Energieänderung.

Dies ist der Satz von der Erhaltung der Energie.

Wir können ihn auch in der Form schreiben

$$\mathfrak{v}\mathfrak{k} = \frac{dk}{d\tau}.$$

Die Energie k ist keine Invariante.

Aus unserer Definition von k:

$$k = mv_t - m$$

folgt: 70

$$k = m(\frac{dt}{d\tau} - 1),$$

[60] The preceding sentence was corrected in an unknown hand from: "wo die Integrationskonstante gesetzt ist."

[61] The preceding sentence was interlineated.

d. h. das Zurückbleiben der Zeit gegen die Eigenzeit ist ein Mass für die Energie. Setzen wir wieder

$$\frac{dt}{d\tau} = \frac{1}{\sqrt{1-\mathfrak{v}^2}},$$

d. h.

$$k = m(\frac{1}{\sqrt{1-\mathfrak{v}^2}} - 1)$$

und entwickeln nach $\mathfrak{v}$, so wird

$$\begin{aligned} k &= m(1 + \frac{1}{2}|\mathfrak{v}|^2 + \frac{3}{8}|\mathfrak{v}|^4 + \cdots - 1) \\ &= \frac{1}{2}m|\mathfrak{v}|^2 + \cdots . \end{aligned}$$

Der Begriff stimmt also in erster Näherung (für kleine Geschwindigkeiten) mit dem der kinetischen Energie aus der gewöhnlichen Mechanik überein.

Es sei hier nochmals ausdrücklich bemerkt, dass unsere bisherigen Untersuchungen nur dazu dienten, unsere aus der Dreiersprache bekannten Grundbegriffe der Mechanik so zu erweitern, dass sie auf die Vierersprache passen. Wir haben also nur Definitionen aufgestellt, keine physikalischen Gesetze. Wir verfahren nun analog mit den Begriffen der Elektrodynamik.

§ 23. Grundbegriffe der Elektrodynamik; Viererpotential, elektrische Viererdichte

Im Mittelpunkt der Mechanik der Kontinua steht der Begriff der Dichte. Die Dichte ist dort eine skalare Grösse. Ihr Analogon in der Elektrodynamik ist die elektrische Dichte. Während wir nun § 21 den Begriff der Dichte direkt auf die
71 Vie|reranalysis übertrugen, wobei wir zu einem Vierervektor der Viererdichte r gelangten, ist es zweckmässig, beim Ausbau der Elektrodynamik von einem tiefer liegenden Begriff auszugehen; die elektrische Dichte werden wir dann später daraus ableiten. Mit Rücksicht auf die gewöhnliche Elektrostatik liegt es nahe, vom Begriff des Potentials auszugehen. Dieses ist dort eine skalare Grösse φ. Wie wir aber in der Mechanik die Dichte ρ zur Viererdichte r erweiterten, wollen wir auch hier einen Vierervektor q,

$$q = (q_x, q_y, q_z, q_l),$$

das Viererpotential, einführen und aus ihm die wichtigsten elektrischen Grundbegriffe ableiten. Durch Differentiation des Potentials φ nach den Koordinaten, bekommen wir dort den elektrischen Vektor

$$\mathfrak{E}_x = -\frac{\partial \varphi}{\partial x},$$
$$\mathfrak{E}_y = -\frac{\partial \varphi}{\partial y},$$
$$\mathfrak{E}_z = -\frac{\partial \varphi}{\partial z}.$$

Wir können also erwarten, dass wir hier den elektrischen Vektor durch Differenzieren von q erhalten. Daher bilden wir $\operatorname{Rot} q$. (Dies ist die einzige Differentaloperation, die in Betracht kommt, da die Divergenz eines Viererverktors bekanntlich ein Skalar ist) und erhalten

$$\operatorname{Rot} q = M,$$

wo M ein Sechservektor ist mit den Komponenten | 72

$$(M_{yz}, M_{zx}, M_{xy}, M_{xl}, M_{yl}, M_{zl})$$
$$= \left(\frac{\partial q_y}{\partial z} - \frac{\partial q_z}{\partial y}, \frac{\partial q_z}{\partial x} - \frac{\partial q_x}{\partial z}, \frac{\partial q_x}{\partial y} - \frac{\partial q_y}{\partial x}, \frac{\partial q_x}{\partial l} - \frac{\partial q_l}{\partial x}, \frac{\partial q_y}{\partial l} - \frac{\partial_l}{\partial y}, \frac{\partial q_z}{\partial l} - \frac{\partial q_l}{\partial z}\right).$$

Nun setzen wir

$$M = (\mathfrak{M}, -i\mathfrak{E}),$$

wo $\mathfrak{M}$ der magnetische, $\mathfrak{E}$ der elektrische Vektor ist. D. h.

$$M = (\mathfrak{M}_x, \mathfrak{M}_y, \mathfrak{M}_z, -i\mathfrak{E}_x, -i\mathfrak{E}_y, -i\mathfrak{E}_z).$$

Dies ist der entscheidende Schritt, den die neue Theorie gemacht hat und der krasseste Ausdruck des Örstedtschen Versuches. Haben wir so die 6 Komponenten von $\operatorname{Rot} q$ zusammengefasst, so ist die weitere Entwicklung selbstverständlich, wenn man das für die Viereranalysis geltende Formelsystem beherrscht. Von hier aus ergibt sich die ganze Elektrodynamik fast zwangsläufig. M ist ein Sechservektor, entstanden aus $\operatorname{Rot} q$. Darum gelten für ihn die Identitäten

$$\operatorname{Div} M^* = 0,$$
$$\text{und} \qquad \operatorname{Div}\operatorname{Div} M = 0.$$

Wir wollen zunächst die erste dieser beiden Identitäten durch die Dreiergrössen $\mathfrak{M}$ und $\mathfrak{E}$ ausdrücken.

$$\mathfrak{M}^* = (-i\mathfrak{E}_x, -i\mathfrak{E}_y, -i\mathfrak{E}_z, \mathfrak{M}_x, \mathfrak{M}_y, \mathfrak{M}_z).$$

Aus

$$\operatorname{Div} F_k = f_k = \frac{\partial F_{xk}}{\partial x} + \frac{\partial F_{yk}}{\partial y} + \frac{\partial F_{zk}}{\partial z} + \frac{\partial F_{lk}}{\partial l}$$

folgt für $\operatorname{Div} M^* = 0$ 73

$$\begin{aligned} \operatorname{rot}\mathfrak{E} + \frac{\partial\mathfrak{M}}{\partial t} &= 0, & &\text{(1. Satz der Maxwellschen} \\ \operatorname{div}\mathfrak{M} &= 0. & &\text{Gleichungen)} \end{aligned}$$

Nunmehr gelangen wir zur Deutung der Identität

$$\operatorname{Div}\operatorname{Div} M = \operatorname{Div}\operatorname{Div}\operatorname{Rot} q = 0.$$

Wir sind ausgegangen von dem Potential und müssen nun einen Begriff einführen, der der Dichte entspricht. In unserer Weltsprache ist diese ein Vierervektor. Aber ein Vierervektor kann nur dann als Dichte r angesprochen werden, wenn er der Kontinuitätsgleichung

$$\operatorname{Div} r = 0$$

genügt. Ein solcher Vierervektor liegt nun gerade hier vor, denn die Divergenz eines Sechservektors M ist ein Vierervektor, und da

$$\operatorname{Div}\operatorname{Div} M = 0$$

ist, so können wir $\operatorname{Div} M$ als elektrische Dichte auffassen. Es ist zweckmässig,

$$\mathrm{M} = -r$$

zu setzen, wo r *die elektrische Viererdichte* ist.
Setzen wir

$$\operatorname{Div} M = -r$$

in die Dreiersprache um, so finden wir:

$$\begin{aligned} \operatorname{rot}\mathfrak{M} - \frac{\partial\mathfrak{E}}{\partial t} &= \rho\mathfrak{v} & &\text{(2. Satz der Maxwellschen} \\ \operatorname{div}\mathfrak{E} &= \rho & &\text{Gleichungen)} \end{aligned}$$

74 Wir sehen also: die Maxwellschen Gleichungen sind eine unmittelbare Folge unserer Definition. Sie passen überaus einfach auf unsere Weltsprache. Historisch haben sie gerade diese Vierersprache veranlasst.[62]

Wie verhält sich das zweite Paar der Maxwellschen Gleichungen, wenn man statt M $\operatorname{Rot} q$ einsetzt? Wir erhalten

$$\operatorname{Div}\operatorname{Rot} q = -r,$$

was, wie sich aus unserem Formelsystem (Seite 52) ergibt, identisch ist mit

$$-\operatorname{Div}\operatorname{Grad} q - \operatorname{Grad}\operatorname{Div} q = -r,$$

[62] For historical accounts of four-dimensional vector calculus, see the references in note 42.

und da

$$\operatorname{Div} \operatorname{Grad} q = -\Box q \quad \text{ist,}$$

folgt hieraus:

$$-\Box q - \operatorname{Grad} \operatorname{Div} q = -r$$

oder

$$\Box q + \operatorname{Grad} \operatorname{Div} q = r.$$

Diese Formel drückt die elektrische Viererdichte aus durch das elektrische Viererpotential. Ihr entspricht in der gewöhnlichen Elektrostatik

$$\Delta\varphi = -\rho.$$

Durch Integration erhalten wir das Viererpotential.

§ 24. Definition der elektromagnetischen Energie und Kraft

In der gewöhnlichen Elektrodynamik definiert man die Energie E folgendermaßen:

$$\begin{aligned} E &= \frac{1}{2}(\mathfrak{E}^2 + \mathfrak{M}^2) \\ &= \frac{1}{2}(\mathfrak{E}_x^2 + \mathfrak{E}_y^2 + \mathfrak{E}_z^2 + \mathfrak{M}_x^2 + \mathfrak{M}_y^2 + \mathfrak{M}_z^2). \end{aligned}$$

Dieser Ausdruck ist in unserer Weltsprache natürlich nicht invariant. Um wieder etwas Invariantes zu erhalten, müssen wir | zu der einen Grösse 15 ande- 75
re hinzufügen. Den so entstehenden 16er Tensor σ bezeichnen wir als die Energie[63]

$$\sigma_{sk} = -\frac{1}{2} M \cdot M \delta_{sk} - M \times M.$$

In dieser Formel bedeutet δ_{sk} den speziellen symmetrischen 16er Tensor, für den gilt:

$$\delta_{sk} = \begin{cases} 1 & \text{für } s = k, \\ 0 & \text{für } s \neq k. \end{cases}$$

Ausführlich geschrieben lautet die obige Formel

$$\sigma_{sk} = -\frac{1}{2} M \cdot M \delta_{sk} + \sum_m M_{sm} M_{km}.$$

Damit ist die Energie definiert. σ stellt sich dar als ein 16er Tensor, dessen Komponenten quadratische Funktionen in den Komponenten von $\mathfrak{E}$ und $\mathfrak{M}$ sind. Er ist ausserdem symmetrisch,

$$\sigma_{ik} = \sigma_{ki},$$

[63]Klein remarked in his excerpts (see note 2): "Wenn nur H. statt Energie wenigstens Zehnerenergie sagen würde."

so dass er nur wesentlich aus 10 Komponenten besteht. Wir wollen die einzelnen Komponenten näher betrachten. Zunächst ergibt sich

$$\sigma_{44} = \frac{1}{2}(\mathfrak{E}^2 + \mathfrak{M}^2).$$

Das ist das, was man in der gewöhnlichen Elektrodynamik Energie nennt.
Für die Grössen σ_{11}, σ_{22}, σ_{33} erhalten wir

$$\begin{aligned}
\sigma_{11} &= -\frac{1}{2}(\mathfrak{E}^2 + \mathfrak{M}^2) + \mathfrak{E}_x^2 + \mathfrak{M}_x^2, \\
\sigma_{22} &= -\frac{1}{2}(\mathfrak{E}^2 + \mathfrak{M}^2) + \mathfrak{E}_y^2 + \mathfrak{M}_y^2, \\
\sigma_{33} &= -\frac{1}{2}(\mathfrak{E}^2 + \mathfrak{M}^2) + \mathfrak{E}_z^2 + \mathfrak{M}_z^2,
\end{aligned}$$

76 ferner ist[64]

$$\begin{aligned}
\sigma_{12} =&\sigma_{21} = \mathfrak{E}_x\mathfrak{E}_y + \mathfrak{M}_x\mathfrak{M}_y, \\
\sigma_{23} =&\sigma_{32} = \mathfrak{E}_y\mathfrak{E}_z + \mathfrak{M}_y\mathfrak{M}_z, \\
\sigma_{13} =&\sigma_{31} = \mathfrak{E}_z\mathfrak{E}_x + \mathfrak{M}_z\mathfrak{M}_x.
\end{aligned}$$

Diese 9 Ausdrücke, die also quadratische Funktionen in $\mathfrak{E}$ und $\mathfrak{M}$ sind, sind die Komponenten eines symmetrischen 9er Tensors, der in der gewöhnlichen Elektrodynamik eine Rolle spielt und dort als elektromagnetischer oder Maxwellscher Spannungstensor[65] bekannt ist.
Von den Komponenten des Energietensors haben wir jetzt nur noch σ_{14}, σ_{24}, σ_{34} zu betrachten. Auch diese begegnen uns in der gewöhnlichen Elektrodynamik. Machen wir sie durch Multiplikation mit $-\frac{1}{i}$ reell, so sind sie die Komponenten von

$$\mathfrak{E} \times \mathfrak{M},$$

und diesen Vektor bezeichnet man dort als Energiefluss oder Poyntingschen Strahlungsvektor. Unser symmetrischer Welttensor σ, den wir als Energie definiert haben, enthält somit in der Sprache der gewöhnlichen Elektrodynamik gesprochen, die Energie[66], den Maxwellschen Spannungstensor und den Poyntingschen Strahlungsvektor oder Energiefluss.[67] Während also in der Dreieranalysis alles unzusammenhängend ist, haben wir hier ein System von 16 (resp. 10 verschiedenen) Grössen, das zusammenhängt und invariant ist.
Definition der Kraft
77 Wir wollen jetzt noch die Divergenz des Energieten|sors[68] berechnen. Dazu

[64] Added by Hilbert in pencil: "Es folgt hieraus noch die Invariante: $\sigma_{11}+\sigma_{22}+\sigma_{33}+\sigma_{44} = 0$."

[65] Interlineated by Hilbert in pencil: "6 verschiedene wegen der Symmetrie".

[66] "die Energie" was corrected from "den Energiesatz".

[67] On top of the words "Energie", "Maxwellschen", and Poytingschen" the numbers "1", "+6", resp., "+3" were written in pencil.

benutzen wir die Identität:

$$\operatorname*{Div}_h(-\frac{1}{2}M \cdot M\delta_{sk} + \sum_m M_{km}M_{sm}) \equiv \sum_h (M_{hs}\operatorname*{Div}_m M_{hm} - M^*\operatorname*{Div}_m M^*_{hm}).$$

Die rechte Seite dieser Identität ist, da nach den Maxwellschen Gleichungen $\operatorname{Div} M^* = 0$ ist, und wir $\operatorname{Div} M = -r$ gesetzt haben

$$= -Mr.$$

Wir erhalten also die überaus einfache Gleichung:

$$\operatorname{Div}\sigma = -Mr.$$

Nun war aber

$$r = \tilde{\rho}v, \qquad \begin{aligned} &\tilde{\rho} = \text{Ruhdichte}, \\ &v = \frac{dx_i}{d\tau}, \qquad i = 1, \cdots, 4. \end{aligned}$$

Setzen wir dieses ein und multiplizieren noch skalar mit v, so verschwindet die rechte Seite, wie aus den Formeln § 18[69] ersichtlich ist. Die Divergenz von σ hat also die Eigenschaft, skalar mit v multipliziert, 0 zu ergeben. Einen solchen Vierervektor aber brauchen wir gerade noch, und da in der Mechanik

$$v \cdot k = 0$$

galt, wo k die Viererkraft war, so definieren wir auch jetzt

$$-Mr$$

als die *elektrische Kraft* und bezeichnen sie mit f.

$$f = -Mr$$

ist die elektromagnetische oder Lorentzkraft. Die Beziehung

$$v \cdot f = 0,$$

die sich in der Mechanik unmittelbar aus der Definition von v | und f ergab, 78
ist bei dieser Definition der elektromagnetischen Kraft stets erfüllt.

[68] Added in a different hand at the top of the page:

$$\begin{aligned} \text{Wenn} \quad \sigma &= F \times F + \frac{1}{4}F \cdot F(\delta) \\ &= \frac{1}{2}\{F \times F - F^* \times F^*\} \\ \text{Dann} \quad \operatorname{Div}\sigma &= -(\operatorname{Div} F) \cdot F + (\operatorname{Div} F^*) \cdot F^* \end{aligned}$$

[69] Interlineated by Hilbert in pencil: "S. 50"; cf. p. 114 above.

Zu bemerken ist noch, dass mit der elektrischen Dichte r die Geschwindigkeit v schon gegeben ist, denn es gilt:[70] $r_i = \tilde{\rho}\frac{dx_i}{d\tau}$, $\sum(\frac{dx_i}{d\tau})^2 = -1$, d. h. $\sum r_i^2 = -\tilde{\rho}^2$ und es wird $v_i = \frac{dx_i}{d\tau} = \frac{1}{\tilde{\rho}}r_i = \frac{r_i}{\sqrt{-\sum r_i^2}} = \frac{r_i}{i\sqrt{r\cdot r}}$.
Damit sind wir mit den Definitionen fertig und können nun dazu übergehen, Naturgesetze aufzustellen. Mit Definitionen allein kann man keine Naturgesetze erklären, es müssen noch Gesetze und Axiome hinzukommen, die man dann durch das Experiment prüft. Das einzige Axiom, das wir bis jetzt aufgestellt haben, ist: die Naturgesetze sollen vom Bezugssystem unabhängig sein. Die übrigen werden wir ableiten. Diese Gesetze werden natürlich im Vergleich zu denen der alten Physik entsprechend komplizierter ausfallen. Wir geben zunächst ein Beispiel aus der Mechanik für den Unterschied zwischen einer Definition und einem Gesetz.

§ 25. Die Bewegungsgleichung eines Massenpunktes unter dem Einfluss einer Kraft; der Bornsche Starrheitsbegriff

Vorerst ist die Kraft nur definiert als Masse mal Beschleunigung. Wenn man aber noch verlangt, dass diese Kraft noch die andere Bedeutung hat, eine konstante Grösse zu sein, so erhält man eine Differentialgleichung und damit ein Naturgesetz. Man muss eben, um ein Gesetz zu erhalten, in der Defini-
79 tionsgleichung $m \cdot b_i = k_i$ der rechten Seite noch eine Bedeu|tung geben. Ein solches Gesetz ist dann an der Erfahrung prüfbar. In der gewöhnlichen Mechanik lautet die entsprechende Aufgabe: Gegeben ist ein Massenpunkt, der sich unter dem Einfluss einer konstanten Kraft bewegt. Nach welchem Gesetz erfolgt die Bewegung?
Nun gilt in der gewöhnlichen Mechanik

$$\mathfrak{k} = m \cdot b.$$

Die gesuchte Differentialgleichung lautet also

$$m\frac{d^2x}{dt^2} = \text{const}\,.$$

Das ist bekanntlich das Gesetz der Schwerkraft. Wir bekommen dieses Gesetz erst durch die Voraussetzung $\mathfrak{k} = \text{const}$, eine Voraussetzung, die wir durch das Experiment bestätigen können.
Dieses Beispiel wollen wir jetzt in der Viereranalysis durchführen. Hier ist die Kraft ein Vierervektor, der stets der Gleichung

$$v \cdot k = 0$$

genügen muss. k kann nicht konstant sein, denn wäre $k = \text{const}$ und $\neq 0$, so müsste $v = 0$ sein, was wegen $v \cdot v = -1$ nicht möglich ist. Wir beschränken

[70] In the following formulas, $\tilde{\rho}$ was corrected from $\bar{\rho}$.

uns bei der Ableitung der Bewegungsgleichung der Einfachheit halber wieder auf eine xt-Ebene, setzen also

$$v_y = v_z = 0,$$

ebenso

$$k_y = k_z = 0.$$

Die Gleichung

$$v \cdot k = 0$$

geht dann über in 80

$$v_x k_x - v_t k_t = 0.$$

Das einfachste Naturgesetz ist nun:[71]

$$\begin{aligned} k_x &= A v_t, \\ k_t &= A v_x, \end{aligned}$$

wo A die gegebene Konstante des Naturgesetzes, etwa die Schwerebeschleunigung g ist. Mit diesem Ansatz gehen wir in die Differentialgleichung

$$m \cdot b = k$$

hinein und erhalten:

$$\begin{aligned} m\frac{d^2x}{d\tau^2} &= k_x = A v_t, \\ m\frac{d^2t}{d\tau^2} &= k_t = A v_x, \end{aligned}$$

Gleichungen, die sich leicht[72] mit Hilfe der Dreiergeschwindigkeit ausdrücken lassen. Es ist

$$\begin{aligned} \mathfrak{v}_x &= \frac{dx}{dt} = \frac{dx}{d\tau}\frac{d\tau}{dt} = \frac{\frac{dx}{d\tau}}{\frac{dt}{d\tau}} \\ &= \frac{v_x}{v_t}, \end{aligned}$$

woraus folgt

$$\frac{d\mathfrak{v}_x}{d\tau} = \frac{v_t \frac{d^2x}{d\tau^2} - v_x \frac{d^2t}{d\tau^2}}{v_t^2},$$

oder wenn wir für $\frac{d^2x}{d\tau^2}$ u. $\frac{d^2t}{d\tau^2}$ bezw. $\frac{Av_t}{m}$ u. $\frac{Av_x}{m}$ einsetzen:

$$\begin{aligned} m\frac{d\mathfrak{v}_x}{d\tau} &= \frac{v_t A v_t - v_x A v_x}{v_t^2} \\ &= A\frac{v_t^2 - v_x^2}{v_t^2}. \end{aligned}$$

Nun ist aber wegen 81

$$v \cdot v = -1,$$
$$v_t^2 - v_x^2 = +1,$$

also

$$m \frac{d\mathfrak{v}_x}{d\tau} = \frac{A}{v_t^2}.$$

Diese Gleichung dividieren wir noch durch $\frac{dt}{d\tau} = \frac{1}{\sqrt{1-\mathfrak{v}^2}} = v_t$ und erhalten

$$m \frac{\frac{d\mathfrak{v}_x}{d\tau}}{\frac{dt}{d\tau}} = m \frac{d\mathfrak{v}_x}{dt} = \frac{A}{v_t^2} = \frac{A}{(1-\mathfrak{v}^2)^{\frac{3}{2}}},$$

oder schliesslich[73]

$$m \frac{d^2x}{dt^2} = A \left(1 - \left(\frac{dx}{dt}\right)^2\right)^{\frac{3}{2}},$$
$$A = m \left(1 - \left(\frac{dx}{dt}\right)\right)^{-\frac{3}{2}} \frac{d^2x}{dt^2}.$$

Dies ist die Bewegungsgleichung für einen Massenpunkt, der sich in der x, t-Ebene bewegt. Die Gleichung ist eine direkte Folge der Relativitätstheorie. Ihr entspricht in der gewöhnlichen Mechanik

$$m \frac{d^2x}{dt^2} = \text{const.}$$

Eine formelle Uebereinstimmung der beiden Gleichungen wird erzielt, wenn man

$$m \left(1 - \left(\frac{dx}{dt}\right)^2\right)^{-\frac{3}{2}} = M$$

setzt, und M die Masse nennt. Ist nun

$$\frac{dx}{dt} = \mathfrak{v}$$

82 sehr klein, so dass man es gegenüber 1 (Lichtgeschwindigkeit) vernachlässigen kann, so sind die Gleichungen tatsächlich identisch. M wird in der Grenze

$$\lim_{\frac{dx}{dt}=0} M = m,$$

[71]"ist nun" was corrected from "würde nun sein".

[72]"leicht" was corrected from "nicht".

[73]In the original, the following formula was erroneously written with A being divided rather than multiplied by the factor $(1 - (dx/dt)^2)^{(3/2)}$. The error was indicated, and in the following equations on this and the next page, the minus signs for the exponent (3/2) were interlineated.

ein Umstand, der dazu geführt hat, m als die Ruhmasse des Punktes zu bezeichnen. Für kleine Geschwindigkeiten ist also das alte Gesetz richtig, aber nicht für grosse. Für grosse Geschwindigkeiten ist die Masse der alten Mechanik keine Konstante, sondern sie hängt von der Geschwindigkeit ab, wie dies durch die Formel

$$M = m\left(1 - \left(\frac{dx}{dt}\right)^2\right)^{-\frac{3}{2}}$$

zum Ausdruck kommt.

Das so erhaltene Gesetz der Bewegung eines Massenpunktes kann man experimentell prüfen, und zwar muss man dazu Körper mit grosser Geschwindigkeit nehmen. Besonders geeignet sind also die Kathodenstrahlen (1/3 Lichtgeschwindigkeit) und die β-Strahlen (9/10 Lichtgeschwindigkeit). Das Experiment hat zu Gunsten der neuen Theorie entschieden.[74] —

Man könnte so die ganze Mechanik aufbauen. Wir wollen hiervon indessen absehen und zwar um so mehr, als man heute die ganze Mechanik nur als ein spezielles Gebiet, einen Spezialfall eines viel allgemeineren und umfassenderen Problems ansieht. Sind auch die Untersuchungen im einzelnen noch nicht vollends abgeschlossen, so kann man heute doch nicht mehr daran zweifeln, dass das Problem der Physik überhaupt mit dem der | Elektrodynamik im 83
letzten Grunde identisch ist.[75] Wir wenden uns daher gleich zu diesem allgemeineren Problem, der Elektrodynamik.

Wir werden dort den Begriff des starren Körpers brauchen und wollen daher jetzt erklären, was man in der neuen Mechanik darunter versteht. Wir erinnern uns an die allgemeine *imaginäre* Darstellung der Weltlinie durch 4 Funktionen x, y, z, l von τ. Stellen wir uns hier alles (auch l) als reell vor, so haben wir in der Eigenzeit τ das vollständige Analogon zur Bogenlänge s im dreidimensionalen Raum. Bei dieser imaginären Darstellung können wir eine beliebige Kurve, deren Neigung gegen die l-Axe kleiner als 45° ist,[76] als Weltlinie ansehen. Bei der Bewegung beschreibt jeder Punkt eines Körpers eine Weltlinie. Sind nun diese von den einzelnen Punkten des Körpers bei der Bewegung beschriebenen Weltlinien äquidistant, so nennen wir mit Born den

[74] Initially, the experimental data on the velocity dependence of the electron's mass, as discussed e.g. by Einstein in his review article *Einstein 1907b*, did not support the special theory of relativity. As admitted by Einstein, these data actually favored alternative theories by Max Abraham and Alfred H. Bucherer. The experimental data remained inconclusive for some years, see *Laub 1910* and *Laue 1911*, pp. 16–18, for contemporary reviews, and were only settled with data reported in *Guye and Lavanchy 1916*, see *CPAE2 1989*, pp. 270–272, for a discussion and further references.

[75] Klein in his excerpts (see note 2 above) remarked: "Glaube, dass die Elektrodynamik die wahre Grundlage der Physik sei." (SUB *Cod. Ms. Klein 22A*, sheet 31). For a historical discussion of Hilbert's position toward the mechanical and electromagnetic worldviews, prevalent at the time, see *Corry 1999b* and *Corry 2004*, ch. 5.

[76] The preceding half-sentence was interlineated.

Körper starr.[77] Da die Weltlinien durch 3 willkürliche Funktionen bestimmt sind, hat der so definierte starre Körper nicht 6 Freiheitsgrade. Die Definition des starren Körpers lässt sich auch folgendermassen fassen: Ein Körper heisst starr, wenn sich in bezug auf ein mitbewegtes System alle seine Punkte gleichzeitig in Ruhe befinden.[78] Hat umgekehrt die Transformation eines beliebigen Punktes eines Körpers auf Ruhe von selbst die Transformation aller Punkte auf Ruhe zur Folge, so haben wir einen starren Körper. In der Tat: Errichtet man eine neue x'-Achse senkrecht zu einem Punkt irgend einer Weltlinie, so stehen alle Weltlinien des starren Körpers senkrecht auf dieser Geraden.

84 § 26. Elektrodynamik auf Grund der atomistischen Hypothese.

Wir wollen nun das Gesetz, von dem die gegenwärtige Elektrodynamik beherrscht wird, aufstellen. Dieses gründet sich im wesentlichen auf zwei Sätze:
1) *Gesetz der Atomistik.*
Man nimmt an, die Elektrizität kommt nur in der Form von Elektronen vor, deren es nur endlich viele gibt, und die sich nur als starre Körper im Bornschen Sinne bewegen können.[79] Transformieren wir ein Elektron auf Ruhe, so hat es die Gestalt einer Kugel, und seine Ruhdichte ist konstant $= \tilde{\bar{\rho}}$. Das ist eine ungeheure Einschränkung für die Dichteverteilung r, und zwar ist diese Forderung der krasseste Ausdruck des Prinzips der Atomistik. Der Bornsche Begriff des starren Körpers ist der strengste dieser Art. Danach hat ein starrer Körper nur 3 Freiheitsgrade. Seine Bewegung ist also bestimmt, wenn x, y, z als Funktionen von t bekannt sind. Wie gross diese Einschränkung für r ist, erkennt man daraus, dass die Bewegung aller n Elektronen charakterisiert ist durch

$$\begin{aligned} x &= x_h(t), \\ y &= y_h(t), \qquad h = 1, \cdots n \\ z &= z_h(t), \end{aligned}$$

also durch $3n$ Funktionen einer einzigen Variablen, während wir für beliebiges r 4 Funktionen von 4 Veränderlichen haben. Gleichzeitig mit r ist auch das Viererpotential q eingeschränkt. So einschneidend aber dieses Weltgesetz

[77] "Born" was corrected from "Vorer". The definition of a rigid body given here is, in fact, due to *Herglotz 1910*. Klein, in his excerpts, remarked: "Born'scher Starrheitsbegriff, der nur 3 Grade d. Fr. läßt. (Bloß erzählt, ohne Bezugnahme auf Herglotz)." (SUB *Cod. Ms. Klein 22A*, sheet 29). For Born's original relativistically invariant definition of a rigid body, see *Born 1909a*. For historical discussions of the debate around the concept of a rigid body in the special theory of relativity, see *Klein 1970*, pp. 152–154, *Stachel 1980*, *Maltese and Orlando 1995*, and *CPAE3 1993*, pp. 478–480.

[78] This is the original definition put forward in *Born 1909a*, see the previous note.

[79] For a historical discussion of different electron models at the time, see *Janssen and Mecklenburg 2007*.

ist, so reicht es doch nicht hin, um als Grundlage der Physik zu dienen. Es ermöglicht uns nicht, Naturgesetze zu postulieren, da | es zur Aufstellung von 85
Gleichungen nicht ausreicht. Dies leistet der zweite Satz:

2) *Das Kraftgesetz*.

Wenn wir r haben, können wir M finden, denn die Integrationstheorie von

$$\begin{aligned} \operatorname{Div} M &= -r, \\ \operatorname{Div} M^* &= 0 \end{aligned}$$

lässt sich vollständig durchführen, da die Differentialgleichungen linear sind. Haben wir aber M und r, so können wir auch f berechnen.[80] Wir wollen nun eine Aussage in bezug auf f machen. Wir fassen den Mittelpunkt eines Elektrons ins Auge und transformieren auf Ruhe. Dann integrieren wir die Kraft über dieses auf Ruhe transformierte Elektron und setzen

$$\int f d\overline{V} = 0,$$

wo $d\overline{V} = \frac{dV}{\sqrt{1-\mathfrak{v}^2}}$ ist.[81]

Die Kraft, über das Ruhvolumen integriert, soll also gleich Null sein; diese Forderung ist ein Analogon der Forderung der gewöhnlichen Mechanik, dass die Summe aller Kräfte $= 0$ sein soll.[82]

Stimmt nun die Zahl der Gleichungen mit der der Unbekannten überein? Das ist tatsächlich der Fall. f ist nämlich Viererkraft, und für diese gilt

$$v \cdot f = 0,$$

d. h.

$$v_x f_x + v_y f_y + v_z f_z = v_t f_t,$$

|[83]eine Gleichung, die sich für $r_i = \text{const.}$[84] innerhalb $\overline{V}$ auch schreibt[85] 86

$$r_t \int f_t d\overline{V} = \sum_{i=1}^{3} r_i d\overline{V}.$$

---

[80]The quantity $f = -M \cdot r$ had been introduced in § 24 above.

[81]Added by an unknown writer: „ist eine invariante Gl system 4 Gl.!“.

[82]Added by Hilbert with pencil: "— genau, wie in der Mechanik starrer Körper".

[83]Deleted: "oder wegen $r = \overline{\rho} v$

$$r_x f_x + r_y f_y + r_z f_z = r_t f_t \quad "$$

[84]"für $r_i = \text{const.}$" was corrected from: "unter Benutzung der Relation für die Gesamtkraft $\int f d\overline{V} = v$"; added by Hilbert with pencil: "wegen $v_i = \text{const}$".

[85]In the following formulas, r was corrected by Hilbert to v with pencil, and the summation index was changed from i to h.

Ist nun
$$\int f_i d\overline{V} = 0 \quad \text{für } i = 1, 2, 3,$$
so folgt
$$v_i \int f_i d\overline{V} = 0 \quad \text{für } i = 1, 2, 3.$$
Also ist
$$v_t \int f_t d\overline{V} = 0,$$
und da $v_t \neq 0$ ist[86], wird
$$\int f_t d\overline{V} = 0,$$
d. h. die vierte Gleichung ist eine Folge der drei anderen. Somit haben wir in der Tat $3n$ Gleichungen für die $3n$ Funktionen der einen Veränderlichen t. Die Gleichungen sind jedoch ausserordentlich kompliziert; sie sind ein Gemisch von Differential-, Integral- und Funktionalgleichungen, weil die Funktionen $x_i(t)$ in M in komplizierter Weise auftreten und $f = -rM$ über das Ruhvolumen integriert wird.

87 Wir wollen nun zeigen, dass dieses Gesetz eine auffällige Analogie zur gewöhnlichen Mechanik enthält. Wir betrachten das $k$te Elektron mit dem Ruhvolumen $\overline{V}_k$. Dann gilt also
$$\int f d\overline{V}_k = 0,$$
integriert über das Ruhvolumen des Elektrons. Nun[87] setzen wir
$$r = r_1 + r_2 + \cdots + r_n,$$
wobei r_i nur im i-ten Elektron $\neq 0$ sein soll. Dabei brauchen für die gegenseitige Bewegung der einzelnen Elektronen noch keinerlei Einschränkungen zu bestehen, die Elektronen können einander vielmehr beliebig durchdringen, d. h. es können am selben Ort mehrere r_i gleichzeitig $\neq 0$ sein.[88] Durch diese Zerlegung geht unser obiges Integral
$$\int r \cdot M d\overline{V}$$
unter Berücksichtigung von
$$f_h = -r_h M$$
über in
$$\int f_i d\overline{V}_k + \cdots + \int f_n d\overline{V}_k = 0.$$

[86] "wegen $v_t = \frac{1}{\sqrt{1-\mathfrak{v}^2}}$." Added by an unknown writer.

[87] "Nun" was corrected from "Dazu".

[88] In the preceding sentence, the words "Dabei brauchen" were corrected from "Es sollen", and the words "noch", "zu", and "vielmehr", as well as the final half-sentence starting with "d. h." were interlineated.

$\int f_k d\overline{V}_k$ ist also die Kraft, die von dem $k$ten Elektron auf sich selbst ausgeübt wird. Auf das $k$te Elektron wirken somit als äussere Kräfte:

$$\sum_h \int f_h d\overline{V}_k, \qquad h \neq k,$$

und als Trägheitskraft:

$$\int f_k d\overline{V}_k.$$

Da das Elektron auf Ruhe transformiert ist, ist die Trägheitskraft gleich der Ruhmasse mal Beschleunigung. Das Gesetz

$$\int f d\overline{V} = 0$$

besagt also, dass die Summe der äusseren Kräfte + Trägheitskraft für jedes 88
Elektron zu jeder Zeit gleich Null sein soll. Dies ist eine wunderbare Beziehung zur gewöhnlichen Mechanik. Sie gilt jedoch nur angenähert und zwar für sehr kleine und sehr weit voneinander entfernte Elektronen. Die Trennung in innere und äussere Kräfte ist nur möglich durch die atomistische Hypothese. — Dass man $\int f_k d\overline{V}_k$ aus der Summe herausgreifen kann, erkannte zuerst Abraham. Aber er hatte den gewöhnlichen Starrheitsbegriff, daher fügte sich die Rechnung nicht so bequem. Sie ist zuerst von Hecke und Behrens durchgeführt worden.[89]

Wir müssen nun noch etwas über die Energie sagen. Das Energiegesetz muss eine Folge unseres Kraftgesetzes sein, das ist auch der Fall, doch gilt der Energiesatz für Elektronen nur in gewisser Annäherung.[90] Wir verschaffen uns eine Impulsenergiegleichung aus dem 16er Tensor σ, indem wir bilden

$$\begin{aligned} 0 = \int f dV &= \int \operatorname{Div} \sigma dV \\ &= \int \left(\frac{\partial \sigma_{1p}}{\partial x_1} + \frac{\partial \sigma_{2p}}{\partial x_2} + \frac{\partial \sigma_{3p}}{\partial x_3}\right) d\overline{V} + \int \frac{\partial \sigma_{4p}}{\partial x_4} dV. \end{aligned}$$

Nach dem Gauss'schen Satz bekommen wir für den ersten Ausdruck ein Oberflächenintegral. Wir nehmen an, dass auf der Oberfläche $\sigma = 0$ ist, dass also

[89]For Abraham's theory of the dynamics of the electron, see *Abraham 1902a*, *Abraham 1902b*, *Abraham 1903*, and *Abraham 1908*. For Behrens's and Hecke's work, see *Behrens and Hecke 1912*. The latter paper was presented for publication in the Göttingen Academy proceedings by Hilbert, and in the introduction the authors mention that the paper was written following Hilbert's lecture on radiation theory of summer 1912 (*Hilbert 1911/12**, this Volume). All three physicists had strong ties to Göttingen. Max Abraham (1875–1922) was *Privatdozent* there from 1900–1909. Wilhelm Behrens (1885–1917) took his Ph.D. with Felix Klein in 1909 (*Behrens 1911*) and obtained his *venia legendi* in Göttingen 1913 (*Behrens 1915*). Erich Hecke (1887–1947) was *Privatdozent* in Göttingen 1912–1915 and Professor 1918–1919.

[90]Klein in his excerpts (cf. note 2 above), remarked: "Wie steht es mit dem Energiesatz? Mangelhaft." (SUB *Cod. Ms. Klein 22a*, sheet 29).

keine Energie aus- und einströmt; dann bleibt also nur das letzte Glied übrig. Wäre nun

$$\int f dV = 0,$$

so würde folgen:

$$\frac{\partial}{\partial x_4} \int \sigma_{4p} dV = 0.$$

89 Unser Weltgesetz gilt nur für das Ruhvolumen[91] $\int f d\overline{V} = 0$ und für jedes einzelne Elektron. Ersetzen wir also das Volumen V eines Elektrons durch das zugehörige Ruhvolumen $\overline{V}$, (und $V = \overline{V}$ gilt angenähert für kleine Geschwindigkeiten), so wird

$$\int \sigma_{4p} dV = \text{const}$$

oder für $p = 4$

$$\int \{\frac{1}{2}(\mathfrak{E}^2 + \mathfrak{M}^2)\} dV = \text{const}\,.$$

Die totale Energie würde damit von der Zeit unabhängig sein. Der Energiesatz für Elektronen verlangt also, dass wir eine Vernachlässigung begehen, die darauf hinauskommt, dass wir nur kleine Geschwindigkeiten zulassen.

Damit haben wir die bis heute[92] acceptierte Elektrodynamik kurz skizziert. Ein grosser Vorzug dieser Theorie ist, dass sie keinen logischen Widerspruch enthält. Andrerseits wäre aber das Causalitätsgesetz unmöglich, da das absolut starre Elektron unendlich-grosse Geschwindigkeiten über die Länge eines Elektronendurchmessers zulässt. Es gäbe also ein Bezugssystem, in dem Ursache und Wirkung vertauscht wären. Die Starrheit erscheint zunächst als eine besonders einfache und natürliche mathematische Vorstellung, indessen ist es etwas sehr künstliches, wie wir noch erkennen werden,[93] diesen Begriff an die Spitze der Theorie[94] zu stellen.

90 Ein weiterer Nachteil dieser Theorie ist, dass der | Energiesatz nur für kleine Geschwindigkeiten gilt. Der tiefere Grund hierfür ist, dass unser Weltgesetz ein Integralgesetz ist, ein Umstand, der es so kompliziert macht, wenn auch andererseits grosse Vorteile entspringen.

Auch physikalisch sind verschiedene Einwände gegen die Theorie zu erheben. Die oben abgeleiteten Gesetze sind nicht richtig, wenn wir in das Innere der Materie eindringen. Die Gesetze der Strahlung stimmen nicht mit unserer Theorie[95], und sowohl Quantentheorie als auch Gravitation fehlen gänzlich und sind auch äusserst schwer diesem System einzuordnen. Wir müssen also

[91] "gilt nur für das Ruhvolumen" was corrected from: "lautet aber".

[92] "bis heute" was corrected from "bisher".

[93] In the following paragraph, Hilbert discusses Gustav Mie's electron theory as an alternative to the rigid electron theory.

[94] "der Theorie" was interlineated.

[95] "mit unserer Theorie" was interlineated.

aus logischen und physikalischen Gründen annehmen, dass es nicht das richtige Weltgesetz ist, und sind so gezwungen, es durch ein anderes zu ersetzen.[96]

§ 27. Die Mie'sche Theorie

In dieser Richtung sind nun in den letzten Jahren grosse Fortschritte gemacht worden, aber die Untersuchungen sind noch nicht abgeschlossen, so dass wir heute noch nicht den Erfolg derselben vollständig übersehen können. Indessen können wir die leitenden Gedanken, die diesen Forschungen zu Grunde liegen, angeben. Wir hatten definiert:

$$\begin{aligned} M &= \operatorname{Rot} q, \\ \operatorname{Div} M &= -r, \\ \sigma_{hk} &= -\frac{1}{2} M \cdot M \delta_{hk} + \sum_s M_{hs} M_{ks}, \\ \operatorname{Div} \sigma &= -Mr = f. \end{aligned}$$

Diese Definitionen sollen nun auch im folgenden bestehen bleiben, dagegen soll unser Kraftgesetz durch ein anderes ersetzt | werden. Auch die im vorigen 91
Paragraphen behandelte atomistische Hypothese wollen wir jetzt fallen lassen, wollen also diese beiden Gesetze durch ein einziges ersetzen. Natürlich wird dieses neue Gesetz von sehr einschneidendem Charakter sein müssen, wenn es der möglichen Dichteverteilung dieselben einschneidenden Beschränkungen auferlegen soll wie die Starrheit. Um nun zu diesem Gesetz zu gelangen, einnern wir uns eines Prinzips, das in der gewöhnlichen Mechanik und auch in der Elektrodynamik eine grosse Rolle spielt, nämlich des Minimalprinzips. Können wir nicht auch hier von einem Minimalprinzip ausgehen, und durch dasselbe der Viererdichte r Einschränkungen auferlegen, die den durch die Starrheit entsprechen? Diese Frage hat sich Mie vorgelegt.[97] Es wird sich also darum handeln, ein Variationsproblem aufzustellen, das dem bekannten Hamiltonschen Prinzip nachgebildet ist. Wir werden in der Tat erkennen, dass die von uns angeführten Begriffe für die Aufstellung eines solchen Prinzips geeignet sind, ja geradezu dazu herausfordern. In unserer Sprache würde das Hamiltonsche Prinzip lauten:

$$\iiiint H dx dy dz dt = \int H dW = \text{Minimum.}$$

H müssen wir dabei so einführen, dass unsere alten Definitionen darin vorkommen. Dazu wird es genügen, dass H eine Funktion von q ist, da wir alle

[96] Cp. similar remarks in Hilbert's notes on the foundations of physics, this Volume pp. 331–333.

[97] *Mie 1912a*, *Mie 1912b*, *Mie 1913*. For historical discussion of Mie's theory and Hilbert's reception of it, see *Vizgin 1994*, pp. 26–38, *Kohl 2000*, *Corry 2004*, ch. 6, *Smeenk and Martin 2007*, and the introduction to this Volume, sec. 3.

übrigen Grössen M, r u. s. w. aus q abgeleitet hatten.[98] Wir setzen also

$$H = Q - f(q \cdot q),$$

wobei

$$Q = M \cdot M = \frac{1}{2} \sum_{kh} M_{kh}^2 = \frac{1}{2} \sum_{kh} (\frac{\partial q_h}{\partial x_k} - \frac{\partial q_k}{\partial x_h})^2$$

92 ist. Dabei wählen wir[99]

$$f(q \cdot q) = (q \cdot q)^3$$

und erhalten speziell für die Hamiltonsche Funktion:

$$H = Q - (q \cdot q)^3.$$

Sie ist also zusammengesetzt aus den beiden Invarianten $\sum_{kh} M_{kh}^2$ und $(q \cdot q)^3$ und ist selbst eine orthogonale Invariante dieser Grössen. Das Hamiltonsche Prinzip lautet somit

$$\int \{Q - (q \cdot q)^3\} dW = \text{Minimum},$$

wo dW das Element eines Weltstückes ist. Wir bilden die Lagrange'schen Ableitungen und erhalten:

$$\sum_l \frac{\partial}{\partial x_l} \frac{\partial H}{\partial q_{kl}} - \frac{\partial H}{\partial q_k} = 0,$$

wobei q_{kl} zur Abkürzung gesetzt ist für

$$\frac{\partial q_k}{\partial x_l}.$$

Das sind 4 Gleichungen für die 4 unbekannten Funktionen q_i von x_1, x_2, x_3, x_4, so dass also unser System bestimmt ist. q_{kl} kommt nun nur in Q vor, also ist

$$\frac{\partial H}{\partial q_{kl}} = \frac{\partial Q}{\partial q_{kl}},$$

und dies ist wieder

$$= M_{lk};$$

denn es ist

$$M_{hk} = q_{hk} - q_{kh}.$$

Unsere Weltgleichungen gehen also über in

$$\text{Div}\, M = \frac{\partial H}{\partial q_k} = \frac{\partial f}{\partial q_k}$$

93 denn q_k kommt nur in f vor.
Nun ist aber nach Definition $\operatorname{Div} M = -r$, also haben wir

$$r_k = \frac{\partial f}{\partial q_k} = b(q \cdot q)^2 q_k \qquad k = 1, 2, 3, 4$$

Die Viererdichte ist also proportional dem Viererpotential:

$$r_1 : r_2 : r_3 : r_4 = q_1 : q_2 : q_3 : q_4$$

Dieses ist nun eine ungeheure und ganz neue Forderung, durch sie wird die Viererdichte mit dem Potential selbst in Beziehung gesetzt u. nicht mit den Ableitungen.[100] Bei der Aufstellung dieses Weltgesetzes haben wir die früheren Definitionen im wesentlichen beibehalten. Die prinzipielle Aenderung der Theorie besteht darin, dass wir an Stelle des früheren Kraftgesetzes, das ein kompliziertes Gemisch von Differential-, Integral- und Funktionalbeziehungen enthielt, ein reines[101] Differentialgesetz an die Spitze der ganzen Elektrodynamik gestellt haben, wie es eleganter und schöner nicht zu denken ist. Es ist abgeleitet aus dem Hamiltonschen Prinzip, befolgt also das Minimalgesetz. Wir haben 4 Gleichungen mit 4 unbekannten Funktionen, so dass alles bestimmt ist bis auf die willkürlichen Funktionen. Mehr aber können wir auch nicht verlangen.[102]

§ 28. Ableitung der Elektronentheorie aus der Mie'schen Hypothese

Das Bewegungsgesetz der Elektronen muss eine Folge unseres Weltgesetzes sein. Ist es nun möglich, die Elektronentheorie aus letzterem abzuleiten? Charakteristisch für die Elektronentheorie war, dass nur in einzelnen Punkten des
Rau|mes, den Elektronen, elektrische Dichte vorhanden war. Nach unserer jet- 94
zigen Auffassung werden wir einen Punkt als ein Elektron bezeichnen, wenn in ihm die Dichte sehr gross ist und ausserhalb desselben sehr schnell abfällt. Im Viererraum wird dann die Dichte auf einer Weltlinie sehr gross und nimmt auf den benachbarten rasch gegen Null ab. Wenn wir einen Grenzübergang

[98] In standard electrodynamics the Lagrangian is taken to be a function of the field variables only. A general dependence on the potential q would allow for theories violating gauge invariance which is the most serious objection against Mie's theory, see, e.g., *Pauli 1921*, pp. 754f.

[99] The following choice of f is the example discussed in *Mie 1912b*, pp. 18–38, and taken up by Hilbert in *Hilbert 1915*, p. 407, (this Volume, p. 44 and note 54 above).

[100] See note 98 above.

[101] "reines" was corrected from "reinstes".

[102] The sentence "Mehr aber können wir auch nicht verlangen." was initially placed before the preceding sentence.

machen, nämlich das Elektron als punktförmig ansehen, so muss die Weltlinie streng in eine singuläre Linie[103] übergehen, so dass die Dichte längs der ganzen Linie unendlich und ausserhalb derselben streng $= 0$ ist. Nun stellte sich Mie die Frage: Ist bei unseren Differentialgleichungen ein solches Integral möglich? d. h. gibt es ein einsames, etwa dauernd ruhendes Elektron?[104] Die Frage ist vollständig zu bejahen. Da das Elektron ruht, ist

$$v_1 = v_2 = v_3 = 0$$

d. h.

$$r_1 = r_2 = r_3 = 0$$

also

$$q_1 = q_2 = q_3 = 0$$

ferner hängt q_4 von t nicht ab. Nun nehmen wir auch noch an, dass das Elektron Kugelgestalt hat. Wir können zwar streng genommen von einer Gestalt des Elektrons überhaupt nicht sprechen, da es ja den ganzen Raum erfüllt. Es hat „Kugelgestalt", soll daher nur heissen: x, y, z sollen in q_4 nur in der Verbindung $\sqrt{x^2+y^2+z^2} = r$ vorkommen. Ist es möglich, die Differentialgleichungen durch eine Funktion zu befriedigen, die nur von r abhängt? Um das zu untersuchen, gehen wir direkt auf das Hamiltonsche Prinzip zurück. Hier lernen wir das Hamiltonsche Prinzip als Weltgesetz schätzen, denn wir
95 müssen nicht die Lagrangeschen Gleichungen auf unser spezielles Pro|blem transformieren, sondern können vielmehr die Gleichungen selber direkt aus dem Hamiltonschen Prinzip aufs neue ablesen. Dieses nimmt jetzt die einfache Gestalt an

$$\delta \iiint \left\{ \left(q_{41}^2 + q_{42}^2 + q_{43}^2\right) - q_4^6 \right\} dx_1 dx_2 dx_3 = 0$$

oder

$$\delta \iiint \left\{ \left(\frac{\partial\varphi}{\partial x}\right)^2 + \left(\frac{\partial\varphi}{\partial y}\right)^2 + \left(\frac{\partial\varphi}{\partial z}\right)^2 - \varphi^6 \right\} dx dy dz = 0,$$

wo $q_4 = \varphi(r)$ eingesetzt ist,
oder, da φ nur von r abhängt:

$$\begin{aligned} 0 &= \delta \iiint \left\{ \varphi'^2 \left[\left(\frac{\partial r}{\partial x}\right)^2 + \left(\frac{\partial r}{\partial y}\right)^2 + \left(\frac{\partial r}{\partial z}\right)^2 \right] - \varphi^6 \right\} r^2 dr \\ &= \delta \iiint \left\{ \varphi'^2 \left(\frac{x^2}{r^2} + \frac{y^2}{r^2} + \frac{z^2}{r^2} \right) - \varphi^6 \right\} r^2 dr \\ &= \delta \iiint \left\{ \varphi'^2 - \varphi^6 \right\} r^2 dr \end{aligned}$$

[103]"streng in eine singuläre Linie" was corrected from "in eine streng singuläre Linie".
[104]See the references in note 97.

Dann lautet die Lagrangesche Gleichung

$$\frac{d}{dr}(\varphi' r^2) + 3r^2\varphi^5 = 0.$$

Dies ist die Differentialgleichung, der das einsame, ruhende Elektron genügen muss. Sie ist nicht linear, darf es auch gar nicht sein, damit eine Knotenstelle im Aether möglich ist.[105] Ihr allgemeines Integral kann man mit Hilfe von elliptischen Funktionen berechnen. Wir wollen hier nur eine einparametrige Schar von Lösungen angeben und zeigen, dass sie unser Problem erfüllt, nämlich

$$\varphi = \frac{\gamma}{\sqrt{r^2+\gamma^4}}$$

wo γ einen Parameter bezeichnet. φ genügt der Differentialgleichung, denn 96

$$\begin{aligned}
3\varphi^5 &= 3\gamma^5\left(r^2+\gamma^4\right)^{-\frac{5}{2}}\\
\frac{2}{r}\varphi' &= -2\gamma\left(r^2+\gamma^4\right)^{-\frac{3}{2}}\\
\varphi'' &= -\frac{\gamma\left(\gamma^4-2r^2\right)}{\left(r^2+\gamma^4\right)^{\frac{5}{2}}}\\
&\text{-- -- -- -- -- -- -- -- -- -- -- --}\\
3\gamma^5 &- 2\gamma^5 - 2\gamma r^2 - \gamma^5 + 2\gamma r^2 = 0
\end{aligned}$$

Dies gibt in der Tat die Summe σ.
Andererseits ist φ im Endlichen überall endlich für $\gamma \neq 0$ d. h. die Dichte ist überall endlich. Für $\gamma = 0$ ist $\varphi = 0$, ausser fuer $r = 0$. Für grosses r ist

$$\varphi = \frac{\gamma}{\sqrt{r^2+\gamma^4}} = \frac{\gamma}{r}\frac{1}{\sqrt{1+\frac{\gamma^4}{r^2}}} = \frac{\gamma}{r}\left\{1 - \frac{1}{2}\frac{\gamma^4}{r^2} + \cdots\right\}$$

Für $\lim r = \infty$ ist $\varphi = \frac{\gamma}{r}$, geht also über in das gewöhnliche elektrostatische Potential. Es gibt aber auch, wie die vollständige Integration der Differentialgleichung zeigt, keine andere Lösung, die diese Bedingungen erfüllt, was besagt, dass das Elektron durch die Differentialgleichung festgelegt ist.

§ 29. Der Energiesatz in der Mie’schen Theorie

Dadurch, dass wir die Forderung: die Elektronen sollen sich nur als starre Körper bewegen, haben fallen lassen, haben wir den Vorteil, dass wir die ganze Energie, (also auch die, die den Kräften entspricht, die das Elektron zusammenhalten, d. h. bewirken, dass es sich nur als starrer Körper bewegen kann)

[105]The preceding half-sentence was interlineated.

gleichzeitig berücksichtigen können. Wir machen daher für die Totalenergie τ den Ansatz

$$\tau_{hk} = \sigma_{hk} + \rho_{hk}$$

97 wo σ_{hk} wie oben definiert ist, während ρ_{hk} ebenfalls einen symmetrischen 16er Tensor bezeichnet, der sich aus $f(q \cdot q) = (q \cdot q)^3$ wie folgt berechnet:

$$\rho_{hk} = \delta_{hk} f - \frac{\partial f}{\partial q_h} q_k$$

ρ_{hk} ist so gewählt, dass die Gesamtenergie die Form hat

$$\tau_{hk} = \delta_{hk} H + \sum_s \frac{\partial H}{\partial M_{hs}} M_{ks} + \frac{\partial H}{\partial q_h} q_k.$$

Dieser zweite Teil ρ der Energie konnte in der alten Theorie nicht berücksichtigt werden, da wir dort die Starrheit als Nebenbedingung hatten. τ_{hk} ist wieder ein symmetrischer 16er Tensor, und es gilt

$$\operatorname{Div} \tau = 0$$

Beweis: (skizziert)[106]

$$\sum_{kl} \frac{\partial H}{\partial q_{kl}} \frac{\partial^2 q_s}{\partial x_k \partial x_l} = 0 \ s = 1 \text{ weil } \frac{\partial H}{\partial q_{kl}} = -\frac{\partial H}{\partial q_{lk}} \tag{1}$$

$$\sum_k \frac{\partial}{\partial x_k} \frac{\partial H}{\partial q_k} = 0 \text{ weil } \operatorname{Div} \operatorname{Div} = 0 \tag{2}$$

$$\sum_k \frac{\partial}{\partial x_k} \frac{\partial H}{\partial q_{ks}} = \frac{\partial H}{\partial q_k} \text{ was aus } \delta \int H dw = 0 \text{ folgt} \tag{3}$$

Dann wird

$$\begin{aligned} \frac{\partial H}{\partial x_s} &= \sum_{kl} \frac{\partial H}{\partial q_{kl}} \frac{\partial (q_{ks} - q_{sk})}{\partial x_l} + \sum_k \frac{\partial H}{\partial q_k} (q_{ks} - q_{sk}) + \sum_k \frac{\partial}{\partial x_k} \left(q_s \frac{\partial H}{\partial q_k} \right) \\ &= \sum_{kl} \frac{\partial}{\partial x_k} \left\{ \frac{\partial H}{\partial q_{kl}} (q_{ks} - q_{sk}) \right\} + \sum_k \frac{\partial}{\partial x_k} \left(q_s \frac{\partial H}{\partial d q_k} \right) \end{aligned}$$

oder

$$\operatorname{Div} \left\{ \delta_{hk} H + \sum_s \frac{\partial H}{\partial M_{hs}} d M_{ks} + \frac{\partial H}{\partial q_h} q_k \right\} = 0 \qquad q.\,e.\,d.$$

98 Dadurch haben wir aber 4 neue Gleichungen gefunden, die den Lagrangeschen Gleichungen äquivalent sind. Es ist übrigens ein altes Prinzip, dass die aus dem Hamiltonschen Prinzip entspringenden Lagrangeschen Gleichungen sich stets auf die Form $\operatorname{Div}(\quad) = 0$ bringen lassen, wenn H die x_i nicht explizit

[106] The rest of the page is written in pencil.

enthält.[107] Umgekehrt ist aber dann auch ersichtlich, dass es so etwas wie Energie geben muss. Der Energiebegriff kommt eben daher, dass man die Lagrangeschen Gleichungen in Divergenzform schreibt, und das, was unter der Divergenz steht, als Energie definiert.
Aus der Bedingung

$$\mathrm{Div}\,\tau = 0$$

wollen wir nun die Folgerung ziehen. Wir betrachten, um die Vorstellung möglichst einfach zu gestalten, die Bewegung eines einzelnen Elektrons und zeichnen die Weltlinie eines in ihm ausgezeichneten Punktes (z. B. des Mittelpunktes, wenn es einen solchen gibt). Alle anderen Punkte des Elektrons beschreiben dann ihrerseits wieder Weltlinien, die jedoch nicht äquidistant zu sein brauchen, da das Elektron ja kein starrer Körper ist. In der so entstehenden dreiparametrigen Schar von Weltlinien hat eine Weltlinie die Eigenschaft, dass die Dichte auf ihr ein Maximum wird, während sie auf den benachbarten sehr schnell gegen Null abfällt. τ sei die Eigenzeit dieses ausgezeichneten Punktes.
Wir begrenzen ein Weltstück W_{12}, in dem die Weltlinie des Elektrons verläuft, nach aussen durch einen Mantel von Weltlinien, die so gewählt sind, dass auf den äussersten Weltlinien mit genügender Annäherung die Dichte als identisch Null | angesehen werden kann. Um ein endliches Weltstück zu erhalten, 99
schneiden wir diese Weltröhre durch zwei zu der ausgezeichneten Weltlinie orthogonale ebene Räume[108] V_1 und V_2 und wenden hierauf den Gauss'schen Satz an, der in unser Vierersprache lautet:[109]

$$\int \mathrm{Div}\, v dW = \int \overline{v}_u dV$$

wo u die äussere Normale bezeichnet, und erhalten

$$\iiiint \mathrm{Div}\, \rho dx_1 dx_2 dx_3 dx_4 = \int \mathrm{Div}\, \rho dW = \int (\rho)_N dG$$

wo G die ganze Begrenzung des Weltstückes bedeutet.
Auf der Mantelfläche verschwindet $(\rho)_N$ identisch, da nach unserer Voraussetzung $q = 0$ sein soll. Auf den beiden orthogonalen Räumen ist

$$(\rho)_N = \text{ bezw. } \begin{cases} +\sum_h \rho_{kh} \frac{dx_h}{d\tau} = \sum_h \rho_{kh} v_h \\ -\sum_h \rho_{kh} \frac{dx_h}{d\tau} = \sum_h \rho_{kh} v_h \end{cases}$$

Dieser Ausdruck ist aber auf den Weltlinien gleich Null, denn bei der Transformation auf Ruhe wird

$$v_1 = v_2 = v_3 = 0$$
$$v_4 = 1$$

[107] The preceding half-sentence was interlineated.

[108] Corrected by Hilbert from: "durch zwei zu den Weltlinien orthogonale Räume".

[109] From this point on, all formulas were written in a different hand.

Die Summe wird also 0 für v_1, v_2, v_3. Nimmt v den Wert $v_4 = 1$ an, so ist aber $\rho_{kh} = 0$, da es ja nicht von τ abhängt. Also haben wir rechts nur noch über V_1 und V_2 zu intergrieren. Wir erhalten

$$\int \operatorname{Div} \rho_{hk} dW = \int (\rho)_N dV = \int (\rho)_N d\overline{V}_2 - \int (\rho)_N d\overline{V}_1$$

100 Setzen wir hierin

$$\begin{aligned}(\rho)_N &= \sum_h \delta_{hk} f v_h - \frac{\partial f}{\partial q_h} q_k v_h \\ &= f v_k - 6 f v_k = -5 f v_k\end{aligned}$$

ein, (denn $q_k v_h = q_h v_k$)
so folgt:

$$\operatorname{Div} \rho_{hk} dW = -5 \left\{ \int f v_k d\overline{V}_2 - \int f v_k d\overline{V}_1 \right\}$$

Wegen

$$\operatorname{Div} \tau = 0$$

ist nun

$$\operatorname{Div} \rho = - \operatorname{Div} \sigma = rM$$

so dass wir haben

$$\int rM dW = \int (-5f) v_k d\overline{V}_2 - \int (-5f) v_k d\overline{V}_1$$

Wir haben also das Resultat:
$(-5f)v_k$ (wo v_k die Vierergeschwindigkeit bezeichnet) integriert über das Ruhvolumen ist eine Funktion von τ

$$p_k(\tau) = \int (-5f) v_k d\overline{V}$$

Bezeichnen wir noch den Mittelpunkt von V_1 mit τ_1, den von V_2 und τ_2, so lautet unsere Gleichung

$$\int_{\tau_1}^{\tau_2} rM dW = p(\tau_2) - p(\tau_1)$$

Das ist der Impulsenergiesatz. Er ist hier ganz streng und allgemein gültig.
101 Ist v wesentlich konstant, so können wir es vor | das Integral ziehen. Die Formel für p geht dann über in

$$p_k(\tau) = v_k \int (-5f) d\overline{V} = v_k \cdot m$$

$m = \int (-5f) d\overline{V}$ nennt man die elektromagnetische Ruhmasse. Unser Ausdruck geht dadurch in die Impulsgleichung der gewöhnlichen Mechanik über. Wir erhalten

$$(\tau_2 - \tau_1) rM = (m v_k)_2 - (m v_k)_1$$

oder, wenn wir durch $\tau_2 - \tau_1$ dividieren und zur Grenze $\tau_2 = \tau_1$ übergehen

$$rM = \frac{dmv}{d\tau}$$

also

$$\text{Impuls} = \frac{\text{Masse} \cdot \text{Geschwindigkeit}}{\text{Zeit}}$$

Für kleine Geschwindigkeiten $h \cdot m\tau = t$ bekommen wir

$$\text{Kraft} = \text{Masse} \cdot \text{Beschleunigung}.$$

§ 30. Vorzüge und Mängel der Mie'schen Theorie

Der Erfolg allein muss und wird zeigen, ob die Mie'sche Theorie richtig ist.[110] Ihre grossen Vorzüge gegenüber der Elektronentheorie wollen wir folgendermassen zusammenfassen.

1. Die Theorie liefert für die Gesetze des physikalischen Geschehens ein System von vier Lagrange'schen *Differentialgleichungen* für die vier unbekannten elektrodynamischen Potentiale, während die Elektronentheorie ein Gemisch von Funktional-, Differential- und Integralgleichungen liefert. In diesem Umstand, dass die Theorie Differentialgleichungen | liefert, kommt (bis auf eine unten zu 102
behandelnde, wesentliche Einschränkung) schon der Verzicht auf jedes Fernwirkungsgesetz in der Physik zum Ausdruck.

2. Da die vier elektrodynamischen Potentiale durch die Differentialgleichungen und durch Rand-(Anfangs-)Bedingungen eindeutig bestimmt sind, ist der Zustand der Welt in jedem künftigen Zeitmoment durch die Angabe der Potentialwerte für irgend einen vorhergehenden Zeitpunkt eindeutig festgelegt, m.a.W. *es gilt das Kausalitätsprinzip*.

3. Es ist in dieser Theorie eine prachtvolle Harmonie vorhanden.

Trotz dieser vielen Vorzüge können wir der Mie'schen Theorie keine Vollkommenheit zusprechen, haften ihr doch noch drei wesentliche Mängel an, die wir erst hervorheben und dann in einer abgeänderten Theorie beseitigen wollen.

1. Bei Mie hat die Lichtgeschwindigkeit noch eine ausgezeichnete, ihr nicht zukommende Bedeutung. In der Tat stellten wir im Vorhergehenden das Axiom an die Spitze: Die Naturgesetze sollen bei allen Transformationen, die die Form

$$x^2 + y^2 + z^2 - t^2$$

in sich überführen, invariant bleiben. Hierin ist aber die Lichtgeschwindigkeit konstant $= 1$ angenommen worden.

[110] See notes 97 and 98 above.

2. In dieser Theorie fehlt noch jede Andeutung von Gravitation, ohne welche wir die Physik nicht aufbauen können.[111]

3. Schliesslich müssen wir bemerken, dass wir implizite immer noch Fernwir-
103 kungsgesetze in unseren Ansätzen stecken | haben. In der Form

$$x^2 + y^2 + z^2 - t^2$$

haben nämlich auch x, y und z denselben, konstanten Koeffizienten $+1$, d. h. die Koordinatenachsen sind starre Gerade, wir haben die Axiome der Euklidischen Geometrie von vornherein zugrunde gelegt. Das dürfen wir aber nicht: denn die Geometrie ist nach unserer Ansicht eine Wissenschaft von Charakter der Physik und das Gauss'sche Experiment, zu prüfen, ob die Winkelsumme im Dreieck 180° beträgt, ein physikalisches Experiment wie jedes andere.

§ 31. Allgemeine Relativitätstheorie[112]

Bei der Aenderung und Erweiterung dieser Theorie der Materie dürfen wir also die Geometrie nicht mehr für endliche Entfernungen, sondern nurmehr im infinitesimalen benutzen. Dies gelingt uns auf Grund des *Einstein*'schen Gedankens von der allgemeinen Relativität des Geschehens. Die Mie'sche Theorie wird dann nur noch asymptotisch gelten, wenn alle Dimensionen unendlich klein sind, gerade so wie auf einer beliebigen Fläche die Gesetze der Ebene nur asymptotisch für sehr kleine Entfernungen gelten. Dann aber werden wir ein reines Differentialgesetz erhalten.

Der Weg, den wir einzuschlagen haben, ist nun ziemlich naheliegend. Wir müssen allgemeine krummlinige Koordinaten einführen, so wie man das schon auf im dreidimensionalen Raum liegenden Flächen zu tun gewohnt ist. Dann haben wir solche Beziehungen aufzusuchen, die von der Wahl des Koordina-
104 tensystems unabhängig sind. Dies sind dann allgemeine Invarianten. Aussagen aber, die bei beliebiger Koordinatentransformation nicht invariant bleiben, sind sinnlos; sie könnten ja durch eine Aenderung des Bezugssystems zerstört werden.

Wir werden also unserer Theorie wiederum eine quadratische Form mit vier Variablen zugrunde legen, die aber so verallgemeinert sein muss, dass nur im Grenzfall der infinitesimalen Entfernungen die neue Form auf die ursprüngliche Lorentz-Form reduziert werden kann. Ganz allgemein schreiben wir diese

[111] Mie discussed the problem of gravitation in the third of his trilogy on the foundations of a theory of matter, see *Mie 1913*, ch. 5, pp. 25–65.

[112] For this section, cf. *Hilbert 1915* (this Volume, pp. 28–46).

neue quadratische Form als[113]

$$\sum_{\mu=1,\nu=1}^{4} g_{\mu\nu} dx_\mu dx_\nu, \quad g_{\mu\nu} = g_{\nu\mu}$$

wobei die 10 Grössen $g_{\mu\nu}$ Funktionen von x_1 bis x_4 sein sollen. In der Tat sind dann für unendlich kleine Strecken die $g_{\mu\nu}$ Konstante und die Form lässt sich dann und nur dann nach bekannten Sätzen der Theorie quadratischer Formen auf die Gestalt[114]

$$\sum_{\nu=1}^{4} dx_\nu^2$$

bringen, so dass für diesen Grenzfall die oben entwickelte Theorie richtig bleibt.

Für endliche Entfernungen, d. h. wenn die $g_{\mu\nu}$ nicht als Konstante angesehen werden dürfen, lässt sich die Form wahrscheinlich als Summe von neun Quadraten schreiben und die Welt liesse sich als Gebilde im 9 dimensionalen Raum deuten.

Die Funktionen $g_{\mu\nu}$ nennt man die Gravitationspotentiale. Sie wurden von uns eingeführt,[115] um den unter 3 er|wähnten Mangel (Vorwegnahme der Geome- 105
trie und Fernwirkungsgesetz) zu beseitigen. Gleichzeitig sind aber auch die Mängel 1 (ausgezeichnete Rolle der Lichtgeschwindigkeit) und 2 (Fehlen der Gravitation) behoben.

Darüber können wir nur einige Andeutungen machen.

An die Spitze stellen wir jetzt das Axiom: *Die physikalischen Gesetze sollen gegenüber allen beliebigen Transformationen invariant bleiben.*

Die alten Koordinaten seien x_1, x_2, x_3, x_4. Dann gehen wir zu neuen über durch

$$\begin{aligned} x_1 &= x_1\,(x_1', x_2', x_3', x_4') \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ x_4 &= x_4\,(x_1', x_2', x_3', x_4') \end{aligned}$$

Setzen wir voraus, dass die ungestrichenen Koordinaten nach den gestrichenen differentiiert werden können (die Differentiierbarkeit wird in der Physik im allgemeinen immer vorausgesetzt) und dass die Funktionaldeterminante

$$\frac{\partial(x_1 \cdots x_4)}{\partial(x_1' \ldots x_4')} \neq 0$$

[113] In Einstein's papers on the subject of the same time, coordinate differentials were written consistently with subscript indices. In the remainder of these course notes, Hilbert reverted to a notation with superscript indices, cf. the discussion below on p. 106 (i.e. p. 156 below).

[114] The following equation should have a minus sign for the term dx_4^2.

[115] In *Hilbert 1915*, p. 395, Hilbert explicitly noted that the gravitational potential $g_{\mu\nu}$ were introduced by Einstein, cf. p. 29 and note 10 above.

ist, so können wir auflösen:

$$\begin{aligned} x_1' &= x_1'\,(x_1, x_2, x_3, x_4) \\ &\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ &\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ x_4' &= x_4'\,(x_1, x_2, x_3, x_4). \end{aligned}$$

Da die Weltgesetze invarianten Charakter haben sollen, so müssen wir nach Ausdrücken suchen, die bei beliebiger Transformation invariant bleiben. Dazu müssen wir uns die Grundbegriffe der Invariantentheorie aneignen:

106 Wir bezeichnen x^1 bis x^4 als *kontravariant*, | wenn

$$dx^\nu = \sum_{\alpha=1}^{4} \left(\frac{\partial x^\nu}{\partial x^{\alpha'}} \right) dx^{\alpha'} \quad (\alpha = 1 \cdots 4)$$

gilt und nennen weiter einen Vektor, dessen Komponenten p^1 bis p^4 sich transformieren nach der Formel

$$p^\nu = \sum_\alpha \left(\frac{\partial x^\nu}{\partial x^{\alpha'}} \right) p^{\alpha'}$$

einen *kontragredienten Vektor*. Die Auflösungen der p^ν nach den $p^{\nu\prime}$ ergibt

$$p^{\nu\prime} = \sum_\beta \left(\frac{\partial x^{\nu\prime}}{\partial x^\beta} \right) p^\beta$$

Die Kontravarianz und Kontragredienz wird durch *oben* angehängte Indizes (z. B. x^1; p^4) bezeichnet. *Unten* angehängt (z. B. q_ν) werden die Indizes beim *kogredienten* Vektor q_1 bis q_r, dessen Komponenten sich transformieren wie

$$q_{\nu'} = \sum_\alpha (\frac{\partial x^\alpha}{\partial x^{\nu'}}) q^\alpha$$

Das einfachste Beispiel eines solchen ist der Gradient $\frac{\partial \Phi}{\partial x^1} \cdots \frac{\partial \Phi}{\partial x^4}$. Als erste Invariante können wir nun das skalare Produkt eines kogredienten in einen kontragredienten Vektor bilden. In der Tat bleibt

$$p \cdot q = \sum_\nu p^\nu q_\nu = \sum_\nu p^{\nu'} q_{\nu'} = S$$

eine invariante Grösse.

Die Tensoren lassen sich nun einteilen in kogrediente $t_{\mu\nu}$, kontragrediente $t^{\mu\nu}$
107 u. gemischte t_μ^ν | je nachdem

$$\begin{aligned} \sum_{\mu\nu} t_{\mu\nu} p^\mu h^v &= y \\ \sum_{\mu\nu} t^{\mu\nu} p_\mu k_v &= y \\ \sum_{\mu\nu} t_\mu^\nu p^\mu q_\nu &= y \end{aligned}$$

eine Invariante ist.

Das totale Differential dy einer Invariante y lässt sich auffassen als skalares Produkt des kogredienten Vektors (Gradienten)

$$\left(\frac{\partial y}{\partial x^1} \cdots \frac{\partial y}{\partial x^4}\right)$$

in den kontragredienten Vektor

$$dx^1 \cdots dx^4$$

denn es gilt

$$dy = \sum_\nu \frac{\partial y}{\partial x^\nu} dx^\nu$$

Die $g_{\mu\nu}$ sind nun die Komponenten eines kogredienten, symmetrischen 16er Tensors; denn es ist

$$\sum_{\mu\nu} g_{\mu\nu} dx^\mu dx^\nu = \text{ Invariante.}$$

Die Eigenschaften der Symmetrie und Schiefsymmetrie erweisen sich – wie die Rechnung zeigt, als invariante. Auch der spezielle gemischte Tensor δ_μ^ν wo $\delta_\mu^\nu = \begin{cases} 0 & \text{für} \quad \mu \neq \nu \\ 1 & \text{für} \quad \mu = \nu \end{cases}$ ist, bleibt invariant. Diese drei Eigenschaften reduzieren die Zahl der Tensorkomponenten auf bez. 10, 6 und 4.

Wir gehen nun an die Modifikation der Mie'schen | Theorie. Wir hatten als 108
Grundgesetz der Physik

$$\delta \int H dw = 0 \qquad dw = dx^1 dx^2 dx^3 dx^4.$$

Hierbei muss H eine allgemeine Invariante sein, die freilich nun ausser den elektrodynamischen Potentialen auch noch die $g_{\mu\nu}$ und ihre Ableitungen zweiter Ordnung enthalten muss. Dass auch diese Ableitungen der $g_{\mu\nu}$ in der Weltfunktion H vorkommen müssen, lehrt die Invariantentheorie. Dort wird nämlich gezeigt, dass es eine Invariante der ersten Ableitungen der $g_{\mu\nu}$ nicht gibt. Wir geben also als logische Erweiterung der Mie'schen Theorie der Hamiltonschen Funktion H die Gestalt

$$H = K + Q - t^3$$

Hierin soll K nur die $g_{\mu\nu}$ und ihre Ableitungen enthalten, während Q und t^3 den früheren Ausdrücken Q und $(qq)^3$ entsprechen, aber nun sämtlich allgemeine Invarianten sein sollen. Dies ist wohl die einfachste Aenderung der Mie'schen Theorie. Wir wollen noch die merkwürdige Tatsache verzeichnen, dass es nur *eine einzige* Invariante gibt, die nur die zweiten Ableitungen der

$g_{\mu\nu}$ enthält. Dies ist die berühmte Riemannsche Invariante, die Krümmung der 4-dimensionalen Welt.[116]

Statt des kogredienten Tensors $g_{\mu\nu}$ führen wir den *kontragredienten Tensor* $g^{\mu\nu}$ ein, den wir erhalten, wenn wir die durch die Determinante $|g_{\mu\nu}|$ dividierten Unterdeterminanten der $g_{\mu\nu}$ nehmen. Dann wird also

$$K = \quad \text{Ausdruck in} \quad g^{\mu\nu}, g_k^{\mu\nu}, g_{kl}^{\mu\nu}$$

wo die unteren Indizes die bez. 1 und 2-malige Differentiation nach x^k bez.
109 x^k; x^l bezeichnen.

Eine weitere merkwürdige Tatsache ist nun, dass der 6er Vektor M der früheren Theorie eine allgemeine, nicht nur eine orthogonale Invariante ist. Wir wollen die elektrodynamischen Potentiale q als Komponenten eines *kogredienten* Vektors auffassen. Dann wird

$$M_{\mu\nu} = \operatorname{Rot} q$$

ein kogredienter 6er Vektor. Ferner wird

$$M^{\mu\nu} = \sum_{\mu\nu} g^{\mu m} g^{\nu n} M_{\mu\nu}$$

ein kontragedienter 6er Vektor (der Vektor M^* der alten Theorie). Die sinngemäss erweiterte Definition für Q und t ist

$$Q = \sum_{k,l,m,n} M_{mn} M_{lk} g^{mk} g^{nl}$$

$$t = \sum_{kl} q_k q_l g^{kl}$$

Da $dw = dx^1 \dots dx^4$ keine Invariante ist, wohl aber $\sqrt{g}dw$, wobei $g = |g_{\mu\nu}|$ ist, so erhalten wir die Gleichungen des physikalischen Geschehens, wenn wir

$$\int H \sqrt{g} dw$$

zu einem Minimum machen, und zwar ergeben sich die 10 Lagrangeschen Differentialgleichungen der Gravitation durch Variation nach $g_{\mu\nu}$ zu

$$\sum_{kl} \frac{\partial^2}{\partial x^k \partial x^l} \frac{\partial \sqrt{g} H}{\partial g_{kl}^{\mu\nu}} - \sum_k \frac{\partial}{\partial x^k} \frac{\partial \sqrt{g} H}{\partial g_k^{\mu\nu}} + \frac{\partial \sqrt{g} H}{\partial g^{\mu\nu}} = 0$$

und die 4 Differentialgleichungen der Elektrodynamik (Maxwellsche Gleichungen) zu

$$\sum_k \frac{\partial}{\partial x^k} \frac{\partial \sqrt{g} H}{\partial q_{hk}} - \frac{\partial \sqrt{g} H}{\partial q_h} = 0$$

110 Dies sind 14 Gleichungen für die 14 unbekannten Funktionen $g^{\mu\nu}$ und q_h ($\mu, \nu, h = 1 \ldots 4$). Das Kausalitätsprinzip kann erfüllt sein, oder nicht (Die Theorie hat diesen Punkt noch nicht aufgeklärt). Jedenfalls lässt sich auf die Gültigkeit dieses Prinzips nicht wie im Falle der Mie'schen Theorie durch einfache Ueberlegungen schliessen. Von diesen 14 Gleichungen sind nämlich 4 (z. B. die 4 Maxwellschen) eine Folge der 10 übrigen (z. B. der Gravitationsgleichungen). Es gilt nämlich der merkwürdige Satz, dass die Zahl der aus dem Hamiltonschen Prinzip fliessenden Gleichungen immer mit der Zahl der unbekannten Funktionen übereinstimmt, ausser in dem hier eintretenden Fall, dass unter dem Integral eine allgemeine[117] Invariante steht.[118]

Zum Schluss machen wir noch ein paar Bemerkungen über die Energie. Sei

$$L = Q - t^3$$

Lassen wir nun das Glied K in H bei Seite, so haben wir als Energietensor

$$-L\delta_k^l + \sum_t \frac{\partial L}{\partial M_{lt}} M_{kt} + \frac{\partial L}{\partial q_l} q_k$$

Die Gravitationsgleichungen heissen in diesem Fall

$$\frac{\partial \sqrt{g} L}{\partial g^{tk}} = 0.$$

Führen wir die Differentiation aus, so erhalten wir obigen Ausdruck, der genau die Form der Energie in der Mie'schen Theorie hat. Wir sagen also:

Lässt man K unberücksichtigt, so liefert die | Hamiltonsche Funktion, nach 111
g^{tk} differentiiert, den Energietensor der Mie'schen Theorie.

Diese Theorie muss nun fortentwickelt werden. Wir haben wieder zunächst die einfachsten Lösungen dieser Gleichungen aufzustellen. Das Mie'sche ruhende Elektron ist eine solche, dazu muss man nur

$$g^{\mu\nu} \quad = \quad \begin{matrix} 0 \\ 1 \end{matrix} \quad \text{für} \quad \begin{matrix} \mu \neq \nu \\ \mu = \nu \end{matrix} \quad \text{setzen.}$$

Dann erhalten wir die ursprünglichen Gleichungen wieder. Bisher wurde nur das Gravitationsfeld eines Massenpunktes nach dieser Theorie berechnet und zwar zuerst genähert von Einstein und dann streng von dem zu früh verstorbenen Schwarzschild.[119]

[116] This uniqueness theorem for the Riemann curvature scalar holds only if one also postulates that the invariant contains second derivatives of the metric only linearly. The theorem was first proven by Hermann Vermeil, at the instigation of Felix Klein, in *Vermeil 1917*. See *Rowe 2001*, pp. 416–418, and *Sauer 2005*, pp. 587–588, for a historical discussion.

[117] "allgemeine" was interlineated.

[118] For further discussion of this point, see *Hilbert 1915*, p. 397, (this Volume, p. 30 and note 19) and the discussion in the Introduction to this Volume, sec. 3.

[119] See *Einstein 1915c*, *Schwarzschild 1916a*, and *Schwarzschild 1916b*. Karl Schwarzschild (1873–1916) was professor and director of the observatory in Göttingen from

1901–1909 before being appointed director of the astrophysical observatory in Potsdam. *Schwarzschild 1916a* was presented to the Prussian Academy on 16 January 1916, and issued on 10 February; *Schwarzschild 1916b* was presented on 24 February, and issued on 4 April. Schwarzschild died on 11 May 1916, at the age of 43, of pemphigus, a metabolic disease of the skin, that he contracted while serving at the Russian front (see *Blumenthal 1918*).

Description of the Text

Collection: Georg-August-Universität Göttingen, Mathematisches Institut, Inv. Nr. 16206 i.

Size: Cover size 22.8 cm ×28.3 cm; page size approx. 22.2 cm ×27.6 cm.

Cover Annotations: On the spine, in gold lettering, is the notation, 'Hilbert, // Grundl. // der // Physik // 1916'.

Composition: 6 signatures of 8 to 26 double pages each; in all 116 sheets, inclusive of front- and end-papers.

Pagination: The title page is not numbered. The following pages are continuously numbered from 1 to 111. The table of contents follows page 111. Its first page is not paginated, its second page is paginated with Roman numeral II.

Original Title: On the title page: 'DIE GRUNDLAGEN DER PHYSIK // Vorlesung // von // D. HILBERT // Göttingen // Sommersemester 1916'.

Text: Typewritten text with equations and occasional additions and emendations in ink. The handwritten material on pp. 1-98 was entered by the person who arranged the script, in all probability R. Bär, since it is the same handwriting as for the handwritten material in *Hilbert 1916/17**, cf. its title page. Handwritten equations and corrections on pp. 98–111 are in ink in a different hand (possibly written by Erich Hecke). Throughout the typescript additional corrections were occasionally made in pencil, some are in Hilbert's hand. If a correction was not made in ink in the hand of the composer of the manuscript, and, in particular, if the handwriting is obviously Hilbert's, these features are pointed out in the annotation.

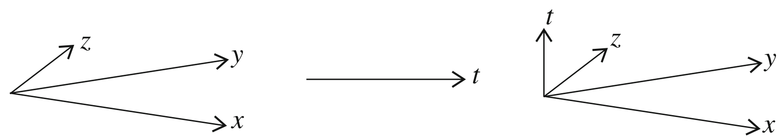
'Die Grundlagen der Physik II'

Inhaltsverzeichnis

Erster Abschnitt

§ 1. Einleitung

Alle physikalischen Vorgänge spielen sich in Raum und Zeit ab, daher müssen wir auch Raum und Zeit als die fundamentalsten Begriffe der Physik ansprechen. Schon seit langem haben die Philosophen ihre Zusammengehörigkeit erkannt, aber erst der modernsten Physik blieb die Erkenntnis vorbehalten, dass Raum ohne Zeit und Zeit ohne Raum gar nicht in Wirklichkeit vorhanden[1] ist. Dies findet darin seinen Ausdruck, dass wir zu den drei Raumdimensionen die eine Zeitdimension addieren und nun von einer vierdimensionalen Mannigfaltigkeit, der Welt des physikalischen Geschehens, sprechen.



In der Mathematik gibt es eine Lehre der vier, ja allgemein der n dimensionalen Mannigfaltigkeiten und zwar die Geometrie des n dimensionalen Raumes. Früher übernahm die Physik die Lehren der Geometrie ohne weiteres. Dies war berechtigt, solange nicht nur die groben, sondern auch die feinsten physikalischen Tatsachen die Lehren der Geometrie bestätigten. Dies war noch der Fall, als Gauss die Winkelsumme im Dreieck experimentell mass und fand, dass sie zwei Rechte beträgt. Dies gilt aber nicht mehr von der neuesten Physik. *Die heutige Physik muss vielmehr die Geometrie mit in den Bereich ihrer Untersuchungen ziehen*. Das ist logisch und natur|gemäss: jede Wissenschaft 2
wächst wie ein Baum, nicht nur die Zweige greifen weiter aus, sondern auch die Wurzeln dringen tiefer.

Vor einigen Jahrzehnten konnte man in der Mathematik eine analoge Entwicklung verfolgen; einen Satz hielt man damals nach Weierstrass dann für bewiesen, wenn er auf Beziehungen zwischen ganzen Zahlen zurückführbar war, deren Gesetze man als gegeben hinnahm. Sich mit diesen zu beschäftigen, wurde abgelehnt und den Philosophen überlassen. Kronecker sagte einmal: „Die ganzen Zahlen hat der liebe Gott geschaffen“.[2] Diese waren damals noch ein noli me tangere der Mathematik. Das ging so fort, bis die logischen Fundamente dieser Wissenschaft selbst zu wanken begannnen. Nun wurden die ganzen Zahlen eines der fruchtbarsten Arbeitsfelder der Mathematik und

[1]“in Wirklichkeit vorhanden” was corrected by Hilbert in pencil from “denkbar”.

[2]In his obituary of Leopold Kronecker (1823–1891), Heinrich Weber wrote: “Manche von Ihnen werden sich des Ausspruchs erinnern, den er in einem Vortrag bei der Berliner Naturforscher-Versammlung im Jahre 1886 that: „Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk“.” *Weber 1893*, p. 19.

speziell der Mengenlehre (Dedekind). Der Mathematiker wurde also gezwungen, Philosoph zu werden, weil er sonst aufhörte, Mathematiker zu sein.

§ 2. Geometrie und Physik; Axiome

So ist es auch jetzt wieder: *der Physiker muss Geometer werden*, weil er sonst Gefahr läuft, aufzuhören, Physiker zu sein und umgekehrt. Die Trennung der Wissenschaften in Fächer und Fakultäten ist eben etwas Anthropologisches, und der Wirklichkeit Fremdes; denn eine Naturerscheinung frägt nicht danach, ob sie es mit einem Physiker oder mit einem Mathematiker zu tun hat. Aus diesem Grunde dürfen wir die Axiome der Geometrie nicht übernehmen. Darin könnten ja Erfahrungen zum Ausdruck kommen, die den ferneren Experimen-
3 ten widersprä|chen. Wenn wir die Elemente der Geometrie danach durchsehen, was wir von ihnen brauchen können, so werden wir zwar prinzipiell vier Dimensionen nötig haben; wir beschränken uns zuerst aber der Einfachheit halber und auch aus pädagogischen Gründen auf nur zwei Dimensionen, *wir haben uns also mit der zweidimensionalen Geometrie zu befassen.*

Diese lässt sich sehr anschaulich auf der Fläche im dreidimensionalen Raume deuten. Daher müssen wir als mathematisch wichtigstes Hilfsmittel die Prinzipien der

I. Flächentheorie

studieren. Man lese das Buch: Hilbert, Grundlagen der Geometrie, 4. Auflage[3] und man wird gewahr werden, wie mannigfaltig die Geometrie begründet werden kann. Im Anhang 4 wird dort die Stetigkeit zur Begründung der Geometrie an die Spitze gestellt.[4] Um die Bedürfnisse und Interessen der Physik wahrzunehmen, werden auch wir, soweit es irgend wünschenswert ist, die *Stetigkeit und Differentiierbarkeit* voraussetzen. Die Atomistik und Quantentheorie in der Physik wird uns davon nicht abschrecken, die *Stetigkeitsaxiome* speziell von Raum und Zeit vorauszusetzen, dies sind fundamentale Forderungen der Physik. Wir akzeptieren auch die in diesem Buche gegebene *Definition der Ebene*: diese ist gegeben als Mannigfaltigkeit von Dingen (Punkten), die wir durch Zahlenpaare charakterisieren (zweidimensionale Zahlenmannigfaltigkeit). Ebenso übernehmen wir die *Definition der Deckung*[5]: die Deckung[6] ist eine Decktransformation[7] dieser Zahlenpaare in sich.

$$x' = f(x, y), \qquad y' = h(x, y).$$

4 Diese Gleichungen seien auflösbar

[3] *Hilbert 1913c*.
[4] *Hilbert 1913c*, Anhang 4.
[5] "Deckung" was corrected by Hilbert in pencil from "Bewegung".
[6] "Deckung" was corrected by Hilbert in pencil from "Bewegung".
[7] "Decktransformation" was corrected by Hilbert in pencil from "Transformation".

$$x = f'(x'y'), \qquad y = h'(x'y'),$$

d. h. es muss

$$\begin{vmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \end{vmatrix} \neq 0$$

sein oder

$$\frac{\partial(x'y')}{\partial(x\ y)} = \frac{1}{\frac{\partial(x\ y)}{\partial(x'y')}} \neq 0.$$

Gehen wir die in dem erwähnten Buch gegebene Begründung der Geometrie weiter durch, so stossen wir auf das *erste Axiom*: die Decktransformationen[8] bilden eine Gruppe, d. h. zwei Transformationen nacheinander angewendet, geben wieder eine Transformation. Dies läuft bei den anderen Begründungen der Geometrie auf die bekannten Kongruenzsätze hinaus, das Axiom ist also ein sehr einschneidendes. Für uns ist es sogar eine viel zu tief eingreifende Massnahme, wir behaupten vielmehr, dieses Axiom (d.h. die Kongruenzsätze) gelten in der Wirklichkeit nicht! Hätten wir Kongruenzsätze, so würden wir mit einem Vorurteil an die Natur herantreten. Wir lehnen sie also ab, d. h. die sogenannten Euklidischen oder Nichteuklidischen Geometrien sind für unsere Zwecke nicht allgemein genug. *Diese Kongruenzsätze sind nämlich etwas ganz Unphysikalisches*. Sie sind Fernbeziehungen, setzen also Fernwirkungsgesetze voraus; denn sie lassen sich nur durch starre Gebilde realisieren, während wir *Stetigkeit* und *Differentialbeziehungen* haben wollen. Wir können nicht annehmen, dass es in der Natur starre x und y Achsen gäbe. Denken wir uns die zweidimensionale Mannigfaltigkeit als Fläche im dreidimensionalen Raum, so gelten | die Kongruenzsätze nur dann, wenn die Fläche überall dasselbe 5
konstante Krümmungsmass hat. Hieraus sieht man wohl, wie einschneidend die durch das oben erwähnte Axiom gestellte Forderung für die Physik wäre.

§ 3. Axiom von der Existenz einer Länge

Was für Axiome werden wir nun brauchen? Auf unendlich viele Weisen können wir in unserer zweidimenisonalen Mannigfaltigkeit eine eindimensionale konstruieren. Diese nennen wir *Kurve*, gegeben durch

$$x = x(p), \qquad y = y(p)$$

(x und y sind hier also auch als Symbole für Funktionen gebraucht). Jeder Kurve und zwar zwei beliebigen Punkten auf der Kurve ordnen wir nun den Begriff der

Kurvenlänge

[8]“Decktransformationen” was corrected by Hilbert in pencil from “Bewegungen”.

zu, definiert durch

$$\begin{aligned}
s &= \int_{p_1}^{p_2} \sqrt{g_{11}\left(\frac{dx_1}{dp}\right)^2 + 2g_{12}\frac{dx_1}{dp}\frac{dx_2}{dp} + g_{22}\left(\frac{dx_2}{dp}\right)^2}\,dp \\
&= \int_{p_1}^{p_2} \sqrt{\sum_{\mu\nu} g_{\mu\nu}\frac{dx_\mu}{dp}\frac{dx_\nu}{dp}}\,dp \qquad (1) \\
&= \int_1^2 \sqrt{g_{11}dx_1^2 + 2g_{12}dx_1dx_2 + g_{22}dx_2^2} \\
&= \int_1^2 \sqrt{\sum_{\mu\nu} g_{\mu\nu}dx_\mu dx_\nu}; \qquad \begin{matrix} g_{\mu\nu} = g_{\nu\mu} \\ \mu, \nu = 1, 2 \end{matrix}
\end{aligned}$$

wobei $g_{\mu\nu}$ gegebene Funktionen von x und y sind. Dabei ist statt x x_1 und statt y x_2 geschrieben. Auf der zweidimensionalen Mannigfaltigkeit liegen also drei Funktionen g_{11}, g_{12}, g_{22} ausgebreitet. In der Euklidischen Geometrie ist

$$s = \int_{p_1}^{p_2} \sqrt{\left(\frac{dx_1}{dp}\right)^2 + \left(\frac{dx_2}{dp}\right)^2}\,dp,$$

d. h.

$$g_{11} = 1, \qquad g_{12} = 0, \qquad g_{22} = 1.$$

Jedem Punkt der Mannigfaltigkeit kommen also drei Koeffizienten einer quadratischen, homogenen Differentialform zu.[9] Diese Koeffizienten sind aber
6 noch so allgemein, dass | weder die Euklidische noch die Nichteuklidische Geometrie gelten kann. Wir setzen also das

Axiom von der Existenz einer Länge

voraus: es soll drei Funktionen g_{11}, g_{12}, g_{22} geben, die die *Länge* einer eindimensionalen Mannigfaltigkeit durch das Integral (1) definieren. Um in einer zweidimensionalen Mannigfaltigkeit Geometrie treiben zu können, müssen diese drei Funktionen $g_{\mu\nu}$ gegeben sein. Eine spezielle Form des Linienelements ist dann notwendig, dass die Kongruenzsätze gelten, z. B. $g_{11} = 1$, $g_{12} = 0$, $g_{22} = 1$. Das Quadrat des Linienelementes

$$ds^2 = g_{11}dx_1^2 + 2g_{12}dx_1dx_2 + g_{22}dx_2^2$$

ist[10] eine homogene, quadratische Differentialform.

Wir fragen uns: *sind diese $g_{\mu\nu}$ völlig willkürlich?* Sie sind natürlich stetig differentiierbar; weiter stellen wir vor der Hand die Forderung, die wir später in der Physik fallen lassen müssen: jede Kurve hat eine *reelle* Länge. Dazu muss die quadratische Form *positiv definit* sein, sie darf also nicht verschwinden,

[9] Added by Hilbert in pencil: “die sich also zuerst als die einfachsten u. wichtigsten Potentiale, P, einstellen”.

[10] “ist” was corrected from “ist dann”.

d. h. es darf keine Kurve von der Länge Null auf der Fläche vorkommen. Dies sind lauter Einschränkungen für unsere $g_{\mu\nu}$. Um diesen Forderungen gerecht zu werden, ist notwendig und hinreichend, daß[11]

$$g_{11}g_{22} - g_{12}^2 > 0, \qquad g_{11} > 0 \qquad \text{wird.}$$

Dies sind aber nur Ungleichungen. Die $g_{\mu\nu}$ sind also im wesentlichen doch noch willkürlich, während die Kongruenzsätze den $g_{\mu\nu}$ sehr einschneidende Beschränkungen auferlegten.

§ 4. Zusammenhang mit der Flächentheorie

Nun untersuchen wir den *Zusammenhang mit der Flä|chentheorie*: 7
Gegeben sei eine Fläche im dreidimensionalen Raum durch

$$x = x(x_1, x_2), \qquad y = y(x_1, x_2), \qquad z = z(x_1, x_2).$$

Durch Elimination erhalten wir die Flächengleichung auch in der Form $f(x, y, z) = 0$. Die Länge der Kurve auf dieser Fläche wird

$$\int \sqrt{dx^2 + dy^2 + dz^2},$$

wobei

$$\begin{aligned}
dx &= \frac{\partial x}{\partial x_1} dx_1 + \frac{\partial x}{\partial x_2} dx_2, \\
dy &= \frac{\partial y}{\partial x_1} dx_1 + \frac{\partial y}{\partial x_2} dx_2, \\
dz &= \frac{\partial z}{\partial x_1} dx_1 + \frac{\partial z}{\partial x_2} dx_2, \\
ds^2 &= g_{11} dx_1^2 + 2g_{12} dx_1 dx_2 + g_{22} dx_2^2.
\end{aligned}$$

Hierin bedeutet

$$\begin{aligned}
g_{11} &= \left(\frac{\partial x}{\partial x_1}\right)^2 + \left(\frac{\partial y}{\partial x_1}\right)^2 + \left(\frac{\partial z}{\partial x_1}\right)^2, \\
g_{12} &= \frac{\partial x}{\partial x_1}\frac{\partial x}{\partial x_2} + \frac{\partial y}{\partial x_1}\frac{\partial y}{\partial x_2} + \frac{\partial z}{\partial x_1}\frac{\partial z}{\partial x_2}, \\
g_{22} &= \left(\frac{\partial x}{\partial x_2}\right)^2 + \left(\frac{\partial y}{\partial x_2}\right)^2 + \left(\frac{\partial z}{\partial x_1}\right)^2.
\end{aligned}$$

Auf der Fläche ist die von uns definierte Länge die Kurvenlänge im Sinne der Euklidischen Geometrie. Auf dieser Fläche sind dann auch die obigen

[11] Added by Hilbert in the page margin with pencil: "Zudem werden wir später untersuchen, was es bedeutet, wenn diese Ungl. nicht statthaben".

Ungleichungen erfüllt und es gibt keine Nulllinien. Ist die Fläche z. B. eine *Kugel*, so wird

$$\begin{aligned} x &= \cos\vartheta, & dx &= -\sin\vartheta d\vartheta, \\ y &= \sin\vartheta\cos\varphi, & dy &= \cos\vartheta\cos\varphi d\vartheta - \sin\vartheta\sin\varphi d\varphi, \\ z &= \sin\vartheta\sin\varphi, & dz &= \cos\vartheta\sin\varphi d\vartheta + \sin\vartheta\cos\varphi d\varphi. \end{aligned}$$

Das Quadrat des Linienelementes wird

$$ds^2 = d\vartheta^2 + \sin^2\vartheta d\varphi^2.$$

Dies ist die Riemann-Helmholtzsche (elliptische) Geomeotrie, in welcher die Kongruenzsätze natürlich gelten. Es ist

$$g_{11} = 1, \qquad g_{12} = 0, \qquad g_{22} = \sin^2\vartheta.$$

8 § 5. Die wichtigsten Begriffe der Invariantentheorie

Ein weiteres wichtiges technisches Hilfsmittel ist die Transformations- oder

II. Invariantentheorie.

Es sei

$$x_1' = f_1(x_1, x_2), \qquad x_2' = f_2(x_1, x_2)$$

oder

$$x_1 = f_1'(x_1', x_2'), \qquad x_2 = f_2'(x_1', x_2').$$

Die Transformation spielt eine fundamentale Rolle in der Physik; denn *die Koordinaten sind nichts der Natur Eigentümliches, sie sind vielmehr die Namen, die wir den Dingen geben.*[12] Daher werden wir uns davon frei machen müssen. *Die Naturgesetze müssen eben ungeändert bleiben, wenn wir diese Benennungen der Dinge ändern.* Durch diese Ueberlegung werden wir auf das Hilfsmittel der Transformationen geführt. Wir fragen: wie drückt sich die Kurvenlänge in dem neuen System aus? Wir transformieren

$$\begin{aligned} ds^2 = g_{11} & \left(\frac{\partial x_1}{\partial x_1'} dx_1' + \frac{\partial x_1}{\partial x_2'} dx_2' \right)^2 \\ & + 2g_{12} \left(\frac{\partial x_1}{\partial x_1'} dx_1' + \frac{\partial x_1}{\partial x_2'} dx_2' \right) \left(\frac{\partial x_2}{\partial x_1'} dx_1' + \frac{\partial x_2}{\partial x_2'} dx_2' \right) \\ & + g_{22} \left(\frac{\partial x_2}{\partial x_1'} dx_1' + \frac{\partial x_2}{\partial x_2'} dx_2' \right)^2 \\ = \quad & g_{11}' dx_1'^2 + 2g_{12}' dx_1' dx_2' + g_{22}' dx_2'^2, \end{aligned}$$

[12] Added by Hilbert in the left margin in pencil: "Vielmehr Prinzip der Objektivität".

wobei also

$$
\begin{aligned}
g'_{11} &= g_{11}\left(\frac{\partial x_1}{\partial x'_1}\right)^2 + 2g_{12}\frac{\partial x_1}{\partial x'_1}\frac{\partial x_2}{\partial x'_1} + g_{22}\left(\frac{\partial x_2}{\partial x'_1}\right)^2,\\
g'_{12} &= g_{11}\frac{\partial x_1}{\partial x'_1}\frac{\partial x_1}{\partial x'_2} + g_{12}\left(\frac{\partial x_1}{\partial x'_1}\frac{\partial x_2}{\partial x'_2} + \frac{\partial x_1}{\partial x'_2}\frac{\partial x_2}{\partial x'_1}\right) + g_{22}\frac{\partial x_2}{\partial x'_1}\frac{\partial x_2}{\partial x'_2},\\
g'_{22} &= g_{11}\left(\frac{\partial x_1}{\partial x'_2}\right)^2 + 2g_{12}\frac{\partial x_1}{\partial x'_2}\frac{\partial x_2}{\partial x'_2} + g_{22}\left(\frac{\partial x_2}{\partial x'_2}\right)^2.
\end{aligned}
$$

Wir erhalten dann

$$s = \int_1^2 \sqrt{g'_{11}d{x_1}'^2 + 2g'_{12}dx'_1dx'_2 + g_{22}d{x_2}'^2}.^{\langle 13\rangle}$$

Wir können auch schreiben

$$g'_{\mu\nu} = \sum_{\alpha\beta} g_{\alpha\beta}\frac{\partial x_\alpha}{\partial x'_\mu}\frac{\partial x_\beta}{\partial x'_\nu}.$$

Weiter setzen wir

$$g = \begin{vmatrix} g_{11} & g_{12}\\ g_{21} & g_{22}\end{vmatrix}, \qquad g^{11} = \frac{g_{22}}{g}, \quad g^{12} = -\frac{g_{21}}{g}, \quad g^{22} = \frac{g_{11}}{g}.$$

Diese Ausdrücke werden wir später nötig haben. 9

Die Koeffizienten $g'_{\mu\nu}$ der transformierten quadratischen Form haben wir so definiert, dass

$$\sum_{\mu\nu} g'_{\mu\nu}dx'_\mu dx'_\nu = \sum_{\mu\nu} g_{\mu\nu}dx_\mu dx_\nu$$

wird. Wir sagen daher, die quadratische Differentialform $\sum_{\mu\nu} g_{\mu\nu}dx_\mu dx_\nu$ ist eine *Kovariante* der Transformation. Das Koeffizientensystem $g_{\mu\nu}$ derselben nennt man deswegen einen *kovarianten Tensor*.[14] Die Transformationsformeln eines solchen sind also

$$g'_{\mu\nu} = \sum_{\alpha\beta}\frac{\partial x_\alpha}{\partial x'_\mu}\frac{\partial x_\beta}{\partial x'_\nu}g_{\alpha\beta}.$$

Die eben definierten Grössen $g^{\mu\nu}$, die sich gemäss[15]

$${g'}^{\mu\nu} = \sum_{\alpha\beta}\frac{\partial x'_\mu}{\partial x_\mu}\frac{\partial x'_\beta}{\partial x_\nu}g^{\alpha\beta}$$

transformieren, nennen wir einen *kontravarianten Tensor*. Die Kovarianz und Kontravarianz deuten wir dadurch an, dass wir die Indizes unten bezw. oben

[13] “g_{22}” should be “g'_{22}”.

[14] Added by Hilbert in the left margin in pencil: “Das erste Potential ist also ein Tensor, nicht ein Skalar! $A'_{\mu\nu} = \sum \frac{\partial x_\alpha}{\partial \ldots}\frac{\partial x_\beta}{\ldots}A_{\mu\nu}$”.

[15] In the left margin in pencil: “$A^{\mu\nu\prime}$”.

an den Koeffizeienten anbringen.[16] Wir wollen noch den *gemischten Tensor* g^μ_ν defineren, dessen Koeffizienten sich nach der Formel[17]

$$g'^{\mu}_{\ \nu} = \sum_{\alpha\beta} \frac{\partial x'_\mu}{\partial x'_\alpha} \frac{\partial x_\beta}{\partial x'_\nu} g^\alpha_\beta$$

transformieren. Durch unsere Betrachtungen wurden wir gleich auf eine *quadratische* kovariante Differentialform geführt. Wenden wir uns nun dem einfacheren Fall der *linearen* Kovarianten zu, so gelangen wir zum Begriff des *kovarianten Vektors*. Die lineare Differentialform sei $\sum_\mu q_\mu dx_\mu$, wobei die q_μ Funktionen von x_1 und x_2 sind. Unterwerfen wir diese Form der Transformation

$$x_1 = x_1(x'_1 x'_2), \qquad x_2 = x_2(x'_1 x'_2),$$

so erhalten wir

$$\sum_\mu q_\mu dx_\mu = \sum_{\mu\nu} q_\mu \frac{\partial x_\mu}{\partial x'_\nu} dx'_\nu = \sum_\mu q'_\mu dx'_\mu,$$

wenn

$$q'_\mu = \sum_\alpha \frac{\partial x_\alpha}{\partial x'_\mu} q_\alpha$$

gesetzt wird. Transformieren die Koeffizienten der linearen Differentialform sich so, dann ist also $\sum_\mu q_\mu dx_\mu$ eine Kovariante und das Koeffizientensystem q_μ wird deswegen ein kovarianter Vektor genannt. Das einfachste Beispiel eines solchen ist der *Gradient* $\frac{\partial \Phi}{\partial x_\mu}$ der skalaren Funktion $\Phi(x_1, x_2)$. Analog der obigen Betrachtung nennen wir einen Vektor q^μ dann einen *kontravarianten Vektor*, wenn er sich transformiert gemäss

$$q'^{\mu} = \sum_\alpha \frac{\partial x'_\mu}{\partial x_\alpha} q^\alpha.$$

Bemerkung: Für die Differentiale der Variabeln gelten die Transformationsformeln

$$dx'_\nu = \sum_\alpha \frac{\partial x'_\nu}{\partial x_\alpha} dx_\alpha,$$

d. h. *die Differentiale sind die Komponenten eines kontravarianten Vektors*[18] und wir sollten sie folgerichtig mit dx^ν bezeichnen. Das totale Differential $d\Phi$ der skalaren Funktion Φ

$$d\Phi = \sum_\alpha \frac{\partial \Phi}{\partial x_\alpha} dx_\alpha$$

[16] However, Hilbert denotes coordinate differentials with subscript indices, see the following note.

[17] In the left margin in pencil: "A'^{μ}_{ν}". Subscript indices for coordinate differentials were used by Einstein, see, e.g. *Einstein 1916a*.

[18] In the left margin Hilbert wrote with pencil, and then deleted it: "man müsste also die Indices oben anbringen!"

kann also aufgefasst werden als skalares Produkt des kovarianten Vektors $\frac{\partial\Phi}{\partial x_\alpha}$ (Gradient) in den kontravarianten Vektor dx_α. Dies Produkt ist eine Kovariante. In der projektiven Geometrie entsprechen den Kovarianten und Kontravarianten die bezw. Punkt = und Linien = (Ebenen) Koordinaten.

§ 6. Geodätische Linien

Nun kommen wir zu einer neuen Fragestellung. Wir haben die Elemente der
Geometrie daraufhin durchzusehen, was | wir von ihnen brauchen können. Da 11
suchen wir nun zu einem Begriff zu gelangen, der demjenigen der geraden Linie in der ebenen Geometrie entspricht. Wir hatten gefordert, dass jedem beliebigen Kurvenstück unserer Mannigfaltigkeit eine Länge entsprechen soll. Die Gerade ist die kürzeste Linie, die zwei gegebene Punkte verbindet, *also fragen wir jetzt nach der kürzesten Verbindungslinie zweier Punkte*. Dies ist ein Problem der

III. Variationsrechnung,

des dritten wichtigen mathematischen Hilfsmittels, das wir benötigen.

Während die Differentialrechnung nur Zahlen variiert und daher die Hilfsmittel liefert, um einen Funktionswert gegenüber den benachbarten Werten zu einem Minimum zu machen, variiert die Variationsrechnung die Funktionen selber. Diese Disziplin ermöglicht es also, Funktionen von Funktionen zum Minimum zu machen, indem die Funktionen, die als Argument stehen, variiert werden. Die kürzeste Verbindungslinie zweier Punkte zu finden, ist also eine solche Aufgabe. In der Tat ist die Länge einer Kurve eine Funktion von zwei Funktionen $x_i(p)$, wir haben nämlich

$$s = \int_{p_1}^{p_2} \sqrt{g_{11}\left(\frac{dx_1}{dp}\right)^2 + 2g_{12}\frac{dx_1}{dp}\frac{dx_2}{dp} + g_{22}\left(\frac{dx_2}{dp}\right)^2}\, dp$$

zum Minimum zu machen. Unter dem Integral kommen die zwei unbekannten Funktionen $x_1(p)$ und $x_2(p)$ vor. Wir erhalten entsprechend auch zwei Gleichungen zu deren Bestimmung. Wir schreiben abgekürzt

$$s = \int_{p_1}^{p_2} \sqrt{\varphi}dp,$$

wobei $\varphi = \sum_{\mu\nu} g_{\mu\nu}\dot{x}_\mu\dot{x}_\nu$, $\dot{x}_\mu = \frac{dx_\mu}{dp}$ bedeutet. Unser Problem ist 12

$$\int_{p_1}^{p_2} \sqrt{\varphi}dp = \text{ Minimum}$$

zu machen. Gesucht sind $x_1(p)$ und $x_2(p)$. Dies ist ein Spezialfall des allgemeineren Problems

$$\int_{p_1}^{p_2} F(x_1 x_2 \dot{x}_1 \dot{x}_2 p)dp = \text{ Minimum}$$

zu machen. Die Variationsrechnung lehrt, dass die beiden *Lagrangeschen Gleichungen* erfüllt sein müssen:

$$\frac{d}{dp}\frac{\partial F}{\partial \dot{x}_1} - \frac{\partial F}{\partial x_1} = 0, \qquad \frac{d}{dp}\frac{\partial F}{\partial \dot{x}_2} - \frac{\partial F}{\partial x_2} = 0.$$

In unserem Falle wird[19]

$$\frac{d}{dp}\frac{\partial \sqrt{\varphi}}{\partial \dot{x}_\kappa} - \frac{\partial \sqrt{\varphi}}{\partial x_\kappa} = 0, \qquad \kappa = 1, 2.$$

Hierin ist

$$\begin{aligned}
\frac{\partial \sqrt{\varphi}}{\partial \dot{x}_\kappa} &= \varphi^{-\frac{1}{2}} \sum_\nu g_{\kappa\nu} \dot{x}_\nu, \\
\frac{d}{dp}\frac{\partial \sqrt{\varphi}}{\partial \dot{x}_\kappa} &= -\frac{1}{2}\varphi^{-\frac{1}{2}} \frac{d\varphi}{dp} \sum g_{\kappa\nu}\dot{x}_\nu + \varphi^{-\frac{1}{2}} \frac{d}{dp}\left(\sum g_{\kappa\nu}\dot{x}_\nu\right), \\
\frac{\partial \sqrt{\varphi}}{\partial x_\kappa} &= \frac{1}{2}\varphi^{-\frac{1}{2}} \sum_{\mu\nu} g_{\mu\nu\kappa}\dot{x}_\mu \dot{x}_\nu,
\end{aligned}$$

wobei zur Abkürzung $g_{\mu\nu\kappa} = \frac{\partial g_{\mu\nu}}{\partial x_\kappa}$ gesetzt ist. Die Langrangeschen Gleichungen werden dann zu

$$\frac{1}{2}\varphi^{-\frac{1}{2}} \left\{ -\frac{d\varphi}{dp} \sum_\nu g_{\kappa\nu}\dot{x}_\nu + 2\varphi \frac{d}{dp}\left(\sum_\nu g_{\kappa\nu}\dot{x}_\nu\right) - \varphi \sum_{\mu\nu} g_{\mu\nu\kappa}\dot{x}_\mu \dot{x}_\nu \right\} = 0, \ \kappa = 1, 2. \tag{2}$$

Diese beiden Differentialgleichungen sind nun aufzulösen, um die kürzeste Linie zu finden. *Das Problem lässt sich aber erheblich vereinfachen*, wobei uns folgender Umstand von Nutzen ist: Wenn man nämlich p durch eine beliebige Funktion von p' ersetzt, so müssen die Gleichungen (2) für diesen neuen Parameter p' wieder erfüllt sein; in der Tat wird dann

$$dp = \frac{dp}{dp'}dp', \qquad x'_i = \frac{dx_i}{dp'}\frac{dp'}{dp},$$

also wird

$$\int_{p_1}^{p_2} \sqrt{\varphi(p)}dp = \int_{p'_1}^{p'_2} \sqrt{\varphi(p')}\frac{dp'}{dp}\frac{dp}{dp'}dp' = \int_{p'_1}^{p'_2} \sqrt{\varphi(p')}dp',$$

13 d. h. das Variationsproblem bleibt ganz ungeändert, und daher auch die Lagrangeschen Gleichungen. Nun werde der Parameter p so normiert, dass *p im wesentlichen die Bogenlänge s darstellt*. Die Bogenlänge ist definiert durch

$$\left(\frac{ds}{dp}\right)^2 = \varphi.$$

[19]To the left of the following equation, Hilbert wrote in pencil: “$\frac{1}{2}\varphi^{-\frac{1}{2}} \frac{d}{dp}\frac{\partial\varphi}{\partial\dot{x}_k} - \frac{\partial\varphi}{\partial x_k}$”.

Ist $p = s$, so wird $\varphi = 1$.

Umgekehrt: Wird der Parameter so gewählt, dass $\varphi = 1$ wird, so ist p die Bogenlänge. Wir wollen hier p so wählen, dass $\varphi = c$ wird. Wollen wir später doch noch die Bogenlänge als Parameter haben, so müssen wir nur

$$p = \frac{p'}{\sqrt{c}}$$

setzen. Dann ist der neue Parameter p' die Bogenlänge s. Jetzt ist die Parameterdarstellung der Kurve nicht mehr willkürlich. Andererseits hat der von uns gewählte Parameter mit Bezug auf die Bogenlänge die Bedeutung

$$p = \frac{s}{\sqrt{c}}.$$

§ 7. Einführung der Bogenlänge als Parameter

Wie vereinfachen sich nun unsere Gleichungen? Wir schliessen den Fall $\varphi = 0$ vorerst aus. Im zweidimensionalen Raum haben wir φ noch definit > 0 angenommen. Wir können also auch keine reellen Lösungen $x_i(p)$ erwarten, die $\varphi = 0$ machen. Im vierdimensionalen Raum dagegen und diesen werden wir später nötig haben, ist φ gar nicht mehr positiv definit. Aber auch in der zweidimensionalen Geometrie werden wir für $\varphi = 0$ Kurven erhalten; nur sind es imaginäre Linien von der Länge Null. Ist also

$$\varphi = c \neq 0,$$

so fällt in der Differentialgleichung (2) der Faktor $\frac{1}{2}\varphi^{-\frac{1}{2}}$ | weg, und wir erhalten 14

$$2\frac{d}{dp}\sum_{v} g_{k\nu}\dot{x}_\nu - \sum_{\mu\nu} g_{\mu\nu k}\dot{x}_\mu\dot{x}_\nu = 0, \qquad k = 1, 2 \tag{3}$$

$$\varphi = c \tag{4}$$

zur Berechnung der zwei Funktionen $x_i(p)$. Hier haben wir also *den paradoxen Fall, drei Gleichungen zu haben zur Bestimmung von nur zwei Funktionen* x_1(p) und $x_2(p)$. Da wir nichts Unerlaubtes getan haben, müssen diese Gleichungen miteinander verträglich sein. In der Tat können wir zeigen, dass aus den zwei Gleichungen (3) die dritte Gleichung (4) folgt. Hierzu multiplizieren wir (3) mit bezw. $\dot{x}_k$ und addieren:

$$2\frac{d}{dp}\left(\sum_{k\nu} g_{k\nu}\dot{x}_k\dot{x}_\nu\right) - 2\sum_{k\nu} g_{k\nu}\ddot{x}_k\dot{x}_v - \sum_{\mu\nu k} g_{\mu\nu k}\dot{x}_\mu\dot{x}_\nu\dot{x}_k = 0,$$

oder

$$2\frac{d\varphi}{dp} - \frac{d\varphi}{dp} = 0, \quad \varphi = \text{const.} \quad \text{q. e. d.}$$

$\varphi = \text{const}$ *ist also ein Integral* dieser beiden Differentialgleichungen und gar kein Widerspruch zu diesen, wie wir befürchten mussten. Damit haben die Differentialgleichungen der geodätischen Linie schon die einfache Form (3) erhalten:

$$2\frac{d}{dp}\sum_{\nu} g_{k\nu}\dot{x}_\nu - \sum_{\mu\nu} g_{\mu\nu k}\dot{x}_\mu\dot{x}_\nu = 0, \qquad k = 1, 2.$$

Diese beiden Gleichungen haben zunächst den Charakter von Lagrangeschen Gleichungen verloren. Ein ganz merkwürdiger Umstand ist es nun, dass auch *diese Gleichungen wieder Lagrangesche sind* und zwar des viel einfacheren, aber vom vorhergehenden ganz verschiedenen Variationsproblems

$$\int_{p_1}^{p_2} \varphi dp = \text{Minimum.} \tag{5}$$

In der Tat sind dessen Lagrangesche Gleichungen

$$\frac{d}{dp}\frac{\partial\varphi}{\partial\dot{x}_k} - \frac{\partial\varphi}{\partial x_k} = 0, \qquad k = 1, 2.$$

Dies sind aber unsere Gleichungen (3)! Wir können also unser[20] | Variationsproblem ersetzen durch das Problem (5). Dann ist aber der Parameter so festgelegt, dass $\varphi = 0$ wird. Dieses letztere Problem ist nun nicht mehr invariant, bei beliebiger Wahl des Parameters. Wir haben also den *Satz: wenn man* $\int_{p_1}^{p_2}\sqrt{\varphi}dp = \text{Min}$ *machen will und gleichzeitig die einschränkende Nebenbedingung* $\varphi = c$ *stellt, die den Parameter festlegt, hat dies Variationsproblem dieselben Lösungen wie das andere* $\int_{p_1}^{p_2}\varphi dp = \text{Min}$.

§ 8. Normalform der Differentialgleichungen der geodätischen Linien

Wir wollen nun die Differentialgleichungen (3) der geodätischen Linien auf ihre *Normalform* bringen. Führen wir die Differentiation nach p aus, so erhalten wir

$$\sum_{\nu} g_{\kappa\nu}\ddot{x}_\nu + \sum_{\nu h} g_{\kappa\nu h}\dot{x}_\nu\dot{x}_h - \frac{1}{2}\sum_{\mu\nu} g_{\mu\nu\kappa}\dot{x}_\mu\dot{x}_\nu \langle = 0.\rangle$$

Nun ist

$$\sum_{\nu h} g_{\kappa\nu h}\dot{x}_\nu\dot{x}_h = \sum_{\nu h} g_{\kappa h\nu}\dot{x}_\nu\dot{x}_h,$$

weil ja über ν und h summiert wird, also auch

$$\sum_{\nu h} g_{\kappa\nu h}\dot{x}_\nu\dot{x}_h = \frac{1}{2}\sum_{\nu h}(g_{\kappa\nu h} + g_{\kappa h\nu})\dot{x}_\nu\dot{x}_h.$$

[20] At the bottom of the page, Hilbert wrote in pencil: “$\sum_k \dot{x}_k \frac{d}{dp}\frac{\partial\varphi}{\partial\dot{x}_k} - \sum_k \dot{x}_k\frac{\partial\varphi}{\partial x_k} = 0$ d. h. $\left(\frac{d}{dp}\sum_k \dot{x}_k\frac{\partial\varphi}{\partial\dot{x}_k} - \sum_k \ddot{x}_k\frac{\partial\varphi}{\partial x_k}\right) - \sum_k \dot{x}_k\frac{\partial\varphi}{\partial x_k} = 0$ d. h. $2\frac{d\varphi}{dp} - \frac{d\varphi}{dp} - \frac{d\varphi}{dp} = 0$”.

Diesen Ausdruck setzen wir ein:

$$\sum_{\nu h} g_{\kappa\nu}\ddot{x}_\nu + \frac{1}{2}\sum_{\mu\nu}(g_{\kappa\nu\mu} + g_{\mu\kappa\nu} - g_{\nu\mu\kappa})\dot{x}_\mu\dot{x}_\nu = 0. \tag{6}$$

Um $\ddot{x}_\nu$ frei von Faktoren zu machen, multiplizieren wir diese Gleichungen mit bezw. $g^{h\kappa}$ und summieren über κ. Wir benutzen die Relationen

$$\sum_{\mu} g^{h\kappa} g_{\kappa\nu} = \begin{cases} 0, & h \neq \nu \\ 1, & h = \nu \end{cases}$$

die aus Definitionsgleichungen der $g^{h\kappa}$ folgen, und die nichts anderes aussagen, als dass die ersten Unterdeterminanten der i-ten Zeile einer Determinante multipliziert mit den entsprechenden Elementen der κ-ten Zeile 0 oder den Wert der Determinante selbst ergeben, je nachdem $i \neq k$ bezw. $i = k$ ist. 16
Dann wird aus (6)

$$\ddot{x}_h + \frac{1}{2}\sum_{\mu\nu k} g^{hk}\left(g_{k\nu\mu} + g_{\mu k\nu} - g_{\nu\mu k}\right)\dot{x}_\mu\dot{x}_\nu = 0 \qquad h = 1, 2$$

oder als *definitive Form der Differentialgleichungen der geodätischen Linie*

$$\ddot{x}_h + \sum_{\mu\nu}\begin{Bmatrix}\mu\nu \\ h\end{Bmatrix}\dot{x}_\mu\dot{x}_\nu = 0, \tag{7}$$

wobei die $\begin{Bmatrix}\mu\nu \\ h\end{Bmatrix}$ die berühmten Ausdrücke sind:

$$\begin{Bmatrix}\mu\nu \\ h\end{Bmatrix} = \frac{1}{2}\sum_{k} g^{hk}\left(g_{k\nu\mu} + g_{\mu k\nu} - g_{\nu\mu k}\right),$$

die wir die *g-Klammern* nennen wollen. Sie wurden zuerst von Christoffel[21] eingeführt. Es sind homogene, lineare[22] Ausdrücke in den ersten Ableitungen der $g_{\mu\nu}$. Sie sind in den beiden oberen Indizes symmetrisch, weil $g_{k\nu\mu} = g_{\nu k\mu}$ ist. Wir werden später noch von ihnen Gebrauch zu machen haben, da sie in der Gravitationstheorie eine grosse Rolle spielen.

§ 9. Geodätische Linien auf der Kugel (als Beispiel)

Aus den Gleichungen (7) wollen wir jetzt die geodätischen Linien in zwei Fällen berechnen.

---

[21] See *Christoffel 1869*.

[22] "lineare" was corrected from "quadratische".

1) Man sieht sofort, dass in der *Euklidischen Geometrie die gerade Linie auch die geodätische ist*. Dann ist nämlich $g_{\mu\nu} = \text{const}$, also verschwinden sämtliche g-Klammern und die allgemeinen Integrale von (6) werden zu

$$x_h = a_h p + b_h,$$

woraus wir durch Elimination von p eine *lineare* Gleichung zwischen den x_h erhalten.[23]

2) Wir hatten oben (S. 7) als Beispiel einer Fläche, auf der wir Geometrie treiben wollten, die *Kugel* herangezogen. Wir wollen also im Euklidischen
17 Raum eine Kugel vom Radius Eins | gegeben annehmen durch die Gleichungen

$$x = \cos\vartheta, \qquad y = \cos\varphi\sin\vartheta, \qquad z = \sin\varphi\sin\vartheta. \tag{8}$$

Dann sind $x_1 = \vartheta$ $x_2 = \varphi$ die Koordinaten, die einen Punkt auf der Kugeloberfläche festlegen. Das Linienelement hatten wir zu

$$ds^2 = d\vartheta^2 + \sin^2\vartheta d\varphi^2$$

gefunden, also $g_{11} = 1$ $g_{12} = 0$ $g_{22} = \sin^2\vartheta$. Hier wollen wir Folgendes bemerken: Um in der Praxis die g-Klammern zu finden, wird man sie nicht aus den eben bestimmten $g_{\mu\nu}$ berechnen, sondern vielmehr vom Variationsproblem $\int_{p_1}^{p_2} \varphi dp = \text{Min.}$ ausgehen und die Lagrangeschen Gleichungen in der Form

$$\ddot{x}_\nu + \text{homogene quadratische Form der } \dot{x}_\mu = 0$$

anschreiben. Dann sind *die Koeffizienten dieser quadratischen Form die gesuchten g-Klammern*, z. B. $\frac{1}{2}\begin{Bmatrix}\lambda\mu\\ \nu\end{Bmatrix}$ der Koeffizient von $\dot{x}_\lambda\dot{x}_\mu$. Dies wollen wir am Beispiel der Kugel auch durchführen, trotzdem wir es zur Lösung der Aufgabe, die geodätischen Linien auf der Kugel zu bestimmen, gar nicht nötig haben. Wir erhalten die g-Klammern nämlich als ein Zwischenresultat dieser Aufgabe. Das Problem lautet also

$$\int_{p_1}^{p_2} (\dot{\vartheta}^2 + \sin^2\vartheta\dot{\varphi}^2)dp = \text{Min.}$$

Gesucht sind dabei ϑ und φ als Funktionen des Parameters p. Die Lagrangeschen Gleichungen werden hier zu:

$$\begin{cases} \frac{d}{dp}(2\dot{\vartheta}) - 2\sin\vartheta\cos\vartheta\dot{\varphi}^2 = 0, \\ \frac{d}{dp}(2\sin^2\vartheta\dot{\varphi}) = 0. \end{cases} \tag{9}$$

18 Führen wir die Differentiation nach p aus, so erhalten wir

$$\ddot{\vartheta} - \sin\vartheta\cos\vartheta\dot{\varphi}^2 = 0,$$
$$\ddot{\varphi} + 2\operatorname{cotg}\vartheta\dot{\vartheta}\dot{\varphi} = 0.$$

[23] In the left margin, there is a reader's mark (⊥).

Hieraus lassen sich die g-Klammern ablesen:

$$\begin{Bmatrix}11\\1\end{Bmatrix}=0,\qquad \begin{Bmatrix}12\\1\end{Bmatrix}=0,\qquad \begin{Bmatrix}22\\1\end{Bmatrix}=-\sin\vartheta\cos\vartheta,$$
$$\begin{Bmatrix}11\\2\end{Bmatrix}=0,\qquad \begin{Bmatrix}12\\2\end{Bmatrix}=+\operatorname{cotg}\vartheta,\qquad \begin{Bmatrix}22\\2\end{Bmatrix}=0.$$

Da wir diese g-Klammern aber augenblicklich gar nicht brauchen, wenden wir uns wieder unserer eigentlichen Aufgabe zu und lösen die Gleichungen (9). Die zweite Gleichung lässt sich sofort einmal integrieren:

$$\sin^2\vartheta\dot{\varphi}=\text{const}=a. \tag{10}$$

Die erste Gleichung multiplizieren wir mit $\dot{\vartheta}$ und ersetzen $\dot{\varphi}^2$ durch seinen aus (10) zu berechnenden Wert $\frac{a^2}{\sin^4\vartheta}$. Dann erhalten wir durch Integration

$$\dot{\vartheta}^2+\frac{a^2}{\sin^2\vartheta}=b.$$

Dieses Integral hätten wir nun auch ohne Rechnung finden können; denn nach unserer allgemeinen Theorie muss $\varphi=\text{const}$ ein Integral sein, dies gibt aber wegen (10)

$$\varphi=\dot{\vartheta}^2+\sin^2\vartheta\dot{\varphi}^2=\dot{\vartheta}^2+\frac{a^2}{\sin^2\vartheta}=\text{const}.$$

Nun können wir auch $\frac{d\vartheta}{d\varphi}$ berechnen. Es ist nämlich

$$\left(\frac{d\vartheta}{d\varphi}\right)^2=\frac{\dot{\vartheta}^2}{\dot{\varphi}^2}=\frac{b-\frac{a^2}{\sin^2\vartheta}}{\frac{a^2}{\sin^4\vartheta}}=\sin^2\vartheta(A\sin^2\vartheta-1).$$

Diese Differentialgleichung ist zu integrieren. Es ist[24]

$$\frac{d}{d\varphi}\operatorname{cotg}\vartheta=-\frac{1}{\sin^2\vartheta}\frac{d\vartheta}{d\varphi}.$$

Sei also $\operatorname{cotg}\vartheta=\Theta$ so wird

$$\left(\frac{d\Theta}{d\varphi}\right)^2=A-\frac{1}{\sin^2\vartheta}=A-1-\frac{cos^2\vartheta}{\sin^2\vartheta}=A-1-\Theta.$$

Diese Gleichung kann man allgemein integrieren und erhält

$$\Theta=\alpha(\sin\varphi+\beta)=\operatorname{cotg}\vartheta,\qquad \alpha,\beta=\text{Integrationskonstanten.}$$

Damit haben wir als Lösung unserer Aufgabe 19

$$\cos\vartheta=\sin\vartheta(C\sin\varphi+D\cos\varphi)$$

erhalten. Dies ist in unserem Euklidischen Raum gedeutet, zufolge der Gleichungen (8), $x=Cz=Dy=$ Ebene durch den Mittelpunkt der Kugel. Wir erhalten das bekannte Resultat: *die geodätischen Linien auf der Kugel sind grösste Kreise*.

[24]"Es ist" was corrected from "Wir erhalten".

§ 10. Riemannsche Koordinaten

Die geodätischen Linien werden das Fundament der Geometrie auf beliebigen Flächen sein, so wie die Gerade das Fundament der ebenen Geometrie ist. Dies müssen wir nun noch mehr herausarbeiten. *Die gerade Linie ist analytisch in geignet gewählten Koordinaten ihrer ganzen Ausdehnung nach als eine lineare Gleichung* zwischen diesen Koordinaten gegeben und nicht nur im *Infinitesimalen*, wie dies bei Kurven der Fall ist. Hat man einmal dieses Resultat gewonnen, so lassen sich alle Sätze über gerade Linien sofort beweisen. Wir fragen jetzt: gibt es in der allgemeinen Geometrie einer zweidimensionalen Mannigfaltigkeit ein Analogon? Aus der Theorie der Grundlagen der Geometrie wissen wir, dass nur in der Euklidischen und in den beiden Nichteuklidischen Geometrien die Geraden oder kürzesten Linien durch *lineare* Beziehungen zwischen zwei geeigneten Parametern ausgedrückt werden können. In der Flächentheorie drückt man dies dadurch aus, dass man sagt: man kann nur auf einer Fläche konstanter Gauss'scher Krümmung krummlinige Koordinaten derart einführen, dass die geodätischen Linien durch lineare
20 Beziehungen zwischen den beiden Parametern ausgedrückt werden. Dabei sind natürlich *Flächen, die durch Verbiegung auseinander hervorgehen, einander äquivalent;* wir treiben ja nur auf der Fläche selbst Geometrie und brauchen uns daher um die Lage dieser Fläche im Raum gar nicht zu kümmern. Auf der Kugel speziell ist das Gauss'sche Krümmungsmass konstant und positiv, auf ihr gilt die Riemann-Helmholtzsche Geometrie. Aequivalent mit der Kugel sind dann alle diejenigen Flächen, die durch Verbiegung aus ihr hervorgehen. Die Kugel lässt sich übrigens nur dann verbiegen, wenn Teile aus ihr herausgeschnitten sind. Auch für die Lobatscheffsky-Boljay'sche Geometrie konstanter negativer Krümmung gilt wieder, dass die geodätischen Linien durch lineare Beziehungen passend gewählter Parameter ausgedrückt werden können.

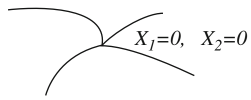
Von diesen Beziehungen ist bei unserer allgemeinen Geometrie keine Rede mehr. Die geodätischen Linien können sicher *nicht* durch lineare Beziehungen zwischen den Parametern ausgedrückt werden. Und doch besteht ein tiefgreifendes Analogon, das auch für die spätere physikalische Anwendung von grosser Bedeutung sein wird. Unser Programm ist es, dieses Analogon aufzudecken.

Auch bei uns gilt die Euklidische Geometrie im Infinitesimalen, d. h. in der Umgebung eines Punktes, und auf die von einem Punkt ausgehenden geodätischen Linien werden wir unsere Betrachtungen übertragen können. Es werden
21 sich in der Tat Parameter so einführen lassen, dass *die von einem | Punkt*[25] *ausgehenden geodätischen Linien* durch geeignete Parameter ausgedrückt, *sich als lineare Beziehungen* derselben *darstellen*. Diese geeigneten Parameter heissen wir

die Riemannschen Koordinaten.

[25]"die von einem Punkt" was corrected from "in der Umgebung eines Punktes".

Ein beliebiger Punkt unserer Mannigfaltigkeit sei $x_1 = 0 \quad x_2 = 0$. Wir können ihn auch als Punkt auf unserer Fläche deuten. Die geodätischen Linien, die von ihm ausgehen, müssen dann eine einparametrige Schar sein, die die Fläche lückenlos überdeckt. Die Gleichungen (7)



$$\ddot{x}_h + \sum_{\mu\nu} \left\{ {\mu\nu \atop h} \right\} \dot{x}_\mu \dot{x}_\nu = 0, \qquad h = 1, 2$$

sind also in der Umgebung des Punktes $x_i = 0$ zu integrieren. Wir zählen auch p vom Punkte $x_i = 0$ aus. Nach der allgemeinen Theorie der Differentialgleichungen zweiter Ordnung dürfen wir in diesem Punkte noch die Werte von $\dot{x}_i = \xi_i$ beliebig vorschreiben und haben also Randbedingungen zu erfüllen

$$(x_1)_{p=0} = 0, \quad (x_2)_{p=0} = 0, \quad (\dot{x}_1)_{p=0} = \xi_1, \quad (\dot{x}_2)_{p=0} = \xi_2. \tag{11}$$

Nun sind die vier willkürlichen Integrationskonstanten des Systems von Differentialgleichungen als 0, 0, ξ_1, ξ_2 festgelegt. Also müssen jetzt die x_h völlig bestimmt sein:

$$x_h = x_h(p, \xi_1, \xi_2).$$

Die Funktionen x_h lassen sich in Reihenform explizit anschreiben, es wird in der Umgebung der Stelle $p = 0$

$$x_h = (x_h)_{p=0} + (\dot{x}_h)_{p=0} p + \frac{1}{2} (\ddot{x}_h)_{p=0} p^2 + \cdots, \qquad h = 1, 2,$$

also wegen (7) und (11)

$$x_h = 0 + \xi_h p - \frac{1}{2} \sum_{\mu\nu} \left\{ {\mu\nu \atop h} \right\} \xi_\mu \xi_\nu p^2 + \cdots, \qquad h = 1, 2. \tag{12}$$

Scheinbar gehen entsprechend den zwei in x_i vorkommenden willkürlichen 22
Konstanten ∞^2 geodätische Linien[26] durch einen festen Punkt. Das ist aber unmöglich, denn dann würden ja zwei beliebige Punkte unserer Mannigfaltigkeit noch durch ∞^1 geodätische Linien verbunden werden. Dieser Widerspruch ist leicht zu lösen. Wir gingen aus vom Problem

$$\delta \int_{p_1}^{p_2} \sqrt{\varphi}\, dp = 0, \qquad \varphi = c.$$

Dabei hatte p die Bedeutung $p = \frac{s}{\sqrt{c}}$ (s = Bogenlänge). Nun ist diese Konstante aber noch ganz willkürlich, d. h. ein Parameter p', der φ den konstanten Wert c' erteilt, ist ebenfalls zulässig. Dann wird $p' = \frac{s}{\sqrt{c'}}$ oder $p' = p\sqrt{\frac{c}{c'}}$. Der Parameter p ist also nur bis auf eine multiplikative Konstante μ bestimmt. Setzen wir $p' = \mu p$ in $x_h(p, \xi_1, \xi_2)$ ein, so muss x_h immer noch eine geodätische

[26] "∞^2" was interlineated between "willkürlichen" and "Konstanten".

Linie durch den Punkt $x_h = 0$ darstellen, aber freilich nicht mehr dieselbe. Dazu müssen wir viel mehr die Anfangsbedingungen verändern. Diese sind für den Parameter p': $(x_i)_{p'=0} = 0$ $(\dot{x}_i)_{p'=0} = \xi'_i$. Durch den Parameter p ausgedrückt werden sie zu

$$(x_i)_{p=0} = 0, \qquad (\dot{x}_i)_{p=0} = \frac{1}{\mu}\left(\frac{dx_i}{dp}\right)_{p=0} = \frac{1}{\mu}\xi_i.$$

Setzen wir auch diese abgeänderten Anfangsbedingungen ein, so müssen wir wieder die nämliche geodätische Linie erhalten, d. h.

$$x_h(p, \xi_1, \xi_2) = x_h\left(\mu p, \frac{\xi_1}{\mu}, \frac{\xi_2}{\mu}\right). \tag{13}$$

Die Richtigkeit dieser Gleichung kann man übrigens an (12) leicht verifizieren. Wir können also sagen:

Aendert man ξ_1 und ξ_2 so, dass ihr Quotient denselben Wert beibehält, so
23 stellen die Funktionen x_h noch die | nämliche geodätische Linie dar; doch entsprechen den gleichen Parameterwerten nicht mehr dieselben Punkte auf der Kurve. Die Gleichungen der geodätischen Linie $x_h = x_h(p, \xi_1, \xi_2)$ sind also nur vom Verhältnis $\frac{\xi_1}{\xi_2}$ der Konstanten abhängig, und es gehen, wie es auch sein muss, nur ∞^1 solcher Kurven durch einen Punkt.

Nun wollen wir die Riemannschen Koordinaten definieren. Wir führen folgende neuen Funktionen f_1 f_2 von ξ_1 und ξ_2 ein, indem wir in $x_h(p\xi_1\xi_2)$ dem Parameter p den Wert eins erteilen:

$$x_h(1\xi_1\xi_2) = f_h(\xi_1\xi_2) = \xi_h - \frac{1}{2}\sum_{\mu\nu}\begin{Bmatrix}\mu\nu\\ h\end{Bmatrix}\xi_\mu\xi_\nu + \cdots. \tag{14}$$

Diese Gleichungen denken wir uns nach ξ_1 und ξ_2 aufgelöst. Die hierbei als Funktionen von x_1 und x_2 bestimmten ξ_1 und ξ_2 nennen wir die *Riemannschen Koordinaten* des Punktes x_1, x_2 unserer Mannigfaltigkeit. *Damit haben wir die Koordinaten*, die doch bis jetzt ganz willkürliche Namen waren, die wir den Punkten gaben, wenigstens *in Bezug auf einen ausgezeichneten* 0 *Punkt auf eine Normalform gebracht*. Wir haben also den *Punkten* der Menge gewissermassen „Normalnamen“ gegeben, allerdings nur Normalnamen relativ zu einem ausgezeichneten Punkt. Wir wollen nun untersuchen, inwiefern diese neuen Koordinaten ausgezeichnete sind. Wir betrachten dabei nur die geodätischen Linien durch den 0-Punkt und behaupten:

§ 11. Gleichungen der geodätischen Linie in Riemannschen Koordinaten

Die Gleichungen der geodätischen Linien durch den 0-Punkt sind in diesen neuen Koordinaten

$$\xi_1 = c_1 p, \quad \xi_2 = c_2 p, \tag{15}$$
$$c_1, c_2 = \text{Integrationskonstanten}.$$

Dann stellt jede lineare Beziehung zwischen den ξ_i z. B. $c_2\xi_1 - c_1\xi_2 = 0$ *geo-* 24
dätische Linien durch den 0*-Punkt dar* und umgekehrt. Zum Beweise müssen wir nur die Werte von ξ_i aus (15) in (14) einsetzen und zeigen, dass die so bestimmten x_h als Funktionen von p den Differentialgleichungen der geodätischen Linie genügen. Dann wird

$$x_i = f_i(c_1 p, c_2 p) = x_i(1, c_1 p, c_2 p)$$

und wegen (13)

$$x_i = x_i(p, c_1, c_2).$$

Dies ist aber die Gleichung einer geodätischen linie durch den 0-Punkt. Wir können die durch (15) ausgedrückte Tatsache auch folgendermassen ausdrükken: Legt man statt der x_i die Riemannschen Koordinaten ξ_i zugrunde, so bricht die Potenzreihe (12) der geodätischen Linie mit dem ersten Gliede ab. Für die nicht durch den Nullpunkt gehenden geodätischen Linien gilt der Satz, dass längs ihr eine lineare Beziehung zwischen geeigneten Parametern besteht, aber nicht, und kann auch gar nicht gelten, wie wir oben erläutert haben. Wir haben damit unsere Analogie zur Geometrie auf einer Fläche konstanten Gaussschen Krümmungsmasses soweit, als dies überhaupt möglich ist, getrieben.

§ 12. Die noch vorhandene Willkür in der Definition der Riemannschen Koordinaten

Die Riemannschen Koordinaten haben nun immer noch eine gewissse Willkür. Setzt man nämlich

$$\xi_i' = \text{lineare, homogene Funktion der } \xi_i,$$

so stellt dies wieder eine geodätische Linie durch den 0-Punkt dar. Wir hätten die Riemannschen Koordinaten also ebenso gut | als lineare, homogene 25
Funktionen der ξ_i definieren können. Wir wollen aber zeigen, dass von diesen projektiven Transformationen oder von dieser *projektiven Willkür* abgesehen,

die Riemannschen Koordinaten eindeutig festgelegt sind. Wir führen irgend welche andere Koordinaten ein durch

$$x'_i = x'_i(x_1, x_2)$$

und suchen für diese x'_i-Koordinaten die zugehörigen Riemannschen Koordinaten ξ'_i in Bezug auf den dem Punkte $x_1 = 0$, $x_2 = 0$ entsprechenden ausgezeichneten Punkt $x'_1 = a$, $x'_2 = b$. Es ist keine Einschränkung, wenn wir annehmen, dass dies der Punkt $x'_1 = 0$, $x'_2 = 0$ sei, da dies nur einer Parallelverschiebung unserer Punktmannigfaltigkeit entspricht. Die ξ'_i sind dann wieder Funktionen der ξ_i, und wir wollen beweisen, dass es *lineare Funktionen* derselben sind. Diese Willkür freilich muss bleiben. Zum Beweise gehen wir auf unsere Regel zurück: Man erhält die ξ_i, indem man die Gleichungen (7) der geodätischen Linie mit den Randbedingungen $(x_i)_{p=0} = 0$, $(\dot{x}_i)_{p=0} = \xi_i$ integriert und hierauf $p = 1$ setzt. Wir machen dasselbe mit den gestrichenen Koordinaten x'_i, die Funktionen des Parameters p' sein sollen. Das System von Differentialgleichungen bleibt nun nach der Transformation $x'_i = x'_i(x_1, x_2)$ in Bezug auf die gestrichenen Koordinaten dasselbe; denn geodätische Linien bleiben auch nach der Transformation solche. Wir müssen also nur überall in den Differentialgleichungen einen $'$ zufügen. Die Striche kann man aber auch weglassen. Nur müssen dann die Anfangsbedingungen auf die ungestrichenen Koordinaten transformiert werden. Diese Bedingungen sind

$$(x'_i)_{p'=0} = 0 \qquad (\dot{x}'_i)_{p'=0} = \xi'_i.$$

26 Nun sind die x'_i festgelegt. Wie drücken sich diese Bedingungen in den ungestrichenen Koordinaten aus?

Es ist keine Einschränkung anzunehmen, dass dem Parameterwert $p' = 0$ der Wert $p = 0$ entspricht; dann wird

$$\left(\frac{dx'_i}{dp'}\right)_{p'=0} = \left(\frac{\partial x'_i}{\partial x_1}\frac{dx_1}{dp}\frac{dp}{dp'} + \frac{\partial x'_i}{\partial x_2}\frac{dx_2}{dp}\frac{dp}{dp'}\right)_{p'=0} = a_{i1}\xi_1 + a_{i2}\xi_2,$$

d. h. $\xi'_i = a_{i1}\xi_1 + a_{i2}\xi_2$. Setzt man also $x'_i = f'_i(\xi'_1\xi'_2)$, so ist dies dasselbe, als ob man die ursprünglichen Differentialgleichungen mit den Randbedingungen

$$(x_i)_{p=0} = 0, \qquad (\dot{x}_i)_{p=0} = a_{i1}\xi_1 + a_{i2}\xi_2$$

integriert hätte. Wir haben also den Satz:

Ersetzt man die x_i-Koordinaten durch willkürliche andere Funktionen $x'_i = x'_i(x_1, x_2)$, so erhält man als Riemannsche Koordinaten ξ'_i lineare homogene Funktionen der ξ_i.

§ 13. Zurückführung der allgemeinen Invarianten auf projektive; Darstellung der $g_{\mu\nu}$ als Potenzreihen Riemannscher Koordinaten

Die Riemannschen Koordinaten entsprechen den rechtwinkligen Koordinaten in der Euklidischen Geometrie. Sie beziehen sich aber nur auf *einen* Punkt. Geht man zu einem *anderen* Punkt über, so sind diese neuen Riemannschen Koordinaten komplizierte, sicher aber *nicht* lineare Funktionen der ursprünglichen ξ_i; denn sonst hätten wir ja die Flächen konstanten Krümmungsmasses vor uns. Wir sehen jetzt schon den enormen Vorteil dieser Koordinaten. Die Riemannschen Koordinaten sind schon so stabil oder invariant, dass sie bei willkürlichen Transformationen der x_i, die nur einen Punkt in sich überführen, sich | nur noch projektiv verändern. Damit ist der Hauptschritt zu unserem 27
Ziel, *allgemeine Invarianten* aufzufinden, schon getan. *Alle projektiven Invarianten der ξ_i sind nämlich allgemeine Invarianten der x_i.* Solche sind z. B., wie die Invariantentheorie lehrt

$$\sum_{\mu\nu} g_{\mu\nu}\xi_\mu\xi_\nu$$

$$\text{kogredient zu} \quad \xi_k \begin{cases} \sum_{\mu\nu\kappa} g_{\mu\nu\kappa}\xi_\mu\xi_\nu\eta_\kappa & \text{wobei } g_{\mu\nu\kappa} = \frac{\partial g_{\mu\nu}}{\partial \xi_\kappa} \\ \sum_{\mu\nu\kappa l} g_{\mu\nu\kappa l}\xi_\mu\xi_\nu\eta_\kappa\eta_l & \text{"} \quad g_{\mu\nu\kappa l} = \frac{\partial g_{\mu\nu}}{\partial \xi_\kappa \partial \xi_l}. \end{cases}$$

Die ξ_i entsprechen, wie wir schon sagten, den rechtwinkligen Koordinaten in der Euklidischen Geometrie. Wir wollen dies zum Ausdruck bringen, indem wir die $g_{\mu\nu}$ nun für die ξ_i-Koordinaten bilden und untersuchen, wie und ob wir den $g_{\mu\nu}$ anmerken können, dass sie sich auf die Normalkoordinaten beziehen. Wir legen unserer Betrachtung also Riemannsche Koordinaten zugrunde. Die Gleichungen der geodätischen Linie sind dann

$$x_h = \xi_h = c_h p, \qquad h = 1, 2.$$

Setzen wir diese Ausdrücke in die Differentialgleichungen (3) der geodätischen Linie ein, so müssen diese also identisch erfüllt sein und zwar in den c_h und in p. Diese Gleichungen lauten in Riemannschen Koordinaten

$$2\frac{d}{dp}\left(\sum_\nu g_{\kappa\nu}\dot{\xi}_\nu\right) - \sum_{\mu\nu} g_{\mu\nu\kappa}\dot{\xi}_\mu\dot{\xi}_\nu = 0.$$

Für $\xi_\kappa = c_\kappa p$ erhalten wir $2\frac{d}{dp}\sum_\nu g_{\kappa\nu}c_\nu - \sum_{\mu\nu} g_{\mu\nu\kappa}c_\mu c_\nu = 0$, und da $\frac{d}{dp}\left(\sum_\nu g_{\kappa\nu}c_\nu\right) = \sum_{\mu\nu} g_{\kappa\nu\mu}c_\mu c_\nu$ ist, so wird

$$\sum_{\mu\nu} \left(2g_{\kappa\nu\mu} - g_{\mu\nu\kappa}\right) c_\mu c_\nu = 0.$$

Die $g_{\mu\nu}$ sind jetzt Funktionen der ξ_i, die wir nach steigenden Potenzen der ξ_i entwickeln,

28 $$g_{\mu\nu} = a_{\mu\nu} + \sum_{\kappa} a_{\mu\nu\kappa}\xi_\kappa + \frac{1}{2}\sum_{\kappa\rho} a_{\mu\nu\kappa\rho}\xi_\kappa\xi_\rho + \cdots; \qquad a_{\mu\nu\kappa} = a_{\nu\mu\kappa} \quad \text{etc.}$$

Diese Entwicklung gilt in der Umgebung der Stelle $\xi_i = 0$. Wir setzen $\xi_\mu = c_\mu p$ ein und erhalten aus $\sum_{\mu\nu}(2g_{\kappa\nu\mu} - g_{\mu\nu\kappa})c_\mu c_\nu = 0$

$$\sum_{\mu\nu}\left(2\left[a_{\kappa\nu\mu} + \sum_{\rho} a_{\kappa\nu\mu\rho}c_\rho p + \cdot\right] - \left[a_{\mu\nu\kappa} + \sum_{\rho} a_{\mu\nu\kappa\rho}c_\rho p + \cdot\right]\right)c_\mu c_\nu = 0.$$

Für $p = 0$ folgt dann $\sum_{\mu\nu}(2a_{\kappa\nu\mu} - a_{\mu\nu\kappa})c_\mu c_\nu = 0$. Da die Gleichungen der geodätischen Linie in c_μ und p identisch erfüllt sein sollen, so muss der Koeffizient von $c_\mu c_\nu$ verschwinden, also $2(a_{\kappa\nu\mu} + a_{\kappa\mu\nu}) - (a_{\mu\nu\kappa} + a_{\nu\mu\kappa}) = 0$, d. h.

$$2(a_{\mu\nu\kappa} + a_{\kappa\mu\nu} + a_{\nu\kappa\mu}) = 4a_{\mu\nu\kappa}.$$

Die linke Seite dieser Gleichung ist in den drei Indizes symmetrisch, also muss es auch die rechte Seite sein, d.h. die $a_{\mu\nu\kappa}$ sind einander gleich und wir erhalten $6a_{\mu\nu\kappa} = 4a_{\mu\nu\kappa}$ oder

$$a_{\mu\nu\kappa} = 0.$$

Wir haben *das merkwürdige Resultat*: Führen wir solche Koordinaten ein, dass die durch einen Punkt gehenden geodätischen Linien durch lineare Beziehungen zwischen diesen Koordinaten dargestellt werden, so haben die $g_{\mu\nu}$ in der Umgebung dieses Punktes die Form: $g_{\mu\nu} = a_{\mu\nu} + \frac{1}{2}\sum_{\kappa\rho} a_{\mu\nu\kappa\rho}\xi_\kappa\xi_\rho + \cdots$. Diese Koordinaten sind, wie wir sahen, nur bis auf eine projektive Transformation bestimmt; diese kann man dazu verwenden, dass $a_{\mu\nu} = \delta_{\mu\nu}$ wird. In der Umgebung des 0-Punktes haben wir also erreicht, dass *in den neuen Koordinaten die Euklidische Geometrie*, soweit dies möglich ist, nämlich *in den Gliedern nullter und erster Näherung, Geltung hat*. Die oben eingeführten g-Klammern[27] waren lineare, homogene Funktionen der ersten Ableitungen der $g_{\mu\nu}$. Sie müssen daher für die neuen Koordinaten sämtlich verschwinden. Dann wird die
29 Gleichung (7) | der geodätischen Linie zu $\ddot{\xi}_h = 0$, und dies stimmt wieder damit überein, dass die geodätischen Linien nun durch $\xi_h = c_h \cdot p$ dargestellt werden.

Erst *in den Gliedern zweiter Ordnung zeigt sich die Abweichung von der Euklidischen Geometrie*. Diese Glieder sind nämlich nicht mehr Null, doch bestehen auch zwischen ihnen noch Relationen. Wir haben als Gleichung dafür, dass der Koeffizient von p verschwindet

$$\sum_{\mu\nu\rho}(2a_{\kappa\nu\rho\mu} - a_{\mu\nu\kappa\rho})c_\mu c_\nu c_\rho = 0.$$

[27] See [p. 16] above.

Hierin ist $\frac{\partial^2 g_{\mu\nu}}{\partial\xi_\kappa \partial\xi_\rho} = a_{\mu\nu\kappa\rho} = a_{\mu\nu\rho\kappa}$. Wir müssen wieder den Faktor von $c_\mu c_\nu c_\rho$ berechnen und gleich Null setzen. Wir berücksichtigen, dass die Grössen $a_{\mu\nu\kappa\rho}$ in dem ersten und zweiten Paar von Indizes symmetrisch sind und erhalten

$$2(a_{\kappa\mu\nu\rho} + a_{\kappa\rho\mu\nu} + a_{\kappa\nu\rho\mu}) = a_{\mu\nu\rho\kappa} + a_{\rho\mu\nu\kappa} + a_{\nu\rho\mu\kappa}.$$

Dies ist die einzige Bedingung, die die $a_{\mu\nu\rho\kappa}$ erfüllen müssen. Sie ist notwendig und hinreichend dafür, dass die Differentialgleichungen durch unseren Ansatz noch in zweiter Näherung[28] erfüllt werden. Wir schreiben diese Gleichung in leicht verständlicher Abkürzung in der Form $2 \cdot \mathrm{I} = \mathrm{II}$. Dann wird $2(\mathrm{I} + \mathrm{II}) = 3\mathrm{II}$, und da $\mathrm{I} + \mathrm{II}$ ungeändert bleibt, wenn man das erste Indizespaar mit dem zweiten vertauscht,[29] so folgt auch $2(\mathrm{I} + \mathrm{II}) = 3\mathrm{I}$, d. h. $\mathrm{I} = \mathrm{II} = 0$ oder

$$a_{\kappa\mu\nu\rho} + a_{\kappa\rho\mu\nu} + a_{\kappa\nu\rho\mu} = 0. \tag{15}$$

Wir schreiben diesselbe Relation nochmals hin, indem wir κ mit ν vertauschen und addieren diese beiden Gleichungen. Dann wird $a_{\kappa\mu\nu\rho} + a_{\kappa\rho\mu\nu} + a_{\kappa\nu\rho\mu} + a_{\nu\mu\kappa\rho} + a_{\nu\rho\mu\kappa} + a_{\nu\kappa\rho\mu} = 0$ oder

$$a_{\kappa\mu\nu\rho} + a_{\kappa\rho\mu\nu} + a_{\kappa\nu\rho\mu} + a_{\nu\mu\kappa\rho} + a_{\nu\rho\mu\kappa} + a_{\mu\rho\kappa\nu} = a_{\mu\rho\kappa\nu} - a_{\kappa\nu\mu\rho}.$$

Linker Hand steht ein in den vier Indizes symmetrischer Ausdruck, also muss 30
auch der rechtsstehende Ausdruck symmetrisch sein, und da er sein Vorzeichen[30] bei Vertauschung der Indizespaare ändert, muss er verschwinden, d. h.

$$a_{\mu\nu\kappa\rho} = a_{\kappa\rho\mu\nu}.$$

Damit haben wir eine dritte wichtige Symmetrieeigenschaft der $a_{\mu\nu\kappa\rho}$ gefunden. Sie ist eine Folge der Differentialgleichung und auch wieder für sich notwendig und *hinreichend*[31] dafür, dass die Gleichung in zweiter[32] Annäherung erfüllt wird. Zu diesen Resultaten ist im wesentlichen schon Riemann gelangt.

§ 14. Aufsuchen neuer Invarianten; die Krümmung

Wir haben nun die Aufgabe, diejenigen Eigenschaften der zweidimensionalen Punktmannigfaltigkeit aufzusuchen, die von der Benennung der Punkte unabhängig sind. Solche Eigenschaften können nur durch Invarianten ausgedrückt

[28]"noch in zweiter Näherung" is a typed interlineation.

[29]"das erste Indizespaar mit dem zweiten vertauscht" was corrected from "die Indizes paarweise vertauscht".

[30]"sein Vorzeichen" was corrected by Hilbert from "seinen Wert".

[31]"hinreichend" was underlined in pencil. On the left hand page, Hilbert wrote in pencil: "nicht hinreichend!"

[32]"zweiter" was corrected from "erster".

werden. Darum müssen wir uns noch kurz mit der Frage befassen, wie man zu solchen Invarianten gelangt. Wir wissen schon, dass die quadratische Form

$$\sum_{\mu\nu} g_{\mu\nu} dx_\mu dx_\nu = \sum_{\mu\nu} g'_{\mu\nu} dx'_\mu dx'_\nu \left(= \sum_{\mu\nu} g_{\mu\nu} \dot{x}_\mu \dot{x}_\nu dp^2 \right)$$

eine Invariante, oder besser eine Kovariante ist, ebenso, wenn q_μ einen kovarianten Vektor bedeutet, die Linearform

$$\sum_{\mu} q_\mu dx_\mu = \sum_{\mu} q'_\mu dx'_\mu .$$

Die allgemeinste Definition einer Invariante sind wir aber oben schuldig geblieben und wollen sie nun nachholen. Die allgemeine Invariante ist, wenn sie keine höheren als zweite Ableitungen enthalten soll, eine solche Funktion I der Argumente

$$g_{\mu\nu}, g_{\mu\nu\kappa}, g_{\mu\nu\kappa\rho}, q_\mu, q_{\mu\rho},$$

31 dass sie ungeändert bleibt, wenn man statt derselben die gestrichenen Grössen $g'_{\mu\nu}$ etc. in Bezug auf neue Veränderliche $x'_i(x_1, x_2)$ einführt, d. h. dass

$$I\left(g_{\mu\nu}, g_{\mu\nu\kappa}, g_{\mu\nu\kappa\rho}, q_\mu, q_{\mu\rho}\right) = I\left(g'_{\mu\nu}, g'_{\mu\nu\kappa}, g'_{\mu\nu\kappa\rho}, q'_\mu, q'_{\mu\rho}\right)$$

wird. Man muss die analytische Technik der Invariantentheorie beherrschen, um Invarianten bilden zu können, speziell in unserem Fall, um die Tatsachen der Geometrie und später diejenigen der Physik überhaupt ausdrücken zu können; Gesetze und Tatsachen nämlich, die nur für ein spezielles Koordinatensystem gelten, interessieren uns gar nicht.

Wir haben übrigens alle Hilfsmittel bereitgestellt, um *allgemeine Invarianten zu bilden.* Der wichtigste, aber auch der schwierigste Schritt war derjenige der Einführung der Riemannschen Koordinaten, *da er unser Problem* auf das viel einfachere *zurückführte, projektive Invarianten aufzufinden. Nun wollen wir aus der wichtigsten uns bekannten Invariante, nämlich* $\sum g_{\mu\nu} dx_\mu dx_\nu$ *neue Invarianten ableiten.* Zu diesem Zweck führen wir Riemannsche Koordinaten ξ_1, ξ_2 ein. Dann wird unsere Invariante zu $\sum a_{\mu\nu} d\eta_\mu d\eta_\nu$, wobei $a_{\mu\nu} = \sum_{\alpha\beta} \frac{\partial x_\alpha}{\partial \xi_\mu} \frac{\partial x_\beta}{\partial \xi_\nu} g^*_{\alpha\beta}$ und $g^*_{\alpha\beta}(\xi_1, \xi_2) = g_{\alpha\beta}(x_1, x_2)$ ist, und η_μ kogredient zu ξ_μ sein soll. Nun entwickeln wir nach steigenden Potenzen der ξ_μ und erhalten

$$\sum_{\mu\nu} g_{\mu\nu} dx_\mu dx_\nu = \sum_{\mu\nu} \left(a_{\mu\nu} + \sum_{\kappa} a_{\mu\nu\kappa} \xi_\kappa + \frac{1}{2} \sum_{\kappa\rho} a_{\mu\nu\kappa\rho} \xi_\kappa \xi_\rho + \cdots \right) d\eta_\mu d\eta_\nu . \tag{15*}$$

Dieser ganze Ausdruck ist eine projektive Invariante, also muss, da die einzelnen Glieder rechter Hand sich bei projektiver Transformation nicht gegenseitig vermischen können, weil sie in den Variabeln von verschiedenem Grad sind,
32 jeder Summand für sich eine | Invariante sein. *Somit haben wir eine unendliche*

Reihe projektiver Invarianten erhalten. Die erste derselben erkennen wir wieder als diejenige, von der wir ausgingen; die zweite Invariante ist die Null. Sie ist in der Tat eine Invariante, freilich keine, die neue Tatsachen über unsere zweidimensionale Mannigfaltigkeit aussagt. Die dritte Invariante dagegen ist *die berühmte Riemannsche Invariante, die Krümmung*:

$$K = \sum_{\mu\nu\kappa\rho} a_{\mu\nu\kappa\rho}\xi_\mu\xi_\nu d\eta_\kappa d\eta_\rho. \tag{16}$$

Die hierdurch ausgedrückte Tatsache können wir auch so in Worte fassen, dass wir sagen: Die $a_{\mu\nu\kappa\rho}$ bilden in den μ, ν bezw. κ, ρ einen quadrato-quadratischen, d. h. einen 4×4-er[33] Tensor. Derselbe steht *in engstem Zusammenhang* mit dem Begriffe, den man in der Flächentheorie die *Gaußische*[34] *Krümmung* nennt, und die bekanntlich bei Verbiegungen der Fläche ungeändert bleibt.

Wir können nun im Folgenden nicht mehr alle Ueberlegungen explizit durchführen, sondern müssen uns mit einer Skizzierung des Gedankenganges begnügen. Wir wollen in (16) als Variabeln statt der willkürlichen $\xi_\mu\xi_\nu$ die wegen (15^*)[35] zu denselben kogredienten $(g^{\mu\nu}) = a^{\mu\nu}$ setzen. Durch Summation über $\kappa\rho$ erhalten wir aus dem 4×4-er Tensor den gewöhnlichen 4-Tensor

$$K_{\mu\nu} = \sum_{\kappa\rho} a_{\mu\nu\kappa\rho}(g^{\kappa\rho})_{\xi_i=0}.$$

Durch abermalige Summation über $\mu\nu$ erhalten wir dann die Invariante

$$K = \sum_{\mu\nu\kappa\rho} a_{\mu\nu\kappa\rho}\,(g^{\mu\nu})_{\xi_i=0}\,(g^{\kappa\rho})_{\xi_i=0} \;= (K_{\mu\nu}g^{\mu\nu})_{\xi_i=0}\,.^{\langle 36\rangle}$$

Wir wollen folgende Terminologie festsetzen. Es heisse

$$a_{\mu\nu\kappa\rho} = \text{Riemannscher Tensor}$$

33 $$K_{\mu\nu} = \text{Krümmungstensor}$$

$$K = \text{Krümmung}$$

Jetzt haben wir zwar eine, ja sogar, wie die Invariantentheorie lehrt,[37] die *einzige Invariante* aufgefunden, *die keine höheren, als zweite Ableitungen und diese nur linear enthält*, aber wir sind immer noch nicht am Ziel. Wir wollen nämlich diese *projektive Invariante* in den Riemannschen Koordinaten als *allgemeine Invariante* in den ursprünglichen Korrdinaten ausdrücken, d. h. wir wollen die Krümmung K als Funktion der $g_{\mu\nu}(x_1, x_2)$ und deren Ableitungen

[33]"4×4-er" was corrected from "16×16-er".

[34]"Gaußische" was corrected from "totale".

[35]"wegen (15^*)" was interlineated.

[36]The last term was added by Hilbert in pencil.

[37]*Rowe 2001*, p. 417, points out that this result was, in fact, only published in *Vermeil 1917*, see also note 40 below.

nach den ursprünglichen Variabeln x_1, x_2 darstellen. Die Riemannschen Koordinaten selbst sollen uns nur ein Hilfsmittel zur Auffindung dieser allgemeinen Invarianten sein. Wir hatten K definiert als[38]

$$\sum_{\mu\nu\kappa\rho}\left(\frac{\partial^2 a_{\mu\nu}}{\partial\xi_\kappa\partial\xi_\rho}g^{\mu\nu}g^{\kappa\rho}\right)_{\xi_i=0}.$$

Hierin ist $a_{\mu\nu}=\sum_{\alpha\beta}g^*_{\alpha\beta}(\xi_i)\frac{\partial x_\alpha}{\partial\xi_\mu}\frac{\partial x_\beta}{\partial\xi_\nu}$. Nun benutzen wir die Gleichung (14):

$$x_\kappa=\xi_\kappa-\frac{1}{2}\sum_{\mu\nu}\left(\left\{\begin{array}{c}\mu\nu\\ \kappa\end{array}\right\}\right)_0\xi_\mu\xi_\nu-\frac{1}{6}\sum_{\mu\nu\lambda}\left(\frac{\partial}{\partial x_\lambda}\left\{\begin{array}{c}\mu\nu\\ \kappa\end{array}\right\}\right.$$
$$\left.-2\sum_\sigma\left\{\begin{array}{c}\sigma\nu\\ \kappa\end{array}\right\}\left\{\begin{array}{c}\mu\lambda\\ \sigma\end{array}\right\}\right)_0\xi_\mu\xi_\nu\xi_\lambda+\cdots.$$

Also wird

$$\frac{\partial x_\alpha}{\partial\xi_\mu}=\delta_{\alpha\mu}-\sum_\rho\left(\left\{\begin{array}{c}\rho\mu\\ \alpha\end{array}\right\}\right)_0\xi_\rho+\frac{1}{3}\sum_{\rho\sigma}\left(-\frac{\partial}{\partial x_\sigma}\left\{\begin{array}{c}\rho\mu\\ \alpha\end{array}\right\}-\frac{1}{2}\frac{\partial}{\partial x_\mu}\left\{\begin{array}{c}\rho\sigma\\ \alpha\end{array}\right\}\right.$$
$$\left.+\sum_\lambda\left[\left\{\begin{array}{c}\lambda\mu\\ \alpha\end{array}\right\}\left\{\begin{array}{c}\rho\sigma\\ \lambda\end{array}\right\}+2\left\{\begin{array}{c}\lambda\sigma\\ \alpha\end{array}\right\}\left\{\begin{array}{c}\mu\rho\\ \lambda\end{array}\right\}\right]_0\right)\xi_\rho\xi_\sigma+\cdots.$$

Diesen Ausdruck setzen wir ein in

$$\frac{\partial^2 g^*_{\alpha\beta}(\xi_1,\xi_2)}{\partial\xi_\rho}=\sum_l\frac{\partial g_{\alpha\beta}(x_1,x_2)}{\partial x_l}\frac{\partial x_l}{\partial x_\rho}$$

und finden

$$\frac{\partial g^*_{\alpha\beta}(\xi)}{\partial\xi_\rho}\equiv\left(\frac{\partial g_{\alpha\beta}(x)}{\partial x_\rho}\right)_0-\sum_{\lambda\mu}\left(\frac{\partial g_{\alpha\beta}(x)}{\partial x_\lambda}\left\{\begin{array}{c}\mu\rho\\ \lambda\end{array}\right\}\right)_0\xi_\mu\ ,\ (\xi^2),\qquad(16')$$

wobei $\equiv,(\xi^2)$ bedeutet, daß nur Glieder ersten Grades in ξ_i berücksichtigt
33a. wurden.[39] | Hieraus folgt für[40]

$$\frac{\partial^2 g^*_{\alpha\beta}(\xi)}{\partial\xi_\rho\partial\xi_\sigma}\equiv-\sum_\iota\left(\frac{\partial g_{\alpha\beta}}{\partial x_\iota}\left\{\begin{array}{c}\rho\sigma\\ \iota\end{array}\right\}\right)_0+\left(\frac{\partial^2 g_{\alpha\beta}}{\partial x_\rho\partial x_\sigma}\right)_0\ ,\ (\xi).\qquad(16'')$$

Nun entwickeln wir $g_{\alpha\beta}(x)=g^*_{\alpha\beta}(\xi)$ nach steigenden Potenzen der ξ_i und erhalten aus

$$g^*_{\alpha\beta}(\xi)=\left(g^*_{\alpha\beta}\right)_0+\sum_\rho\left(\frac{\partial g^*_{\alpha\beta}}{\partial\xi_\rho}\right)_0\xi_\rho+\frac{1}{2}\sum_{\rho\sigma}\left(\frac{\partial^2 g^*_{\alpha\beta}}{\partial\xi_\rho\partial\xi_\sigma}\right)_0\xi_\rho\xi_\rho+\cdots$$

[38] In the following formula “$a_{\mu\nu}$” was corrected to “$\gamma_{\mu\nu}$” in pencil.

[39] The preceding half-sentence was interlineated.

[40] At the top of the page, Hilbert wrote in pencil: “Diese Rechnung ist durch Ueberlegungen zu ersetzen vgl. Vermeil Annalen”, see *Vermeil 1917*, *Vermeil 1918*, and *Rowe 2001*, pp. 417–418.

wegen (16') und (16")

$$g^*_{\alpha\beta}(\xi) = \left(g^*_{\alpha\beta}\right)_0 + \sum_\rho \left(\frac{\partial g_{\alpha\beta}}{\partial x_\rho}\right)_0 \xi_\rho \qquad (16"')$$
$$+ \sum_{\rho\sigma} \left(\frac{1}{2} \left(\frac{\partial^2 g_{\alpha\beta}}{\partial x_\rho \partial x_\sigma}\right)_0 - \cdot \frac{3}{2} \sum_\lambda \left(\frac{\partial g_{\alpha\beta}}{\partial x_\lambda} \left\{ \begin{matrix} \rho\sigma \\ \lambda \end{matrix} \right\} \right) \right)_0 \xi_\rho \xi_\sigma + \cdots .$$

Jetzt können wir aus (16'), (16") und (16"')

$$a_{\alpha\beta} = \sum_{\mu\nu} g^*_{\mu\nu}(\xi) \frac{\partial x_\mu}{\partial \xi_\alpha} \frac{\partial x_\nu}{\partial \xi_\beta}$$

als Potenzreihe in den ξ_i-Koordinaten berechnen. Dann finden wir in der Tat, dass die Koeffizienten der ersten Potenzen verschwinden. Die Koeffizienten der zweiten Potenzen sind die gesuchten Grössen

$$a_{\alpha\beta\rho\sigma} = \frac{\partial^2 a_{\alpha\beta}}{\partial \xi_\rho \partial \xi_\sigma},$$

und zwar sind es komplizierte Ausdrücke, welche die zweiten Ableitungen $\frac{\partial^2 g_{\mu\nu}}{\partial x_\kappa \partial x_\lambda}$ nur linear und die ersten Ableitungen nur quadratisch enthalten. Aus denselben bilden wir den Krümmungstensor $K_{\mu\nu} = \sum_{\rho\sigma} a_{\mu\nu\rho\sigma} g^{\rho\sigma}$ und finden

$$K_{\mu\nu} = \sum_\kappa \left(\frac{\partial}{\partial x_\nu} \left\{ \begin{matrix} \mu\kappa \\ \kappa \end{matrix} \right\} - \frac{\partial}{\partial x_\kappa} \left\{ \begin{matrix} \mu\nu \\ \kappa \end{matrix} \right\} \right) \qquad (16^*)$$
$$+ \sum_{\kappa\lambda} \left(\left\{ \begin{matrix} \mu\kappa \\ \lambda \end{matrix} \right\} \left\{ \begin{matrix} \lambda\nu \\ \kappa \end{matrix} \right\} - \left\{ \begin{matrix} \mu\nu \\ \kappa \end{matrix} \right\} \left\{ \begin{matrix} \kappa\lambda \\ \lambda \end{matrix} \right\} \right), \qquad (2)$$

und schliesslich erhalten wir

$$K = \sum_{\mu\nu} K_{\mu\nu} g^{\mu\nu}$$

als *allgemeine Invariante.*

Diese Invariante nimmt, weil sie die einfachste ist, die wir aus der bekann- 34
ten Invariante $\sum_{\mu\nu} g_{\mu\nu} dx_\mu dx_\nu$ als neue ableiteten, sowohl hier in der zweidimensionalen Geometrie, als auch später in der vierdimensionalen Physik eine ausgezeichnete Stellung ein.

Ueber ihre Bauart ist zu bemerken, dass die erste in ihr auftretende Summe linear ist in den zweiten Ableitungen der $g_{\mu\nu}$, und dass sie keine ersten Ableitungen enthält. Die zweite Summe dagegen ist quadratisch in den ersten Ableitungen. Sie enthält wiederum keine zweiten Ableitungen.

Der Gedankengang, der uns zu dieser neuen Invariante führte, war verhältnismässig einfach. Der gewaltige Formalismus und Rechenapparat entsteht erst dadurch, dass man die Riemannschen Koordinaten wieder eliminieren muss, weil sie eben nur ein Hilfsmittel sind, um diese Invariante aufzufinden. Gauss hat in seiner „curvatura integra“ auf ganz anderem Wege, aber nur für den Fall der Fläche, d. h. der zweidimensionalen Mannigfaltigkeit die Krümmung schon berechnet.[41] Die Riemannsche Rechnungsweise dagegen, welcher wir uns angeschlossen haben, lässt sich für eine n-dimensionale Mannigfaltigkeit durchführen. Sie findet sich in Riemanns „ungekrönter Preisarbeit“: „Commentatio nova etc.“.[42]

§ 15. Beispiel: Berechnung der Krümmung der Kugel

Die allgemeine Theorie wollen wir uns nun an einem Beispiel klar machen. Die Euklidische Geometrie ist hierzu nicht geeignet; denn für sie wird alles trivial einfach. Dann | sind nämlich alle g-Klammern Null, also werden auch alle $K_{\mu\nu} = 0$ und daher verschwindet die Invariante K ebenfalls identisch. Ein anderes Beispiel, das wir oben oft herangezogen haben, war
die Kugel: Für sie ist, wenn wir wieder $x_1 = \vartheta, x_2 = \varphi$ setzen

$$\begin{array}{llll} g_{11} = 1; & g_{12} = 0; & g_{22} = \sin^2\vartheta; & g = \sin^2\vartheta, \\ g^{11} = 1; & g^{12} = 0; & g^{22} = \frac{1}{\sin^2\vartheta}. & \end{array}$$

Die von Null verschiedenen g-Klammern sind

$$\left\{ \begin{array}{c} 22 \\ 1 \end{array} \right\} = -\sin\vartheta\cos\vartheta, \qquad \left\{ \begin{array}{c} 12 \\ 2 \end{array} \right\} = \left\{ \begin{array}{c} 21 \\ 2 \end{array} \right\} = \frac{\cos\vartheta}{\sin\vartheta}.$$

Hieraus berechnen wir

$$\begin{aligned} K_{11} &= \frac{\partial}{\partial\vartheta}\left\{ \begin{array}{c} 12 \\ 2 \end{array} \right\} + \left\{ \begin{array}{c} 12 \\ 2 \end{array} \right\}\left\{ \begin{array}{c} 21 \\ 2 \end{array} \right\} = \frac{\partial}{\partial\vartheta}\frac{\cos\vartheta}{\sin\vartheta} + \frac{\cos^2\vartheta}{\sin^2\vartheta} = 1, \\ K_{12} &= 0, \\ K_{22} &= -\frac{\partial}{\partial\vartheta}\left\{ \begin{array}{c} 22 \\ 1 \end{array} \right\} + \left\{ \begin{array}{c} 21 \\ 2 \end{array} \right\}\left\{ \begin{array}{c} 22 \\ 1 \end{array} \right\} + \left\{ \begin{array}{c} 22 \\ 1 \end{array} \right\}\left\{ \begin{array}{c} 12 \\ 2 \end{array} \right\} - \left\{ \begin{array}{c} 22 \\ 1 \end{array} \right\}\left\{ \begin{array}{c} 12 \\ 2 \end{array} \right\} \\ &= \frac{\partial}{\partial\vartheta}(\sin\vartheta\cos\vartheta) - \frac{\cos\vartheta}{\sin\vartheta}\sin\vartheta\cos\vartheta = -\sin^2\vartheta. \end{aligned}$$

Hiermit sind alle Komponenten des Krümmungstensors aufgefunden, und wir erhalten für die Invariante K den Wert

$$K = -1 \cdot 1 - \sin^2\vartheta \cdot \frac{1}{\sin^2\vartheta} = -2.$$

[41] See *Gauss 1828*.

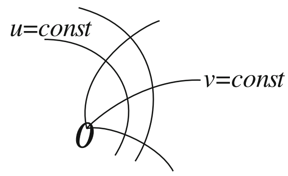
[42] *Riemann 1861* was submitted to the Academy in 1861 in response to a prize question concerning heat conduction. It did not win the prize which was withdrawn in 1868. Riemann's submission was only published posthumously in his *Collected Works*. For historical discussion and an English translation, see *Farwell and Knee 1990*.

Im geometrischen Sinne hat die Krümmung der Kugel vom Radius eins den Wert $+1$. Man hätte also, um die geometrische Krümmung zu finden, an dem K unserer Theorie den Faktor $-\frac{1}{2}$ anzubringen. Da dieser Faktor aber für unsere späteren Zwecke bedeutungslos ist, lassen wir ihn weg. Das Wesentliche, dass nämlich $K = \text{const.}$ wird, dass also auf der Kugel die Kongruenzsätze gelten, ist auch so zu erkennen.

§ 16. Gaussische Koordinaten 36

Nun wenden wir uns wieder der allgemeinen Theorie zu. Wir sind nämlich in der glücklichen Lage, im Falle der zweidimensionalen Mannigfaltigkeit die Formeln ganz *allgemein* noch wesentlich durchsichtiger gestalten zu können, und zwar greifen wir wieder zu dem bewährten Hilfsmittel der *Einführung neuer Koordinaten*. Schon eine flüchtige Ueberlegung lehrt, dass wir wahrscheinlich eine beträchtliche Vereinfachung werden erzielen können. Es sind uns doch drei willkürliche Funktionen $g_{\mu\nu}$ von x_1 und x_2 gegeben, die sich bei einer Transformation $u = u(x_1, x_2)$ und $v = v(x_1, x_2)$ wie ein kovarianter Tensor verhalten. An der Allgemeinheit der $g_{\mu\nu}$ dürfen wir natürlich vorerst nicht rütteln, ohne unerlaubte Spezialisierungen zu machen; wohl aber stehen uns noch die zwei Funktionen u und v von x_1 und x_2 zur freien Verfügung. Wir fragen also, wie sind dieselben zu wählen, damit die $g'_{\mu\nu}$ möglichst einfache Werte annehmen? Wir dürfen wohl schliessen, und die Rechnung wird die Richtigkeit dieser Ueberlegung alsbald bestätigen, *dass wir zwei Funktionen $g'_{\mu\nu}$ durch geeignete Wahl der frei verfügbaren Funktionen u und v ganz spezielle Werte erteilen können*: und zwar versuchen wir $g'_{11} = 1$, $g'_{12} = 0$ zu machen, worauf $g'_{22} = g'_{22}(u, v)$ immer noch eine ganz willkürliche Funktion von u und v bleibt. Der wesentliche Unterschied dieser Transformation gegenüber derjenigen der Einführung Riemannscher Koordinaten ist also der, dass wir jetzt versuchen, zwei Funktionen $g'_{\mu\nu}$ besonders einfach zu gestalten, während wir damals die Gleichungen der geodätischen Linie auf eine besonders ein|fache Form zu bringen trachteten. Dies sind zwei ganz verschiedene 37
Probleme. Trotzdem werden die jetzt einzuführenden Koordinaten mit den Riemannschen eine gewisse Aehnlichkeit haben.

Um die Rechnung durchzuführen, greifen wir aus unserer Geometrie einen beliebigen Punkt als 0-Punkt heraus und ziehen durch ihn alle geodätischen Linien. Solange die Punkte der Mannigfaltigkeit noch ihre alten Namen x_1, x_2 haben, sei die Gleichung der durch den 0-Punkt gehenden geodätischen Linie von der Form $f_2(x_1, x_2) = \text{const.}$ Führen wir also diese Funktion $f_2 = v$ als die eine neue Koordinate ein, so stellt $v = \text{const.} = a$, wenn wir die Konstante a noch willkürlich lassen, die Schar der durch den 0-Punkt gehenden geodätischen Linien dar. Wir werden sehen, dass durch diese Wahl der einen Koordinate $g'_{11} = \text{const.}$ gemacht wird. Die zweite Veränderliche $u = f_1(x_1, x_2)$ oder — wenn wir aus $v = f_2(x_1, x_2)$ die Funktion $x_2 = \varphi(v, x_1)$ berechnen und in u einsetzen — die unbekannte Funktion $u = U(x_1, v)$ bezw. $x_1 = \Phi(u, v)$ bestimmen wir dann aus der partiellen Differentialgleichung $g'_{12} = (u, v) = 0$. Die hierbei auftretende willkürliche Funktion legt die Kurve in der zweidimensionalen Mannigfaltigkeit fest.



Wir werden zeigen, dass $g'_{12}(u, v) = 0$ darauf hinauskommt, dass die Kurven $u =$ const. die orthogonalen Trajektorien der Kurven $v =$ const. sind. Doch müssen wir hierzu erst den *Begriff des*
38 *Winkels* definieren. | Hiermit haben wir *ein Koordinatensystem* eingeführt, *das dem der Polarkoordinaten* $r(= u)$, $\varphi(= v)$ in der Euklidischen Geometrie *analog ist*. Wir müssen noch beweisen, dass für dieses Bezugssystem in der Tat $g'_{11}(u, v) =$ const. wird:

Die Gleichung der geodätischen Linie bestimmt sich in der Form $v = f(u)$ aus dem Variationsproblem

$$\int_{u_1}^{u_2} \sqrt{g'_{11}(u, f(u)) + g'_{22}(u, f(u)) \left(\frac{dv}{du}\right)^2}\, du = \text{Min.}$$

Hierbei sind also g'_{11} und g'_{22} Funktionen von u allein.[43] Geodätische Linien, für welche $u =$ const. wird, können wir freilich auf diese Weise nicht erhalten. Ein Integral der zu diesem Problem gehörigen Lagrangeschen Differentialgleichung muss aber $v =$ const. sein, da dies geodätische Linien durch den 0-Punkt darstellt. Wir bilden also die Lagrangesche Ableitung und erhalten so als Differentialgleichung der geodätischen Linie

$$\frac{d}{du}\left(\frac{\partial}{\partial \frac{dv}{du}} \sqrt{g'_{11} + g'_{22}\left(\frac{dv}{du}\right)^2}\right) - \frac{\partial}{\partial f}\sqrt{g'_{11} + g'_{22}\left(\frac{dv}{du}\right)^2} = 0.$$

Führen wir die Differentiation aus, so ergibt sich

$$\begin{aligned} &\frac{d}{du}\left(\frac{1}{2}\frac{1}{\sqrt{g'_{11} + g'_{22}\left(\frac{dv}{du}\right)^2}} 2g'_{22}\frac{dv}{du}\right) \\ &\qquad - \frac{1}{2}\frac{1}{\sqrt{g'_{11} + g'_{22}\left(\frac{dv}{du}\right)^2}}\left(\frac{\partial g'_{11}}{\partial f} + \left(\frac{dv}{vu}\right)^2 \frac{\partial g'_{22}}{\partial f}\right) = 0. \end{aligned}$$

Hierin setzen wir die Lösung $v =$ const., d.h. $\frac{dv}{du} = 0$ ein und erhalten $\frac{\partial g'_{11}}{\partial f} = 0$, d. h. $g'_{11} = \varphi(U)$. Nun hat also das Linienelement die gewünschte Gestalt, wobei $\int^u \sqrt{\varphi(u)}du = U$ ist,

$$ds^2 = \left\{1 + \hat{g}_{22}(U, \hat{f}(U))\left(\frac{dv}{dU}\right)^2\right\} dU^2$$

[43] In the left margin, Hilbert wrote in pencil: "' ' überall weglassen der Einfachheit halber".

oder, wenn die Kurve wieder in der Parameterdarstellung $u = u(p)$, $v = v(p)$ gegeben ist,

$$ds^2 = \left\{ u^2 + g'_{22}(u,v)\dot{v}^2 \right\} dp^2,$$

worin g'_{22} wieder Funktion von u und v ist. Trotzdem das | Linienelement 39
nun eine ganz spezielle Form hat, haben wir also, um das nochmals zu betonen, keine Spezialisierung unserer Geometrie vorgenommen. Die hier skizzierte Rechnung wurde schon von *Gauss* durchgeführt.[44] Derselben Schlussweise bedient man sich übrigens auch in der projektiven Geometrie, um eine Form auf ihre einfachste Gestalt zu bringen. Setzen wir hier speziell $g_{22} = \sin^2 u$, so erhalten wir wieder die Geometrie auf der Kugel.

In der Physik werden zwar ähnliche Ueberlegungen wie die hier angestellten auch möglich sein, doch kann man dort leider durch spezielle Wahl des Koordinatensystems nicht entfernt dieselbe Vereinfachung erzielen. In der Tat haben wir dort entsprechend den vier Dimensionen 10 willkürliche Funktionen $g_{\mu\nu}$ und nur 4 Funktionen $x'_i(x_1, x_2, x_3, x_4)$ zu unserer freien Verfügung.

Nun werden sich *die Komponenten $K_{\mu\nu}$ des Krümmungstensors* besonders *einfach berechnen* lassen, so dass wir den Wert dieser Transformation, die die Rechnung so sehr vereinfacht, zu ermessen vermögen. Wir haben also $ds^2 = du^2 + g_{22}dv^2$ und müssen hieraus zunächst wieder die g-Klammern bilden, in denen wir sozusagen die Bausteine des Krümmungstensors erkannt haben. Dieselben müssen sich nun durch g_{22} und dessen beide Ableitungen nach u und v ausdrücken lassen. Wir benutzen wieder die Methode, die Gleichungen der geodätischen Linie in ihre Normalform

$$\left.\begin{matrix} \ddot{u} \\ \ddot{v} \end{matrix}\right\} + \text{homogene quadratische Form in } \dot{u} \text{ und } \dot{v} = 0 \text{ zu bringen.}$$

Aus

$$\int_{p_1}^{p_2} \left(\dot{u}^2 + g_{22}\dot{v}^2\right) dp = \text{Minimum}$$

erhalten wir als Lagrangesche Ableitungen 40

$$\begin{gathered} \ddot{u} - \frac{1}{2}\frac{\partial g_{22}}{\partial u}\dot{v}^2 = 0, \\ \frac{d}{dp}\left(2g_{22}\dot{v}\right) - \frac{\partial g_{22}}{\partial v}\dot{v}^2 = 0 \end{gathered}$$

oder $(u = x_1, v = x_2)$

$$\ddot{u} - \frac{1}{2}g_{221}\dot{v}^2 = 0, \qquad \ddot{v} + \frac{1}{g_{22}}g_{221}\dot{u}\dot{v} - \frac{1}{2g_{22}}g_{222}\dot{v}^2 = 0.$$

[44] See *Gauss 1828*.

Hieraus kann man die g-Klammern ablesen:

$$\left\{ \begin{matrix} 22 \\ 1 \end{matrix} \right\} = -\frac{1}{2} g_{221},$$

$$\left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\} = \left\{ \begin{matrix} 21 \\ 2 \end{matrix} \right\} = \frac{1}{2g_{22}} g_{221}; \qquad \left\{ \begin{matrix} 22 \\ 2 \end{matrix} \right\} = \frac{1}{2g_{22}} g_{222}.$$

Damit haben wir alle Stücke, aus denen die $K_{\mu\nu}$ aufgebaut sind. Wir finden

$$K_{11} = \frac{\partial}{\partial u} \left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\} + \left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\} \left\{ \begin{matrix} 21 \\ 2 \end{matrix} \right\} = \frac{1}{2} \frac{g_{22} g_{2211} - g_{221}^2}{g_{22}^2} + \frac{1}{4} \frac{g_{221}^2}{g_{22}^2},$$

$$K_{11} = \frac{1}{2g_{22}} \left\{ g_{2211} - \frac{1}{2g_{22}} g_{221}^2 \right\}.$$

K_{12} braucht zwar gar nicht berechnet zu werden, weil $g^{12} = 0$ ist, so dass K_{12} in der Formel für die Krümmung nicht auftritt. Tun wir es trotzdem, so finden wir

$$K_{12} = 0.$$

Schliesslich wird

$$\begin{aligned} K_{22} = {} & \frac{\partial}{\partial v} \left\{ \begin{matrix} 22 \\ 2 \end{matrix} \right\} - \frac{\partial}{\partial u} \left\{ \begin{matrix} 22 \\ 1 \end{matrix} \right\} - \frac{\partial}{\partial v} \left\{ \begin{matrix} 22 \\ 2 \end{matrix} \right\} + \left\{ \begin{matrix} 21 \\ 2 \end{matrix} \right\} \left\{ \begin{matrix} 22 \\ 1 \end{matrix} \right\} \\ & + \left\{ \begin{matrix} 22 \\ 1 \end{matrix} \right\} \left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\} + \left\{ \begin{matrix} 22 \\ 2 \end{matrix} \right\} \left\{ \begin{matrix} 22 \\ 2 \end{matrix} \right\} - \left\{ \begin{matrix} 22 \\ 1 \end{matrix} \right\} \left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\} \\ & - \left\{ \begin{matrix} 22 \\ 2 \end{matrix} \right\} \left\{ \begin{matrix} 22 \\ 2 \end{matrix} \right\} = \frac{1}{2} g_{2211} - \frac{1}{4g_{22}} g_{221}^2 = g_{22} K_{11}, \end{aligned}$$

so dass wir für

$$K = K_{11} \cdot 1 + K_{22} \cdot \frac{1}{g_{22}} = 2K_{11} = \frac{1}{g_{22}} \left\{ g_{2211} - \frac{1}{2g_{22}} g_{221}^2 \right\}$$

erhalten. Nun kann K_{11} noch auf eine einfachere Form gebracht werden. Wir setzen $g_{22} = \gamma^2$ dann wird $g_{221} = 2\gamma\gamma_u$ $g_{2211} = 2\gamma_v^2 + 2\gamma\gamma_{uu}$ dann wird

$$K_{11} = \frac{1}{2} \frac{2\gamma_u^2 + 2\gamma\gamma_{uu}}{\gamma^2} - \frac{1}{4} \frac{4\gamma^2\gamma_u^2}{\gamma^2} = \frac{1}{\gamma} \gamma_{uu}$$

und wir erhalten als *Schlussresultat*

$$K = \frac{2}{\gamma} \frac{\partial^2 \gamma}{\partial u^2}. \tag{17}$$

41 Setzen wir statt $\frac{1}{\gamma}\gamma_{uu}$ seinen Wert $\frac{1}{2}K$ in die Ausdrücke | für die Komponenten des Krümmungstensors ein, so verifizieren wir die Formel

$$K_{\mu\nu} = \frac{1}{2} K g_{\mu\nu},$$

d. h. der Krümmungstensor ist dem Tensor $g_{\mu\nu}$ selbst proportional. Dies ist aber nur im binären Gebiet der Fall!

§ 18. Zusammenhang der Invariante K mit der Gaussischen Krümmung der Fläche

Dier hiermit abgeleitete einfachste Ausdruck für K setzt uns in den Stand, den innigen

Zusammmenhang zwischen der Gauss'schen Krümmung und der Invariante K

aufzuklären. Wir können nämlich nachweisen, dass der Wert von K *für irgend einen Punkt der Fläche*, auf welcher wir Geometrie treiben, bis auf den konstanten Faktor -2 *identisch* ist *mit dem* Wert des *Gauss'schen Krümmungsmasses* auf dieser Fläche. Damit sehen wir ein, dass wir berechtigt sind, die Invariante K als Krümmung unserer zweidimensionalen Mannigfaltigkeit anzusprechen. Ein gewisser Zusammenhang zwischen den beiden Begriffen wurde uns schon oben dadurch nahegelegt, dass wir für die Kugel, von der wir wissen, dass ihre Gauss'sche Krümmung konstant ist, auch die Konstanz der Invariante K nachgewiesen haben.

Wir denken uns also eine beliebige Fläche in ihrer einfachsten Form $Z = Z(x, y)$ im Euklidischen Raum gegeben. Die Länge S der Verbindunglinie zweier Flächenpunkte ist dann definiert als das Integral ($u, v =$ Koordinaten auf der Fläche)

$$S = \int \sqrt{dx^2 + dy^2 + dz^2} = \int \sqrt{g_{11} + 2g_{12}\frac{dv}{du} + g_{22}\left(\frac{dv}{du}\right)^2}\, du.$$

Die Fläche selber denken wir uns nun im Raum in solcher Lage, dass der Punkt, in dem die Krümmung berechnet werden soll, in | den 0-Punkt des 42
räumlichen x, y, z-Koordinatensystems fällt, und dass die x, y-Ebene in jenem Punkt die Tangentialebene der Fläche ist. Um die z-Achse, d. h. die Flächennormale im Punkt als Achse werde die Fläche dann noch so gedreht, dass in der Gleichung der Fläche, entwickelt nach steigenden Potenzen von x und y für die Umgebung des Punktes das Glied mit $x \cdot y$ verschwindet. Dann hat die Gleichung der Fläche, sofern sie keine Ebene ist, immer die Gestalt $Z = \frac{a}{2}(x^2 \pm y^2) + \cdots$, so dass *die einzigen, eine beliebige Fläche in irgend einem ihrer Punkte beschreibenden Grössen die Konstante a* und *das Vorzeichen von y^2* sind. Setzen wir noch

$$x = u, \qquad y = v, \qquad Z = \frac{a}{2}(u^2 \pm v^2) + \cdots\,, \tag{18}$$

so haben wir in der Umgebung dieses Punktes die Fläche in der Parameterdarstellung $x_i = x_i(u, v)$. Dann wird gemäss unseren allgemeinen Formeln

$$g_{11} = \sum \left(\frac{\partial x_i}{\partial u}\right)^2 = 1 + a^2u^2 + \cdots\,, \qquad g_{12} = \sum \frac{\partial x_i}{\partial u}\frac{\partial x_i}{\partial v} = \pm a^2uv + \cdots\,,$$

$$g_{22} = \sum \left(\frac{\partial x_i}{\partial v}\right)^2 = 1 + a^2v^2 + \cdots\,,$$

und wir erhalten

$$g_{11} + 2g_{12}\frac{dv}{du} + g_{22}\left(\frac{dv}{du}\right)^2 = 1 + a^2\left(u \pm v\frac{dv}{du}\right)^2 + \left(\frac{dv}{du}\right)^2.$$

Von diesem Ausdruck bildet man die Lagrangesche Ableitung und kommt so zu den g-Klammern. Aus diesen und aus ihren Ableitungen kann man dann in bekannter Weise die Tensorkomponenten $K_{\mu\nu}$ aufbauen. Wir sehen übrigens, dass im 0-Punkt selber die sämtlichen Ableitungen der $g_{\mu\nu}$ und daher auch alle g-Klammern verschwinden, so dass nur die Ableitungen der g-Klammern Beiträge zum Krümmungstensor liefern. Ferner ersieht man, dass
43 der Wert der Invariante K im willkürlichen 0-Punkt | allein eine Funktion von a und dem Vorzeichen von $a^2uv\frac{dv}{du}$ ist, also vollkommen unabhängig davon, wie die Fläche ausserhalb der Umgebung des Punktes beschaffen sein mag. Um herauszufinden, *was für eine Funktion* von a und diesem Vorzeichen sie ist, werden wir daher eine ganz spezielle, möglichst einfache Fläche auswählen. Die Ebene hatten wir schon oben ausgeschlossen, und so ziehen wir wieder *die Kugel* in unsere Betrachtung.

Für die Kugel vom Radius eins hat K, wie wir schon wissen, den Wert $K = -2$. Für die Kugel vom Radius r wollen wir K von neuem berechnen, indem wir den einfachen Ausdruck (17) benutzen, um zu zeigen, wie sehr wir uns die Rechnung nun gegen früher erleichtert haben. Das allgemeine räumliche Linienelement hat in Polarkoordinaten die Form

$$ds^2 = dr^2 + r^2dv^2 + r^2\sin^2\vartheta d\varphi^2,$$

also ist auf der Kugel vom Radius r $ds^2 = r^2d\vartheta^2 + r^2\sin^2\vartheta d\varphi^2$. Damit K durch (17) dargestellt wird, muss $g_{11} = 1$ werden. Wir setzen also $r\vartheta = u$, $\varphi = v$ und erhalten

$$ds^2 = du^2 + r^2\sin^2\frac{u}{r}dv^2.$$

Hierin ist $\gamma = \sqrt{g_{22}} = r\sin\frac{u}{r}$ und es wird

$$K = \frac{2}{r\sin\frac{u}{r}}r\frac{\partial^2}{\partial u^2}\left(\sin\frac{u}{r}\right) = -\frac{2}{r^2}.$$

Nun ist noch der *Zusammenhang zwischen dem a der allgemeinen Fläche und dem r der Kugel* festzustellen. Dies geschieht, indem wir die Gleichung der Kugel in der Umgebung eines Punktes auf die Gestalt (18) bringen. Die Kugel, die im 0-Punkt die x, y-Ebene zur Tangentialebene hat, wird durch die Gleichung

$$x^2 + y^2 + (z - r)^2 = r^2$$

44 dargestellt. Hieraus folgt $x^2 + y^2 + z^2 - 2rz = 0$. | Dies ist eine quadratische Gleichung für z. Ihre Wurzeln sind $z = r - \sqrt{r^2 - x^2 - y^2}$. Für $x = 0$, $y = 0$

soll $z = 0$ sein, also gilt nur das negative Zeichen vor der Wurzel, und wir erhalten, wenn wir nach steigenden Potenzen von x und y entwickeln

$$Z = r \pm r \left(1 - \frac{x^2 + y^2}{r^2}\right)^{\frac{1}{2}} = \frac{1}{2}\frac{x^2 + y^2}{r^2}.$$

Durch den Vergleich mit (18) finden wir $a = \frac{1}{r}$ und wir erhalten

$$K = -2 \cdot a^2$$

als die gesuchte Funktion $K(a)$ in einem beliebigen Punkt einer willkürlichen Fläche, in dem in (18) das positive Vorzeichen gilt. Ohne Beweis teilen wir mit, dass

$$K = -2(-a^2)$$

wird in einem Flächenpunkt mit negativem Vorzeichen von y^2.

§ 19. Die Hauptkrümmungsradien einer Fläche

Der Grösse $\pm a^2$ kann man eine anschauliche *geometrische Bedeutung* geben, wenn man den Begriff der *Hauptkrümmungsradien* in einem Flächenpunkte einführt, den wir folgendermassen erläutern wollen. Wir denken uns wieder die x, y-Ebene als Tangentialebene der Fläche in dem betrachteten Punkt und die Z-Achse wieder als Normale in dem selben. Dann schneidet jede Ebene des durch die Z-Achse gehenden Bündels die Fläche in einer Kurve. Der die Kurve in diesem Punkte oskulierende Kreis habe den Radius ρ, dessen Grösse eine Funktion des Winkels ist, welchen die betreffende Schnittebene mit der x-Achse bildet. Das Produkt seiner beiden Extremwerte[45] $\rho_1 \cdot \rho_2$ findet man, wenn die Fläche in der Form $z = z(x, y)$ gegeben ist, nach einem Satze der Flächentheorie zu 45

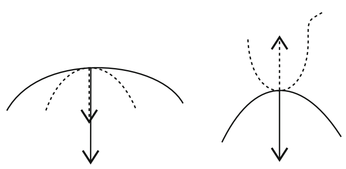
$$\rho_1\rho_2 = \frac{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}{\frac{\partial^2 z}{\partial x^2}\frac{\partial^2 z}{\partial y^2} - \frac{\partial^2 z}{\partial x \partial y}}.$$

Den Ausdruck rechter Hand können wir in der Umgebung eines Punktes (18) berechnen und erhalten $\rho_1\rho_2 = \pm\frac{1}{a^2}$ also ergibt sich das wichtige Endresultat

$$K = -2\frac{1}{\rho_1\rho_2}. \tag{19}$$

Der Wert des Produktes $\rho_1\rho_2$ ist positiv oder negativ, je nachdem

[45]“Das Produkt seiner beiden Extremwerte” was corrected from “Seine beiden Extremwerte”.



die beide Hauptkrümmungradien in dem Flächenpunkt nach derselben oder nach verschiedenen Seiten der Fläche wiesen. Im ersten Fall gilt in (18) das positive, im zweiten das negative Vorzeichen. *Die Fläche heisst* in diesem Punkte bezw. *elliptisch* und *hyperbolisch gekrümmt*. Bezeichnen wir die Krümmung im Gauss'schen Sinne mit $G = \frac{1}{\rho_1 \rho_2}$, so hat die Riemannsche Invariante K den Wert $K = -2G$. Damit haben wir ein ganz *merkwürdiges Resultat* erhalten:

Die Gauss'sche Krümmung G hat die geometrische Bedeutung: $\frac{1}{\rho_1 \cdot \rho_2}$; sie ist also ein Begriff, den man mit der Fläche erst dann verbinden kann, wenn man sie sich im dreidimensionalen Euklidischen Raum gegeben denkt. Diese Krümmung haben wir nun, und dies ist ein höchst *überraschendes Ergebnis*, durch den Tensor der $g_{\mu\nu}$ und deren Ableitungen nach den beiden Parametern u, v der Fläche darstellen können. Dies sind aber lauter Grössen, die mit der *Lage* der Fläche im Raum nicht das Mindeste zu tun haben, die vielmehr der „geo-
46 metria intrinsica“ angehören, d. h. der Geometrie der Fläche, wenn wir sie als zweidimensionale Punktmannigfaltigkeit deuten, in welcher jedem Punkt u, v drei Funktionswerte $g_{\mu\nu}$ zugeordnet sind. Durch diese drei Funktionen wird der Begriff der Länge einer zwei Punkte der Mannigfaltigkeit verbindenden Kurve definiert. Nun wird es klar dass *die Gauss'sche Krümmung bei Verbiegungen der Fläche invariant* bleibt. Verbiegung einer Fläche nennt man nämlich jede solche Lagenänderung der Fläche im Raum, dass eine zwei Flächenpunkte verbindende Kurve ihre Länge beibehält. Zwei Flächen sind dann und nur dann ineinander verbiegbar, wenn es durch Einführung geeigneter Parameter gelingt, dass die Funktionen $g_{\mu\nu}$ in entsprechenden Flächenpunkten dieselben Werte erhalten. *Vom Standpunkt der Geometrie auf zweidimensionalen Punktmannigfaltigkeiten sind aber zwei Flächen*, denen dieselben Funktionen $g_{\mu\nu}$ eigen und *die* daher *ineinander verbiegbar sind, einander vollkommen identisch.* Für sie hat natürlich die Riemannsche Invariante K denselben Wert. Vom Standpunkt desjenigen aber, der Flächentheorie treibt, der in der Fläche also ein im dreidimensionalen Raum liegendes zweidimensionales Gebilde sieht, sind zwei Flächen, die durch Verbiegung auseinander hervorgehen, ganz verschieden. Auch die Hauptkrümmungsradien haben in entsprechenden Punkten jeder einzeln durchaus nicht denselben Wert, ihr Produkt aber bleibt bei Verbiegungen ungeändert; denn es ist allein eine Funktion von Grössen der geometria intrinsica. Von deren Standpunkt betrachtet bedeutet aber eine Verbiegung überhaupt nichts, da die $g_{\mu\nu}$ ja ungeändert bleiben. Sie ist
47 nicht einmal eine Koordinatentransformation, | da sogar die Koordinaten eines Punktes auf der Fläche dieselben bleiben. — Der Beweis der Invarianz der

Krümmung bei Verbiegung einer Fläche ist der Hauptsatz der schon zitierten „curvatura integra" von Gauss.[46]

In der Flächentheorie ist es von Wichtigkeit zu wissen, ob man jede durch drei willkürlich vorgegebene Funktionen $g_{\mu\nu}$ bestimmte zweidimensionale Geometrie als Geometrie auf einer Fläche im Euklidischen Raum deuten kann. Diese Frage ist in der Tat, aber nur für den Fall der zweidimensionalen Punktmannigfaltigkeit (drei Funktionen $g_{\mu\nu}$), zu bejahen. D. h. also, man kann die Fundamentalgrössen $g_{\mu\nu}$ der Fläche beliebig vorgeben, und es wird *die Theorie der Geometrie der zweidimensionalen Punktmannigfaltigkeit identisch mit der Flächentheorie* im dreidimensionalen Raum. Für uns ist das aber ganz gleichgültig, da wir die Fläche nur zur Veranschaulichung herangezogen haben.

§ 20. Die Flächen konstanter Krümmung

Wir machen eine weitere Anwendung von der einfachen Form, auf die wir die Krümmung gebracht haben und fragen: Welches sind *die allgemeinsten Geometrien, bei denen*[47]

$$K = \text{const} = 2k^2$$

wird? Dann muss $\frac{1}{\gamma}\frac{\partial^2\gamma}{\partial u^2} = k^2$ sein. Für $k^2 \neq 0$ unterscheiden wir die zwei wesentlichen Fälle:

1) Für $k^2 = -1$ wird $\gamma = \sin u$. Dies gibt das Linienelement der *Kugel* und aller duch Verbiegung aus ihr hervorgehenden Flächen.

2) Für $k^2 = +1$ wird $\gamma = e^u$ bezw. $= \sinh u$. | In diesem Fall, den wir noch nicht 48
betrachtet haben, ist die Gauss'sche Krümmung der Fläche negativ konstant. Flächen, die diese Geometrie realisieren, sind sattelförmig. *Die Rotationsfläche der Traktrix*, auf der diese Geometrie auch gilt, nennt man wegen ihres negativen konstanten Krümmungsmasses wohl auch *Pseudosphäre*.[48]

Dass die Krümmung auf der Kugel konstant sein muss, wissen wir; denn die Kugel ist ja in sich transformierbar. Es frägt sich nun, ob bei der Pseudosphäre ebenfalls eine solche Transformation in sich möglich ist, worauf es selbstverständlich wird, dass die Krümmung derselben konstant ist. Diese Frage ist zu bejahen. In der Tat ist deren Linienelement $ds^2 = du^2 + \sinh^2 u dv^2$ das Linienelement der berühmten Bolyai-Lobatscheffskyschen Geometrie, auf die wir noch zu sprechen kommen werden.[49]

[46] Cf. Gauss's *theorema egregium*: "Si superficies curva in quamcunque aliam superficiem explicatur, mensura curvaturae in singulis punctis invariata manet." ("If a curved surface is developed upon any other surface whatever, the measure of curvature in each point remains unchanged.") *Gauss 1828*, § 12, *Gauss 1965*, p. 20.

[47] In the following formula, $K = \text{const}$ is underlined.

[48] See § 23 below.

[49] See § 22 below.

§ 21. Definition des Winkels

In jeder Geometrie spielt neben der Geraden oder kürzesten Linie der Begriff des *Winkels* die wichtigste Rolle, und so müssen wir erst noch definieren, was wir darunter verstehen wollen. Als wir oben (S. 36 ff.) solche Koordinaten auf der Fläche einführten, dass $g_{12} = 0$ wurde, hätten wir zur geometrischen Interpretation dieser Differentialgleichung diesen Begriff schon kennen sollen. Wir wollen nun zeigen, dass der *Winkel um einen Punkt herum* trotz der wenigen Voraussetzungen, die wir zur Begründung der Geometrie nötig hatten, einwandfrei definiert werden kann. Zu diesem Zwecke greifen wir den willkür-
49 lichen Punkt $x_i = 0$ unserer Mannigfaltigkeit heraus. | Das Linienelement hat, in der Ausdrucksweise der Zahlentheorie geschrieben, in der Umgebung dieses Punktes die Form

$$ds^2 \equiv (g_{11})_0 dx_1^2 + 2(g_{12})_0 dx_1 dx_2 + (g_{22})_0 dx_2^2 + (x_\lambda dx_\mu dx_\nu).$$

Es ist also im wesentlichen Euklidisch. Durch eine homogene, lineare Transformation erteilen wir $g_{\mu\nu}$ den Wert $\delta_{\mu\nu}$, so dass

$$ds^2 \equiv dx_1^2 + dx_2^2 + (x_\lambda dx_\mu dx_\nu) \tag{20}$$

wird. Dabei ist von der Voraussetzung, dass die quadratische Form $\sum_{\mu\nu} g_{\mu\nu}$ $dx_\mu dx_\nu$ positiv definit ist, d. h. dass $g_{11} > 0 \quad g_{11}g_{22} - g_{12}^2 > 0$ ist, Gebrauch gemacht. Wir behaupten, dass *durch Substitution* $x_1 = 2\cos\varphi \quad x_2 = 2\sin\varphi$ der *Winkel* φ *eindeutig*, d. h. unabhängig von dem zugrundegelegten Koordinatensystem x_i *definiert* ist. Für kleine Werte von r geht (20) nun über in

$$ds^2 \equiv dr^2 + r^2 d\varphi^2 + (rdr^2, r^2 drd\varphi, r^3 d\varphi^2).$$

Um unsere Behauptung zu beweisen, müssen wir eine beliebige Transformation der $x_i = x_i(x_1' x_2')$ durchführen, bei der der 0-Punkt ungeändert bleibt, und zeigen, dass sich dabei φ nicht verändert. Durch eine solche Transformation wird

$$r = r(r', \varphi'), \qquad \varphi = \varphi(r', \varphi')$$

und wir haben nachzuweisen, dass $\varphi' = \varphi$ ist. Da die Transformation nur für kleine r' ausgeführt wird, entwickeln wir r und φ nach Potenzen von r'.

$$\begin{aligned} r &= f_1(\varphi')r' + f_2(\varphi')r'^2 + \cdots, \\ \varphi &= \varphi_0(\varphi') + \varphi_1(\varphi')r' + \varphi_2(\varphi')r'^2 + \cdots, \end{aligned}$$

wobei also f_i und φ_i Funktionen von φ' allein sind. Wir bilden

$$\begin{aligned} dr^2 &= \left\{ (f_1 + 2f_2 r' + \cdots)\, dr' + \left(\frac{\partial f_1}{\partial \varphi'} r' + \frac{\partial f_2}{\partial \varphi'} \partial\varphi' r'^2 + \cdots \right) d\varphi' \right\}^2, \\ d\varphi^2 &= \left\{ (\varphi_1 + 2\varphi_2 r' + \cdots)\, dr' + \left(\frac{\partial \varphi_0}{\partial \varphi'} + r' \frac{\partial \varphi_1}{\partial \varphi'} + \frac{\partial \varphi_2}{\partial \varphi'} r'^2 + \cdots \right) d\varphi \right\}^2. \end{aligned}$$

50 Die zweite Gleichung wird mit r^2 multipliziert, rechter Hand für r^2 der Ausdruck in r' und φ' eingesetzt und zur ersten addiert. Dann sammeln wir, was mit dr'^2 multipliziert ist, und dies muss $\equiv 1 \cdot dr'^2 + (r'dr'^2, r'^2dr'd\varphi', r'^3d\varphi'^2)$ ⟨sein⟩. Diese Gleichung liefert für $r' = 0$: $f_1^2(\varphi') = 1$, d. h. $f_1 = +1$, weil ja $r' > 0$ ist. Nun sammeln wir noch, was mit $d\varphi'^2$ mulitpliziert ist. Dies muss $\equiv r'^2 d\varphi'^2 + (r'dr'^2, r'^2dr'd\varphi', r'^3d\varphi'^2)$ ⟨sein.⟩ Die erste Gleichung liefert wegen $\frac{\partial f_1}{\partial \varphi'} = 0$ keinen Beitrag. Wir erhalten also nur $r'^2 \left(\frac{\partial \varphi_0}{\partial \varphi'}\right)^2 = r'^2$, d. h. $\left(\frac{\partial \varphi}{\partial \varphi'}\right)^2_{r'=0} = 1, \quad \varphi' = \pm\varphi$. Der Winkel φ ist also *eindeutig festgelegt bis auf den Richtungssinn* und seine Definition ist invariant, w.z.b.w. Nun kann man auch rechte Winkel definieren. Ferner kann man jetzt zeigen, dass *für* $g_{12} = 0$ die Koordinatenkurven *orthogonale Trajektorien* zueinander sind.

Wir wollen noch beweisen, dass die hier gegebene Definition des Winkels, wenn wir unsere Geometrie durch eine Fläche im Raum gewinnen, mit der *Euklidischen Definition des Winkels* auf der Fläche übereinstimmt: die Flächengleichung sei wieder in der Umgebung des betrachteten Punktes $z = \frac{a}{2}(x^2 + y^2) + \cdots$. Wir setzen $x = u$, $y = v$, $z = \frac{a}{2}(u^2 \pm v^2) + \cdots$, dann wird $ds^2 = dx^2 + dy^2 + dz^2 \equiv du^2 + dv^2 + (udu^2, \cdots)$. Dies ist aber die Gleichung (20), aus der wir den Winkel gewannen. Man sieht also, dass durch die drei Funktionen $g_{\mu\nu}$ mit einem Schlage Länge und Winkel um einen Punkt herum festgelegt sind. φ hat, dies liegt in seiner Definition, die | Periode 2π. Nun sind 51
wir mit der Begründung der Geometrie zu Ende. Wir haben eine Wissenschaft aufgebaut, die an der Erfahrung prüfbar ist. Die Experimente, die das Kind von Anfang an macht, sind schon eine solche Prüfung. Sie sind vom selben Charakter wie diejenigen des Physikers, nur lassen sie sich ohne komplizierte Apparate ausführen, und man bezeichnet sie daher als „Anschauung“. Wir haben freilich nicht die Anschauung herangezogen, wir sind vielmehr axiomatisch vorgegangen, haben uns auch über alle Schwierigkeiten hinweggesetzt, dadurch, dass wir einerseits absolute Stetigkeit und Differenzierbarkeit aller Funktionen vorausgesetzt haben und andererseits die ganze Analysis als gegeben hinnahmen.

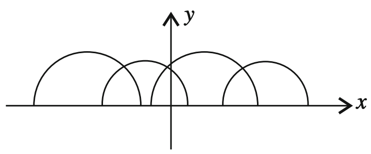
§ 22. Deutung der Bolyai-Lobatscheffsky'schen Geometrie in der Gaussischen Zahlenebene

Einzig in der Euklidischen ($G = 0$) und in den beiden Nichteuklidischen ($G = \pm 1$) Geometrien hat man die Möglichkeit, die zweidimensionalen Punktmannigfaltigkeiten in sich zu transformieren. Dann gehen geometrische Figuren von endlicher Ausdehnung in sich über. Auf einer Fläche, die eine solche Geometrie realisiert, kann man also ein biegbares Blechstück überall verschieben, ohne dass es verzehrt wird. Für die Kugel ($G = \pm 1$) ist dies unmittelbar klar. Wir wollen nun noch *analytisch* zeigen, dass auch

auf der *die Bolyai-Lobatscheffskysche Geometrie* darstellenden Pseudosphäre ($G = -1; ds^2 = du^2 + e^{2u} dv^2$) *Transformationen in sich möglich sind*, besonders, da dies auch in der Funktionentheorie von Interesse ist. Wir setzen $x = v$,
52 $y = e^{-u}$ und setzen nachher $z = x + iy$. | Dann wird $dx = dv$, $dy = -e^{-u} du$; $dx^2 + dy^2 = dv^2 + e^{-2u} du^2$, also ist $du^2 + e^{2u} dv^2 = (dx^2 + dy^2)e^{2u} = \frac{dx^2+dy^2}{y^2}$. Die Länge eines endlichen Kurvenstückes wird zu

$$\int ds = \int \frac{\sqrt{dx^2 + dy^2}}{y} = \int \frac{|dz|}{y}.$$

Das ist die bekannte Form des Nichteuklidischen Linienelementes bei einer Interpretation in der oberen Hälfte der Gauss'schen Zahlenebene.



Die Winkel, unter denen sich zwei Kurven auf der Pseudosphäre schneiden, stimmen mit denen der entsprechenden Bildkurven in der Ebene überein. Dagegen ist die Länge der Kurven nicht die wirkliche. Hier repräsentiert die x-Achse ($y = 0$) das Unendlichferne. Auf ihr liegen also die Enden aller ins Unendliche laufenden Kurven. Die x-Achse selber gehört daher nicht mehr zur Geometrie. Diese Deutung der Nichteuklidischen Geometrie verdankt ihre Berühmtheit dem Umstand, dass sich die *geodätischen Linien* in der komplexen Zahlenebene besonders einfach darstellen. Um ihre Gleichung aufzufinden, haben wir $\int \frac{\dot{x}^2+\dot{y}^2}{y^2} dp =$ Minimum zu machen. Zwei Integrale der zugehörigen Lagrangeschen Differentialgleichungen sind

$$\frac{\dot{x}^2 + \dot{y}^2}{y^2} = c, \qquad \frac{\dot{x}}{y^2} = C.$$

Hieraus erhalten wir $1 + \left(\frac{dy}{dx}\right)^2 = \frac{a}{y^2}$ oder

$$y^2 = -x^2 + \alpha x + \beta.$$

Dies sind die geodätischen Linien und zwar stellt diese Gleichung die zweifach
53 unendliche Mannigfaltigkeit der zur | *x-Achse orthogonalen Kreise* dar. Zwei Punkte der Mannigfaltigkeit sind also durch eine und nur eine gerade Linie zu verbinden.

Setzt man $x^2 + y^2 = u$, $x = v$, so ist die Gleichung aller zur x-Achse orthogonalen Kreise in diesen Koordinaten $u = \alpha v + \beta$ eine *lineare*, d. h. *sämtliche* geodätischen Linien der Geometrie, und nicht nur die durch einen Punkt gehenden werden durch eine *lineare Gleichung* zwischen *denselben* geeignet gewählten Parametern dargestellt.

Die Geraden dieser Geometrie haben eine unendliche Länge, weil sie bis an die x-Achse heranreichen; ganz im Gegensatz zur Riemann-Helmholtzschen

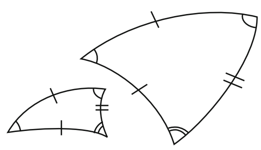
Geometrie, deren Gerade als grösste Kreise auf der Kugel in sich zurücklaufen und eine endliche Länge haben. Die *Kongruenzsätze* gelten hier, weil die Transformation in sich

$$z' = \frac{\alpha z + \beta}{\gamma z + \delta}$$

möglich ist. Trennen wir Real- und Imaginärteil, so folgt

$$x' = \frac{1}{(\gamma x + \delta)^2 + \gamma^2 y^2} \left\{ \alpha\gamma(x^2 + y^2) + (\alpha\delta + \beta\lambda)x + \alpha\delta \right\}, \tag{21}$$

$$y' = \frac{\alpha\delta - \beta\gamma}{(\lambda x + \delta)^2 + \lambda^2 y^2} y.$$



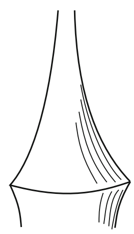
Die neue x'-Achse hat die Gleichung $y' = 0$, d. h. $y = 0$, wobei $\alpha\delta - \beta\gamma \neq 0$ vorausgesetzt ist. Die y'-Achse ist in der z-Ebene der Kreis $\alpha\gamma(x^2 + y^2) + (\alpha\delta + \beta\gamma)x + \alpha\delta = 0$. Kreise, die vor der Transfor-
mation die x-Achse orthogonal | schnitten, behalten 54
die Eigenschaft bei, d. h. Gerade bleiben Gerade. Die Transformation kann durch geeignete Wahl von α, β, γ, δ so eigerichtet werden, dass zwei Seiten eines beliebigen Dreiecks und der von ihnen eingeschlossene Winkel zur Deckung kommen mit gleich großen entsprechenden Stücken eines andern Dreiecks.[50] Dann stimmen auch die beiden anderen Winkel und die dritte Seite überein. Dies folgt aus der Invarianz des Integrals $\int_{p_1}^{p_2} ds$ bei dieser Transformation. Diese Invarianz ist leicht zu beweisen. Es ist nämlich

$$\frac{|dz'|}{|dz|} = \frac{\alpha\delta - \beta\gamma}{|(\gamma z + \delta)^2|} = \frac{\alpha\delta - \beta\gamma}{(\gamma x + \delta)^2 + \gamma^2 y^2} = \frac{y'}{y'},$$

also folgt

$$\frac{|dz'|}{y'} = \frac{|dz|}{y}.$$

§ 23. Deutung dieser Geometrie auf der Pseudosphäre



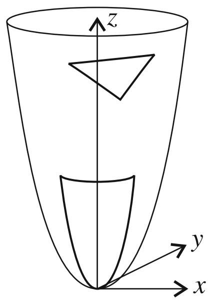
Die *Darstellung der Nichteuklidischen Geometrie auf der Pseudosphäre*, d. h. auf der Rotationsfläche der Traktrix hat gegenüber der funktionentheoretischen Darstellung in der komplexen Zahlenhalbebene den Nachteil, dass auf der Fläche nur ein Teil der in der Halbebene liegenden Punktmannigfaltigkeit abgebildet werden kann, der durch die Singularitäten der Fläche (Kante und im Unendlichen Spitze) begrenzt wird. Man wird so auf die Frage geführt, ob es im Raum eine singularitätenfreie

[50]"mit gleich großen entsprechenden Stücken eines andern Dreiecks" was corrected from "mit den entsprechenden Stücken eines kongruenten Dreiecks".

Fläche gibt, die die ganze Nichteuklidische Geometrie realisiert. Diese Frage muss, wie ohne Beweis erwähnt werden möge, verneint werden. Man kann die Fläche zwar so verbiegen, dass die ursprünglichen Singularitäten aufhören, solche zu sein, doch treten dann notwendigerweise neue auf.

55 § 24. Die Geometrie auf dem Rotationsparaboloid

Wir haben jetzt alle drei Typen von Geometrien mit konstanter Krümmung diskutiert, und wollen nun noch ein einfaches Beispiel einer *Fläche* geben, *deren Krümmung nicht konstant ist*, auf der also die Kongruenzsätze nicht gelten.



Als solche wählen wir das *Rotationsparaboloid*, dessen Scheitelpunkt im Ursprung des Koordinatensystems liegen und dessen Rotationsachse die z-Achse sein möge. Seine Gleichung ist $z = \frac{1}{2}(x^2 + y^2)$ Das Quadrat des Linienelementes wird

$$dx^2 + dy^2 + dz^2 = dx^2 + dy^2 + (xdy + ydy)^2,$$

d. h. es ist $g_{11} = 1 + x^2$, $g_{12} = xy$, $g_{22} = 1 + y^2$. Um das Linienelement auf die Form $du^2 + \gamma^2(u, v)dv^2$ zu bringen, müssen wir wieder verallgemeinerte Gauss'sche Polarkoordinaten (geodätische Linien durch irgend einen Punkt und deren orthogonale Trajektorien) einführen. Besonders einfach gestaltet sich die Rechnung, wenn wir als diesen Punkt den Scheitel des Paraboloids wählen. Die *geodätischen Linien* sind dann nämlich die Meridiankurven, deren Gleichung $\frac{x}{y} = \text{const}$ ist, und die orthogonalen Trajektorien sind die Schnittkreise der Ebenen $z = \text{const}$ mit dem Paraboloid. Wir setzen also $u = \frac{1}{2}(x^2 + y^2)$, $v = \frac{x}{y}$ und erhalten

$$\begin{aligned} du &= xdx + ydy, \\ y^2| \quad dv &= \frac{ydx - xdy}{y^2}. \end{aligned}$$

Die zweite Gleichung multiplizieren wir mit y^2, dann quadrieren wir die beiden Gleichungen, addieren sie zueinander und erhalten

$$du^2 + y^4 dv^2 = (x^2 + y^2)(dx^2 + dy^2).$$

56 Hieraus folgt

$$ds^2 = dx^2 + dy^2 + dz^2 = \frac{du^2 + y^4 dv^2}{x^2 + y^2} + du^2 = (1 + \frac{1}{2u})du^2 + \frac{y^4}{2u}dv^2.$$

Setzen wir noch $u' = \int \sqrt{1 + \frac{1}{2u}} du$, so erhält ds^2 die gewünschte Form $du' + \gamma^2(u', v) dv^2$. Wenn man hieraus K berechnet, so folgt in der Tat $K \neq \text{const.}$ Dass *von Kongruenzsätzen* nun *keine Rede* mehr ist, kann man sich leicht plausibel machen. Man denke sich nur ein rechtwinkliges Dreieck, dessen rechter Winkel im Scheitel des Paraboloid liegt, und desssen Schenkel sehr lang sein mögen. Dann wird die Hypothenuse, bei hinreichender Länge der Schenkel des rechten Winkels kleiner sein als eine Kathete. Denkt man sich dagegen einen rechten Winkel irgendwo auf dem Paraboloid weit entfernt vom Scheitel aufgetragen, so hat man wenigstens angenähert die Massverhältnisse der Euklidischen Geometrie, d. h. die Hypothenuse wird länger sein als eine Kathete.

§ 25. Das indefinite Linienelement der Pseudogeometrie

Wir haben nun die Geometrie der zweidimensionalen Punktmannigfaltigkeit unter der Voraussetzung, dass das Quadrat des Linienelements positiv definit ist, zu einem gewissen Abschluss gebracht. Da uns die zweidimensionale Geometrie aber nur ein Hilfsmittel sein soll, um später Physik oder die vierdimensionale Geometrie um so leichter diskutieren zu können, so müssen wir dieses Hilfsmittel unseren späteren Bedürfnissen möglichst anpassen. In der alten Physik hat nun das Linienelement die Form $dx^2 + dy^2 + dz^2 - dt^2$ und da auch in der neuen Physik | die alte als ein besonders einfacher Spezialfall, der 57
in der Wirklichkeit nur mit einer gewissen Annäherung realisiert sein wird, enthalten ist, wollen wir nun eine zweidimensionale Geometrie, die der vierdimensionalen im Falle, dass $g_{11} = 1$, $g_{22} = 1$, $g_{33} = 1$, $g_{44} = -1$ ist, möglichst ähnlich ist⟨,⟩ untersuchen. Bevor wir also von zwei Dimensionen zu drei und vier hinaufsteigen, müssen wir unsere Voraussetzung, dass $g_{11}g_{22} - g_{12}^2 > 0$ ist, fallen lassen, bezw. durch die andere $g_{11}g_{22} - g_{12}^2 > 0$ $g_{11} < 0$[51] ersetzen. Wir stellen uns daher das Problem, zu untersuchen, wie sich die zweidimensionale Geometrie modifiziert, wenn

das Quadrat des Linienelements indefinit[52]

ist. Wir kommen so zu ganz wesentlichen Aenderungen gegenüber früher, insbesondere müssen wir unsere Vorstellung von der Massbestimmung der Länge in dieser Pseudogeometrie[53] vollkommen ändern. Die *Kurvenlänge* definieren wir zwar wieder durch das Integral $\int_{p_1}^{p_2} \sqrt{\sum_{\mu\nu} g_{\mu\nu} dx_\mu dx_\nu}$. Die Wurzel wird nun aber nicht immer reell zu ziehen sein, das Vorzeichen der quadratischen Form in einem Kurvenpunkt $x_1(p)$, $x_2(p)$ hängt vielmehr von $\dot{x}_1(p)$, $\dot{x}_2(p)$ ab. Vor allen Dingen kann es jetzt auch *Kurven von der Länge Null*

[51] Should be "$g_{11}g_{22} - g_{12}^2 < 0$" and "$g_{11} > 0$".

[52] "indefinit" was corrected from "negativ".

[53] "in dieser Pseudogeometrie" was interlineated.

geben, während das früher nicht möglich war, ausser für $\dot{x}_1 = 0$, $\dot{x}_2 = 0$, und dies kann nur für diskrete Punkte eintreten. Jetzt erhält man diese Kurven, wenn man die Differentialgleichung $\sum_{\mu\nu} g_{\mu x}\dot{x}_\mu \dot{x}_\nu = 0$[54] integriert. Sind also $\sigma_1(x_1, x_2)$ und $\sigma_2(x_1, x_2)$ die beiden Wurzeln der quadratischen Gleichung $g_{11}\sigma^2 + 2g_{12}\sigma + g_{22} = \sigma$, so hat man die beiden Differentialgleichungen erster Ordnung

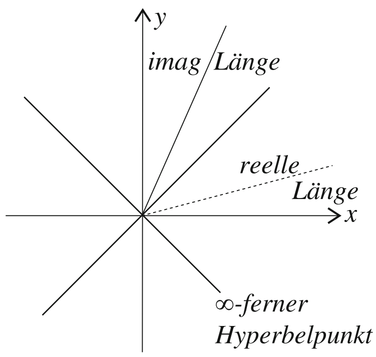
$$\dot{x}_1 - \sigma_i(x_1, x_2)\dot{x}_2 = 0 \qquad i = 1, 2 \tag{22}$$

58 zu lösen. Zu jedem Punkt x_1, x_2 unserer zweidimensionalen Mannigfaltigkeit gehören zwei bestimmte *reelle* Werte $\sigma_1(x_1, x_2)$ und $\sigma_2(x_1, x_2)$. Durch jeden Punkt der Mannigfaltigkeit gehen also zwei bestimmte Kurven $x_1^{(\sigma_1)}(p)$, $x_2^{(\sigma_1)}(p)$; $x_1^{(\sigma_2)}(p)$, $x_2^{(\sigma_2)}(p)$, die man aus (22) erhält, wenn man den beiden dort auftretenden Integrationskonstanten feste Werte erteilt. Entsprechend diesen beiden Integrationskonstanten gibt es zwei Scharen von Kurven, die (22) erfüllen. Es sind dies *reelle* Linien von der Länge Null, die wir daher *Nulllinien* nennen wollen. Wir haben also in unserer Geometrie gegenüber früher ein *total verändertes Bild*, da diese Nulllinien für den Fall des positiv definiten Linienelementes imaginär werden, ein Fall, der den uns von der Euklidischen Geometrie her geläufigen Anschauungen viel näher kommt.

§ 26. Die Pseudoeuklidische Geometrie

Wir wollen uns nun nur beim einfachsten Fall aufhalten und setzen $g_{11} = 1$, $g_{12} = 0$, $g_{22} = -1$. Dieser Fall ist auch insofern der einfachste, als für ihn wieder die Riemannsche *Invariante K verschwindet*. Insofern kommen wir hiermit der Euklidischen Geometrie am nächsten und nennen deswegen diese Geometrie die *pseudoeuklidische*. Dann wird die Kurvenlänge

$$s = \int_{p_1}^{p_2} \sqrt{dx^2 - dy^2}. \tag{23}$$



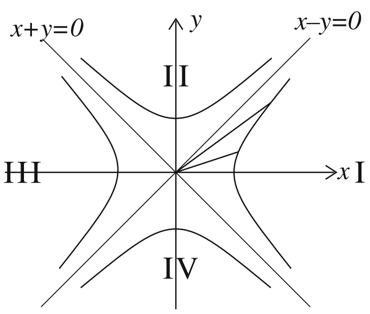
Die quadratische Gleichung zur Bestimmung von σ_1 und σ_2 lautet jetzt $\sigma^2 - 1 = 0$ und liefert als Wurzeln statt Funktionen von x_1, x_2 die Konstanten $\sigma_1 = +1$, $\sigma_2 = -1$.

Die Differentialgleichungen (22) werden zu $\frac{dx}{dp} \pm \frac{dy}{dp} = 0$ oder $x \pm y = \text{const}$. Dies sind *zwei Scharen von Geraden*, die aufeinander senkrecht stehen und die x-Achse un-
59 ter dem Winkel $\frac{\pi}{4}$ bezw. $\frac{3\pi}{4}$ | schneiden. Die *Länge* irgend einer anderen Kurve ist in einem ihrer Punkte nach (23) *reell oder imaginär*, je nachdem $\frac{dy}{dx} < 1$

[54] “$g_{\mu x}$” should be “$g_{\mu\nu}$”.

bezw. $\frac{dy}{dx} > 1$ ist, d. h. je nachdem die Kurve in dem betrachteten Punkte *flacher oder steiler als* $45°$ *gegen die* x*-Achse* ansteigt.

Wir fragen nun, welches sind die Kurven, die vom 0-Punkt eine konstante Entfernung haben? Es sei irgend ein Punkt $x = \xi$, $y = \eta$ gegeben. Durch denselben legen wir vom 0-Punkt aus eine Gerade, & berechnen die Länge der Strecke vom Ursprung bis zu diesem Punkt. Die Bedingung dafür, dass diese Länge eine (reelle oder imaginäre) Konstante ist, liefert die gesuchte Kurve. Die Gerade durch den 0-Punkt und den Punkt ξ, η hat die Gleichung $x\eta - y\xi = 0$. Also muss $\int_0^\xi \sqrt{1 - (\frac{dy}{dx})^2} dx = \text{const.}$ sein und wegen $\frac{dy}{dx} = \frac{\eta}{\xi}$ erhalten wir $\sqrt{1 - \frac{\eta^2}{\xi^2}} \int_0^\xi dx = \text{const.}$ oder $\xi^2 - \eta^2 = \pm r^2$. Diese Gleichung stellt *die Schar der gleichseitigen Hyperbeln* dar mit den gemeinsamen Asymptoten $\xi - \eta = 0$ und $\xi + \eta = 0$. Der Hyperbelschar entsprechen in der Euklidischen Geometrie die *konzentrischen Kreise* um den 0-Punkt. Den gemeinsamen Asymptotenrichtungen entsprechen dort die imaginären Kreispunkte.



([55] In der Euklidischen Geometrie hatten wir als Winkel φ unter dem sich zwei durch den 0-Punkt gehende Gerade schneiden,
das Stück des Einheitskreises | definiert, 60
das zwischen diesen Geraden lag.[56] Analog werden wir hier als *Winkel* die *Länge des Stückes*, das von zwei Geraden aus den beiden *Einheitshyperbeln* $x^2 - y^2 = \pm 1$ herausgeschnitten wird, ansprechen. Da die Einheitshyperbel $x^2 - y^2 = +1$, die im ersten und dritten Quadranten liegt, gegen die x-Achse eine Neigung von mehr als $45°$ hat, so ist ihre Länge imaginär, und wir bezeichnen in diesem Quadranten daher als Winkel den Faktor von $\sqrt{-1}$. Der Winkel φ wird also im ersten und dritten Quadranten definiert durch $x = r\cosh\varphi$, $y = r\sinh\varphi$, weil wegen $\cosh^2\varphi - \sin^2\varphi = 1$, dann $x^2 - y^2 = +1$ ist. Bekanntlich ist $\sinh\varphi = \frac{e^\varphi - e^{-\varphi}}{2}$, $\cosh\varphi = \frac{e^\varphi + e^{-\varphi}}{2}$. Die Gleichung der x-Achse ist $y = 0$ oder $e^\varphi = e^{-\varphi}$, d. h. $e^{2\varphi} = 1$, woraus $\varphi = 0$ folgt. Der Winkel φ wird also von der x-Achse aus gezählt. Er wird $-\infty$ für $x + y = 0$, weil dann $e^\varphi = 0$ sein muss und $+\infty$ für $x - y = 0$ wegen $e^{-\varphi} = 0$. Im zweiten und vierten Quadranten setzt man $x = r\sinh\psi$, $y = r\cosh\psi$, dann ist $x^2 - y^2 = -r^2$. Der Winkel wird von der y-Achse ausgezählt. In der Tat muss für $x = 0$ $e^\psi = e^{-\psi}$ also $\psi = 0$ sein.

In der Euklidischen Geometrie stehen zwei Gerade $Ax + By + C = 0$ und $A_1x + B_1x + C_1 = 0$ aufeinander senkrecht, wenn $AA_1 + BB_1 = 0$ ist. Hier

[55] Opening bracket added in pencil. There is no closing bracket.

[56] At the top of the page, Hilbert wrote in pencil: "$\cos\phi = \cos i\phi$, $\sin\phi = \frac{\sin i\phi}{i}$." On the left hand page, Hilbert wrote in pencil: "$\varphi = \frac{1}{2}\log\frac{x+y}{x-y}$".

definieren wir die *Orthogonalität* durch

$$AA_1 - BB_1 = 0. \tag{24}$$

Diese Gleichung bedeutet, dass Radius-Vektor und Tangente oder dass zwei *konjugierte Durchmesser* bei unseren gleichzeitigen Hyperbeln aufeinander
61 senkrecht stehen, so wie das in der Eu|klidischen Geometrie für Kreise der Fall ist. In der Tat hat die durch den Koordinatennullpunkt gehende, zur Tangente im Punkte $\xi\eta$ an die Hyperbel $x^2 - y^2 = \pm r^2$ parallele Gerade die Gleichung $x\xi - y\eta = 0$. Die durch den 0-Punkt und den Punkt $\xi\eta$ gehende Gerade wird durch $x\eta - y\xi = 0$ dargestellt. Für diese beiden Geraden ist (24) in der Tat erfüllt. Da diese beiden konjugierten Durchmesser also aufeinander senkrecht stehen und da daher dasselbe von irgend einem Durchmesser einer Hyperbel und den Tangenten in den beiden Schnittpunkten gilt, so stehen auch die beiden Tangenten in den Punkten,[57] in den konjugierte Durchmesser irgend zwei Hyperbeln in zwei nebeneinanderliegenden Quadranten schneiden, aufeinander senkrecht.

Der durch (24) ausgedrückten Bedingung, dass die durch den 0-Punkt und die beiden Punkte A, B und A_1, B_1 gehenden Geraden *aufeinander senkrecht stehen* kann man noch eine andere, *anschaulichere Form* geben. Da die beiden Punkte in verschiedenen Quadranten liegen müssen, ist

$$\begin{aligned} A &= r\cosh\varphi, & B &= r\sinh\varphi, \\ A_1 &= r\sinh\psi, & B_1 &= r\cosh\psi.^{\langle 58\rangle} \end{aligned}$$

Damit (24) erfüllt ist, muss also $\varphi = \psi$ sein.

Eine beliebige Transformation unserer Punktmannigfaltigkeit in sich ist neben der Parallelverschiebung die Drehung

$$x' = r\cosh(\varphi + \psi), \qquad y' = r\sinh(\varphi + \psi),$$

und da

$$\begin{aligned} \cosh(\varphi + \psi) &= \cosh\varphi\cosh\psi + \sinh\varphi\sinh\psi, \\ \sinh(\varphi + \psi) &= \sinh\varphi\cosh\psi + \cosh\varphi\sinh\psi \end{aligned}$$

ist, so haben wir[59]

$$\begin{aligned} x' &= x\cosh\psi + y\sinh\psi, \\ y' &= x\sinh\psi + y\cosh\psi. \end{aligned}$$

62 Dies sind sämtliche (lineare) Transformationen, die die Form $x^2 - y^2$ d. h. die

[57]"in den Punkten" was interlineated.

[58]Added by Hilbert in pencil: "φ = Länge auf der Hyperbel/Abstand".

[59]At the bottom of the page, Hilbert added with pencil: "Entf^2 =Differenz der Quadrate/Alle Entfern⟨ungen⟩ invariant. Dass φ proportional der Länge folgt. Dies Beispiel ist prinzipiell wegen der infin. genommen Euklid."

Entfernung $\pm r^2$ des Punktes x, y vom Ursprung in $x'^2 - y'^2$ überführen. Diese *Gruppe der Drehungen* in der pseudoeuklidischen Geometrie spielt bekanntlich in der alten Elektrodynamik und in der kleinen Relativitätstheorie eine fundamentale Rolle.

§ 27. Von der allgemeinen zweidimensionalen Pseudogeometrie

Wie diese Untersuchungen für eine *allgemeine Pseudogeometrie* $g_{11} > 0$, $\begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix} < 0$ zu verallgemeinern sind, ist eigentlich vorgeschrieben. Wir haben schon erwähnt, wie die Länge einer Kurve und wie die Nulllinien dann zu definieren sind, wir wollen nur noch bemerken, dass für die zweidimensionale Mannigfaltigkeit *alle Nulllinien zugleich geodätische Linien* sind, während das für alle höheren Mannigfaltigkeiten nicht mehr der Fall ist. Im binären Gebiet wird die geodätische Linie in der Form $y = y(x)$ aus einer Lagrangeschen Differentialgleichung gefunden. Die Bedingung für die Nulllinie ist $\sum_{\mu\nu} g_{\mu\nu}\dot{x}_\mu\dot{x}_\nu = 0$. Diese Gleichung ist ein Integral der Lagrangeschen Differentialgleichung, aus ihr lässt sich die Funktion $y^{(0)} = y^{(0)}(x)$, welche die geodätische Nullinie darstellt, berechnen. Für drei- und mehrdimensionale Mannigfaltigkeiten erhält man aber mehrere Lagrangesche Differentialgleichungen, während man immer nur eine Bedingungsgleichung für die Nullinie erhält. Die Funktionen $y^{(0)}(x)$, $z^{(0)}(x) \cdots$ u. s. w., welche diese eine Differentialgleichung erfüllen, brauchen dann nicht auch Lösungen aller Lagrangeschen Gleichungen sein.

Auch die *Winkel* um einen Punkt herum sind in der | allgemeinen Pseudo- 63
geometrie wie oben zu definieren, da ja im Infinitesimalen die $g_{\mu\nu}$ Konstante sind, mithin dort die pseudoeuklidische Geometrie gilt.

§ 28. Die Invarianten der dreidimensionalen Geometrie

Den Fall der

dreidimensionalen Geometrie

wollen wir ausführlich behandeln.[60] In der Tat treten hier schon alle Schwierigkeiten auf, die wir bei vier Dimensionen antreffen werden; die vierdimensionale Punktmannigfaltigkeit, die wir in der Physik vor uns haben, können wir dann um so kürzer erledigen.[61] Der einzelne Punkt oder das einzelne Ding der Menge wird hier durch ein Zahlentripel x_1, x_2, x_3 bezw. x, y, z charakterisiert. Um Geometrie treiben zu können, muss man nun 6 Funktionen $g_{\mu\nu}$ kennen,

[60] "ausführlich behandeln" was corrected from "in Kürze erledigen".

[61] The preceding half-sentence was corrected from: "ist aber gerade diejenige, die wir in der Physik vor uns haben."

von denen wir wieder absolute Stetigkeit und Differenzierbarkeit beliebig hoher Ordnung voraussetzen, wie wir dies auch später in der Physik annehmen werden. Nach unserer Auffassung sind ja alle Unstetigkeiten, die man in der Natur antrifft, nur scheinbare, d. h. ein Grenzfall der Stetigkeit, also etwas Sekundäres.

Wir legen unserer Betrachtung die homogene quadratische Form

$$G(X_1, X_2, X_3) = g_{11}X_1^2 + g_{22}X_2^2 + g_{33}X_3^2 + 2g_{12}X_1^2X_2^2 + 2g_{23}X_2X_3 + 2g_{31}X_3X_1.$$

zugrunde, in der die $g_{\mu\nu}$ Funktionen von x_1, x_2, x_3 sind. Die Länge einer Kurve $x_i(p)$ definieren wir wie früher durch $s = \int_{p_1}^{p_2} \sqrt{\sum_{\mu\nu} g_{\mu\nu}\dot{x}_\mu\dot{x}_\nu dp}$. Es sind wieder zwei Fälle zu unterscheiden, die *eigentliche Geometrie*, in der G
64 *positiv definit* ist und in welcher daher alle reellen Kurven eine reelle, von | Null verschiedene Länge besitzen, und die *Pseudogeometrie* mit *indefinitem* G, in der eine reelle Kurve im allgemeinen eine komplexe Länge, im besonderen auch die Länge Null haben kann.

I) Die *analytische Bedingung* für den ersten Fall ist

$$g_{11} > 0, \qquad g_{11}g_{22} - g_{12}^2 > 0, \qquad \begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix} > 0.$$

Die Definition des Winkels um einen Punkt herum geschieht wie oben. In diesem Punkt, den wir zum 0-Punkt machen, wird nämlich $G = (g_{11})_0 X_1^2 + \cdots + 2(g_{23})_0 X_2 X_3$ durch eine lineare Substitution auf die Form $X_1'^2 + X_2'^2 + X_3'^2$ gebracht. Setzen wir $X_1' = r\cos\varphi$, $X_2' = r\sin\varphi$, $X_3' = 0$, so ist φ der Winkel, den die Projektion der Geraden durch den 0-Punkt und den Punkt X_1', X_2', X_3' auf die X_1', X_2'-Ebene mit der X_1'-Achse bildet.

Genau wie im binären Gebiet, fragen wir wieder, welche *Eigenschaften* der Geometrie *von der Koordinatenwahl unabhängig* sind. Setzen wir $x_i = x_i(x_1', x_2', x_3')$, so finden wir, dass die quadratische Form $\sum_{\mu\nu} g_{\mu\nu} dx_\mu dx_\nu$ dann eine Invariante ist, wenn die $g_{\mu\nu}$ einen kovarianten Tensor $g'_{\mu\nu} = \sum_{\alpha\beta} \frac{\partial x_\alpha}{\partial x'_\mu}\frac{\partial x_\beta}{\partial x'_\nu} g_{\alpha\beta}$ bilden. So treten hier wieder dieselben Begriffe auf, nämlich ausser dem erwähnten noch der kontravariante Tensor $h'^{\mu\nu} = \sum_{\alpha\beta} \frac{\partial x'_\mu}{\partial x_\alpha}\frac{\partial x'_\nu}{\partial x_\beta} h^{\alpha\beta}$ und der gemischte Tensor $k'^{\ \nu}_{\mu} = \sum_{\alpha\beta} \frac{\partial x_\alpha}{\partial x'_\mu}\frac{\partial x'_\nu}{\partial x_\beta} k^\beta_\alpha$. Diese drei Tensoren brauchen natürlich in den beiden Indizes durchaus nicht symmetrisch zu sein. Die Symmetrie reduziert die Zahl der Tensorkomponenten auf sechs, die Antisymmetrie auf drei. Als Gebilde erster Stufe oder Vektoren haben wir den kovarianten Vektor $q'_\mu = \sum_\alpha \frac{\partial x_\alpha}{\partial x'_\mu} q_\alpha$ und den kontravarianten $p'^\nu = \sum_\beta \frac{\partial x'_\nu}{\partial x_\beta} p^\beta$ zu erwähnen.

65 Da die *Länge* einer Kurve die einfachste Invariante ist, so spielen die *geodätischen Linien* und die mit ihnen zusammenhängenden Riemannschen Koordinaten hier dieselbe ausgezeichnete Rolle wie in der zweidimensionalen Geometrie. Die ganze Theorie der geodätischen Linien ist übrigens wörtlich

aus dem binären Gebiet zu übertragen. Man kann die kürzesten Linien wieder aus dem Variationsproblem[62] $\int_{p_1}^{p_2} G(x_i(p)dp = \text{Min}$ finden. Ein Integral der zugehörigen Lagrangeschen Differentialgleichungen

$$\ddot{x}_h + \sum_{\mu\nu} \left\{ \begin{matrix} \mu\nu \\ \alpha \end{matrix} \right\} \dot{x}_\mu \dot{x}_\nu = 0$$

ist $G = c$,[63] und der Parameter p hat dann die Bedeutung $p = \frac{\text{Kurvenlänge}}{\sqrt{c}}$. Die einzige Abänderung ist, dass nun durch jeden Punkt ∞^2 geodätische Linien gehen. Man erhält zwar drei Differentialgleichungen für drei Funktionen $x_i(p)$, doch bleibt die geodätische Linie unverändert, wenn man p durch eine Funktion $\varphi(p)$ ersetzt, so dass also die geodätische Linie durch zwei Funktionen bestimmt ist, z. B., wenn x_1 als Parameter gewählt wird, durch $x_2(x_1)$ und $x_3(x_1)$. Diese beiden Funktionen werden aus zwei[64] gewöhnlichen Differentialgleichungen zweiter Ordnung gewonnen. Mithin treten vier Integrationskonstanten auf. Soll die Kurve durch den Punkt $\hat{\xi}_1$, $\hat{\xi}_2$, $\hat{\xi}_3$ gehen, so bestimmen sich zwei Konstanten aus $x_2(\hat{\xi}_1) = \hat{\xi}_2$ und $x_3(\hat{\xi}_1) = \hat{\xi}_3$, so dass noch zwei Konstanten willkürlich bleiben, entsprechend den ∞^2 geodätischen Linien.

§ 29. Riemannsche Koordinaten

Die *Riemannschen Koordinaten* ξ_i in Bezug auf einen festen Punkt werden wieder so eingeführt, dass $\xi_i = ap$ die | Gleichungen der durch diesen Punkt 66
gehenden geodätischen Linien darstellt. Eliminiert man p aus ihnen, so findet man: *zwei lineare Gleichungen* zwischen den ξ_i *bestimmen eine geodätische Linie* durch den Punkt. Führt man die Riemannschen Koordinaten anstelle der ursprünglichen ein, so werden die ξ_i Funktionen von x_1, x_2, x_3 bezw. $x_i = x_i(\xi_1, \xi_2, \xi_3)$. Das Quadrat des Linienelementes geht dann in $\sum_{\mu\nu} a_{\mu\nu}(\xi_1, \xi_2, \xi_3) d\xi_\mu d\xi_\nu$ über, wobei man die $a_{\mu\nu}$ erhält, indem man erstens die neuen $g_{\mu\nu}$ gemäss ihren Transformationsformeln berechnet und in denselben dann die x_i durch ihre Funktionen in den ξ_i ersetzt. Bezeichnet man $g_{\mu\nu}$ als Funktion der ξ_i mit $g^*_{\mu\nu}$, so ist also

$$a_{\mu\nu}(\xi_1, \xi_2, \xi_3) = \sum_{\alpha\beta} \frac{\partial x_\alpha}{\partial \xi_\mu} \frac{\partial x_\beta}{\partial \xi_\nu} g^*_{\alpha\beta}(\xi_1, \xi_2, \xi_3).$$

Es verschwinden wieder alle ersten Ableitungen $a_{\mu\nu h}$ der $a_{\mu\nu}$ nach den ξ_h, und man erhält

$$a_{\mu\nu} = (a_{\mu\nu})_0 + \frac{1}{2} \sum_{\alpha\beta} (a_{\mu\nu hk})_0 \xi_h \xi_k + \cdots .$$

[62] In the following equations "G" was corrected in pencil to "φ".

[63] "G" was corrected in pencil to "φ".

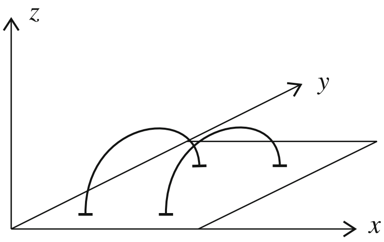
[64] "zwei" was corrected from "einer".

Das Linienelement ist jetzt wieder so euklidisch wie möglich, d. h. in den Gliedern nullter und erster Näherung während dies in gewöhnlichen Koordinaten nur in den Gliedern nullter Näherung der Fall ist. Aus den zweiten Ableitungen $a_{\mu\nu hk}$ kann man sich dann eine neue Invariante bilden, die *Krümmung* der dreidimensionalen Geometrie und zwar ist wieder

$$K_{\mu\nu} = \sum_{hk} a_{\mu\nu hk} a^{hk}, \qquad K = \sum_{\mu\nu} K_{\mu\nu} a^{\mu\nu}.$$

§ 30. Die drei Geometrien konstanter Krümmung

In den ξ_i-Koordinaten für einen festen Punkt der Mannigfaltigkeit drücken sich wiederum nur dann *alle* geodätischen Linien und nicht nur die durch den betreffenden Punkt gehenden durch lineare Gleichungen aus, wenn Transfor-
67 mationen | der Mannigfaltigkeit in sich möglich sind, oder, was damit völlig äquivalent ist, wenn die Kongruenzsätze gelten. Dieselben gelten nur in drei Fällen, nämlich in der *Euklid*ischen und in den beiden Nichteuklidischen Geometrien, der *Riemann-Helmholtz*schen und der *Bolyai-Lobatscheffky*schen. Die erstere kann man als diejenige deuten, die auf einer dreidimensionalen Overfläche einer im vierdimensionalen Euklidischen Raum liegenden vierdimensionalen Kugel gilt. Der letzteren kann man wieder eine *anschauliche Deutung* geben, wenn man das Linienelement auf die Form $ds^2 = \frac{dx^2+dy^2+dz^2}{z^2}$ bringt, wo x, y, z rechtwinklige, räumliche Koordinaten darstellen.



Dann werden nämlich die kürzesten Linien durch die Gesamtheit der Kreise gebildet, die im oberen Halbraum ($z > 0$) liegen und die auf der (x, y)-Ebene ($z = 0$) senkrecht stehen. Diese Ebene stellt das Unendlichferne der dreidimensionalen Geometrie dar. Durch einen Punkt der Mannigfaltigkeit gehen, wie es sein muss, ∞^2 Kreise, die auf der Ebene senkrecht stehen, durch zwei Punkte nur ein einziger.

Als einfachstes Beispiel einer Geometrie, in der die Kongruenzsätze nicht gelten, ist wieder die dreidimensionale Oberfläche des vierdiimensionalen Paraboloids zu nennen:

$x_4 =$ homogene quadratische Funktion von x_1, x_2, x_3.

§ 31. Gaussische Koordinaten

Die *Einführung der Gauss'schen Koordinaten* u, v, welche $g_{11} = 1$, $g_{12} = 0$, $g_{22} = \gamma^2(u, v)$ machten, brachte im binären Gebiet den enormen Vorteil, dass

68 nur noch die eine willkürliche | Funktion $\gamma^2(u, v)$ stehen blieb. Jetzt haben wir *sechs willkürliche Funktionen* $g_{\mu\nu}$ und nur *drei Funktionen* $x_i(u, v, w)$ *zur freien Verfügung*, welche wir dazu verwenden wollen $g_{11} = 1$ $g_{12} = 0$, $g_{13} = 0$ zu machen. Dann werden zwar immer noch drei willkürliche Funktionen übrig bleiben, und wir können somit die Krümmung nicht mehr auf die einfache Form (17) bringen. Wir führen also neue Koordinaten u, v, w so ein, dass $v = e_1$, $w = e_2$ die Schar der durch einen beliebigen Punkt gehenden geodätischen Linien darstellt und $u =$ const. die dazu orthogonale Flächenschar. Damit die letzte Bedingung erfüllt ist, muss das Linienelement die Form $ds^2 = \gamma du^2 + \gamma_{11} dv^2 + \gamma_{22} dw^2 + 2\gamma_{12} dv dw$ annehmen, wobei γ und $\gamma_{\mu\nu}$ noch Funktionen von u, v, w sein werden. Führen wir als Kurvenparameter u selber ein, so wird irgend eine Linie durch $v = v(u)$, $w = w(u)$ dargestellt. Um die geodätischen Linien zu finden, hat man

$$\delta \int_{u_1}^{u_2} \sqrt{G^*} du = \delta \int_{u_1}^{u_2} \sqrt{\gamma + \gamma_{11} \left(\frac{dv}{du}\right)^2 + \gamma_{22} \left(\frac{dw}{du}\right)^2 + 2\gamma_{12} \frac{dv}{du}\frac{dw}{du}} du = 0$$

zu machen. Man erhält die beiden Lagrangeschen Differentialgleichungen

$$\begin{aligned} 0 = \frac{d}{du} & \left\{ \frac{1}{\sqrt{G^*}} \left(\gamma_{11} \frac{dv}{du} + \gamma_{12} \frac{dw}{du} \right) \right\} \\ & - \frac{1}{\sqrt{G^*}} \left(\frac{\partial \gamma}{\partial v} + \frac{\partial \gamma_{11}}{\partial v} \left(\frac{dv}{du}\right)^2 + \frac{\partial \gamma_{22}}{\partial v} \left(\frac{dw}{du}\right)^2 + 2 \frac{\partial \gamma_{12}}{\partial v} \frac{dv}{du}\frac{dw}{du} \right), \\ 0 = \frac{d}{du} & \left\{ \frac{1}{\sqrt{G^*}} \left(\gamma_{22} \frac{dw}{du} + \gamma_{12} \frac{dv}{du} \right) \right\} \\ & - \frac{1}{\sqrt{G^*}} \left(\frac{\partial \gamma}{\partial w} + \frac{\partial \gamma_{11}}{\partial w} \left(\frac{dv}{du}\right)^2 + \frac{\partial \gamma_{22}}{\partial w} \left(\frac{dw}{du}\right)^2 + 2 \frac{\partial \gamma_{12}}{\partial w} \frac{dv}{du}\frac{dw}{du} \right). \end{aligned}$$

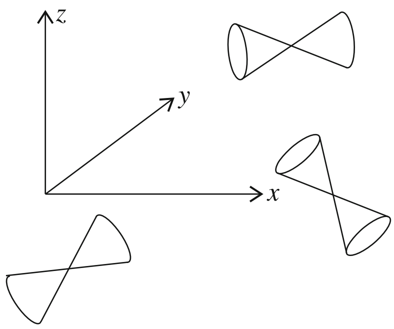
Diese Gleichungen müssen für $v = c_1$, $w = c_2$ erfüllt sein, da dies geodätische Linien darstellt. Setzt man dies ein, so folgt aus der ersten Gleichung $\frac{\partial \gamma}{\partial v} = 0$, aus der zweiten $\frac{\partial \gamma}{\partial w} = 0$, so dass $\gamma = \varphi(u)$ wird. Führt man schliesslich noch u' durch $\frac{du'}{du} = \sqrt{\varphi(u)}$ ein, so hat man es erreicht, dass $g_{11} = 1$ wird. Es gilt auch die Umkehrung: Hat man solche Koordinaten eingeführt, dass 69
$g_{11} = \varphi(u)$, $g_{12} = 0$, $g_{13} = 0$ wird, so sind $v = c_1$, $w = c_2$ Integrale der Lagrangeschen Differentialgleichung, also geodätische Linien, und es stellt $u =$ const. die zu denselben orthogonale Flächenschar dar. *Das Linienelement ist* hiermit *auf eine Normalform gebracht*, und es sind jetzt diese Funktionen $\gamma_{\mu\nu}$ *transzendente Invarianten* der $g_{\mu\nu}$ und ihrer Ableitungen, im Gegensatz zur Invariante K, die eine rationale Invariante dieser Funktion ist.

§ 32. Mongesche Differentialgleichung

II) Um die *Pseudogeometrie* des ternären Gebietes zu studieren, ersetzen wir unsere Voraussetzungen über die $g_{\mu\nu}$ durch[65]

$$g_{11} > 0, \qquad g_{11}g_{22} - g_{12}^2 \gtreqless 0, \qquad \begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix} < 0.$$

Dann wird $G(X_1, X_2, X_3)$ indefinit. Es existieren jetzt reelle[66] Kurven, die die Länge Null haben. $G(X_1 - x_1, X_2 - x_2, X_3 - x_3) = 0$ *stellt einen reellen Kegel* mit der Spitze im Punkte x_1, x_2, x_3 *dar*. Bei veränderlichen x_1, x_2, x_3 definiert diese Gleichung in jedem Punkte der Mannigfaltigkeit einen solchen Kegel.



Sie stellt, wie wir sagen wollen, ein *Kegelfeld* dar. Da die $g_{\mu\nu}$ Funktionen von x_1, x_2, x_3 sind, hat die Achse des Kegels in jedem Punkte der Mannigfaltigkeit eine andere Richtung. Eine Kurve in unserer Geometrie ist durch $x_i = x_i(p)$ gegeben. Die darunter befindlichen *Nulllinien* erhalten wir als Lösungen der Differentialgleichung $G(\dot{x}_1(p), \dot{x}_2(p), \dot{x}_3(p)) = 0$. Dies ist
70 eine Differen|tialgleichung zur Bestimmung von zwei und nicht von drei Funktionen, da dieselbe Kurve dargestellt wird, wenn man p durch eine Funktion $\varphi(p)$ ersetzt. Diese eine Differentialgleichung für zwei unbekannte Funktionen einer Veränderlichen ist eine sogenannte *Monge'sche* (diophantische) *Gleichung*. Man kann daher eine der beiden Funktionen willkürlich wählen und hat dann eine gewöhnliche Differentialgleichung zur Bestimmung der anderen Funktion. Ein System von Differentialgleichungen, in dem die Zahl der unbekannten Funktionen grösser ist als die Zahl der Gleichungen, heisst ein *unterbestimmtes System*. Ein solches haben wir hier vor uns. Wegen der auftretenden willkürlichen Funktion gibt es durch einen beliebigen Punkt der Geometrie nun eine viel grössere Mannigfaltigkeit von Nulllinien als im binären Gebiet. Während diese Nulllinien aber dort alle zugleich geodätische waren, ist dies hier nicht der Fall; denn zur Bestimmung der geodätischen Linien erhält man ja auch hier wieder aus dem Variationsprinzip ein *bestimmtes System*, nämlich zwei Differentialgleichungen für die beiden unbekannten Funktionen $y(x)$ und $z(x)$. Hier sind vielmehr *nur diejenigen Nulllinien* auch *geodätische, die die beiden Lagrangeschen Differentialgleichungen und die Nebenbedingung* $G(\dot{x}, \dot{y}, \dot{z}) = 0$ *erfüllen*.

[65] Added by Hilbert in the left margin in pencil: "statt $G = \varphi$!".

[66] "reelle" was interlineated.

Der letzteren kann man wieder eine geometrische Interpretation geben. Seien nämlich α_1, α_2, α_3 drei Werte, die $G(\alpha_1, \alpha_2, \alpha_3) = 0$ erfüllen, so stellt $X_1 - x_1 : X_2 - x_2 : X_3 - x_3 = \alpha_1 : \alpha_2 : \alpha_3$ diejenige Erzeugende des Kegels $G(X_1 - x_1, X_2, x_2, X_3 - x_3) = 0$ im Punkte x_1, x_2, x_3 dar, welche | mit den drei 71
Koordinatenachsen die durch die Richtungskosinusse α_1, α_2, α_3 bestimmten Winkel bildet. $G(\dot{x}_i(p)) = 0$ bedeutet also, dass die Kurve $x_i(p)$ im Punkte x_i die Erzeugende $X_1 - x_1 : X_2 - x_2 : X_3 : x_3 = \dot{x}_1(p) : \dot{x}_2(p) : \dot{x}_3(p)$ des Kegels zur Tangente haben soll. Die Differentialgleichung $G(\dot{x}_i(p)) = 0$ lösen, heisst daher, man soll eine solche Kurve (Nulllinie) $x_i(p)$ finden, die in jedem ihrer Punkte den zu demselben gehörigen Kegel des Kegelfeldes tangiert. Wir drücken dies kürzer aus, indem wir sagen: *eine Nulllinie „passt" auf das Kegelfeld.*

Man kann wi⟨e⟩der wie im R_2 schliessen: suche die Lösung der Differentialgl. der geodät. die in Richt⟨un⟩g einer Erzeugenden laufen. Dann ist für diese im Anfangsp. $\varphi = 0$ und also überall $\varphi = 0$, d. h. sie ist Nulllinie.[67]

§ 33. Die Hamiltonsche partielle Differentialgleichung

In der Ebene war das entsprechende Problem bedeutend einfacher. Dort waren die Nulllinien in jedem Punkte x, y durch zwei Richtungen $(\frac{dy}{dx})_i$ gegeben. Wir hatten ein Richtungenfeld und mussten eine Kurve suchen, die auf dies Richtungsfeld passt. Dies führte sofort auf eine gewöhnliche Differentialgleichung $(\frac{dy}{dx})_i = f_i(x, y)$. Im Raume gibt es nun noch eine ganz andere Aufgabe, die sich in der Ebene auf dieses nämliche Problem reduziert: *Gesucht seien alle Flächen, die* in jedem ihrer Punkte *den zugehörigen Kegel berühren*, die, wie wir wieder sagen wollen, auf das Kegelfeld passen. Um die Aufgabe analytisch zu formulieren, müssen wir *Ebenenkoordinaten* einführen. Ist die Fläche in der Form $z = z(x, y)$ gegeben, so hat ihre Tangentialebene im Punkte x, y, z die Gleichung

$$(X - x)\,\frac{\partial z}{\partial x} + (Y - y)\,\frac{\partial z}{\partial y} - (Z - z) = 0.$$

Diese Ebene soll nun Tangentialebene an den Kegel $G(X_i - x_i) = 0$ sein. Bezeichnen wir die Ebenenkoordinaten mit U_i, so ist die | *Bedingung dafür*, 72
dass die Ebene U_i *Tangentialebene des Kegels* ist,

$$\begin{vmatrix} g_{11} & g_{12} & g_{13} & U_1 \\ g_{21} & g_{22} & g_{23} & U_2 \\ g_{31} & g_{32} & g_{33} & U_3 \\ U_1 & U_2 & U_3 & 0 \end{vmatrix} = 0,$$

[67] The preceding paragraph was interlineated by Hilbert in pencil.

oder

$$H(x_1, x_2, x_3, U_1, U_2, U_3) = \sum_{\mu\nu} g^{\mu\nu}(x_1, x_2, x_3) U_\mu U_\nu = 0.$$

Da die Ebenenkoordinaten der Tangentialebene der Fläche proportional mit[68] $\frac{\partial z}{\partial x}$, $\frac{\partial z}{\partial y}$, -1 sind, so ist

$$H\left(x, y, z, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}, -1\right) = 0$$

identisch in x und y die Bedingung dafür, dass die Fläche auf das Kegelfeld passt. *Die Funktion* $z(x, y)$ *muss* also einer *partiellen Differentialgleichung erster Ordnung*, der sogenannten *Hamiltonschen* Differentialgleichung *genügen*.

Diese Gleichung ist das *duale Gegenstück* zur Mongeschen und dieser letzteren durchaus gleichwertig. In der Tat bestimmen die Grössen $g^{\mu\nu}$, die in ihr auftreten, und die natürlich Funktionen von x, y, z sind, die Geometrie der dreidimensionalen Mannigfaltigkeit ebenso gut wie die $g_{\mu\nu}$. Ferner definiert auch diese Gleichung das Kegelfeld. Sie setzt nämlich zwischen den Richtungscosinussen $\frac{\partial z}{\partial x}$, $\frac{\partial z}{\partial y}$, -1 der Normalen zur Tangentialebene in jedem Punkte eine Beziehung fest. Die Gesamtheit der Ebenen, deren Normalen dieser Beziehung in einem Punkte genügen, umhüllt dann einen Kegel dessen Spitze in diesem Punkte liegt.[69] Das Abbild dieser Differentialgleichung ist also wieder ein Kegelfeld.

Eine Funktion $z = z(x, y)$, die der partiellen Differentialgleichung genügt,
73 heisst eine *Integralfläche*. So wie bei | der gewöhnlichen Differentialgleichung erster Ordnung die Integralkurve festgelegt ist, wenn ein Punkt derselben vorgegeben ist, ist bei der partiellen Differentialgleichung erster Ordnung die Integralfläche im allgemeinen festgelegt, wenn eine Raumkurve vorgegeben wird, durch welche die Fläche gehen soll. Die Aufgabe, eine Fläche $z = z(x, y)$ zu finden, die der Differentialgleichung genügt und durch eine vorgegebene Kurve geht, heisst das *Cauchysche Problem*.

§ 34. Charakteristikentheorie

Wir wollen nun die *Hauptresultate der Integrationstheorie* dieser partiellen Differentialgleichung mitteilen. Auf alle Beweise müssen wir verzichten, da uns dies zu weit führen würde. Der fundamentale Begriff, der in dieser Theorie auftritt, ist derjenige der *Charakteristik*. Er entspricht demjenigen der Determinante der Algebra und dem der determinierenden Fundamentalgleichung in der Theorie der gewöhnlichen Differentialgleichungen. Wir sagten schon, dass *im allgemeinen* die Integralfläche der Differentialgleichung $H(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}, -1) = 0$

[68] "proportional mit" was interlineated by Hilbert with pencil.

[69] The preceding half-sentence was corrected from: "umhüllt dann den Kegel des Feldes, der zu diesem Punkte gehört."

durch eine willkürlich vorgegebene Kurve festgelegt ist⟨,⟩ und fragen nun, ob es gewisse *ausgezeichnete Kurven* gibt, *welche die Integralfläche nicht bestimmen*. Solche Kurven, die tatsächlich vorhanden sind, heissen *Charakteristiken*. Sie sind die *Unbestimmtheitscurven*[70] *des Cauchyschen Problems*. Um sie zu finden, versucht man, den Beweis für die Eindeutigkeit einer Lösung der partiellen Differentialgleichung zu geben, welche eine vorgegebene Kurve enthält. Man findet, dass dieser Beweis sich gerade dann nicht erbringen lässt, wenn diese Raumkurve selber das Integral | einer gewissen gewöhnlichen Differenti- 74
algleichung ist. Diese Differentialgleichung hat eine einparametrige Schar von Kurven auf der Integralfläche zu Lösungen. Durch diese Schar kann sogar, wie die Theorie zeigt, die Integralfläche selber erzeugt werden. Durch jeden Punkt der Fläche geht nur eine solche Kurve.

Der springende Punkt der Theorie ist es nun, diese *Charakteristiken* allein *aus der Differentialgleichung zu finden*, also ohne dass die Integralfläche bekannt ist. Dies gelingt in der Tat und zwar sind diese Kurven dann Integrale eines Systems von gewöhnlichen Differentialgleichungen. Diese letzeren fliessen, wie wir ohne Beweis mitteilen müssen, auch aus folgendem *Variationsproblem*: gesucht sind in unserer dreidimensionalen Mannigfaltigkeit die Kurven kürzesten Falles in der Richtung der z-Achse, die zugleich Nulllinien sind, die also $\delta \int_{p_1}^{p_2} \dot{z} dp = 0$ mit der Nebenbedingung $G(\dot{x}, \dot{y}, \dot{z}) = 0$ erfüllen. Dieses Variationsproblem ist nach einer bekannten Regel der Variationsrechnung äquivalent mit

$$\delta \int_{p_1}^{p_2} (\dot{z} + \lambda G) dp = 0,$$

wobei λ als Funktion von x, y, z so zu bestimmen ist, dass die Nebenbedingung $G(\dot{x}, \dot{y}, \dot{z}) = 0$ erfüllt wird. Die zugehörigen Lagrangeschen Gleichungen sind

$$\frac{d}{dp}\left(1 + \frac{\partial \lambda G}{\partial \dot{z}}\right) - \frac{\partial \lambda G}{\partial z} = 0; \quad \frac{d}{dp}\frac{\partial \lambda G}{\partial \dot{x}_i} - \frac{\partial \lambda G}{\partial x_i} = 0, \qquad i = 1, 2, \text{ also}$$

$$\frac{d}{dp}\frac{\partial \lambda G}{\partial \dot{x}_i} - \frac{\partial \lambda G}{\partial x_i} = 0, \qquad i = 1, 2, 3.$$

Diese nämlichen Lagrangeschen Gleichungen erhält man aber auch aus dem Problem $\delta \int_{p_1}^{p_2} \lambda G dp = 0$, wobei λ immer noch durch $G = 0$ festgelegt ist. Da die Kurve $x_i = x_i(p)$ | die dieses Problem löst, invariant bleibt, wenn p 75
durch eine Funktion $\varphi(p)$ ersetzt wird, so bestimmen wir $\varphi(p)$ noch so, dass $\lambda(x(p), y(p), z(p)) \equiv 1$ wird für jeden Wert von p. Damit ist das ursprüngliche Variationsproblem übergeführt in das folgende

$$\delta \int_{p_1}^{p_2} G(\dot{x}, \dot{y}, \dot{z})\, dp = 0 \text{ mit der Nebenbedingung } G(\dot{x}, \dot{y}, \dot{z}) = 0.$$

Dies ist aber das Variationsproblem der geodätischen Nulllinien. Wir haben also den

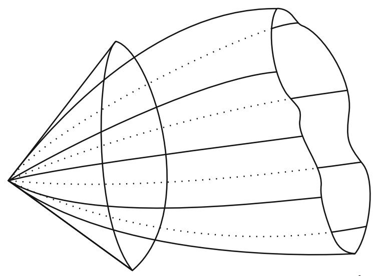
[70] "Unbestimmtheitsstellen" was corrected in pencil to "Unbestimmtheitscurven".

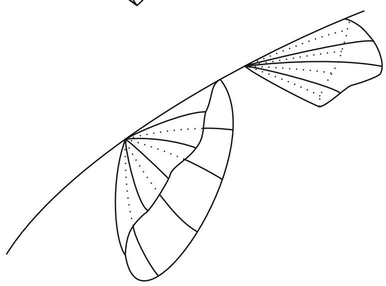
Satz: *Die Charakteristiken sind die geodätischen Nulllinien der Pseudogeometrie.*

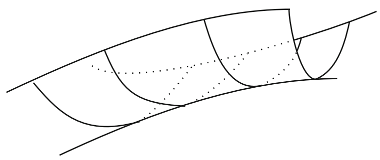
Damit haben wir unsere Aufgabe gelöst, die Charakteristiken direkt aus der vorgelegten partiellen Differentialgleichung $H = 0$ zu berechnen, ohne deren Integralflächen zu kennen; und zwar findet man sie, wie wir sehen, als Integrale eines Systems gewöhnlicher Differentialgleichungen.

Es ergibt sich jetzt das andere Problem, durch eine vorgegebene Kurve die Integralfläche der Differentialgleichung zu finden, wenn die Charakteristiken bekannt sind; anders formuliert: *wie kann man sich aus den Charakteristiken*, die ja in jedem Punkt als geodätische Nulllinien die Kegel des Feldes berühren, *eine Integralfläche* (die auch diese Eigenschaft haben muss) *zusammensetzen*? Wir wollen nun nur zwei besonders wichtige und interessante Spezialfälle dieser Aufgabe erwähnen. Ohne den Beweis dafür zu erbringen, bemerken wir nämlich

1) Die einparametrige Schar der durch einen beliebigen Punkt x_i gehenden
76 Charakteristiken bildet eine solche In|tegralfläche.







Man kann diese Fläche als eine *transzendente Kegelfläche* bezeichnen. Errichtet man nämlich in dem Punkte x_i den durch die Differentialgleichung definierten Kegel, so sind dessen Erzeugende Tangenten dieser Integralfläche. Jetzt lässt sich der Grund, warum gerade die geodätischen Nulllinien die Unbestimmtheitsstellen des Cauchyschen Problems sind, klar erkennen. Ist nämlich die vorgegebene Kurve, die die Integralfläche bestimmen sollte, eine Charakteristik, so kann man durch jeden ihrer Punkte eine solche transzendente Kegelfläche legen. Eine Erzeugende dieser Fläche ist dann immer die vorgegebene Kurve, liegt also tatsächlich in der Fläche. Andererseits gehört *zu jedem Punkt der Charakteristik* ein anderer *transzendenter Kegel*.

2) Statt eine beliebige Raumkurve zu wählen, wie es das allgemeine Cauchysche Problem verlangt, nehmen wir eine *gewöhnliche* Nulllinie. Dann ist die durch die Enveloppe der diese Nulllinie tangierenden *geodätischen* Nulllinien gebildete Fläche eine Integralfläche der Differentialgleichung.

Integralfläche und Nulllinie stehen im allgemeinen in einem *umkehrbar eindeutigen Verhältnis* zueinander. Hat man nämlich eine Integralfläche, so konstruiere man auf ihr die geodä|tischen Nulllinien. Gehen dieselben nicht, wie 77
im Falle 1 alle durch einen Punkt, so haben sie eine Enveloppe. Diese ist dann eine gewöhnliche Nulllinie, also ein Integral der Mongeschen Gleichung. Die *Lösungen von* $G = 0$ *und* $H = 0$ *gehen* also eindeutig *Hand in Hand* und zwar lösen die *geodätischen Nulllinien* beide Probleme. Es besteht eine vollständige Reziprozität zwischen den beiden Differentialgleichungen.

§ 35. Die Hamilton-Jacobische Theorie

Dieser Dualismus hat zur Folge, dass wir die Charakteristiken noch aus einem anderen Variationsproblem erhalten können. Das von uns schon genannte Problem $\int_{p_1}^{p_2} \dot{z}dp = \text{Min.}$ hatte nämlich die Nebenbedingung $G = 0$ (Mongesche Gleichung). Wir dürfen erwarten, dass die geodätischen Nulllinien auch aus einem *Variationsproblem* fliessen, *das die Hamiltonsche Differentialgleichung zur Nebenbedingung* hat. Dies ist in der Tat der Fall, und zwar lautet dieses Problem

$$\int_{p_1}^{p_2} (u_1\dot{x}_1 + u_2\dot{x}_2 + u_3\dot{x}_3)dp = \text{Min},$$

wenn zugleich $H(u_1, u_2, u_3) = 0$ erfüllt ist. Hierin sind $x_i(p)$ und $u_i(p)$ sechs unbekannte Funktionen, die aus sechs Lagrangeschen Differentialgleichungen zu bestimmen sind. Natürlich kommen die Funktionen $x_i(p)$ auch in H vor, weil ja die $g^{\mu\nu}$ Funktionen der x_i sind. Dieses Variationsproblem ist von grösster Wichtigkeit, weil es die Differentialgleichungen der Charakteristiken in ihrer *kanonischen Gestalt*, in der sie in der Mechanik gebraucht werden, liefert. Man wird so auf die *Hamilton-Jacobische Theorie* geführt, die sich mit der Herleitung dieser kanonischen Differentialgleichungen aus den Prinzipien der Mechanik und ihrer Integration beschäftigt.

Wir verfahren wieder nach den Regeln der Variationsrechnung und bilden die 78
Lagrangeschen Differentialgleichungen des Problems

$$\delta \int_{p_1}^{p_2} (u_1\dot{x}_1 + u_2\dot{x}_2 + u_3\dot{x}_3 + \lambda H)dp = 0,$$

wobei λ durch $H = 0$ bestimmt wird. Wir denken uns noch p durch eine solche Funktion $\varphi(p)$ ersetzt, dass $\lambda = -1$ wird und erhalten dann durch Lagrangesche Differentiation

$$\left.\begin{array}{ll} \text{nach } x_i: & \dot{u}_i = -\dfrac{\partial H}{\partial x_i} \\ & \qquad\qquad i = 1, 2, 3. \\ \text{nach } u_i: & \dot{x}_i = +\dfrac{\partial H}{\partial u_i} \end{array}\right\} \quad \textit{kanonische Gestalt der Gleichungen der Mechanik!}$$

Unter den Lösungen greifen wir wieder nur diejenigen heraus, die $H = 0$ erfüllen und zeigen, dass dieselben wieder die geodätischen Nulllinien oder Charakteristiken sind.

Zum Beweise müssen wir nochmals auf die *Entstehung der Gleichung* $H = 0$ *aus* $G = 0$ zurückkommen. Die Ebene mit den Koordinaten $\langle u_1, u_2, u_3 \rangle$ ist Tangentialebene an den Kegel $G \equiv \sum_{\mu\nu} g_{\mu\nu}\dot{x}_\mu\dot{x}_\nu = 0$, wenn $\frac{\partial G}{\partial \dot{x}_1} = u_1$, $\frac{\partial G}{\partial \dot{x}_2} = u_2$, $\frac{\partial G}{\partial \dot{x}_3} = u_3$ erfüllt ist. Hierzu kommt noch die Gleichung des Ineinanderliegens $\dot{x}_1 u_1 + \dot{x}_2 u_2 + \dot{x}_3 u_3 = 0$. Da G in den $\dot{x}_\mu$ homogen vom zweiten Grade ist, wird $\frac{\partial G}{\partial \dot{x}_\mu}$ in den $\dot{x}_\mu$ linear. Wir haben also vier in $\dot{x}_1$, $\dot{x}_2$, $\dot{x}_3$, $+1$ lineare Gleichungen aufzulösen. Das Eliminationsresultat ergibt die oben erwähnte Determinante $H = 0$. Aus den drei in $\dot{x}_\mu$ linearen Gleichungen $\frac{\partial G}{\partial \dot{x}_\mu} = u_\mu$ kann man umgekehrt die

$$\dot{x}_\mu = \text{Funktion der } u_\nu \text{ und } x_\lambda$$

berechnen. Diese Funktionen werden sogar in den u_ν linear sein. Setzt man diese Ausdrücke für $\dot{x}_\mu$ in $G = 0$ ein, so erhält man ebenfalls $H = 0$. Da
79 G in den $\dot{x}_\mu$ homogen quadratisch ist, | wird $\sum_\mu \dot{x}_\mu \frac{\partial G}{\partial \dot{x}_\mu} = 2G$ oder $G = \sum_\mu \dot{x}_\mu \frac{\partial G}{\partial \dot{x}_\mu} - G$. Will man hieraus die Funktion H erhalten, so hat man rechter Hand für $\dot{x}_\mu$ die Funktionen der u_ν und x_λ zu setzen und ausserdem $\frac{\partial G}{\partial \dot{x}_\mu} = u_\mu$ einzusetzen. Dann erhält man

$$H = \left(\sum_\mu \dot{x}_\mu u_\mu - G\right)_{\dot{x}_\mu = \text{ Funktion von } u_\nu \text{ und } x_\lambda}.$$

Hieraus ergibt sich

$$\frac{\partial H}{\partial X_i} = \sum_\mu u_\mu \frac{\partial \dot{x}_\mu}{\partial x_i} - \frac{\partial G}{\partial x_i} - \sum_\mu \frac{\partial G}{\partial \dot{x}_\mu}\frac{\partial \dot{x}_\mu}{\partial x_i},$$

und wegen $\frac{\partial G}{\partial \dot{x}_\mu} = u_\mu$ wird $\frac{\partial H}{\partial x_i} = -\frac{\partial G}{\partial x_i}$. Nun machen wir von den Gleichungen Gebrauch, denen die u_i genügen müssen und erhalten

$$\dot{u}_i = \frac{d}{dp}u_i = \frac{d}{dp}\frac{\partial G}{\partial \dot{x}_i} = \frac{\partial G}{\partial x_i}$$

oder

$$\frac{d}{dp}\frac{\partial G}{\partial \dot{x}_i} - \frac{\partial G}{\partial x_i} = 0,$$

wobei wegen $H = 0$ nur diejenigen Lösungen genommen werden dürfen, die $G = 0$ erfüllen. Damit ist gezeigt, dass auch das Variationsproblem

$$\delta \int_{p_1}^{p_2} (u_1\dot{x}_1 + u_2\dot{x}_2 + u_3\dot{x}_3)dp = 0, \qquad H = 0$$

auf die geodätischen Nulllinien führt und die *vollkommene Aequivalenz von Mongescher und Hamiltonscher Gleichung* ist klargelegt.

Machen wir uns die Verhältnisse am *Beispiel der Pseudoeuklidischen Geometrie* klar. Hier ist $G = \dot{x}_1^2 + \dot{x}_2^2 - \dot{x}_3^2$, also $u_i = \frac{\partial G}{\partial \dot{x}_i} = \begin{array}{cc} +2\dot{x}_i & i = 1,2 \\ -2\dot{x}_i & i = 3 \end{array}$. Hieraus folgt $H = \frac{1}{4}(u_1^2 + u_2^2 - u_3^2)$. Die Lagrangeschen kanonischen Gleichungen liefern $\dot{u}_i = -\frac{\partial H}{\partial x_i} = 0$, also wird $\ddot{x}_i = \frac{1}{2}\dot{u}_i = 0$, d. h. die geodätischen Linien[71] sind Gerade. Als geodätische Nulllinien findet man aus $G = 0$ oder $H = 0$
diejenigen, welche mit der x_3-Achse einen | Winkel von 45° bilden. 80

§ 36. Die vierdimensionale eigentliche u. Pseudogeometrie

Die vierdimensionale Geometrie

und die vierdimensionale Pseudogeometrie, mit der wir es in der Physik zu tun haben, unterscheiden sich in allem Wesentlichen von der eben behandelten dreidimensionalen so wenig, dass wir uns nun ganz kurz fassen können. In der Tat haben alle Betrachtungen, die wir im binären und ternären Gebiet angestellt hatten, die Eigenschaft, dass sie sich dem Inhalt und der Form nach ohne weiteres auf vier Dimensionen übertragen lassen. Um in der vierdimensionalen Mannigfaltigkeit x_1, x_2, x_3, x_4 Geometrie treiben zu können, benötigt man zehn Funktionen $g_{\mu\nu}$. Als *eigentliche* Geometrie bezeichnen wir wieder eine solche, in der $G(X_1, X_2, X_3, X_4)$ positiv definit ist, und haben als *analytische Bedingung* dafür

$$g_{11} > 0; \qquad \begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix} > 0; \qquad \begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix} > 0; \qquad g > 0;$$

wobei unter g die Determinante der $g_{\mu\nu}$ verstanden wird. In einem festen Punkt der Mannigfaltigkeit lässt sich dann G auf eine Summe von vier positiven Quadraten bringen.

Der einzig *wesentliche Unterschied gegen früher* ist, dass es nun *zwei verschiedene Pseudogeometrien* gibt: die in der Physik vorliegende, in welcher das Linienelement in einem Punkt der Mannigfaltigkeit auf die Form $\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^3 - \dot{x}_2^4$[72] gebracht werden kann (dies ist bekanntlich das Linienelement der kleinen
Relativitätstheorie) und die uns nicht weiter in|teressierende Geometrie, in 81
welcher zwei negative Vorzeichen auftreten $\dot{x}_1^2 + \dot{x}_2^2 - \dot{x}_3^2 + \dot{x}_4^2$. Dieser Unterschied gegenüber der dreidimensionalen Mannigfaltigkeit kommt daher, dass die Zahl 3 eben nur auf eine Weise in zwei Summanden zerlegt werden kann $3 = 2 + 1$, während 4 aus $3 + 1$ und $2 + 2$ entsteht. Um die Pseudogeometrie der Physik $4 = 3 + 1$ zu erhalten, muss man nur zur dreidimensionalen Pseudogeometrie eine solche Dimension hinzufügen, dass das hinzukommende Glied im Linienelement das positive Vorzeichen hat. Deswegen lassen sich

[71]"Linien" was corrected from "Nulllinien".

[72]Should read: "$\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2 - \dot{x}_4^2$"

auch alle Resultate des ternären Gebietes ohne weiteres auf die Physik übertragen, insbesondere auch die Hamilton-Jacobische Theorie. In der Mongeschen Differentialgleichung stehen nun drei unbekannte Funktionen $x_2(x_1)$, $x_3(x_1)$, $x_4(x_1)$; in der Hamiltonschen Gleichung eine Funktion von drei unabhängigen Veränderlichen $x_4(x_1, x_2, x_3)$. Eine Integralfläche derselben ist also noch eine dreidimensionale Mannigfaltigkeit.

Die Theorie der quadratischen Formen lehrt, dass das *Kriterium dafür*, dass wir die *eigentliche Geometrie* bezw. die eine oder die andere *Pseudogeometrie* vor uns haben, die Differenz s der Zahl der Vorzeichenfolgen und der Zahl der Vorzeichenwechsel in der Reihe der Determinanten

$$\begin{array}{ccccc} I & II & III & IV & V \\ 1, & g_{11}, & \begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix}, & \begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix}, & g \end{array}$$

ist. Der Fall von vier Vorzeichenfolgen liefert die eben genannte Bedingung für die eigentliche Geometrie. Dasselbe erhalten wir noch, wenn nur II und IV negativ sind. In der Tat werden dann alle Vorzeichen positiv, wenn nur $g_{\mu\nu}$
82 durch $-g_{\mu\nu}$ | ersetzt wird. Die in der Physik vorliegende Geometrie (Differenz $s = \pm 2$) wird realisiert, wenn nur V negativ ist, ferner wenn IV und V, endlich wenn III und V negativ sind und in den drei weiteren Fällen, die man erhält, wenn $g_{\mu\nu}$ durch $-g_{\mu\nu}$ ersetzt wird. Wir wollen als *Bedingung für die physikalische Pseudogeometrie*

$$g_{11} > 0, \qquad \begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix} > 0, \qquad \begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix} > 0, \qquad g < 0$$

festhalten.

§ 37. Zusammenhang der Theorie mit der Wirklichkeit

Wir haben nun alle Fragen erledigt, die sich auf die in der Physik herrschende Pseudogeometrie beziehen. Wir haben alle Sorgfalt aufgewendet auf die Konstruktion eines wunderbaren Fachwerks der Begriffe,[73] aber wir haben noch eine fundamentale Frage unbeantwortet gelassen, die den Physiker vor allem interessiert und die wir daher jetzt behandeln wollen, nämlich den

Zusammenhang unserer Theorie mit der Wirklichkeit,

[73]For a similar use of the phrase "Fachwerk der Begriffe", see Hilbert's lecture on *Epistemological Questions of Physics*, this Volume, p. 419 and passim.

mit den zu beobachtenden Erscheinungen. In der Tat soll sich die Summe alles physikalischen Geschehens in unserer Geometrie darstellen lassen. Jedes Ereignis und jeder Zusammenhang zwischen den Ereignissen soll sein Abbild in dieser vierdimensionalen Mannigfaltigkeit haben; jede physikalische Erscheinung muss hier Platz finden.

Da die physikalischen Ereignisse durch die zehn Funktionen $g_{\mu\nu}$, die sogenann- 83
ten Gravitationspotentiale, beschrieben werden, so muss sich der Physiker vor allem fragen, wie er *diese zehn Funktionen durch Experimente finden* kann. Erst wenn diese Frage gelöst ist, ist unsere Theorie nicht nur eine den Mathematiker durch ihre Einfachheit erfreuende Gedankenkonstruktion, sondern das, was sie sein will, eine die physikalischen Erscheinungen beschreibende Theorie. Erst dann haben wir den Zusammenhang mit der Wirklichkeit hergestellt; denn erst dann kann die Theorie an der Erfahrung geprüft werden.[74]

Unsere Theorie zeigt die *Welt als ein starres Gebilde*, in dem der Ablauf des Geschehens vorgeschrieben ist. Wir müssen uns also einen Geist denken, der, ohne die Welt der Physik zu beeinflussen, sich doch an jede beliebige Stelle derselben, d. h. an jeden Ort und zu jeder Zeit hinbegeben kann, insbesondere an solche Stellen, wo gerade einfache, der Messung und dem Experiment zugängliche Verhältnisse vorliegen.

Bevor wir die Frage beantworten, welche Apparate der Physiker besitzen muss, um die $g_{\mu\nu}$ in seiner vierdimensionalen Pseudogeometrie messen zu können, fragen wir, wie *diese Funktionen in einer zweidimensionalen Geometrie zu finden* sind. Erst die Kenntnis derselben setzt uns in Stand, Längen und Winkel zu messen, d. h. in unserer Ebene *Kilometersteine* anbringen zu können. Der einfachste Fall ist wieder der, dass $g_{\mu\nu} = \delta_{\mu\nu}$ ist. Der Physiker kann freilich erst durch Messung entscheiden, ob dieser Fall wirklich realisiert wird. Da er aber, wie wir behaupten, immer wenigstens angenähert dieses Resul-
tat finden wird, | und da in der alten Physik die Gültigkeit der Euklidischen 84
Geometrie vorausgesetzt wird, so wollen wir, um an Bekanntes anzuknüpfen, diesen Fall als Beispiel behandeln. Dann gelten die *Kongruenzsätze* als Folgen der Transformierbarkeit der Geometrie in sich. Ein Instrument, das diese Transformation realisiert, ist *der starre Körper*. Wir stellen also das Axiom auf: es gibt Körper in der Physik, die die Transformation der Ebene in sich durch einen einfachen Vorgang realisieren.[75] Diese Körper nennen wir starr und den Vorgang nennen wir Bewegung des Körpers. Dieser axiomatisch postulierte starre Körper setzt erst den Physiker in Stand, Längen und Winkel zu messen. Der starre Körper ist also der *Apparat*, den der Physiker *in einer Euklidischen Welt* benutzen müsste, *um die Theorie an der Erfahrung zu prüfen*, indem er zeigt, das dieser Körper bewegt werden kann. Die Möglichkeit des Vorhandenseins desselben beruht aber allein auf der Transformierbarkeit der Mannigfaltigkeit in sich. Ist die Transformation in sich in einer Geometrie

[74] In the left margin, Hilbert wrote in pencil: "von Ewigkeit zu Ewigkeit".

[75] For another discussion of the concept of a rigid measuring rod, see *Hilbert 1916a**, p. 3, (this Volume, p. 83).

unmöglich, so gibt es auch keine starren Körper mehr, und man kann gar nicht sagen, was man überhaupt unter Bewegung verstehen soll. Als anschauliches Beispiel einer solchen Geometrie hatten wir die Oberfläche des Paraboloids $Z = \frac{1}{2}(x^2+y^2)$ im Euklidischen Raum herangezogen. Auf dieser Fläche gelten die Kongruenzsätze bekanntlich nicht, und ein starres Centimetermass kann man nicht mehr anlegen; denn die Fläche ist ja krumm. Der Apparat, der in der Euklidischen Geometrie zum Messen diente, ist hier also unbrauchbar. Hierfür müssen wir Ersatz schaffen.

85 § 38. Axiomatische Begründung der eigentlichen Geometrie; der Massfaden

Zwei Gedanken weisen uns auf den richtigen Weg. Einmal bemerken wir, dass auch jetzt noch, d. h. im Falle der allgemeinen eigentlichen Geometrie, in der die Transformation in sich unmöglich ist, doch noch eine eindimensionale Mannigfaltigkeit, eine Kurve in sich transformiert, d. h. in sich bewegt werden kann. Ebenso kann eine Kurve in eine beliebige andere von gleicher Länge transformiert werden. Wir werden unserem *Massstab* also, damit er in der Punktmannigfaltigkeit — d. h. im Raum, in dem wir Physik treiben — frei beweglich ist, die *Gestalt eines Fadens* geben müssen. Die andere Bemerkung ist die, dass auch jetzt noch *im Infinitesimalen*, wo die $g_{\mu\nu}$ konstant sind, die *Euklidische Geometrie* gilt. Anders ausgedrückt: je kleiner ein Körper ist, desto eher kann er starr sein. Diese beiden Gedanken müssen wir verwerten, um einem Physiker, der in dieser eigentlichen Geometrie experimentiert, durch axiomatische Postulate ein Instrument in die Hand zu geben, mit dem er zunächst die Möglichkeit hat, festzustellen, dass die Massbestimmung festlegende Funktionen $g_{\mu\nu}$ vorhanden sind und mit dem er weiter diese $g_{\mu\nu}$ auch messen kann, um nachzuweisen, dass diese Funktionen wirklich gewisse Differentialgleichungen erfüllen, die wir erst weiter unten aus unserer Theorie der Materie ableiten wollen.

Wir führen folgende *Definitionen* ein:

Einen zwei-(drei-)dimensionalen Raum, in dem das Linienelement positiv definit ist, nennen wir einen *Streckenraum*.[76]

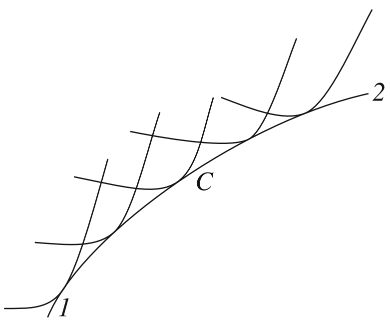
Der Apparat, mit dem der Physiker in diesem Raume die $g_{\mu\nu}$ messen kann,
86 und den wir nun durch seine Eigenschaften | axiomatisch definieren, heisse ein *Massfaden*. Der Zweck unserer Ueberlegungen ist nun der, durch Postulierung einfacher Axiome — wir werden mit zweien auskommen — es zu erreichen, dass die Funktionen $g_{\mu\nu}$ und ein Apparat, der längs irgend einer Kurve angelegt, die Zahl $\int_{p_1}^{p_2} \sqrt{\sum_{\mu\nu} g_{\mu\nu} \frac{dx_\mu}{dp} \frac{dx_\nu}{dp}}\, dp$ angibt, existieren. Wir wollen also

[76] For another definition of the concept of a “Streckenraum”, see *Hilbert 1917*, p. 58, (this Volume, p. 52).

eine *axiomatische Begründung* sowohl *unserer Geometrie* als auch des Messapparates geben. *Ein Physiker*, der nur diesen Massstab besitzt, *kennt von vornherein von seinem Streckenraum noch gar nichts*, weder die Zahl der Dimensionen desselben, noch die Koordinaten der Punkte, noch die Winkel, unter denen sich zwei Kurven schneiden. *Dies alles kann er* aber, wie wir zeigen werden, *experimentell finden*. Wir machen noch die Annahme, dass der Streckenraum[77] ruhen soll, und definieren diesen letzteren nun durch zwei Eigenschaften eindeutig auf folgende Weise.

§ 39. Die Enveloppeneigenschaft

1) Die *Enveloppeneigenschaft*: Gegeben sei eine beliebige, zwei Punkte 1 und 2 verbindende Kurve C des zweidimensionalen Streckenraums und eine Kurvenschar, deren Enveloppe die Kurve C ist.



Längs der Kurve werde der Massfaden angelegt und zeige eine bestimmte Zahl S. Ist die Kurve C in der Form $x = x(p)$, $y = y(p)$ gegeben, so wird also

$$S = \int_{p_1}^{p_2} F(x, y, \dot{x}, \dot{y}, \ddot{x}, \ddot{y}, \cdots, p)dp,$$

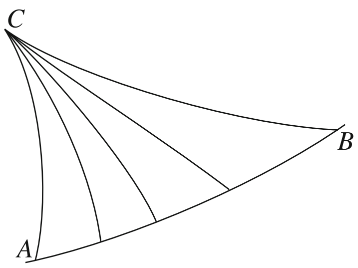
wobei F von vornherein beliebig hohe Ableitungen von x und y enthalten kann. Nun fordern wir aber *axiomatisch*, dass der | Massfaden dieselbe Zahl S anzeigt, 87
wenn er statt längs der gegebenen Kurve angelegt zu werden, längs aller infinitesimalen, die Kurve C berührenden Kurvenstücke ds der Kurvenschar angelegt wird. Diese Eigenschaft kann man auch als die Möglichkeit des Abrollens, als die *Abrollbarkeit des Massfadens* bezeichnen. Wir teilen ohne Beweis mit, dass dies darauf hinausläuft, dass die Funktion F ausser p, x und y selbst nur die ersten Ableitungen x und y[78] als Argumente enthält, oder, wie wir uns auch ausdrücken können: F *hängt* in jedem Kurvenpunkt *nur von der Richtung der Kurve*, nicht aber von ihrer Krümmung *ab*; denn die Krümmung enthält ja schon zweite Ableitungen. Wir können sogar sagen, dass F *eine in $\dot{x}$ und $\dot{y}$ homogene Funktion vom ersten Grad* sein muss und p nicht explizit enthalten darf. Diese Forderung ist notwendig, weil die Kurve von ihrer Parameterdarstellung unabhängig ist und also die Zahl S nicht geändert werden darf, wenn p durch eine Funktion $\varphi(p)$ ersetzt wird.

[77]"Streckenraum" was corrected from "Streckenraum in Bezug auf den Massfaden".
[78]Should be "$\dot{x}$ und $\dot{y}$".

§ 40. Gültigkeit des Pythagoräischen Lehrsatzes im Infinitesimalen

2) Wird der Apparat an ein *infinitesimales rechtwinkliges Dreieck* (wie wir ein solches definieren und konstruieren können, werden wir gleich zeigen) angelegt, so soll er die *Gültigkeit des Pythagoräischen Lehrsatzes* anzeigen. Diese Eigenschaft ist, um es nochmals zu betonen, ebenso wie die erste, insofern auch eine Eigenschaft des Streckenraumes, als dieser eben von einer bestimmten Beschaffenheit sein muss, damit er die Möglichkeit der Existenz eines solchen Apparates zulässt. Die Eigenschaft 2) ist äquivalent damit, dass im Infinitesimalen die Euklidische Geometrie gilt.

88 Bevor wir zeigen, dass diese beiden Eigenschaften | die Funktion F eindeutig explizit bestimmen, wollen wir erklären, wie mit *Hilfe des Apparates ein rechtwinkliges Dreieck konstruiert* werden kann. Wir greifen aus unserer Mannigfaltigkeit zwei beliebige Punkte A und B heraus. Dann spannen wir den Massfaden zwischen diesen beiden Punkten straff an. An demselben können wir nun eine Zahl ablesen, nämlich die Länge der durch den Faden selbst gebildeten Verbindungskurve AB. Da der Massfaden aber straff angezogen wurde, d. h. so, dass diese Zahl ein Minimum wird, so ist die Kurve, an die der Faden sich anschmiegt und deren Länge er misst, gerade die *geodätische Linie* AB.

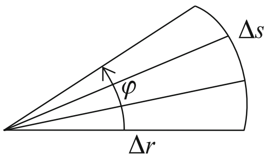


Durch den Punkt A legen wir nun mit Hilfe derselben Konstruktion eine beliebige andere geodätische Linie AC. Wann steht diese Linie auf AB senkrecht? Um dies festzustellen, greifen wir einen beliebigen Punkt der Kurve AC heraus, legen durch ihn alle geodätischen Linien nach den Punkten von AB und lesen auf dem Massfaden die zugehörigen Zahlen ab. Ist unter allen diesen Zahlen die längs der Kurve AC selber gemessene die kleinste, so sagen wir, *die geodätische Linie AC steht senkrecht* auf AB.

§ 41. Axiom von der reziproken Orthogonalität

Die *analytische Bedingung* für das Aufeinandersenkrechtstehen müssen wir ohne Beweis aus der Variationsrechnung übernehmen. Dort wird gezeigt, dass das Integral $\int_A^C F(x, y, \dot{x}, \dot{y})\, dp$ erstreckt über eine Kurve, die den festen Punkt C mit dem auf der Kurve $X = X(p)$, $Y = Y(p)$ liegenden beweglichen Punkt
89 A verbindet, dann ein Minimum wird, wenn die Orthogonalitätsbedin|gung (dort: Transversalitätsbedingung) $\frac{\partial F}{\partial \dot{x}}\dot{X} + \frac{\partial F}{\partial \dot{y}}\dot{Y} = 0$ erfüllt ist. Diese Bedingung ist in den beiden Richtungen $\dot{x}_i$ und $\dot{X}_i$ im allgemeinen unsymmetrisch.

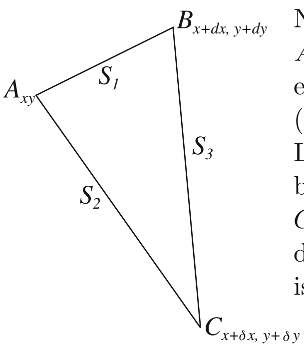
Es folgt also im allgemeinen aus $AB \perp AC$ nicht auch $AC \perp AB$. Damit dies im 3 oder mehrdimensionalen Raum[79] der Fall ist, muss vielmehr F eine bestimmte Gestalt haben und zwar gerade $F = \sqrt{\sum_{\mu\nu} g_{\mu\nu}\dot{x}_\mu\dot{x}_\nu}$ (siehe W. Blaschke, Ueber räumliche Variationsprobleme mit symmetrischer Transversalitätsbedingung, Leipziger Berichte, Math. phys. Kl. 27, (1916) S. 50).[80] Wir könnten also durch die *axiomatische Forderung*, dass die *Orthogonalität* zweier Kurven immer *eine gegenseitige* sein soll, die gewünschte Form von F, nämlich $\sqrt{\sum_{\mu\nu} g_{\mu\nu}\dot{x}_\mu\dot{x}_\nu}$ folgern. Wenn wir hier dasselbe aus der Forderung des Pythagoräischen Lehrsatzes im Unendlichkleinen folgern, so geschieht es nur, weil dieser Beweis kürzer zu führen ist.



Nachdem wir wissen, was ein rechter Winkel ist, können wir auch einen beliebigen Winkel mit dem Massfaden definieren und messen. Gegeben seien nämlich zwei geodätische Linien, die sich schneiden. Dann ziehen wir durch den Schnittpunkt nach allen Richtungen geodätische Linien von der Länge Δr und konstruieren die zu denselben orthogonalen Trajektorien. Deren Länge bestimmen wir mit dem Massfaden als Δs. Dann bezeichnen wir als Winkel

$$\varphi = \lim_{\Delta r} \frac{\Delta S}{\Delta r}.$$

§ 42. Bestimmung der Funktion $F^2 = \sum_{\mu\nu} g_{\mu\nu}\dot{x}_\mu\dot{x}_\nu$ aus 2 Axiomen



Nun wollen wir die *Funktion F aus* den beiden *Axiomen 1) und 2) bestimmen.* Wir konstruieren ein beliebiges infinitesimales rechtwinkliges Dreieck (dessen Seiten sind als |unendlich kleine geodätische 90
Linien einfach Gerade). Die drei Eckpunkte mögen bezw. die Koordinaten $A(x, y), B(x+dx, y+dy)$ und $C(x+\delta x, y+\delta y)$ haben. Jetzt legen wir den Massfaden an die Seite AB an. Die Gleichung dieser Seite ist in Parameterdarstellung

$$x(p) = x + pdx, \qquad \text{also} \quad \dot{x} = dx$$
$$y(p) = y + pdx, \qquad \text{"} \quad \dot{y} = dy.$$

[79]“im 3 oder mehrdimensionalen Raum” was interlineated.

[80]*Blaschke 1916*.

Dabei durchläuft p die Werte von 0 bis 1. Statt $F(x(p), y(p), \dot{x}(p), \dot{y}(p))$ wird der Wert im Punkte A: $F(x, y, dx, dy)$ genommen. Dann liefert der erste Mittelwertsatz:

$$S_1 = \int_A^B F(x(p), y(p), \dot{x}(p), \dot{y}(p))\, dp = F(x, y, dx, dy).$$

Für die Seite AC ergibt sich ebenso

$$S_2 = \int_A^C F(x(p), y(p), \dot{x}(p), \dot{y}(p)) = F(x, y, \delta x, \delta y).$$

Die Gleichung der dritten Seite ist

$$\begin{aligned} x(p) &= x + dx + p(\delta x - dx), \\ y(p) &= y + dy + p(\delta y - dy), \end{aligned}$$

also wird

$$S_3 = \int_B^C F(x(p), y(p), \dot{x}(p), \dot{y}(p))dp = F(x, y, \delta x - dx, \delta y - dy).$$

Nun wenden wir den Pythagoräischen Lehrsatz an

$$F^2(dx, dy) + F^2(\delta x, \delta y) = F^2(\delta x - dx, \delta y - dy).$$

Dies ist eine *Funktionalgleichung für* F^2, die wir leicht auflösen können, wenn wir die Taylorentwicklung nach dx und dy anwenden. Wir erhalten

$$\begin{aligned} F^2(\delta x - dx, \delta y - dy) &= F^2(\delta x, \delta y) - \left(\frac{\partial F}{\partial \delta x} dx + \frac{\partial F}{\partial \delta y} dy\right) \\ &+ \frac{1}{2}\left\{\frac{\partial^2 F}{\partial (\delta x)^2} dx^2 + 2\frac{\partial^2 F}{\partial \delta x \partial \delta y} dx dy + \frac{\partial^2 F}{\partial (\delta y)^2} dy^2\right\} + \cdots . \end{aligned}$$

91 Wegen der Orthogonalitätsbedingung wird $\frac{\partial F}{\partial \delta x} dx + \frac{\partial F}{\partial \delta y} dy = 0$. Ferner ist F homogen vom ersten Grade in dx und dy, also fallen alle Glieder mit dx^3, $dx^2 dy, \cdots$ usw. weg. Setzen wir den so gefundenen Ausdruck in die Funktionalgleichung ein, so wird $F^2 = \frac{1}{2}\left\{\frac{\partial^2 F}{\partial (\delta x)^2} + 2\frac{\partial^2 F}{\partial \delta x \partial \delta y} dx dy + \frac{\partial^2 F}{\partial (\delta y)^2} dy^2\right\} = \sum_{\mu\nu} g_{\mu\nu} dx_\mu dx_\nu$. Wir erhalten $S = \int_{p_1}^{p_2} \sqrt{\sum_{\mu\nu} g_{\mu\nu} \frac{dx_\mu}{dp} \frac{dx_\nu}{dp}}\, dp$ q.e.d.

§ 43. Uebertragung der Resultate auf die Pseudogeometrie

Diese Resultate haben wir nun auf den vierdimensionalen physikalischen Raum, die Welt, in welcher das Linienelement nicht mehr positiv definit ist, und die also in unserer Ausdrucksweise keinen Streckenraum darstellt,

zu übertragen. Wir haben bei der Definition des Massfadens gefordert, dass der Streckenraum ruhe. Der Begriff der Ruhe ist aber etwas relatives, vom zu Grunde gelegten Koordinatensystem abhängiges. Wir wollen aber gerade allgemein invariante Eigenschaften der Pseudogeometrie messen können. Daher müssen wir den Apparat ganz allgemein anwenden können. Wir fordern also wieder axiomatisch: *Der Massfaden soll sich über irgend eine Weltkurve spannen lassen*, d. h. über eine Kurve, deren einzelne Punkte zu verschiedenen Zeiten existieren. *Da die Welt eine Pseudogeometrie* der Form $3+1$ darstellt, *steht dem Physiker* neben dem Massfaden *noch ein anderer Apparat* gleichberechtigt *zur Verfügung*;[81] denn in einer x, t-Ebene gilt ja im Infinitesimalen gar nicht Euklidische, sondern die Pseudoeuklidische Geometrie. Es muss also in der Welt auch solche Kurven geben, dass der Massfaden, wenn er längs derselben angelegt wird, keine Zahl angibt, dass | also der physikalischen 92
Apparat in diesem Falle versagt. Ganz allgemein wird der Physiker *folgende Verhältnisse realisiert* finden:

Sind mit Hilfe des Massfadens in einem Weltpunkt drei zueinander senkrecht stehende Richtungen aufgefunden, so kann mit diesem Instrument, trotzdem die Welt noch eine vierte Dimension hat, keine neue Richtung aufgefunden werden, die auf diesen drei Richtungen senkrecht steht. Wir sagen weiter: Werden ⟨wir⟩ *zwei Ereignisse* durch eine geodätische Weltlinie verbinden, die mit dem Massfaden gemessen werden kann, so sind dieselben *voneinander unabhängig*. Kann diese geodätische Linie aber nicht mehr mit dem Massfaden gemessen werden, so könnnen die Ereignisse *im Verhältnis von Ursache und Wirkung* zueinander stehen. Die Richtungen der durch einen Punkt gehenden geodätischen Linien, auf welchen der Massfaden anwendbar ist, kann man auch durch Rechnung finden. In diesem Punkt habe das Linienelement die Form $\sum_{\mu\nu}(g_{\mu\nu})_0 dx_\mu dx_\nu$. Dann lässt sich eine solche lineare Transformation (auf Hauptachsen) angeben, dass diese quadratische Form zu $dx^2 + dy^2 + dz^2 - dt^2$ wird. Alle geodätischen Linien durch den Nullpunkt nun, die mit der neuen t-Achse einen Winkel von mehr als 45° bilden, sind mit dem Massfaden zu messen. Beträgt der Winkel 45°, so zeigt der Faden die Länge Null und gehen wir darüber hinaus, so kommen wir in das Gebiet, in welchem dieser physikalische Apparat versagt.

Um auf diesen Weltlinien die Länge zu messen, muss der Physiker also ein anderes Instrument,

die *Lichtuhr* 93

benutzen, die wir wieder durch ihre Eigenschaften axiomatisch definieren.

Die *Länge der Weltlinie*, die mit der Lichtuhr gemessen wird, nennen wir die *Eigenzeit* τ. Die Weltlinie selber nennen wir dann eine *Zeitlinie*. Von der Lichtuhr verlangen wir wieder

[81] "Verfügung" was corrected from "Seite".

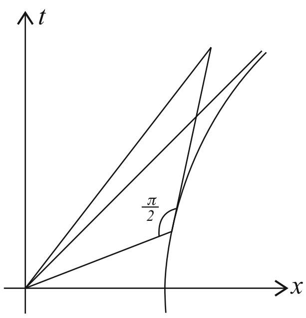
1) die Enveloppeneigenschaft.

Für Zeitlinien in einer x, t-Ebene folgt daraus

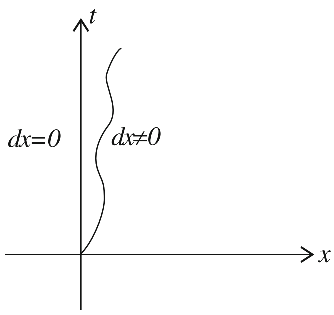
$$\tau = \int_{p_1}^{p_2} F(x, t, \dot{x}, \dot{t})dp,$$

wo F eine *homogene Funktion ersten Grades* von $\dot{x}$ und $\dot{t}$ allein ist.

2) Da nun im Infinitesimalen die Pseudoeuklidische Geometrie, d. h. das kleine Relativitätsprinzip gilt, weil ja $x^2 - t^2$ eine Invariante der *Lorentztransformation* ist, so *fordern wir*:



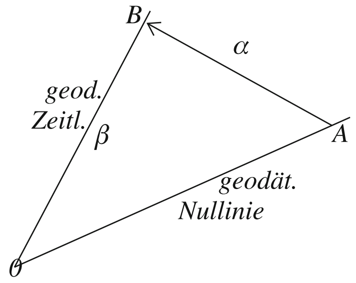
Wird die Lichtuhr an ein infinitesimales rechtwinkliges Dreieck angelegt, so soll der Pseudopythgoräische Lehrsatz (Differenz der Quadrate der Katheten=Quadrat der Hypotenuse) gelten. In der Pseudoeuklidischen Geometrie stehen zwei Gerade aufeinander senkrecht, wenn die eine Gerade ein Durchmesser der gleichseitigen Hyperbel $t^2 - x^2 =$ const. und die andere Gerade die konjugierte Tangente ist. Die eine Kathete muss also mit
94 dem Massfaden, die andere mit der | Lichtuhr gemessen werden. Die Hypotenuse wird bezw. mit dem Massfaden und mit der Lichtuhr gemessen, wenn sie einen Winkel von bezw. weniger als 45° und mehr als 45° mit der x-Achse bildet.



Geodätische Linien, die einen Winkel von 45° mit der x-Achse bilden, können wir eben noch mit dem Massfaden und eben noch mit der Lichtuhr messen.Ersteres Instrument zeigt $s = 0$, letzteres zeigt $\tau = 0$. Auch die *Orthogonalität* kann *festgestellt*, also können wieder Winkel d. h. Geschwindigkeiten[82] gemessen werden; und zwar steht eine Zeitlinie auf einer Weltlinie senkrecht, wenn die Zahl, die die Lichtuhr anzeigt, ein Maximum ist. In der Tat wird $d\tau = \sqrt{dt^2 - dx^2}$ für $dx = 0$, d. h. für die t-Achse ein Maximum.

Wir wollen dem zweiten Axiom noch eine *andre Deutung* geben, in der es physikalisch anschaulich wird und in der es durch das Experiment bewiesen wurde.

[82] "d. h. Geschwindigkeiten" was interlineated.



Wir konstruieren von einem Punkt aus eine solche geodätische Weltlinie, dass dieselbe mit der Lichtuhr gemessen werden kann und eine andere, auf der die Lichtuhr gerade versagt. Von einem beliebigen Punkt dieser zweiten geodätischen Linie konstruieren wir das Lot auf der erste. Seine Länge sei α. Den Fusspunkt des Lotes bezeichnen wir mit B und die mit der Lichtuhr gemessene Länge OB mit β. Nun fordern wir im Infinitesimalen axiomatisch $\lim\limits_{\alpha=0}\alpha = \beta$. Hieraus folgt dann wieder $F^2 = \sum_{\mu\nu} g_{\mu\nu} \frac{dx_\mu}{dp} \frac{dx_\nu}{dp}$.[83]
Dieses Axiom wird durch den Michelsonschen Versuch tatsächlich an der Er- 95
fahrung *bestätigt*. Es ist also eine aus der Physik stammende kontrollierte Tatsache, die darauf hinausläuft, dass sich das Licht von einem bewegten Körper mit einer nach allen Richtungen hin konstanten Gechwindigkeit ausbreitet.

Die beiden von uns skizzierten Apparate sind nun freilich in der Wirklichkeit nicht realisierbar, aber sie werden doch mit einer gewissen Annäherung durch den Physiker zu konstruieren sein, und die physikalischen Institute sind dazu da, um an den tatsächlich verwendeten Instrumenten diese Korrektionen anzubringen.

§ 46. Bestimmung der Gravitationspotentiale mit Massfaden und Lichtuhr

Jetzt können wir die $g_{\mu\nu}$ an jeder Stelle der Welt *finden*, und zwar zeigen wir, dass *jedes der beiden Instrumente* Massfaden und Lichtuhr *für sich ausreicht*, um mit seiner Hilfe die Werte der $g_{\mu\nu}$ als Funktionen von x_i zu berechnen, sobald nur ein bestimmtes Raum-Zeit-Koordinatensystem x_i eingeführt worden ist. In der Tat wählen wir irgend 10 Strecken aus, die sämtlich längs verschiedenen Richtungen in den nämlichen Weltpunkt x_s einlaufen, so dass diesem Endpunkt jedesmal derselbe Parameterwert p zukommt, dann ergibt sich für jede der 10 Strecken im Endpunkt die Gleichung

$$\left(\frac{ds^{(h)}}{dp}\right)^2 = \sum_{\mu\nu} g_{\mu\nu} \frac{dx_\mu^{(h)}}{dp} \frac{dx_\nu^{(h)}}{dp} \qquad (h = 1, 2, \cdots 10),$$

hier sind die linken Seiten bekannt, sobald wir die Längen s mittelst des *Massfadens* bestimmt haben: Setzen wir nun zur Abkürzung $\sum_{\mu\nu} g_{\mu\nu} X_\mu X_\nu =$
u und: | 96

[83]The preceding sentence was interlineated.

$$D(u) = \begin{vmatrix} \left(\frac{dx_1^{(1)}}{dp}\right)^2, & \frac{dx_1^{(1)}}{dp}\frac{dx_2^{(1)}}{dp}, & \cdots & \left(\frac{dx_4^{(1)}}{dp}\right)^2, & \left(\frac{ds^{(1)}}{dp}\right)^2 \\ \vdots & \vdots & & \vdots & \vdots \\ \left(\frac{dx_1^{(10)}}{dp}\right)^2, & \frac{dx_1^{(10)}}{dp}\frac{dx_2^{(10)}}{dp}, & \cdots & \left(\frac{dx_4^{(10)}}{dp}\right)^2, & \left(\frac{ds^{(10)}}{dp}\right)^2 \\ X_1^2, & X_1 \cdot X_2, & \cdots & X_4^2, & u \end{vmatrix},$$

so wird $D(0) + u\frac{\partial D}{\partial u} = 0$, also

$$\sum_{\mu\nu} g_{\mu\nu} X_\mu X_\nu = -\frac{D(0)}{\frac{\partial D}{\partial u}}, \tag{25}$$

wodurch sich zugleich für die Richtungen der ausgewählten 10 Strecken im Punkte $x_i(p)$ die *Bedingung*

$$\frac{\partial D}{\partial u} \neq 0$$

als *notwendig* herausstellt, damit sich alle 10 Funktionen berechnen lassen.

Ist $\sum g_{\mu\nu} X_\mu X_\nu$ nach (25) berechnet, so würde die Anwendung des Verfahrens auf irgend eine 11te Strecke, die in $x_i(p)$ endigt, die Gleichung

$$\left(\frac{ds^{(11)}}{dp}\right)^2 = \sum_{\mu\nu} g_{\mu\nu} \frac{dx_\mu^{(11)}}{dp}\frac{dx_\nu^{(11)}}{dp}$$

liefern und diese Gleichung wäre dann sowohl eine Kontrolle für die Richtigkeit des Massfadens als auch eine experimentelle Bestätigung dafür, dass die Voraussetzungen der Theorie für die wirkliche Welt zutreffen.

Für die Lichtuhr gilt die entsprechende Ueberlegung.

97 § 47. Ursache und Wirkung

Wir wenden uns nun zur Beantwortung einer anderen Frage, die ebenfalls für den Physiker von fundamentalster Wichtigkeit ist, wir wollen nämlich

das Kausalitätsprinzip in der Physik

behandeln. Dazu müssen wir an Betrachtungen anschliessen, die wir oben (S. 82) angestellt hatten. Dort wurden die Bedingungen dafür aufgestellt, dass die quadratische Form $\sum g_{\mu\nu} X_\mu X_\nu$, wenn sie in einem Punkt auf die Normalform einer Summe von vier Quadraten gebracht wird, drei positive und ein negatives Vorzeichen hat. Wir werden sehen, dass *obige Bedingungen*, die durch Ungleichungen ausgedrückt waren, *allein nicht hinreichen, um die vierdimensionale Mannigfaltigkeit* x_1, x_2, x_3, x_4 mit ihren 10 die Massbestimmungen festlegenden Funktionen $g_{\mu\nu}$ *zu einer physikalisch zulässigen zu machen.*

In der alten Physik wurde die Gültigkeit der Euklidischen Geometrie a priori angenommen. Die Zeit als vierte Dimension stand mit den drei Raumdimensionen in gar keinem Zusammenhang. Einen grossen Schritt weiter in der *Verschmelzung von Raum und Zeit* in vier gleichberechtigten Dimensionen ging die sogenannte “kleine Relativitätstheorie”, welche bekanntlich lehrt, dass die Naturgesetze in einer dreifach unendlichen Mannigfaltigkeit von Bezugssystemen ihren einfachsten Ausdruck erhalten, die durch solche lineare Transformationen auseinander hervorgehen, dass die quadratische Form $x^2 + y^2 + z^2 - t^2$ dabei unverändert bleibt. Die allgemeine Relativitätstheorie geht | nun noch viel weiter in der Vereinheitlichung von Raum und Zeit. Sie 98
sieht von vornherein alle Bezugssysteme, die durch beliebige Transformation aus einem gegebenen hervorgehen, als gleichberechtigt an. Dadurch wird der Zeit ihre Sonderstellung unter den vier Dimensionen der Welt des Physikers vollkommen genommen. Mit der Forderung dieser allgemeinsten Relativität sind wir aber, wie wir nun zeigen wollen, doch ein wenig zu weit gegangen. Auch jetzt noch bleibt eine *kleine Trennung zwischen Raum und Zeit* bestehen, eine Trennung freilich, die nur noch ein Schatten der früheren ist.

Wenn wir sagen, dass das physikalische Geschehen in einer vierdimensionalen Mannigfaltigkeit stattfindet, so meinen wir damit nur, dass es in einer solchen Mannigfaltigkeit gedeutet werden kann, dass uns *diese Mannigfaltigkeit* also *ein Bild der Wirklichkeit* liefert. Von diesem Bilde müssen wir dann aber verlangen, das es alle uns von der Welt des Physikers bekannten Eigenschaften ebenfalls aufweist. Eine solche ist in erster Linie die *Kausalität*, oder zeitliche Aufeinanderfolge von Ursache und Wirkung. Betrachten wir die physikalischen Ereignisse von einem bestimmten Bezugssystem aus, so können wir ein gewisses Punktereignis B nur dann als Folge eines anderen A ansprechen, wenn es durch Lichtsignale von A aus erreicht werden kann. Wir sind gewohnt, diese *kausale Folge* der Ereignisse als etwas *vom Bezugssystem Unabhängiges* anzusehen, und wir werden sie nicht ohne Not aufgeben, trotzdem sie keine logische Forderung ist, d. h. eine Forderung, deren Auf|gabe zu Widersprüchen 99
des Denkens führen würde. Daher werden wir auch von unserem Bilde verlangen, dass das Verhältnis von Ursache und Wirkung zwischen zwei Ereignissen daselbst bestehen bleibt.

Als Zeitlinie haben wir oben eine solche Weltlinie bezeichnet, deren Länge mit der Lichtuhr gemessen werden muss. Die Bedingung dafür, dass $x_i = x_i(p)$ eine Zeitlinie darstellt ist $\sum g_{\mu\nu}\dot{x}_\mu(p)\dot{x}_\nu(p) < 0$. *Zwei Ereignisse* nun, *die im Verhältnis von Ursache und Wirkung stehen, können immer durch eine Zeitlinie verbunden werden.* Wir müssen also verlangen, dass zwei Punkte einer Zeitlinie, die in Bezug auf das gegebene Koordinatensystem verschiedene Zeiten hatten, durch keine Transformation auf Gleichzeitigkeit gebracht werden können. Würde man nun alle Transformationen als erlaubt ansehen, so könnte man aber unendlich viele angeben, für welche diese beiden Ereignisse gleichzeitig stattfinden. Wir kommen so dazu, dass wir den Koordinatensystemen, die wir als Raum und Zeit ansehen können, *gewisse Beschränkungen* auferlegen müssen. Im kleinen Relativitätsprinzip tritt diese Schwierigkeit nicht auf; dort sind wegen der grossen Beschränktheit der zulässigen Transformationen von selbst alle diejenigen ausgeschlossen, welche Ursache und Wirkung vertauschen können.

§ 48. Bedingungen für erlaubte (eigentliche) Raum-Zeit-transformationen

Wir wollen nun zeigen, dass auch in der allgemeinen Relativitätstheorie *zwei Punktereignisse einer Zeitlinie nicht auf Gleichzeitigkeit transformiert* werden können durch alle Transformationen, die den durch die *Ungleichungen* (S. 82)

$$g_{11} > 0, \qquad \begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix} > 0, \qquad \begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix} > 0, \qquad g < 0$$

100 ausgedrückten Beschränkungen genügen. Dieselben sind in der Tat eine Einschränkung der erlaubten Transformationen, denn wir haben auf S. 82 gesehen, dass es 6 Systeme von 4 Ungleichungen gibt, die alle die in der Physik herrschende Pseudogeometrie realisieren. Durch eine beliebige Transformation wird sich im allgemeinen unser System von Ungleichungen in ein anderes transformieren. Alle solchen Transformationen sind also unzulässig. Um zu beweisen, dass erlaubte Transformationen Ursache und Wirkung nicht auf Gleichzeitigkeit bringen können, haben wir nun zu bemerken, dass *für eine Zeitlinie* $\dot{x}_4(p) \neq 0$ sein muss. Für eine solche Linie gilt nämlich nach Definition $G = \sum g_{\mu\nu}\dot{x}_\mu\dot{x}_\nu < 0$. Wäre nun $\dot{x}_4(p) = 0$, so folgte aus unseren Ungleichungen $G > 0$ im Widerspruch zur Voraussetzung.

Die Ungleichung $\dot{x}_4(p) \neq 0$ kann durch Einführung eines anderen Parameters $p = p(p')$ sofort zerstört werden. Dies ist nur scheinbar ein Widerspruch gegen den Satz, dass die Zeitlinie von der Parameterdarstellung unabhängig ist. In der Tat wird $\dot{x}_4(p') = \dot{x}_4(p)\frac{dp}{dp'} = 0$ nur, wenn $\frac{dp}{dp'} = 0$ ist. Dann wird aber auch $\dot{x}_1 = \dot{x}_2 = \dot{x}_3 = 0$. Während also p' eine Reihe von Werten durchläuft, bleiben sämtliche $x_i = \text{const}$. Eine solche Parameterdarstellung ist unzulässig.

Wir schliessen aus $\dot{x}_4(p) \neq 0$, dass auf einer Zeitlinie $x_4(p)$ entweder immer zunehmen oder immer abnehmen muss. Da aber der Begriff der Zeitlinie etwas Invariantes ist (die Ungleichung $G < 0$ ist unabhängig vom Bezugssystem), so muss auch für das neue Koordinatensystem $\dot{x}_4(p)$ entweder immer
101 zunehmen oder immer abnehmen, d. h. *zwei Punkte einer Zeitlinie können nicht auf Gleichzeitigkeit transformiert werden.*[84] Hieraus folgt weiter, dass

[84] At the bottom of the page, the composer of the manuscript added: "Ein anderer wichtiger Satz über Zeitlinien findet sich auf S. 140."

das Verhältnis von Ursache und Wirkung zweier Ereignisse durch Transformation auch nicht umgekehrt werden kann. Es kann wohl vorkommen, dass $x_4(p)$ vor der Transformation auf der Zeitlinie bei wachsendem p immer zunimmt, während nach derselben das neue $x'_4(p)$ bei wachsenden p immer abnimmt. Da aber die Reihenfolge der Ereignisse auf der Zeitlinie durch die Transformation nicht geändert wird, wenn man den Parameter p beide Male dieselben Werte durchlaufen lässt, so bleibt die Kausalität doch erhalten. Man muss nur in diesem neuen Bezugssystem denjenigen von zwei Punkten einer Zeitlinie, welcher das grössere x'_4 hat, als Ursache und den mit kleinerem x'_4 als Wirkung ansprechen.

Beispiel: Wir wollen annehmen, dass in der x, t-Ebene die Pseudogeometrie $g_{11} = 1$, $g_{14} = 0$, $g_{44} = -1$ herrschen soll. Eine Gerade durch den Nullpunkt kann in Parameterdarstellung durch $x = \alpha p$, $t = p$ gegeben werden. Sie ist dann eine Zeitlinie, $G < 0$, oder Raumlinie (Strecke), $g > 0$, je nachdem $\alpha < 1$ oder $\alpha > 1$ ist. In der Tat wird für dieselbe

$$G = \alpha^2 - 1 \begin{array}{lll} < 0 & \text{für} & \alpha < 1, \\ > 0 & \text{für} & \alpha > 1. \end{array}$$

Wir versuchen die Punkte dieser Geraden auf Gleichzeitigkeit zu transformieren, indem wir sie zur neuen x'-Achse, d. h. $t' = 0$ machen. Wir setzen $x' = x$, $t' = x - \alpha t$ und erhalten

$$g'_{11} = 1 - \frac{1}{\alpha^2}, \quad g'_{14} = \frac{1}{\alpha^2}, \quad g'_{44} = -\frac{1}{\alpha^2}, \quad g' = -\frac{1}{\alpha^2}.$$

Die Transformation ist, wie wir sahen, erlaubt, wenn $g'_{11} > 0$ und $g' < 0$ ist. 102
Die erste Bedingung ist nun nur erfüllt, wenn $\alpha > 1$ ist, d. h. wenn die Gerade eine Strecke ist.

Aus der Tatsache, dass sich längs einer Zeitlinie x_4 unabhängig vom Bezugssystem nur in einer Richtung verändert, folgt weiter, dass es keine geschlossenen Zeitlinien geben kann. Wir haben ferner den

Satz: *Kann man die Punkte einer Weltlinie* alle *auf Gleichzeitigkeit transformieren, so ist diese Weltlinie eine Strecke*, d. h. eine Linie, die mit dem Massfaden gemessen werden kann, für welche also $G > 0$ ist. In der Tat gilt für Gleichzeitigkeit $\dot{x}_4(p) = 0$ also ist, wie es sein muss, $G > 0$. Durch diesen Satz wurde aber *nicht* bewiesen, dass man die Punkte einer Strecke *immer* auf Gleichzeitigkeit transformieren kann.

§ 49. Weitere Einschränkung: $g_{44} < 0$

Wir müssen nun, um mit der Wirklichkeit in Uebereinstimmung zu bleiben, noch *eine weitere einschränkende Forderung* an unser Koordinatensystem stellen, die bewirkt, *dass x_4 den Charakter der physikalischen Zeit hat*. Wir müssen nämlich fordern, dass zwei Punkte der *Zeitachse* selber auch nicht auf

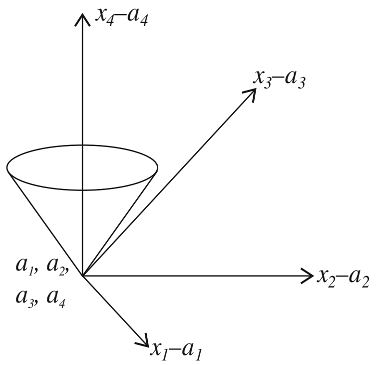
Gleichzeitigkeit transformiert werden können, d. h. dass die Zeitachse selbst ebenfalls eine Zeitlinie ist. Die Zeitachse hat die Gleichung $x_1 = 0$, $x_2 = 0$, $x_3 = 0$, also wird $G = g_{44}\dot{x}_4(p)$ und damit $G < 0$ wird, muss auch $g_{44} < 0$ sein. Diese Ungleichung

$$g_{44} < 0$$

103 zusammen mit den vier Ungleichungen

$$g_{11} > 0, \qquad \begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix} > 0, \qquad \begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix} > 0, \qquad g < 0$$

charakterisiert erst *die physikalisch zulässigen Bezugssysteme*. Hiermit haben wir nun auch jene kleine, oben erwähnte Trennung der drei Raumdimensionen von der Zeitdimension erhalten; denn wir haben die vier Dimensionen x_1, x_2, x_3, x_4, von denen wir von vornherein keine auszeichnen können, so zu benennen, dass obige Ungleichungen erfüllt sind. Dann muss diejenige Veränderliche, deren Koeffizient g_{44} ist, als Zeit angesprochen werden. Diese Ungleichungen stellen eine Beschränkung der in der Physik durch die $g_{\mu\nu}$ a priori möglichen Massbestimmungen dar. Koordinatensysteme, die dieselben erfüllen, wollen wir *eigentliche Raumzeitkoordinaten* nennen. *Eine Transformation, bei der dieselben bestehen bleiben, heisse eine eigentliche Raumzeittransformation.*



Um uns die *Bedeutung* dieser Ungleichungen *geometrisch* klar zu machen, betrachten wir im Weltpunkt a_1, a_2, a_3, a_4 den zugehörigen Nullkegel

$$\sum g_{\mu\nu}(X_\mu - a_\mu)(X_\nu - a_\nu) = 0.$$

Für die Punkte des dreidimensionalen Raumes $X_4 = a_4$ wird infolge unserer Ungleichungen $\sum_{\mu\nu} g_{\mu\nu}(X_\mu - a_\mu)(X_\nu - a_\nu) > 0$. *Dieser Raum liegt* also ganz *ausserhalb des Nullkegels. Die Gerade* $X_1 = a_1$, $X_2 = a_2$, $X_3 = a_3$ liegt dagegen ganz *innerhalb des Kegels*, weil $g_{44}(X_4 - a_4) < 0$
104 | wird. Die Zusatzbedingung $g_{44} < 0$ heisst geometrisch: die Zeitachse muss im Innern des zum 0-Punkt gehörigen Nullkegels liegen.[85]

[85] Added by Hilbert in pencil: "Gaussches Koordinatensystem ist ein eigentliches."

Zweiter Abschnitt

§ 50. Der Sinn der Frage: Gilt die Euklidische Geometrie?

Wir hatten schon mehrfach Gelegenheit, zu bemerken, dass es eine *unerlaubte Annahme der alten Physik war, von vornherein die Gültigkeit der Euklidischen Geometrie anzunehmen*. Auch die alte Physik verwarf schon alle Fernwirkungstheorien. Die Euklidische Geometrie ist aber eine solche. Die alte Physik perhorreszierte also die Euklidische Geometrie und legte sie trotzdem ihren Naturgesetzen zugrunde. Wenn wir jetzt wieder die alte Frage der Geometrie aufwerfen: *Gilt die Euklidische Geometrie oder gilt sie nicht*? so müssen wir uns klar sein, dass man dieser Frage zwei ganz verschiedene Bedeutungen beilegen kann. Hat die Frage den Sinn: ist die Euklidische Geometrie in sich widerspruchsfrei? so ist sie eine rein mathematische und von den Mathematikern schon längst bejahend beantwortet worden. Wenn der Mathematiker sagt, dass die Euklidische Geometrie existiere, so meint er damit auch nur, dass sie nicht zu logischen Widersprüchen führt. In der Tat kann jede mathematische Existenzbehauptung — z. B. jede algebraische Gleichung hat eine Wurzel — so formuliert werden: die Annahme, | dass jede algebraische 105
Gleichung eine Wurzel hat, führt nicht auf logische Widersprüche. Mit einer solchen Beantwortung der Frage können wir uns in der Physik nicht zufrieden geben. Was nützt uns das schönste, in sich widerspruchslose physikalische System, wenn es in der Natur nicht realisiert ist?

Für uns hat also die Frage der Gültigkeit der Euklidischen Geometrie einen ganz anderen Sinn wie für den Mathematiker. Wir fragen: *ist die Euklidische Geometrie in der Welt des Physikers realisiert*? Um diese Frage zu beantworten, müssen wir uns nun einem neuen Abschnitt unserer Vorlesung zuwenden, *indem wir wirkliche Physik treiben. Diese Frage kann* eben *durch blosses Nachdenken nicht entschieden werden*, mit ihr treten wir nämlich in das Gebiet der Physik hinein. Damit würde mir auch die neue Auffassung von der Welt des Physikers vollkommen Recht geben; denn hier gehört die Geometrie zur Physik im Gegensatz zur alten Auffassung.

Bis jetzt hatten wir uns im wesentlichen nur mit den geometrischen Sätzen der Massbestimmung beschäftigt und dabei nur das Vorhandensein gewisser Ungleichungen zwischen den $g_{\mu\nu}$ vorausgesetzt. *Nun müssen wir die Differentialgleichungen aufstellen*, *denen die $g_{\mu\nu}$ genügen sollen*, und haben dann zu entscheiden, ob und unter welchen Voraussetzungen diese Gleichungen die Lösungen $g_{\mu\nu} = \delta_{\mu\nu}$ zulassen. Wir bemerken, dass wir im Folgenden statt der Zeit t als vierte Veränderliche $x_4 = it$ benutzen wollen. Dann haben wir ein positiv definites Linienelement, also die eigentliche und nicht die | Pseu- 106
dogeometrie vor uns. Diese homogenisierende, imaginäre Bezeichnungsweise ist bei allen Rechnungen der reellen Schreibweise vorzuziehen, während die

reelle bei prinzipiellen Fragen, wo es gerade auf den noch bestehenden kleinen Unterschied zwischen Raum und Zeit ankommt, besser anzuwenden ist.
Wir wollen das Resultat unserer Rechnung vorwegnehmen: *unsere physikalischen Grundgleichungen haben im allgemeinen keineswegs $g_{\mu\nu} = \delta_{\mu\nu}$ zu Lösungen*. Dies ist meiner Meinung nach ein positives Resultat der Theorie; denn wir können der Natur die Euklidische Geometrie durch andere Deutung der Experimente durchaus nicht aufzwingen. Vorausgesetzt nämlich, dass meine zu entwickelnden physikalischen Grundgleichungen wirklich die richtigen sind, so ist auch keine andere Physik möglich, d. h. die Wirklichkeit kann nicht anders aufgefasst werden. Andererseits werden wir sehen, dass unter gewissen sehr spezialisierenden Voraussetzungen — vielleicht ist das Fehlen von Materie im ganzen Raum dazu schon hinreichend — die einzigen Lösungen der Differentialgleichungen $g_{\mu\nu} = \delta_{\mu\nu}$ sind. Auch dies muss ich als eine Stütze meiner Theorie ansehen; denn das Gauss'sche Experiment der Messung der Winkelsumme im Dreieck hat gezeigt, dass *die Euklidische Geometrie in der Wirklichkeit* sicher *mit grosser Annäherung erfüllt* ist.[86]

§ 51. Die leitenden Gesichtspunkte bei der Herleitung der physikalischen Grundgleichungen

Wir wollen nun daran gehen, *die physikalischen Grundgleichungen* abzuleiten. Dazu muss ich Ihnen meine Anschauungen vom Wesen der Materie darlegen.
107 Ich bin der Ansicht, |[87] dass es in der physikalischen Welt *ausser Elektrizität keine weitere Materie gibt*⟨.⟩ Um nun den elektrischen Zustand in jedem Weltpunkt zu beschreiben, brauchen wir neue, unbekannte Funktionen, die den $g_{\mu\nu}$ der Pseudogeometrie gleichberechtigt an die Seite treten. Da die Weltgesetze invariant sein sollen, müssen sich auch diese Funktionen zu einer Invariante zusammensetzen lassen. Die einfachste Invariante, die wir haben, ist der Vierervektor. Daher führen wir *noch 4 unbekannte Funktionen* q_1, q_2, q_3, q_4 von x_1, x_2, x_3, x_4 ein, die den elektrischen Zustand der Welt in jedem Punkte beschreiben sollen. Dies ist sicher die einfachste Annahme, die man über das Wesen der Materie machen kann. Mit noch weniger neuen Funktionen lässt sich sicher nicht auskommen; dagegen wäre es von vornherein denkbar, dass man noch mehr Vektoren oder Tensoren einführen muss, um die Natur zu beschreiben. Ich habe aber die feste Ueberzeugung, dass die 4 Funktionen q_i zusammen mit den 10 Gravitationspotentialen $g_{\mu\nu}$ die Welt des Physikers

[86] For another mention of Gauß's experiment, see *Hilbert 1916a**, p. 5, (this Volume, p. 85 and its note 10).

[87] P. 107 is missing in the copy of *Hilbert 1916/17** deposited in Göttingen. The following text is taken from p. 107 of the copy extant in the Hückel papers in the Staatsbibliothek Berlin (Nachl. Hückel 2.11). This text is identical to the corresponding text of pp. 100–101 in the copy owned by the Archives of the California Institute of Technology, and to the corresponding text on the same pages of the copy extant in the Born papers in the Staatsbibliothek Berlin (Nachl. Born 1818), see also the Introduction to this Chapter, p. 77.

vollkommen darstellen. Ob diese Ueberzeugung berechtigt ist, kann nur der Erfolg lehren.

Nach welchen Gesichtspunkten sollen wir nun die Gleichungen aufstellen, denen diese 14 Funktionen zu genügen haben? Dass *diese Gleichungen allgemein invariant sein sollen*, muss unsere erste Sorge sein. Damit erhalten wir schon die Möglichkeit, dieselben aus einem Variationsprinzip abzuleiten, das eine allgemeine Invariante unter dem Integralzeichen stehen hat. Nun spielt schon in der alten Physik das *Hamiltonsche* | *Prinzip* eine hervorragende Rol- 108
le. In diese alte Physik soll aber die neue im Spezialfalle degenerieren. So bleibt uns gar keine Wahl, *wir werden vielmehr zwangsläufig auf ein Hamiltonsches Variationsprinzip geführt*, aus dem unsere Differentialgleichungen fliessen müssen. Unter dem Integralzeichen muss dabei, damit die Gleichungen invariant werden, eine allgemeine Invariante stehen. Dann folgt freilich aus rein mathematischen Gesetzen, dass — entsprechend den 4 unabhängigen Veränderlichen — von den 14 zur Bestimmung der $g_{\mu\nu}$ und q_i notwendigen Differentialgleichungen 4 eine Folge der 10 übrigen sind. Es sind also nur *10 Differentialgleichungen voneinander unabhängig*. Jede weitere Gleichung ist entweder eine Folge derselben oder sie steht mit ihnen im Widerspruch. Wir erhalten also im Lösungssystem noch *4 willkürliche Funktionen* von 4 Veränderlichen. Diese 4 willkürlichen Funktionen rühren daher, dass das den Gleichungen zugrunde gelegte Koordinatensystem willkürlich ist, da wir jederzeit x_i durch $x_i(x'_1, x'_2, x'_3, x'_4)$ ersetzen dürfen. *Die physikalischen Gesetze* selber, ja sogar jede physikalische Aussage,[88] muss freilich, um allgemein invariant zu sein, *von diesen 4 willkürlichen Funktionen unabhängig* sein, andernfalls ist sie *physikalisch sinnlos.*[89]

Bezeichnen wir jetzt mit

$$H(g_{\mu\nu}, g_{\mu\nu h}, \cdots, q_i, q_{ik} \cdots)$$

eine invariante Funktion der 14 Funktionen $g_{\mu\nu}$ und q_i und ihrer Ableitungen von zunächst beliebig hoher Ordnung, so lehrt die Invariantentheorie, dass das Integral $\iiiint H\sqrt{g}dx_1dx_2dx_3dx_4$ eine allgemeine Invariante ist. Wir haben also die 14 Grund|gleichungen der Physik abzuleiten aus dem *Hamiltonschen* 109
Prinzip

$$\delta \iiiint H\sqrt{g}dx_1dx_2dx_3dx_4 = \text{Minimum}, \tag{1}$$

— wobei das vierfache Integral über die ganze Welt zu erstrecken ist, — indem wir die Lagrangeschen Ableitungen nach den $g_{\mu\nu}$ und den q_i bilden. Es wird sich übrigens als praktisch erweisen, die Lagrangeschen Ableitungen nach den $g^{\mu\nu}$ statt nach den $g_{\mu\nu}$ zu nehmen. Mit der Integration dieser 14 allgemeinen Gleichungen werden wir uns später beschäftigen. Vorerst wollen wir nur einen Spezialfall derselben behandeln, der uns Auskunft gibt auf die oben gestellte Frage:

[88] The preceding five words are interlineated.

[89] The last half-sentence was added.

Gilt die Euklidische Geometrie in der Physik?

§ 52. Aufstellung der Grundgleichungen beim Fehlen von Materie

Wir suchen nämlich nach der *einfachsten Form, die die Invariante H annehmen kann*, d. h. wir fragen: welches sind die einfachsten Verhältnisse, die in der Wirklichkeit realisiert sein können? Nur dann dürfen wir hoffen, dass unsere Differentialgleichungen ein so einfaches Lösungssystem, wie es $g_{\mu\nu} = \delta_{\mu\nu}$ darstellt, zulassen. Die einfachste Annahme, die wir machen können, ist *das vollständige Fehlen der Materie oder Elektrizität*. Wir nehmen also $q_i \equiv 0$ an. Dann haben wir eine Invariante H zu suchen, die nur von den $g_{\mu\nu}$ und deren Ableitungen abhängt. Unter diesen Invarianten wollen wir nun diejenige ausfindig machen, die die Ableitungen der $g_{\mu\nu}$ wieder nur in möglichst niederer Ordnung enthält. Nun lehrt die Invariantentheorie, dass es Invarianten, die nur die $g_{\mu\nu}$ und deren erste Ableitungen zu Argumenten haben, nicht gibt.
110 Also | müssen wir eine Invariante wählen, welche die $g_{\mu\nu}$ samt ihren ersten und zweiten Ableitungen linear[90] enthält, und da in der Invariantentheorie gezeigt wird, dass es nur eine einzige solche Invariante gibt,[91] nämlich *die Riemannsche Krümmung K der vierdimensionalen Welt*, so werden wir zwangläufig darauf geführt, *die physikalischen Grundgleichungen im einfachsten Falle* aus

$$\iiiint K\sqrt{g}dx_1dx_2dx_3dx_4 = \text{Minimum}$$

abzuleiten. Hierin ist, wie wir wissen,[92]

$$\begin{aligned} K &= \sum_{\mu\nu} K_{\mu\nu} g^{\mu\nu}, \\ K_{\mu\nu} &= \sum_{\kappa} \left(\frac{\partial}{\partial x_\nu} \left\{ \begin{matrix} \mu\kappa \\ \kappa \end{matrix} \right\} - \frac{\partial}{\partial x_\kappa} \left\{ \begin{matrix} \mu\nu \\ \kappa \end{matrix} \right\} \right) \\ &\quad + \sum_{\kappa\lambda} \left(\left\{ \begin{matrix} \mu\kappa \\ \lambda \end{matrix} \right\} \left\{ \begin{matrix} \lambda\nu \\ \kappa \end{matrix} \right\} - \left\{ \begin{matrix} \mu\nu \\ \kappa \end{matrix} \right\} \left\{ \begin{matrix} \kappa\lambda \\ \lambda \end{matrix} \right\} \right). \end{aligned} \tag{2}$$

Durch Lagrangesche Differentiation erhalten wir die 10 Gravitationsgleichungen in allgemein invarianter Gestalt in der Form

$$[K\sqrt{g}]_{\mu\nu} = 0, \tag{3}$$

[90]"linear" was interlineated.

[91]See notes 37 and 40 above.

[92]In the typescript, in eq. (2) the Christoffel symbol $\left\{ \begin{matrix} \lambda\nu \\ \kappa \end{matrix} \right\}$ erroneously was written as $\left\{ \begin{matrix} \lambda\kappa \\ \kappa \end{matrix} \right\}$.

wobei eben $\left[K\sqrt{g}\right]_{\mu\nu}$ die Bedeutung der Lagrangeschen Ableitung von $K\sqrt{g}$ nach $g^{\mu\nu}$ hat.

Aus invariantentheoretischen Ueberlegungen folgt, dass diese Gleichungen die Form[93]

$$K_{\mu\nu} - \frac{1}{2}Kg_{\mu\nu} = 0$$

haben müssen. Multiplizieren wir diese Gleichungen mit $g^{\mu\nu}$ und summieren über μ und ν, so erhalten wir

$$\sum_{\mu\nu} g^{\mu\nu}K_{\mu\nu} - \frac{1}{2}K\sum_{\mu\nu} g^{\mu\nu}g_{\mu\nu} = 0$$

oder

$$K - \frac{1}{2}K \times 4 = 0, \qquad \text{d. h.} \quad K = 0.$$

Also haben wir *im Falle des Fehlens von Materie die einfachen physikalischen Grundgleichungen*

$$K_{\mu\nu} = 0 \tag{3'}$$

zu integrieren.

Um es nochmals übersichtlich darzustellen, so haben wir zur Aufstellung der Grundgleichungen für den einfachsten denkbaren Fall nur von *drei Prinzipien* Gebrauch gemacht: 111

1) das Hamiltonsche Prinzip,
2) das Prinzip der allgemeinen Invarianz,
3) das Prinzip, dass in der Natur von allen denkbaren Fällen gerade der einfachste realisiert ist.

§ 53. Zwei noch unbewiesene Sätze über die Gültigkeit der Pseudoeuklidischen Geometrie in der Physik

Von den zwei Gleichungen (3) sind, wie wir schon wissen, 4 eine Folge der 6 übrigen. Man verifiziert leicht, dass diese Gleichungen nun tatsächlich ein Lösungssystem $g_{\mu\nu} = \delta_{\mu\nu}$ haben, d. h. aber, dass die pseudoeuklidische Massbestimmung den Differentialgleichungen genügt, oder dass das kleine Relativitätsprinzip $dx^2+dy^2+dz^2-dt^2 =$ Invariante ein Spezialfall unserer allgemeinen Theorie ist. Dies dürfen wir mit gutem Grund als Bestätigung für die Richtigkeit unserer Theorie ansehen; denn schon das kleine Relativitätsprinzip

[93] For a similar comment on the derivation of the explicit equations from the variational formulation on general invariant-theoretic grounds, see *Hilbert 1915*, p. 405, (this Volume, p. 42 and its note 44).

gibt in vielen Fällen mit genügender Genauigkeit die in der Physik realisierten Verhältnisse wieder.

Wir wollen uns jetzt mit der Integration der 6 bezw. 10 Differentialgleichungen (3') beschäftigen. *Es ist möglich, dass folgender Satz richtig ist*:

Satz: Nimmt man alle Elektrizität aus der Welt hinweg (d. h. $q_i = 0$) und verlangt man absolute Regularität, — d. h. Möglichkeit der Entwicklung in eine Potenzreihe —, der Gravitationspotentiale $g_{\mu\nu}$ (eine Forderung, die nach
112 unserer | Auffassung auch im allgemeinen Fall immer erfüllt sein muss), so herrscht in der Welt die Euklidische Geometrie, d. h. die 10 Gleichungen (3) haben $g_{\mu\nu} = \delta_{\mu\nu}$ als *einzige Lösung*.

Von einem Beweise dieses Satzes bin ich sehr weit entfernt. Ein solcher würde jedenfalls eine sehr weitschichtige Theorie erfordern, ja ich bin gar nicht sicher, ob dieser Satz überhaupt richtig ist. Die Funktionen $g_{\mu\nu}$, die die *Massbestimmung* der Geometrie festlegen, haben wir auch gelegentlich *Gravitationspotentiale* genannt. Es ist in der Tat der grosse Gedanke von Einstein, dass das, was man in der Geometrie die Massbestimmung nennt, in der Physik mit Gravitation bezeichnet wird (siehe z. B. A. Einstein, Die Grundlage der allgemeinen Relativitätstheorie, Ann. d. Physik (49) 1916 § 2).[94]

Für sehr wahrscheinlich richtig halte ich folgenden

Satz: Nimmt man alle Elektrizität aus der Welt fort und verlangt von den Gravitationspotentialen ausser der selbstverständlichen Forderung der Regularität noch, dass $g_{\mu\nu}$ von t unabhängig ist, d. h. dass die *Gravitation stille steht*, und schliesslich noch reguläres Verhalten im Unendlichen, so sind $g_{\mu\nu} = \delta_{\mu\nu}$ *die einzigen Lösungen* der Gravitationsgleichungen (3).

Von diesem Satz kann ich schon jetzt so viel beweisen, dass *in der Nachbarschaft der Euklidischen Geometrie* sicher *keine Lösungen* dieser Gleichungen vorhanden sind.[95]

§ 54. Gültigkeit dieser Geometrie bei zentrischer Symmetrie

Der Beweis eines noch spezielleren Satzes dagegen lässt sich vollkommen durchführen und soll nun gegeben werden. Dazu machen wir über die $g_{\mu\nu}$
113 folgende drei Voraussetzungen: |

1) Es sei wieder $g_{\mu\nu}$ unabhängig von t.

2) Es sei $g_{\nu 4} = 0, \nu = 1, 2, 3$ d. h. Gauß'sches Koordinatensystem, das durch Transformation immer eingeführt werden kann.[96] (Orthogonalität der t-Achse auf dem x_1, x_2, x_3-Raum, dem sogenannten Steckenraum)

[94] *Einstein 1916a*, pp. 771–773.

[95] See *Hilbert 1917*, pp. 65–66, (this Volume, pp. 59–61).

[96] The words "d. h. . . . kann" are interlineated.

3) Es gebe einen ausgezeichneten Punkt in der Welt, in Bezug auf welchen *zentrische Symmetrie* vorhanden sein soll, d. h. die Drehung des Koordinatensystems um diesen Punkt ist eine Transformation der Welt in sich.

Nun gilt folgender
Satz: *Erfüllen die Gravitationspotentiale die Bedingungen 1-3, so ist die Euklidische Geometrie die einzige Lösung der physikalischen Grundgleichungen.* Wenn wir uns klar machen, dass der Euklidische Raum in Bezug auf jeden seiner Punkte zentrisch symmetrisch ist, so können wir diesen Satz auch so formulieren: Erfüllen die Funktionen $g_{\mu\nu}$ die Voraussetzungen 1 und 2 und gibt es ausserdem einen ausgezeichneten Weltpunkt, in Bezug auf welchen zentrische Symmetrie herrscht, so ist die Welt in jedem ihrer Punkte zentrisch symmetrisch. Der Beweis dieses Satzes erfordert einen gewissen Aufwand an Rechnung. Immerhin ist dieses Mass von Rechnung jetzt auf ein Minimum reduziert gegenüber den Arbeiten über den gleichen Gegenstand von Einstein (Erklärung der Perihelbewegung der Merkur, Sitzungsber. d. Preuss. Akad. d. Wiss. (47) S. 831 1915) und Schwarzschild (Ueber das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie, dieselbe Zeitschrift (7) S. 189 1916).[97]
Wir haben nun *die Gravitationsgleichungen* aus $\delta \iiiint K\sqrt{g}dx_1dx_2dx_3\, dx_4 = 0$ *abzuleiten*, wobei in K solche | Funktionen $g_{\mu\nu}$ einzuführen sind, die die Vor- 114
aussetzung 1–3 erfüllen. Eine gewaltige Vereinfachung der Rechnung lässt sich durch Einführung neuer Veränderlicher erzielen. Wir werden sehen, dass wir durch diesen Kunstgriff die Zahl der Differentialgleichungen von 10 auf 2 reduzieren und dass die Differentialgleichungen selbst von überraschender Einfachheit sein werden. Wir haben zuerst die Krümmung K für unseren Spezialfall zu berechnen und dies ist eine Rechnung, die sich auch im allgemeinen nie umgehen, sondern nur vereinfachen lässt. Dazu müssen wir das allgemeinste Linienelement $ds^2 = \sum_{\mu\nu} g_{\mu\nu}dx_\mu dx_\nu$ aufstellen, das die Bedingungen 1–3 erfüllt. Dieses Linienelement können wir, da Gravitation und Massbestimmung ein und dasselbe sind, auch auffassen als *das Gravitationsfeld eines ruhenden Massenpunktes*.

§ 55. Das Linienelement bei zentrischer Massenverteilung

Wir verlegen den ausgezeichneten Punkt oder Massenpunkt in den 0-Punkt unseres Koordinatensystems. Dann muss nach Bedingung 3 $\sum_{\mu\nu} g_{\mu\nu}\, dx_\mu dx_\nu$ *bei linearen Substitutionen*, die $x^2+y^2+z^2$ ungeändert lassen, *invariant* bleiben. Die *einzigen Invarianten* dieser Transformation sind nun

$$\begin{gathered} xdx + ydy + zdz, \qquad dx^2+dy^2+dz^2, \\ x^2+y^2+z^2, \qquad t, \qquad dt. \end{gathered}$$

[97] *Einstein 1915c*, *Schwarzschild 1916a*.

Aus diesen haben wir eine, in den Differentialen homogene, quadratische Invariante zusammenzusetzen. Bezeichnet $R = \sqrt{x^2 + y^2 + z^2}$, so setzt sich diese Invariante also additiv zusammen aus

$$\begin{gathered}(xdx + ydy + zdz)^2, \qquad dx^2 + dy^2 + dz^2, \\ dt^2, \qquad (xdx + ydy + zdz)dt,\end{gathered}$$

115 die alle noch mit willkürlichen Funktionen von R und t multipliziert sein können. Um jedoch der Bedingung 1 zu genügen, müssen diese willkürlichen Funktionen von t frei sein, und um die zweite Bedingung zu erfüllen, muss die zuletzt genannte Invariante wegfallen. Wir erhalten also⟨,⟩ $l = it$

$$ds^2 = F(R)(dx^2 + dy^2 + dz^2) + G(R)(xdx + ydy + zdz)^2 + H(R)dl^2$$

oder in Polarkoordinaten

$$\begin{aligned} ds^2 &= F(R)(dR^2 + R^2 d\vartheta^2 + R^2 \sin^2\vartheta d\varphi^2) + G(R)R^2 dR^2 + H(R)dl^2 = \\ &= \Phi(R)dR^2 + \Psi(R)(d\vartheta^2 + \sin^2\vartheta d\varphi^2) + H(R)dl^2. \end{aligned}$$

Nun setzen wir $\sqrt{\Psi(R)} = r$, wobei r wieder die Bedeutung der Entfernung eines Punktes vom 0 Punkt hat, und erhalten für

$$\begin{gathered}\Phi(R)\left(\frac{dR}{dr}\right)^2 = M(r); \qquad H(R) = W(r), \\ ds^2 = M(r)dr^2 + r^2(d\vartheta^2 + \sin^2\vartheta d\varphi^2) + W(r)dl^2. \end{gathered} \tag{4}$$

§ 56. Berechnung der Krümmung; das vereinfachte Variationsprinzip

Hierin sind M und W noch *unbekannte Funktionen einer Veränderlichen*, die aus dem Variationsprinzip jetzt bestimmt werden müssen. Dazu ist nun K *nach Formel (2) zu berechnen* und hierzu müssen wir wieder die zu (4) gehörigen g-Klammern kennen. Diese könnte man zwar aus (4) nach der Formel

$$\left\{ \begin{matrix} \mu\nu \\ \lambda \end{matrix} \right\} = \frac{1}{2} \sum g^{\lambda l}(g_{l\mu\nu} + g_{\nu l\mu} - g_{\mu\nu l}) \tag{5}$$

direkt ausrechnen. Bekanntlich führt aber der Weg, die g-Klammern aus den Differentialgleichungen der geodätischen Linie

$$\ddot{x}_s + \sum_{\mu\nu} \left\{ \begin{matrix} \mu\nu \\ s \end{matrix} \right\} \dot{x}_\mu \dot{x}_\nu = 0 \tag{6}$$

zu entnehmen, viel schneller zum Ziel. Wir bilden also die vier Lagrangeschen Ableitungen von

$$\delta \int_{p_1}^{p_2} \left\{ M(r) \left(\frac{dr}{dp} \right)^2 + r^2 \left(\frac{d\vartheta}{dp} \right)^2 + r^2 \sin^2 \vartheta \left(\frac{d\varphi}{dp} \right)^2 + W(r) \left(\frac{dl^2}{dp} \right)^2 \right\} dp = 0$$

nach r, ϑ, φ, l und erhalten durch Differentiation | 116

$$\begin{aligned}
&\text{nach } r : \ddot{r} + \frac{1}{2} \frac{M'}{M} \dot{r}^2 - \frac{r}{M} \dot{\vartheta}^2 - \frac{r}{M} \sin^2 \vartheta \dot{\varphi}^2 - \frac{1}{2} \frac{W'}{M} \dot{l}^2 = 0, \\
&\text{nach } \vartheta : \ddot{\vartheta} + \frac{2}{r} \dot{r} \dot{\vartheta} - \sin \vartheta \cos \vartheta \dot{\varphi}^2 = 0, \\
&\text{nach } \varphi : \ddot{\varphi} + \frac{2}{r} \dot{r} \dot{\varphi} + 2 \frac{\cos \vartheta}{\sin \vartheta} \dot{\vartheta} \dot{\varphi} = 0, \\
&\text{nach } l : \ \ddot{l} + \frac{W'}{W} \dot{r} \dot{l} = 0.
\end{aligned}$$

Hieraus lesen wir die g-Klammern ab:

$$\begin{aligned}
&\begin{Bmatrix} 11 \\ 1 \end{Bmatrix} = -\frac{1}{2} \frac{M'}{M}; \quad \begin{Bmatrix} 22 \\ 1 \end{Bmatrix} = -\frac{r}{M}; \quad \begin{Bmatrix} 33 \\ 1 \end{Bmatrix} = -\frac{r}{M} \sin^2 \vartheta; \\
&\begin{Bmatrix} 44 \\ 1 \end{Bmatrix} = -\frac{1}{2} \frac{W'}{M}; \\
&\begin{Bmatrix} 12 \\ 2 \end{Bmatrix} = \frac{1}{r}; \quad \begin{Bmatrix} 33 \\ 2 \end{Bmatrix} = -\sin \vartheta \cos \vartheta; \quad \begin{Bmatrix} 13 \\ 3 \end{Bmatrix} = \frac{1}{r}; \\
&\begin{Bmatrix} 23 \\ 3 \end{Bmatrix} = \frac{\cos \vartheta}{\sin \vartheta}; \quad \begin{Bmatrix} 14 \\ 4 \end{Bmatrix} = \frac{1}{2} \frac{W'}{W}.
\end{aligned}$$

Hierauf bilden wir nach Formel (2)

$$\begin{aligned}
K_{11} = & \frac{\partial}{\partial r} \left(\begin{Bmatrix} 11 \\ 1 \end{Bmatrix} + \begin{Bmatrix} 12 \\ 2 \end{Bmatrix} + \begin{Bmatrix} 13 \\ 3 \end{Bmatrix} + \begin{Bmatrix} 14 \\ 4 \end{Bmatrix} \right) - \frac{\partial}{\partial r} \begin{Bmatrix} 11 \\ 1 \end{Bmatrix} \\
& + \begin{Bmatrix} 11 \\ 1 \end{Bmatrix} \begin{Bmatrix} 11 \\ 1 \end{Bmatrix} + \begin{Bmatrix} 12 \\ 2 \end{Bmatrix} \begin{Bmatrix} 21 \\ 2 \end{Bmatrix} + \begin{Bmatrix} 13 \\ 3 \end{Bmatrix} \begin{Bmatrix} 31 \\ 1 \end{Bmatrix} \\
& + \begin{Bmatrix} 14 \\ 4 \end{Bmatrix} \begin{Bmatrix} 41 \\ 1 \end{Bmatrix} \\
& - \begin{Bmatrix} 11 \\ 1 \end{Bmatrix} \left(\begin{Bmatrix} 11 \\ 1 \end{Bmatrix} + \begin{Bmatrix} 12 \\ 2 \end{Bmatrix} + \begin{Bmatrix} 13 \\ 3 \end{Bmatrix} + \begin{Bmatrix} 14 \\ 4 \end{Bmatrix} \right) = \\
= & -\frac{1}{r^2} - \frac{1}{r^2} + \frac{1}{2} \frac{W''W - W'^2}{W^2} + \frac{1}{r^2} + \frac{1}{r^2} + \frac{1}{4} \frac{W'^2}{W} \\
& - \frac{1}{2} \frac{M'}{RM} - \frac{1}{2} \frac{M'}{RM} - \frac{1}{4} \frac{M'W'}{MW}.
\end{aligned}$$

Also wird

$$K_{11} g^{11} = \frac{1}{2} \frac{W''}{MW} + \frac{1}{4} \frac{W'^2}{MW^2} - \frac{M'}{rM^2} - \frac{1}{4} \frac{M'W'}{M^2W}.$$

Weiter erhalten wir

$$
\begin{aligned}
K_{22} = {} & \frac{\partial}{\partial \vartheta} \left\{ \begin{matrix} 33 \\ 3 \end{matrix} \right\} - \frac{\partial}{\partial r} \left\{ \begin{matrix} 22 \\ 1 \end{matrix} \right\} + \left\{ \begin{matrix} 21 \\ 2 \end{matrix} \right\} \left\{ \begin{matrix} 22 \\ 1 \end{matrix} \right\} + \left\{ \begin{matrix} 22 \\ 1 \end{matrix} \right\} \left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\} \\
& + \left\{ \begin{matrix} 23 \\ 3 \end{matrix} \right\} \left\{ \begin{matrix} 32 \\ 3 \end{matrix} \right\} - \\
& - \left\{ \begin{matrix} 22 \\ 1 \end{matrix} \right\} \left(\left\{ \begin{matrix} 11 \\ 1 \end{matrix} \right\} + \left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\} + \left\{ \begin{matrix} 13 \\ 3 \end{matrix} \right\} + \left\{ \begin{matrix} 14 \\ 4 \end{matrix} \right\} \right) = \\
= {} & -\frac{1}{\sin^2 \vartheta} + \frac{1}{M} - \frac{rM}{M^2} - \frac{1}{M} + \frac{\cos^2 \vartheta}{\sin^2 \vartheta} + \frac{1}{2} \frac{rM'}{M^2} + \frac{1}{M} + \frac{1}{2} \frac{rW'}{MW}
\end{aligned}
$$

und hieraus

$$K_{22} g^{22} = -\frac{1}{r^2} - \frac{1}{2} \frac{M'}{rM^2} + \frac{1}{r^2 M} + \frac{1}{2} \frac{W'}{rMW}.$$

Ferner wird

$$
\begin{aligned}
K_{33} = {} & -\frac{\partial}{\partial r} \left\{ \begin{matrix} 33 \\ 1 \end{matrix} \right\} - \frac{\partial}{\partial \vartheta} \left\{ \begin{matrix} 33 \\ 2 \end{matrix} \right\} + \left\{ \begin{matrix} 31 \\ 3 \end{matrix} \right\} \left\{ \begin{matrix} 33 \\ 3 \end{matrix} \right\} + \left\{ \begin{matrix} 32 \\ 3 \end{matrix} \right\} \left\{ \begin{matrix} 33 \\ 2 \end{matrix} \right\} \\
& + \left\{ \begin{matrix} 33 \\ 1 \end{matrix} \right\} \left\{ \begin{matrix} 13 \\ 3 \end{matrix} \right\} + \left\{ \begin{matrix} 33 \\ 2 \end{matrix} \right\} \left\{ \begin{matrix} 23 \\ 3 \end{matrix} \right\} \\
& - \left\{ \begin{matrix} 33 \\ 1 \end{matrix} \right\} \left(\left\{ \begin{matrix} 11 \\ 1 \end{matrix} \right\} + \left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\} + \left\{ \begin{matrix} 13 \\ 3 \end{matrix} \right\} + \left\{ \begin{matrix} 14 \\ 4 \end{matrix} \right\} \right) \\
& - \left\{ \begin{matrix} 33 \\ 2 \end{matrix} \right\} \left\{ \begin{matrix} 23 \\ 3 \end{matrix} \right\} = \\
= {} & \frac{sin^2 \vartheta}{M} - \sin^2 \vartheta \frac{rM'}{M^2} + \cos^2 \vartheta - \sin^2 \vartheta - \frac{\sin^2 \vartheta}{M} - \cos^2 \vartheta + \\
& + \frac{1}{2} \sin^2 \vartheta \frac{M'}{M^2} + \frac{sin^2 \vartheta}{M} + \frac{1}{2} \sin^2 \vartheta \frac{rW'}{MW}
\end{aligned}
$$

117 und hieraus

$$K_{33} g^{33} = -\frac{1}{2} \frac{M'}{rM^2} - \frac{1}{r^2} + \frac{1}{r^2 M} + \frac{1}{2} \frac{W'}{rMW}.$$

Schliesslich bilden wir noch

$$
\begin{aligned}
K_{44} = {} & -\frac{\partial}{\partial r} \left\{ \begin{matrix} 44 \\ 1 \end{matrix} \right\} + \left\{ \begin{matrix} 41 \\ 4 \end{matrix} \right\} \left\{ \begin{matrix} 44 \\ 1 \end{matrix} \right\} + \left\{ \begin{matrix} 44 \\ 1 \end{matrix} \right\} \left\{ \begin{matrix} 41 \\ 4 \end{matrix} \right\} \\
& - \left\{ \begin{matrix} 44 \\ 1 \end{matrix} \right\} \left(\left\{ \begin{matrix} 11 \\ 1 \end{matrix} \right\} + \left\{ \begin{matrix} 12 \\ 2 \end{matrix} \right\} + \left\{ \begin{matrix} 13 \\ 3 \end{matrix} \right\} + \left\{ \begin{matrix} 14 \\ 4 \end{matrix} \right\} \right) \\
= {} & \frac{1}{2} \frac{W''}{M} - \frac{1}{2} \frac{M'W'}{M^2} - \frac{1}{4} \frac{W'^2}{MW^2} + \frac{1}{4} \frac{M'W'}{M^2} + \frac{1}{2} \frac{W'}{rM} + \frac{1}{2} \frac{W'}{rM}
\end{aligned}
$$

und erhalten für

$$K_{44} g^{44} = \frac{1}{2} \frac{W''}{MW} - \frac{1}{4} \frac{M'W'}{M^2 W} - \frac{1}{4} \frac{W'^2}{MW^2} + \frac{W'}{rMW}.$$

Da alle $g^{\mu\nu}$ für $\mu \neq \nu$ verschwinden, so brauchen wir die $K_{\mu\nu}$ für $\mu \neq \nu$ gar nicht zu berechnen, da sie in der Formel (2) für die Krümmung nicht auftreten. Die Krümmung ergibt sich nun zu

$$K = \frac{W''}{MW} - \frac{1}{2}\frac{W'^2}{MW^2} - 2\frac{M'}{rM^2} - \frac{M'W'}{M^2W} - \frac{2}{r^2} + \frac{2}{r^2M} + 2\frac{W'}{rMW}.$$

Weiter ist $\sqrt{g} = \sqrt{MW} r^2 \sin\vartheta$, also wird

$$K\sqrt{g} = \left\{ \left(\frac{r^2W'}{\sqrt{MW}} \right)' - 2\frac{rM'\sqrt{W}}{M^{\frac{3}{2}}} - 2\sqrt{MW} + 2\sqrt{\frac{W}{m}} \right\} \sin\vartheta.$$

Jetzt setzen wir noch $M(r) = \frac{r}{r-m(r)}$, $W(r) = w^2(r)\frac{r-m(r)}{r}$, wobei nun m und w willkürliche Funktionen von r sind, und erhalten

$$K\sqrt{g} = \left\{ \left(\frac{r^2W'}{\sqrt{MW}} \right)' - 2w\frac{rm' - m}{r} - 2w + 2w\frac{r-m}{r} \right\} \sin\vartheta,$$

also als *Schlussresultat* für $K\sqrt{g}$ den einfachen Ausdruck

$$K\sqrt{g} = \left\{ \left(\frac{r^2W'}{\sqrt{MW}} \right)' - 2wm' \right\} \sin\vartheta, \qquad (7)$$

wobei noch zu berücksichtigen ist, dass bei den nun zu bildenden Lagrangeschen Ableitungen nach m und w das Glied $\left(\frac{r^2W'}{\sqrt{MW}}\right)'$, weil es eine totale Ableitung ist, keinen Beitrag liefert, weil die Variationen an den Grenzen verschwinden sollen.[98] Wir erhalten nun aus dem Variationsprinzip

$$\delta \iiiint K\sqrt{g} dr d\vartheta d\varphi dl = 0$$

in diesem Fall $\delta \int mw'dr = 0$, also $m' = 0$ und $w' = 0$ oder $m = \alpha$, $w = 0$. 118
Setzen wir noch $cdt = dt'$ und lassen den Strich wieder weg, so erhalten wir als Schlussresultat

$$ds^2 = \frac{r}{r-\alpha}dr^2 + r^2d\vartheta^2 + r^2\sin^2\vartheta d\varphi^2 - \frac{r-\alpha}{r}dt^2. \qquad (8)$$

§ 57. $g_{\mu\nu} = \delta_{\mu\nu}$ als einzige Lösung

Dies ist das gesuchte Gravitationsfeld eines ruhenden Massenzentrums. Nach unserer Auffassung vom Wesen der Materie können wir als physikalisch realisierbare Lösungen $g_{\mu\nu}$ der Differentialgleichungen $K_{\mu\nu} = 0$ nur diejenigen ansehen, welche regulär und singularitätenfrei sind.

[98] The preceding half-sentence was interlineated.

„Regulär" nennen wir ein Gravitationsfeld oder eine Massbestimmung, — diese Definition war noch nachzutragen — wenn es möglich ist, ein solches Koordinatensystem einzuführen, dass die Funktionen $g_{\mu\nu}$ an jeder Stelle der Welt regulär sind und eine von Null verschiedene Determinante haben.[99] Wir bezeichnen ferner eine einzelne Funktion als regulär, wenn sie mit allen ihren Ableitungen endlich und stetig ist. Dies ist übrigens immer die Definition der Regularität in der Physik, während in der Mathematik von einer regulären Funktion verlangt wird, dass sie analytisch ist.

Für $\alpha > 0$ bez. $\alpha \leqq 0$ *hat die Massbestimmung* (8) and den Stellen $r = 0$ u. $r = \alpha$ bez. $r = 0$ *Singularitäten*. Wenn wir bedenken, dass diese Singularitäten von der Anwesenheit einer Masse herrühren, so erscheint es auch plausibel, dass dieselben durch Koordinatentransformation nicht zu beseitigen sind.[100]
119 Einen strengen Beweis dafür werden wir aber erst weiter unten geben, | indem wir den Verlauf der geodätischen Linien in der Umgebung dieser Punkte untersuchen. Wir müssen also, um singularitätenfreie Lösungen zu erhalten, $\alpha = 0$ annehmen. Dann geht (8) über in das pseudoeuklidische Linienelement

$$ds^2 = dx^2 + dy^2 + dz^2 - dt^2.$$

Wir haben damit den auf S. 113 formulierten Satz bewiesen: *Bei Abwesenheit von Materie* ($q_i \equiv 0$) existiert unter den auf Seite 113 genannten Voraussetzungen 1–3 die pseudoeuklidische Geometrie des kleinen Relativitätsprinzips in der Physik tatsächlich, und für $t =$ const. ist in der Welt die *Euklidische Geometrie wirklich realisiert*. Es ist also gezeigt, wie einerseits im allgemeinen die Massbestimmung von der Massenverteilung in der Welt abhängig ist, und wie andererseits bei Abwesenheit von Materie diese allgemeine Massbestimmung in diejenige der Euklidischen Geometrie degeneriert, wobei wir es noch dahingestellt sein lassen, ob die Bedingungen 1–3 auch notwendig sind, damit wirklich die Euklidische Geometrie Geltung hat.

§ 58. Charakteristische Singularitäten der Massbestimmung

Aus unserem Resultat (8) können wir aber noch viel mehr schöpfen, wenn wir nur $\alpha \neq 0$ annehmen. Dann handeln wir zwar entgegen unserer eigenen Vorschrift, dass wir nur singularitätenfreie Gravitationsfelder als in der Natur realisierbar ansehen wollen. Daher müssen wir *die Annahme* $\alpha \neq 0$ vorher *rechtfertigen*.
120 *Die Integration der 14 allgemeinen physikalischen Grundgleichungen* für $q_i = 0$ *ist ausserordentlich schwierig*, wir sind sogar, wie wir sahen, noch weit davon

[99] For Hilbert's definition of a singularity, see also *Hilbert 1917*, p. 70, (this Volume, p. 66 and its note 57).

[100] The singularity at $r = \alpha$ is, in fact, a coordinate singularity, see *Eisenstaedt 1982* for a historical discussion of the Schwarzschild solution; see also *Hilbert 1917*, p. 70, (this Volume, p. 66 and its note 57).

entfernt, auch nur den einfachen Spezialfall, dass dieselben in $K_{\mu\nu} = 0$ übergehen, allgemein integrieren zu können. Die mathematischen Schwierigkeiten hindern uns z. B. schon an der Konstruktion eines einzigen neutralen Massenpunktes. Könnten wir eine solche neutrale Masse kontruieren, und würden wir den Verlauf der $\langle g_{\mu\nu} \rangle$ in der Umgebung dieser Stelle kennen, so würden die $g_{\mu\nu}$, wenn wir die neutrale Masse immer mehr gegen einen Massenpunkt hin degenerieren lassen, in diesem Punkte eine *Singularität* aufweisen. Eine solche müssten wir als *erlaubt* ansehen in dem Sinne, dass die $g_{\mu\nu}$ ausserhalb der nächsten Umgebung der Singularität den in der Natur wirklich realisierten Verlauf richtig wiedergeben. Eine solche Singularität müssen wir nun in (8) vor uns haben. Im übrigen können wir schon jetzt sagen, dass die Konstruktion eines neutralen Massenpunktes, auch wenn sie später möglich sein wird, sich als so kompliziert erweisen wird, dass man für die Zwecke, in denen man nicht die nächste Umgebung des Massenpunktes betrachtet, mit ausreichender Genauigkeit mit den mit einer Singularität behafteten, angenähert richtigen Gravitationspotentialen wird rechnen können.

§ 59. Aufstellung provisorischer Axiome

Wir behaupten nun Folgendes: Wenn wir die mathematische Entwicklung, die zur Konstruktion eines neutralen Massenteilchens führt, wirklich werden durchführen können, so werden wir dabei vermutlich auf Gesetze stossen, die wir einst|weilen noch *axiomatisch formulieren* müssen, die aber später sich als 121
*Folgen unserer allgemeinen Theorie erg*eben werden, als Folgen freilich, die bestimmt nur durch eine weitschichtige Theorie und komplizierte Rechnung zu begründen sein werden. Diese Axiome, die also nur *provisorische Geltung* haben sollen, fassen wir folgendermassen:

Axiom I.: Die Bewegung eines Massenpunktes im Gravitationsfeld wird durch eine geodätische Linie dargestellt, welche *Zeitlinie ist.*

Axiom II.: Die Lichtbewegung im Gravitationsfeld wird durch eine geodätische *Nulllinie* dargestellt.

Axiom III.: Eine singuläre Stelle der Massbestimmung ist äquivalent einem Gravitationszentrum.

Zu dieser Formulierung der Axiome I u. II[101] ist Folgendes zu bemerken: Dieselben gehen, in der Grenze $(\delta_{\mu\nu})$ in diejenigen der alten Physik über, sind also in der Tat eine *vernünftige Verallgemeinerung* der letzteren. Für $g_{\mu\nu} = \delta_{\mu\nu}$ wird die geodätische Zeitlinie nämlich zu einer Geraden, die mit der t-Achse einen Winkel von weniger als 45° bildet, d. h. die Bahnkurve eines Massenpunktes im gravitationsfreien, dreidimensionalen Streckenraum ist eine Gerade, und die Bewegung in derselben ist gleichförmig und geschieht mit *Unterlichtgeschwindigkeit*. Das zweite Axiom besagt für $g_{\mu\nu} = \delta_{\mu\nu}$, dass

[101]"I u. II" was interlineated.

das Licht sich in einem solchen Raum ebenfalls geradlinig ausbreitet. Da die Bewegung in einer Nulllinie vor sich geht, ist noch $dx^2 + dy^2 + dz^2 - dt^2 = 0$,
122 d. h. $\frac{dr}{dt}$ = Lichtgeschwindigkeit = 1. | Im allgemeinen ist aber im zweiten Axiom *keine Rede mehr von gerader Linie und konstanter Geschwindigkeit*, wie dies in der alten Physik noch[102] der Fall ist.[103]

Wir werden sehen, dass aus diesen Axiomen tatsächlich das Newtonsche Attraktionsgesetz bezw. die aus diesem resultierenden Kepplerschen Gesetze der Planetenbewegung *in erster Näherung* folgen. *Prinzipiell* aber hat dieses neue *Einsteinsche Gesetz gar keine Aehnlichkeit mit dem Newtonschen*. Es ist ungleich komplizierter als das letztere. Wenn wir es trotzdem dem Newtonschen vorziehen, so ist dies darin begründet, dass dieses Gesetz einem tiefliegenden philosophischen Prinzip — dem der allgemeinen Invarianz —, Genüge leistet, und dass es zwei so heterogene Dinge, wie das Newtonsche Gesetz einerseits und die tatsächliche Gültigkeit der Euklidischen Geometrie in der Physik unter gewissen einfachen Voraussetzungen andererseits als Spezialfälle enthält, so dass wir also nicht, wie dies bis jetzt der Fall war, zuerst die Gültigkeit der Euklidischen Geometrie voraussetzen, und dann ein Attraktionsgesetz anflicken müssen.

§ 60. Die Bahnkurve ist von der Masse des sich bewegenden Punktes unabhängig

Aus den eben formulierten drei[104] Axiomen wollen wir nun *die wichtigsten Folgerungen* ziehen. Ein im Nullpunkt ruhendes Gravitationszentrum bewirkt also, dass die Massbestimmung die Form hat $g_{\mu\nu} = 0$ für $\mu \neq \nu$ und $g_{11} = \frac{r}{r-\alpha}$, $g_{22} = r^2$, $g_{33} = r^2 \sin^2 \vartheta$, $g_{44} = -\frac{r}{r-\alpha}$. Die Geometrie ist eindeutig festgelegt, wenn α bestimmt ist, weil α die einzige noch willkürliche Konstante
123 des Feldes ist. | Diese Grösse α muss also *die Schwere der Masse*, die im 0-Punkt angebracht wird, darstellen. (Wir wissen noch nicht, ob $\alpha > 0$, oder ob $\alpha < 0$ angenommen werden muss. Für $\alpha < 0$ ist $r = 0$ die einzige Singularität, da $r \geqq 0$ sein muss. Für $\alpha > 0$ ist auch die Kugel $r = \alpha$ eine Singularität. Dann darf man, wenn man von $r > \alpha$ ausgeht, nicht $r < \alpha$ werden lassen, d. h. man darf nicht vom Aeusseren in das Innere der Kugel hineingehen, sondern man muss das Kugeläussere und das Kugelinnere als zwei getrennte Gravitationsfelder behandeln.)[105]

Aus dem Axiom I ersehen wir schon, dass der *Satz* gilt:

A) *Die Bahnkurve und die Bewegung des Massenpunktes* in derselben *ist von der Masse des sich bewegenden Punktes unabhängig.*

[102] "noch" was interlineated.

[103] In the left margin, there is a reader's mark (⊤).

[104] In the original, "drei" was put into brackets and corrected to "2" in pencil.

[105] The brackets in the preceding paragraph were added in pencil.

Dies stimmt mit der Newtonschen Theorie überein; denn dort ist $m\ddot{x} = -\frac{mMx}{r^3}$, wo m die Masse des sich bewegenden Punktes und M die Masse des Zentrums darstellt. m hebt sich in der Gleichung weg. Diese erste Folgerung A) oder dieses erste Gesetz ist eine strenge Folgerung der von uns aufgestellten Axiomatik; die Uebereinstimmung mit der Newtonschen Mechanik ist also eine exakte und nicht nur eine angenäherte. Um weitere Schlüsse zu ziehen, müssen wir die Differentialgleichungen der Bewegung integrieren. Wir gehen wieder aus vom Variationsproblem

$$\delta \int \left\{ \frac{r}{r-\alpha}\dot{r}^2 + r^2\dot{\vartheta}^2 + r^2 \sin^2\vartheta\dot{\varphi}^2 - \frac{r-\alpha}{r}\dot{t}^2 \right\} dp = 0.$$

Ein Integral der zugehörigen Differentialgleichungen[106] ist zufolge unserer allgemeinen Theorie | (siehe S. 15) 124

$$\frac{r}{r-\alpha}\dot{r}^2 + r^2\dot{\vartheta}^2 + r^2 \sin^2\vartheta\dot{\varphi}^2 - \frac{r-\alpha}{r}\dot{t} = A. \tag{9}$$

A ist die Integrationskonstante und zwar muss für einen Massenpunkt $A < 0$ sein (Bedingung für die Zeitlinie) und für das Licht $A = 0$ (Bedingung für die Nulllinie). Durch Bildung der Lagrangeschen Ableitung[107] nach ϑ erhalten wir

$$(r^2\dot{\vartheta})^{\cdot} - r^2 \sin\vartheta\cos\vartheta\dot{\varphi}^2 = 0. \tag{10}$$

Die durch die Lagrangesche Differentiation nach φ entstehende Gleichung lässt sich sofort einmal integrieren und liefert

$$r^2 \sin^2\vartheta\dot{\varphi} = B. \tag{11}$$

Dasselbe gilt für die Lagrangesche Ableitung nach t:

$$\frac{r-\alpha}{r}\dot{t} = C. \tag{12}$$

B und C sind ebenfalls Integrationskonstanten. Damit haben wir schon drei Integrationen durchgeführt und haben dementsprechend statt 4 Differentialgleichungen 2. Ordnung nur mehr eine Differentialgleichung 2. und 3 Differentialgleichungen 1. Ordnung.

§ 61. Die Bahnkurven liegen in Ebenen durch das Gravitationszentrum

Nun lässt sich eine zweite Uebereinstimmung mit der Newtonschen Theorie angeben, die das System der Differentialgleichungen erheblich zu vereinfachen gestattet. Wir wollen nämlich den *Satz* beweisen:

---

[106] "der zugehörigen Differentialgleichungen" was corrected from "desselben".

[107] "Bildung der Lagrangeschen Ableitung" was corrected by Hilbert with ink from "Differentiation".

B) *Die Bahnkurven liegen in Ebenen, die durch das Gravitationszentrum gehen.*

Zum Beweise eliminieren wir den Parameter p aus den Differentialgleichungen (10) und (11), um so eine einzige Gleichung für ϑ als Funktion von φ zu
125 erhalten. Wir machen | die identische Umformung

$$
\begin{aligned}
(r^2\dot{\vartheta})' &= \left(r^2\frac{d\vartheta}{d\varphi}\dot{\varphi}\dot{\varphi}\right)^{\cdot} = \frac{d}{d\varphi}\left(r^2\frac{d\vartheta}{d\varphi}\dot{\varphi}\right)\dot{\varphi} \\
&= \dot{\varphi}^2\left(2r\frac{dr}{d\varphi}\frac{d\vartheta}{d\varphi}+r^2\frac{d^2\vartheta}{d\varphi^2}\right)+r^2\frac{d\vartheta}{d\varphi}\ddot{\varphi}.
\end{aligned}
$$

Weiter erhalten wir aus (11) durch Differentiation nach p wegen $\frac{d}{dp}=\dot{\varphi}\frac{d}{d\varphi}$

$$\left\{2r\frac{dr}{d\varphi}\sin^2\vartheta+2r^2\sin\vartheta\cos\vartheta\frac{d\vartheta}{d\varphi}\right\}\dot{\varphi}^2+r^2\sin^2\vartheta\ddot{\varphi}=0$$

oder nach Division mit $\sin^2\vartheta$ und Multiplikation mit $\frac{d\vartheta}{d\varphi}$

$$2\left\{r\frac{dr}{d\varphi}+r^2\operatorname{cotg}\vartheta\frac{d\vartheta}{d\varphi}\right\}\dot{\varphi}^2\frac{d\vartheta}{d\varphi}+r^2\frac{d\vartheta}{d\varphi}\ddot{\varphi}=0.$$

Setzen wir diesen Wert von $r^2\frac{d\vartheta}{d\varphi}\ddot{\varphi}$ in den Ausdruck für $(r^2\dot{\vartheta})^{\cdot}$ ein, so erhalten wir aus (10)

$$
\begin{aligned}
\dot{\varphi}^2\left(2r\frac{dr}{d\varphi}\frac{d\vartheta}{d\varphi}+r^2\frac{d^2\vartheta}{d\varphi^2}\right)-2\left(r\frac{dr}{d\varphi}+r^2\operatorname{cotg}\vartheta\frac{d\vartheta}{d\varphi}\right)\dot{\varphi}^2\frac{d\vartheta}{d\varphi}& \\
-r^2\sin\vartheta\cos\vartheta\dot{\varphi}^2=0&
\end{aligned}
$$

oder $\dot{\varphi}^2r^2\frac{d^2\vartheta}{d\varphi^2}-2\dot{\varphi}^2r^2\operatorname{cotg}\vartheta\left(\frac{d\vartheta}{d\varphi}\right)^2-\dot{\varphi}^2r^2\sin\vartheta\cos\vartheta=0$, also

$$\frac{d^2\vartheta}{d\varphi^2}-2\operatorname{cotg}\vartheta\left(\frac{d\vartheta}{d\varphi}\right)^2=\sin\vartheta\cos\vartheta.$$

Das allgemeine Integral dieser Gleichung ist

$$\sin\vartheta\cos(\varphi+a)+b\cos\vartheta=0.$$

Man verifiziert leicht, dass dasselbe eine, und, weil es zwei Konstante enthält, die allgemeinste Lösung der Differentialgleichung ist. In der Tat wird $\cos(\varphi+a)=-b\operatorname{cotg}\vartheta$. Hieraus durch Differentiation nach φ: $-\sin(\varphi+a)=\frac{b}{\sin^2\vartheta}\frac{d\vartheta}{d\varphi}$ Nochmalige Differentiation nach φ ergibt

$$-\cos(\alpha+a)=\frac{b}{\sin^2\vartheta}\left\{\frac{d^2\vartheta}{d\varphi^2}-2\cot g\vartheta\left(\frac{d\vartheta}{d\varphi}\right)^2\right\}.$$

Die linke Seite dieser Gleichung ist gleich $b \operatorname{cotg} \vartheta$ und die rechte wird vermöge der Differentialgleichung $b \operatorname{cotg} \vartheta$. q. e. d.

Die Gleichung $\sin\vartheta\cos(\varphi+a)+b\cos\vartheta$ *stellt eine durch den Nullpunkt gehende Ebene dar*. Es ist nämlich für $x = r\sin\vartheta\cos\varphi$, $y = r\sin\vartheta\sin\varphi$, $z = r\cos\vartheta$
126 die Gleichung $r\sin\vartheta\cos\varphi\cos a - r\sin\vartheta\sin\varphi\sin a + r\cos\vartheta\cdot b = 0$ oder

$$\cos ax - \sin ay + bz = 0, \qquad \text{w. z. b. w.}$$

Wir haben also einen zweiten *Satz* gefunden, der eine strenge Folgerung unserer Axiome ist und in Uebereinstimmung mit der Newtonschen Theorie steht: Die Bahnkurven liegen in Ebenen durch das Gravitationszentrum.

§ 62. Ableitung der Differentialgleichung der Bahnkurven

Nun setzen wir $\vartheta = \frac{\pi}{2}$ d. h. $z = 0$ und erhalten anstelle der Gleichungen (9)–(12)

$$\frac{r}{r-\alpha}\dot{r}^2 + r^2\dot{\varphi}^2 - \frac{r-\alpha}{r}\dot{t}^2 = A, \tag{9'}$$

$$r^2\dot{\varphi} = B, \tag{11'}$$

$$\frac{r-\alpha}{r}\dot{t} = 1 (= C). \tag{12'}$$

Um $C = 1$ zu machen, darf man nicht einfach $Ct = t^*$ setzen, weil sonst im Integral (9') das letzte Glied $\frac{r-\alpha}{r}C^2\dot{t}^{*2}$ lauten würde. Dagegen kann man der Konstanten C den Wert 1 erteilen durch *Normierung des Parameters p*, der (siehe S. 13) nur bis auf eine multiplikative Konstante bestimmt ist. Für $r = \infty$ wird (12') zu $\dot{t} = C$. Setzt man also für $r = \infty$ $\dot{t} = 1$ d. h. $p' = Cp$, so nimmt die Gleichung (12') die gewünschte Form $\frac{r-\alpha}{r}\dot{t} = 1$ an. Aus obigen drei Gleichungen wollen wir nun *die Gleichung der Bahnkurve* in der Form $r = r(\varphi)$ ableiten. Dazu setzen wir den aus (12') berechneten Wert von $\dot{t}$ in (9') ein, ersetzen noch $\dot{r}$ durch $\dot{\varphi}\frac{dr}{d\varphi}$ und erhalten

$$\frac{r}{r-\alpha}\dot{\varphi}^2\left(\frac{dr}{d\varphi}\right)^2 + r^2\dot{\varphi}^2 - \frac{r}{r-\alpha} = A.$$

Berücksichtigen wir nun (11'), so finden wir

$$\frac{r}{r-\alpha}\frac{B^2}{r^4}\left(\frac{dr}{d\varphi}\right)^2 + r^2\frac{B^2}{r^4} - \frac{r}{r-\alpha} = A.$$

Hierin setzen wir $r = \frac{1}{\rho}$, $\frac{dr}{d\varphi} = -\frac{1}{\rho^2}\frac{d\rho}{d\varphi}$ und erhalten

$$\frac{1}{1-\alpha\rho}B^2\left(\frac{d\rho}{d\varphi}\right)^2 + \rho^2 B^2 - \frac{1}{1-\alpha\rho} = A.$$

Hieraus ergibt sich als *Gleichung der Bahnkurve* des Massenpunktes die Dif- 127
ferentialgleichung erster Ordnung

$$\left(\frac{d\rho}{d\varphi}\right)^2 = \frac{1+A}{B^2}\rho - \rho^2 + \alpha\rho^3. \tag{13}$$

Nach Axiom I ist $A < 0$, weil die Bahnkurve eines Massenpunktes eine Zeitlinie sein muss. Hätten wir dagegen die Bewegung des Lichts zu untersuchen, so wären zwar nach Axiom II die Differentialgleichungen dieselben, jedoch wäre dann $A = 0$ zu setzen, weil sich das Licht auf einer Nulllinie fortpflanzt.

§ 63. Das Newtonsche Attraktionsgesetz als erste Näherung

Die Differentialgleichung der Bahnkurve für einen Massenpunkt, der sich nach dem *Newtonschen Attraktionsgesetz* bewegt, ist

$$\left(\frac{d\rho}{d\varphi}\right)^2 = a + b\rho - \rho^2 \tag{14}$$

und hat den Kegelschnitt $\rho = \frac{b}{2} \pm \sqrt{a + \frac{b^2}{4}}\cos\varphi$ als Lösung. Man sieht also, wenn man (13) und (14) vergleicht, dass das *Einsteinsche Gravitationsgesetz im allgemeinen keineswegs auf die Bahnkurven der Newtonschen Theorie führen wird*. Wohl aber ist dies *angenähert unter gewissen Bedingungen*, die wir nun aufsuch⟨en⟩ wollen, der Fall. Die Differentialgleichung (13) geht nämlich für endliche Werte von ρ *angenähert* in (14) über, wenn α sehr klein wird. Damit der Koeffizient $\frac{\alpha A}{B^2}$ von ρ dann den endlichen Wert b der Gleichung (14) annimmt, muss entweder A sehr gross oder B sehr klein werden. Sehr grosse Werte von A sind aber unzulässig, weil sonst der Ausdruck $\frac{1+A}{B^2}$, d. h. die Konstante a der Newtonschen Theorie unendlich gross werden würde. Also muss B sehr klein werden. Dann bleibt wieder $\frac{1+A}{B^2}$, worin ja A negativ
128 sein muss, nur | endlich, wenn A gegen -1 konvergiert. Um also *die Kegelschnitte als Bahnkurven* zu erhalten, d. h. um die Erscheinungen, die durch die Newtonsche Gravitationstheorie erklärt werden, zu realisieren, muss

$$1)\quad \lim\alpha = 0 \qquad 2)\quad \lim B = 0$$

werden. Dann ist $\lim A = -1$ — d. h. der Parameter p ist angenähert gleich der Eigenzeit — von selbst erfüllt als Folge von (9'), (11') und (12'). *Diese Verhältnisse finden wir nun in unserem Sonnensystem* — mit Ausnahme der bekannten kleinen Störung beim Merkur, die wir unten erörtern wollen — wirklich vor. Aus den Gleichungen (11') und (12') folgt nämlich

$$\frac{\dot\varphi}{\dot t} = \frac{d\varphi}{dt} = \frac{r-\alpha}{r^3}B. \tag{15}$$

Linker Hand steht die Winkelgeschwindigkeit, welche also wegen $\lim B = 0$ sehr klein sein muss. Die Himmelskörper bewegen sich mit einer Geschwindigkeit, deren eine Komponente $r\frac{d\varphi}{dt}$ ist und für die Werte zwischen 20 und 200 km $\sec^{-1}$ beobachtet wurden. Da wir, wie wir unten bei der Behandlung der Lichtbewegung zeigen werden, die Lichtgeschwindigkeit im Unendlichen (300000 km $\sec^{-1}$) als Einheit gewählt haben, so ist die Bedingung $\lim B = 0$ tatsächlich erfüllt.

Früher, d. h. bevor man auch nur das kleine Relativitätsprinzip kannte, war es ganz unverständlich, warum man nie mit beliebig grosser, selbst mit Ueberlichtgeschwindigkeit sich bewegende Körper beobachtete. In der kleinen Relativitätstheorie kommt nun, wie auch in unserer allgemeinen Theorie, der *Lichtgeschwindigkeit die Rolle einer Grenzgeschwindigkeit* zu. Dieser | Gedan- 129
ke, an den wir uns durchaus schon gewöhnt haben, schien zu jener Zeit ganz abscheulich. Damals war es aber auch noch unverständlich, warum man für die Lichtgeschwindigkeit immer jene selbe Zahl fand, ob man sie nun von einem ruhenden oder von einem bewegten Bezugssystem aus mass.

Die erste[108] Bedingung $\lim \alpha = 0$ verlangt, dass *die gravitierenden Massen sehr klein* sind. Auf diesen Punkt werden wir unten, wenn wir erkannt haben, dass α nicht anderes ist, als die schwere Masse eines Körpers, noch zurückkommen. Schon jetzt können wir aus der Gleichung (9'), in der $r-\alpha$ als Koeffizient in zwei Gliedern auftritt, entnehmen, dass α die Dimension einer Länge haben muss. Für den Astronomen sind also sehr kleine Werte von α solche, die im Verhältnis zu den Entfernungen, in denen sich die Trabanten um ihr Zentrum bewegen, zu vernachlässigen sind. Dass wir die Newtonsche Mechanik bei der *Elektronenbewegung* nicht mehr realisiert finden, braucht uns nun auch nicht wunderzunehmen. Denn es[109] sind Elektronengeschwindigkeiten bis zu $\frac{9}{10}$ derjenigen des Lichts beobachtet worden.

§ 64. Der Kreis ist eine Bahnkurve: als Lösung der Differentialgleichung der Bahnkurven

Um die Bedeutung der Konstante α zu erfassen, bemerken wir, dass der *Kreis* eine Bahnkurve ist, die gleichzeitig *eine strenge Lösung der Newtonschen Theorie*, d. h. der Gleichung (14) *und der Einsteinschen Theorie*, d. h. der Gleichung (13) ist. Um zu sehen, dass $\rho = \text{const.}$ die Differential|gleichung 130
(14) befriedigt, darf man nicht einfach deren rechte Seite $a+b\rho-\rho^2 = 0$ setzen und hieraus ρ entnehmen, obgleich dann auch die linke Seite $(\frac{d\rho}{d\varphi})^2 = 0$ wird. Auf diesem Wege erhält man zwei *singuläre* Integrale.[110] Man muss vielmehr,

[108]"erste" was corrected from "zweite".

[109]"Denn es" was corrected by Hilbert in pencil from "Einerseits mag dort das α des positiv geladenen schweren Kerns gegenüber der Entfernung der negativen Elektronen von demselben eine nicht zu vernachlässigende Grösse sein, andererseits".

[110]Deleted: "nämlich die beiden Enveloppen der Ellipsenschar".

wie in der Theorie der Differentialgleichungen gezeigt wird, zur Integration der Gleichung $\frac{d\rho}{d\varphi}$ als Potenzreihe $\mathfrak{P}(\rho)$ ansetzen. Ist nun $a+b\rho-\rho^2 = -(\rho-\rho_1)(\rho-\rho_2)$, so wird vermöge der Differentialgleichung $\mathfrak{P}(\rho) = \sqrt{-(\rho-\rho_1)(\rho-\rho_2)}$. Um die Wurzel ziehen zu können, muss ρ_1 *eine Doppelwurzel* der Gleichung $a + b\rho - \rho^2 = 0$ sein. Die Bedingung hierfür ist $a + \frac{b^2}{4} = 0$ *Dann ist die Bahnkurve der Kreis* mit dem Radius $r = \frac{2}{b}$.
Dass der Kreis auch eine mögliche Bahnkurve der Einsteinschen Theorie ist, entnimmt man am einfachsten den Lagrangeschen Differentialgleichungen selbst, aber nicht aus den daraus abgeleiteten Gleichungen (9'), (11'), (12')[111] Wir werden diesen Weg nachher einschlagen; zuerst wollen wir, weil diese Rechnung von mathematischem Interesse ist, den Kreis als mögliche Lösung der Differentialgleichung (13) darstellen. Um die *Doppelwurzel der Gleichung* $a + b\rho - \rho^2 + \alpha\rho^3 = 0$ zu finden, in der jetzt

$$a = \frac{1+A}{B^2}, \qquad b = -\frac{\alpha A}{B^2} \tag{16}$$

gesetzt ist, haben wir nach einer Regel der Algebra die Konstanten a und b aus $a + b\rho - \rho^2 + \alpha\rho^3 = 0$ und der hieraus durch Differentiation nach ρ entstehenden Gleichung $b - 2\rho + 3\alpha\rho^2 = 0$ zu eliminieren. Die zweite Relation
131 liefert | $b = 2\rho - 3\alpha\rho^2$. Hierauf ergibt sich aus der ersten $a = 2\alpha\rho^3 - \rho^2$. Nun ist wegen (16) $\frac{1}{B^2} = a + \frac{b}{\alpha} = 2\alpha\rho^3 - \rho^2 + \frac{2\rho}{\alpha} - 3\rho^2$ und für $r = \frac{1}{\rho}$ $\frac{1}{B^2} = \frac{2\alpha}{r^3} - \frac{4}{r^2} + \frac{2}{\alpha r} = \frac{2}{\alpha r^3}(r-\alpha)^2$ oder

$$B^2 = \frac{\alpha}{2}\frac{r^3}{(r-\alpha)^2}.$$

Aus der Gleichung (15) folgt dann für die Winkelgeschwindigkeit

$$\left(\frac{\dot\varphi}{\dot t}\right)^2 = \left(\frac{d\varphi}{dt}\right)^2 = \frac{\alpha}{2}\frac{1}{r^3} \qquad \textit{III. Kepplersches Gesetz.} \tag{17}$$

Diese Gleichung ist insofern ganz merkwürdig, als sich aus ihr die Integrationskonstanten A und B weggehoben haben. Den Grund hierfür werden wir kennen lernen, wenn wir die Kreisbewegung direkt aus den Lagrangeschen Gleichungen ableiten. *Die Winkelgeschwindigkeit der Kreisbewegung* ist also allein *vom Kreisradius und von der Konstanten* α, d. h. der im Zentrum angebrachten Masse *abhängig*.

§ 65. Die Konstante α

An die Gleichung (17)[112] müssen wir eine wichtige Bemerkung knüpfen. Linker Hand steht eine positive Grösse, rechter Hand ist r ebenfalls positiv, *daher*

[111]"selbst, aber nicht aus den daraus abgeleiteten Gleichungen" was interlineated, "(9'), (11') und (12')" was deleted, the deletion afterwards rescinded.

[112]Interlineated by Hilbert in pencil: "oder schon der drüberstehenden".

muss, damit die in der Natur beobachtete Kreisbahn auch aus unserer Theorie folgt,

$$\alpha > 0$$

sein. Diese Ungleichung folgt, wie wir sehen, nicht schon aus unserer allgemeinen Theorie, sie muss vielmehr aus der Erfahrung gewonnen werden. Wir hätten dieselbe übrigens schon aus der zweiten Gleichung (16) $b = -\frac{\alpha A}{B^2}$ entnehmen können. Damit nämlich[113] die Lösung der Gleichung (14) d. h. der Kegelschnitt $\rho = \frac{b}{2} \pm \sqrt{a + \frac{b^2}{4}} \cos\varphi$ eine Ellipse sein kann, muss $b > 0$ sein.
Also muss in (16), weil $a < 0$ ist, $\alpha \mid$ positiv sein. Nur *unter dieser Bedingung* 132
wirkt die Kraft des Gravitationszentrums *anziehend* auf den sich bewegenden Massenpunkt, d. h. nur dann sind als Bahnkurven, wie die Erfahrung lehrt, alle drei Arten von Kegelschnitten möglich. Weil $\alpha > 0$ ist, haben wir also eine *reelle Gravitationskugel*, auf deren Oberfläche *die Massbestimmung aufhört, regulär zu sein.*

Aus der Gleichung (17) können wir endlich auch die Bedeutung der Konstanten α entnehmen, wenn wir sie mit ihrer Form, in der sie sich aus der Newtonschen Theorie ergibt, vergleichen. Dort folgt aus der Beziehung: Zentrifugalkraft gleich Anziehungskraft, d. h. $\frac{mv^2}{r} = \kappa \frac{mM}{r^2}$ im Falle des Kreises $mr\left(\frac{d\varphi}{dt}\right)^2 = \kappa\frac{mM}{r^2}$. Also wird[114]

$$\left(\frac{d\varphi}{dt}\right)^2 = \kappa M \frac{1}{r^3} \quad \left(= \frac{\alpha}{2}\frac{1}{r^3}\right), \qquad \alpha = 2\kappa M.$$

Die Konstante α ist, wie man sieht, bis auf den Zahlenfaktor 2κ *gleich der gravitierenden Masse des Zentralkörpers*, um den sich der Massenpunkt bewegt. In dieser Gleichung hat κ die Bedeutung der Gravitationskonstante in unserem Masssystem, in welchem die Lichtgeschwindigkeit $= 1$ gesetzt ist. Bezeichnen wir im cm, gr, sec.-Masssystem die Lichtgeschwindigkeit mit c und die Gravitationskonstante mit κ^*, so wird $\kappa = \frac{\kappa^*}{c^2}$.

§ 66. Wie das aus dem Hamiltonschen Prinzip abzulesende Integral der Lagrangeschen Gleichungen aus den letzteren folgt

Wir wollen nun *auf anderem Wege nochmals beweisen*, dass *die Kreisbahn eine strenge Lösung der Einsteinschen Theorie* bei allgemeinem α, A und B ist, und zwar, indem wir direkt von den Lagrangeschen Differentialgleichungen des Variationsproblems

$$\int \left\{ \frac{r}{r-\alpha}\dot{r}^2 + r^2\dot{\varphi}^2 - \frac{r-\alpha}{r}\dot{t}^2 \right\} dp = \text{Min.} \tag{18}$$

[113] Added by Hilbert in pencil: "ein Kreis möglich sein soll als Bahnkurve"

[114] In the following equation, "$\kappa M \frac{1}{r^3}$" was corrected by Hilbert in pencil from "$\kappa M = \frac{1}{r^3}$". Also, both the formula in brackets and the relation for α were added by Hilbert in pencil.

133 ausgehen. Dasselbe unterscheidet sich von demjenigen, das wir früher unserer Untersuchung zugrunde legten, insofern, als wir jetzt, wo wir wissen, dass die Bahnkurven eben sind, schon im Variationsproblem selbst $\vartheta = \frac{\pi}{2}$ setzen. Diese Betrachtung wird uns in doppelter Hinsicht von Nutzen sein. Einmal werden wir erkennen, *dass gerade der Kreis* unter allen Bahnkurven der Einsteinschen Theorie *eine ausgezeichnete Stellung einnimmt*. Dann aber wird sich die wichtige Gleichung (17), aus der wir übrigens noch mehrere interessante Folgerungen zu ziehen haben, nun fast ohne Rechnung ergeben. Die Mühe, die wir oben zur Erzielung des nämlichen Resultates hatten, war trotzdem keineswegs vergebens; denn bei der Diskussion der Bahnkurven, auf denen sich das Licht bewegt, werden wir doch wieder auf jene Methode zurückgreifen müssen.

Wir erhalten aus (18) durch Lagrangesche Differentiation nach r die Gleichung

$$[D] \equiv 2\frac{r}{r-\alpha}\ddot{r} + \left(\frac{r}{r-\alpha}\right)\dot{r}^2 - 2r\dot{\varphi}^2 + \frac{\alpha}{r^2}\dot{t}^2 = 0, \tag{19}$$

deren linke Seite wir zur Abkürzung mit $[D]$ bezeichnen. Durch Variation nach φ und t erhalten wir die schon oben abgeleiteten Gleichungen (11') und (12'), in welchen eine Integration sofort auszuführen ist:

$$[B] \equiv r^2\dot{\varphi} = B, \qquad [C] \equiv \frac{r-\alpha}{r}\dot{t} = C.$$

In den vorangehenden Untersuchungen haben wir, statt diese Lagrangeschen Differentialgleichungen zur Bestimmung der unbekannten Funktionen zu benutzen, anstelle der Gleichung (19) das Integral (9'), das eine Folge aller Lagrangeschen Gleichungen ist, benutzt; unter der stillschweigenden Voraus-
134 setzung, dass | alle Lösungen dieses Gleichungssystems auch Integrale der Lagrangeschen Gleichungen sind. Nun wollen wir prüfen, inwiefern diese Annahme berechtigt war.

Vorher bemerken wir noch, dass man ganz allgemein zeigen kann, dass *die zu einem Variationsproblem gehörigen Differentialgleichungen ein Integral haben*, welches sich aus dem Problem selbst *lediglich durch Differentiationsprozesse* entnehmen lässt, wenn nur dass Variationsproblem die Form

$$\int H(\dot{r}\dot{\varphi}\dot{t})dp = \text{Min}$$

hat, wo H ausser $\dot{r}$, $\dot{\varphi}$ und $\dot{t}$ nur noch r, φ und t selbst, nicht aber die unabhängige Veränderliche p als Argumente enthält. Wir haben diesen Satz als Folge unserer allgemeinen Theorie (S. 11 ff) auf S. 14 erhalten, wenn H die spezielle Gestalt $\sum_{\mu\nu} g_{\mu\nu}\dot{x}_\mu\dot{x}_\nu$ hat. Unter dieser selben Voraussetzung wollen wir ihn nochmals auf kürzerem Wege ableiten. Zu diesem Zwecke *führen wir in das Variationsproblem künstlich eine neue unbekannte Funktion ein*, indem wir den Parameter $p = p(q)$ setzen. Dann erhalten wir — wenn wir die Differentiation nach q mit einem Strich bezeichnen —

$$\int H\left(\frac{r'}{p'}, \frac{\varphi'}{p'}, \frac{t'}{p'}\right) p' dq = \text{Min.}$$

Hieraus folgt, weil H eine homogene, quadratische Funktion von $\dot{r}$, $\dot{\varphi}$ und $\dot{t}$ sein soll,

$$\int H(r', \varphi', t') \frac{dq}{p'} = \text{Min.}$$

Nun können wir eine *neue Differentialgleichung* aufstellen, indem wir die Lagrangesche Ableitung nach der unbekannten Funktion $p(q)$ bilden. Wir erhalten, wenn wir eine Integration gleich ausführen, und dann wieder von der Eigenschaft der | Funktion H, in r', φ' und t' homogen quadratisch zu sein, 135
Gebrauch machen,

$$\frac{H(r', \varphi', t')}{p'^2} = \text{const} = H(\dot{r}, \dot{\varphi}, \dot{t}), \qquad \text{w.z.b.w.}$$

Dieses Resultat, welches in unserem Spezialfall die Gleichung

$$[A] \equiv \frac{r}{r-\alpha}\dot{r}^2 + r^2\dot{\varphi}^2 - \frac{r-\alpha}{r}\dot{t}^2 = A \tag{9'}$$

ergibt, *wollen wir verifizieren*. Durch Differentiation dieser Gleichung nach p fällt die Konstante A weg. Zu der so entstehenden Gleichung addieren wir die beiden anderen, die wir erhalten, wenn wir die Gleichungen (11') und (12') nach p differentiieren und danach mit bezw. $-2\dot{\varphi}$ und $-2\dot{t}$ multiplizieren. Es ergibt sich

$$\begin{array}{rl} \frac{d[A]}{dp} \equiv & 2\frac{r}{r-\alpha}\dot{r}\ddot{r} + \left(\frac{r}{r-\alpha}\right)_r \dot{r}^3 + 2r\dot{r}\dot{\varphi}^2 + 2r^2\dot{\varphi}\ddot{\varphi} - \frac{\alpha}{r^2}\dot{r}\dot{t}^2 - 2\frac{r-\alpha}{r}\dot{t}\ddot{t} = 0, \\ -2\dot{\varphi}\frac{d[B]}{dp} \equiv & -4r\dot{r}\dot{\varphi}^2 - 2r^2\dot{\varphi}\ddot{\varphi} = 0, \\ +2\dot{t}\frac{d[C]}{dp} \equiv & +2\frac{\alpha}{r^2}\dot{r}\dot{t}^2 + 2\frac{r-\alpha}{r}\dot{t}\ddot{t} = 0, \\ \hline & 2\frac{r}{r-\alpha}\dot{r}\ddot{r} + \left(\frac{r}{r-\alpha}\right)_r \dot{r}\dot{r}^2 - 2r\dot{r}\dot{\varphi}^2 + \frac{\alpha}{r^2}\dot{r}\dot{t}^2 = 0. \end{array}$$

Dies ist aber nichts anderes als die Gleichung $\dot{r}[D] = 0$. Wir haben also die *Beziehung*

$$\frac{d}{dp}[A] - 2\dot{\varphi}\frac{d}{dp}[B] + 2\dot{t}\frac{d}{dp}[C] = \dot{r}[D].$$

§ 67. Die ausgezeichnete Stellung des Kreises unter den Bahnkurven

Jetzt können wir die zu Anfang dieser Betrachtung gestellte Frage beantworten: *Sind die Lösungen des Gleichungssystems (9'), (11') und (12') dieselben*

wie diejenigen des Systems der Lagrangeschen Gleichungen *(19), (11') und (12')*; d. h. folgt aus $[A] = A$, $[B] = B$, $[C] = C$ die Gleichung (19) $[D] = 0$. Man sieht, dass dies tatsächlich der Fall ist, wenn nur $\dot{r} \neq 0$ ist. *Diese Bedingung* ist aber gerade *für den Kreis nicht erfüllt*. Wenn wir also untersuchen
136 wollen, ob der | Kreis eine mögliche Bahnkurve ist, dürfen wir *nicht* von den drei Gleichungen $[A] = A$, $[B] = B$, $[C] = C$ ausgehen. Die Integrale dieser drei Gleichungen erfüllen vielmehr nur dann die Lagrangeschen Gleichungen, wenn die Bahnkurve kein Kreis ist. Man sieht also, dass *der Kreis* unter allen Bahnkurven des Variationsproblems (18) eine *ganz besondere Stellung* einnimmt, zum Unterschied von der Newtonschen Mechanik. In der Tat wird durch die Kepplerschen Gesetze der Kreis in keiner Weise vor den anderen Kegelschnitten ausgezeichnet. Vielleicht liegt in diesem Umstand ein Fingerzeig, wo die tiefsten Geheimnisse der Elektrodynamik verborgen sind.

Man sieht leicht, dass man *zu ganz verkehrten Resultaten* kommt, wenn man die Kreisbahn auf Grund der Gleichungen (9'), (11') und (12') diskutiert. Es folgt nämlich aus (11') und (12')

$$\varphi = \frac{B}{r^2}p, \qquad t = \frac{r}{r-\alpha}p.$$

Die Gleichung (9') liefert dann für $\dot{r}^2 = 0$ vermöge (11') und (12')

$$r^2\left(\frac{B}{r^2}\right)^2 - \frac{r-\alpha}{r}\left(\frac{r}{r-\alpha}\right)^2 = A.$$

In dieser Gleichung tritt die willkürliche Konstante A auf, d. h. *bei beliebig vorgegebenem B* kann zu *jedem Radius r* die Konstante A so bestimmt werden, dass die Gleichung erfüllt ist. Die Winkelgeschwindigkeit in der Bahnkurve $\frac{d\varphi}{dt} = B\frac{r-\alpha}{r^3}$ kann dann wieder ganz beliebig vorgegeben werden, weil B noch willkürlich ist. Wir erhalten also das *falsche* Resultat: es sind alle Kreise als Bahnkurven und alle Winkelgeschwindigkeiten als Bewegungen in der Bahnkurve zulässig.

137 Dass wir oben auch für den Kreis aus Gleichung (13) das richtige Ergebnis, nämlich das dritte Kepplersche Gesetz fanden, trotzdem wir die Gleichung (13) aus den in diesem Fall unbrauchbaren Gleichungen (9'), (11') und (12') ableiteten, hat seinen Grund darin, dass wir den Kreis als Grenzfall einer Kurve[115] ansahen, bei der zwei Wurzeln der rechten Seite von (13) $a + b\rho - \rho^2 + \alpha\rho^3 = 0$ einander gleich werden.

§ 68. Die Kreisbewegung als Lösung der Lagrangeschen Differentialgleichungen

Ziehen wir jetzt zur *Diskussion der Kreisbewegung* die Gleichung (19) heran, so erhalten wir die schon bekannten Resultate nochmals in einfachster Weise.

[115]"Kurve" was corrected from "Ellipse".

Aus (19) wird nämlich für $\dot{r} = 0$

$$-2r\dot{\varphi}^2 + \frac{\alpha}{r^3}\dot{t}^2 = 0 \qquad \text{oder} \qquad \left(\frac{d\varphi}{dt}\right)^2 = \frac{\frac{\alpha}{2}}{r^3}.$$

Dies ist aber die Gleichung (17), die die Winkelgeschwindigkeit aus dem Kreisradius zu berechnen gestattet. Ferner wird t bestimmt aus (12'): $t = \frac{r}{r-\alpha}p$. Aus (11') folgt noch $\varphi = \frac{B}{r^2}p$, wobei B aber nicht mehr willkürlich ist, sondern aus

$$\left(\frac{d\varphi}{dt}\right)^2 = \left(B\frac{r-\alpha}{r^3}\right)^2 = \frac{\frac{\alpha}{2}}{r^3}$$

entnommen werden muss. Die Gleichung (11') lehrt eben nichts Neues, sie ist vielmehr eine Folge von (19) und (12'). Aus diesem Grunde kommt die Integrationskonstante B in (17) nicht vor. Dass die Konstante A dort nicht auftritt, ist nun auch selbstverständlich; durften wir doch die Gleichung (9') gar nicht benutzen.

§ 69. Zwei merkwürdige Folgerungen: Untere Grenze für den Kreisradius und obere Grenze für die Winkelgeschwindigkeit

Nun können wir einige *ganz merkwürdige Folgerungen* ziehen,[116] wenn wir nur bemerken, dass nicht allle geodätischen Linien als Bahnkurven für den Massenpunkt zulässig sind, son|dern nur Zeitlinien unter ihnen. Es muss also 138

$$\frac{r}{r-\alpha}\dot{r}^2 + r^2\dot{\varphi}^2 - \frac{r-\alpha}{r}\dot{t}^2 < 0$$

sein. Ist die Bahnkurve ein Kreis, so wird

$$r^2\dot{\varphi} - \frac{r-\alpha}{r}\dot{t}^2 < 0$$

oder wegen (17) $r^2 \frac{\alpha}{2}\frac{1}{r^3} < \frac{r-\alpha}{r}$ d. h.

$$r > \frac{3}{2}\alpha.$$

Dies ist ein durchaus *unerwartetes Resultat*. Es besagt, dass *ein Massenpunkt, der sich in einer Kreisbahn um ein Zentrum* mit der Masse α *bewegt, nur bis zur Entfernung* $\frac{3}{2}\alpha$ *an dasselbe herankommen kann.* Von vornherein sollte man glauben, dass alle Kreise mit $r > \alpha$ erlaubt seien, weil erst für $r = \alpha$ die Massbestimmung singulär wird. Dies ist in der Tat die Grenze, die Schwarzschild (Ueber das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie, Sitzungeber. d. Preuss. Akad. d. Wiss. (7) S 189, 1916)

[116] In the left margin, there is a reader's mark (⊤).

angibt,[117] und die ich als falsch bezeichnen muss; denn für $\alpha < r < \frac{3}{2}\alpha$ wird die Bewegung nicht mehr durch eine Zeitlinie, sondern durch eine Strecke dargestellt. Dies ist sicher keine mögliche Bewegung für einen Massenpunkt.

Nimmt man in obigen Ungleichungen überall das Gleichheitszeichen, so hat man die Bewegung des Lichts vor sich. Man sieht, dass *der „Lichtplanet“* sich um ein Massenzentrum in einem Kreis vom Radius $r = \frac{3}{2}\alpha$ bewegen kann, und dass dies gleichzeitig die *einzig mögliche Kreisbahn des Lichtes* ist. Elektronen können also bis nahe an diese Grenze herankommen. Die absolute Geschwindigkeit in der Kreisbahn ist $v = r\frac{d\varphi}{dt}$ und für dieselbe finden wir
139 die Ungleichung $\mid v^2 \leq 1 - \frac{\alpha}{2} \leq \frac{1}{3}$. *Die Geschwindigkeit des Planeten in der Kreisbahn ist also* $< \frac{1}{\sqrt{3}}$, und für den Lichtplaneten selbst wird sie im Kreis $r = \frac{3}{2}\alpha$ zu $v = \frac{1}{\sqrt{3}}$. Durch das Wirken der Gravitation wird, wie dies Ergebnis zeigt, die Lichtgeschwindigkeit auf dem Kreis $r = \frac{3}{2}\alpha$ auf den ca. $0,7$ Teil ihres Betrages im Unendlichen herabgedrückt.

§ 70. Transformation auf Ruhe bei der Kreisbewegung

Wir wollen an dieser Stelle eine kleine Zwischenbetrachtung einschieben, die uns Gelegenheit gibt, die bei der Formulierung des Kausalitätsprinzips auf S. 97 ff angestellten allgemeinen Ueberlegungen auf einen Spezialfall anzuwenden, und zwar stellen wir uns *die Aufgabe, einen sich in einer Kreisbahn* um die Sonne *bewegenden Planeten auf Ruhe zu transformieren*. Unter den unendlich vielen möglichen Transformationen, die dies leisten, gibt es eine ganz besonders einfache, die wir nun daraufhin prüfen wollen, ob sie eine erlaubte, d. h. eine "eigentliche Raum-Zeittransformation" ist oder nicht. Die Weltlinie des Planeten im ruhenden (ungestrichenen) System ist, wenn wir die konstante Winkelgeschwindigkeit mit $\frac{1}{w}$ bezeichnen und als Parameter die Zeit t nehmen:

$$r = r_0, \quad \varphi = \varphi_0 + \frac{1}{w}t.$$

Die erwähnte *Transformation* des Planeten *auf Ruhe* lautet dann:

$$r = r', \quad \varphi = \varphi' + t', \quad t = wt'.$$

In der Tat wird im mitbewegten (gestrichenen) Koordinatensystem in welchem sich also die Sonne und der Fixsternhimmel drehen, die Weltlinie des Planeten zu

$$r' = r_0, \quad \varphi' = \varphi - \frac{1}{w}t = \varphi_0, \quad t' = \frac{1}{w}t.$$

140 Wir bilden nun die Gravitationspotentiale im gestrichenen System und erhal-

[117] *Schwarzschild 1916a.*

ten

$$\begin{aligned}ds^2 &= \left\{\frac{r}{r-\alpha}r'^2 + r^2\dot\varphi^2 - \frac{r-\alpha}{r}\dot t^2\right\}dp^2 = \\ &= \left\{\frac{r'}{r'-\alpha}\dot r'^2 + \dot r'^2(\dot\varphi' + \dot t')^2 - \frac{r'-\alpha}{r'}w^2\dot t'^2\right\}dp^2 = \\ &= \left\{\frac{r'}{r'-\alpha}\dot r'^2 + r'^2\dot\varphi'^2 + 2r'^2\dot\varphi'\dot t' - \left(\frac{r'\alpha}{r'}w^2 - r'^2\right)\dot t'^2\right\}dp^2.\end{aligned}$$

Sind nun die 5 auf S 102f. aufgestellten Ungleichungen,[118] d. h. die Bedingungen für ein eigentliches Raum-Zeit-Koordinatensystem erfüllt? Es wird

$$I.\quad g'_{11} = \frac{r'}{r'-\alpha} > 0, \qquad III.\quad g'_{11}g'_{33} - {g'_{13}}^2 = \frac{r'}{r'-\alpha}r'^2 > 0,$$

$$IV.\quad g'_{44} = r'^2 - \frac{r'-\alpha}{r'}w^2 < 0, \qquad V.\quad g' = -r'^2w^2 < 0.$$

Alle Ungleichungen, mit Ausnahme der IVten sind, wie man sieht, unter allen Umständen erfüllt. *Damit auch die IVte Ungleichunge* $r'^2 - \frac{r'-\alpha'}{r'}w^2 < 0$ *erfüllt ist, muss* $\left(\frac{d\varphi}{dt}\right)^2 = \frac{1}{w^2} < \frac{r'-\alpha}{r'}$ werden. Dies ist aber nichts anderes als die schon oben genannte Ungleichung $A < 0$ oder die Bedingung dafür, dass der Kreis eine Zeitlinie ist. Unsere Transformation ist also tatsächlich eine eigentliche Raum-Zeit-Transformation, und jeder sich in einer erlaubten Kreisbahn sich bewegende Planet kann durch dieselbe in erlaubter Weise auf Ruhe transformiert werden.

§ 71. Satz: Jeder Massenpunkt kann auf Ruhe transformiert werden

Wir dürfen sogar die viel allgemeinere Behauptung aufstellen:
Jeder sich auf einer Zeitlinie bewegende Massenpunkt kann durch eine eigentliche Raum-Zeit-Transformation auf Ruhe transformiert werden.
Zum Beweise denken wir uns die Weltlinie des sich | bewegenden Massenpunk- 141
tes in der Form $x_i = w_i(x_4)$ $(i = 1, 2, 3)$ gegeben und setzen $x'_1 = x_1 - w_1$, $x'_2 = x_2 - w_2$, $x'_3 = x_3 - w_3$, $x'_4 = x_4$. Dann wird die Weltlinie im gestrichenen Bezugssystem dargestellt durch $x'_1 = 0$, $x'_2 = 0$, $x'_3 = 0$. *Der Punkt befindet sich* also tatsächlich *in Ruhe. Um zu erkennen, dass diese Transformation* eine — aber natürlich nicht die einzige — *erlaubte ist*, bilden wir die $g'_{\mu\nu}$. Wir erhalten, weil $x_i = x'_i + w_i(x'_4)$, $x_4 = x'_4$ $(i = 1, 2, 3)$, ist $g'_{\mu\nu} = \frac{\partial x_\mu}{\partial x'_\mu}\frac{\partial x_\nu}{\partial x'_\nu}g_{\mu\nu} = g_{\mu\nu}$ $(\mu, \nu = 1, 2, 3)$. Also sind die drei ersten Ungleichungen tatsächlich erfüllt. Die Vte Ungleichung $g' < 0$ ist sicher erfüllt und zwar als Folge der drei ersten. Unsere vierdimensionalte Pseudogeometrie ist ja vom Typus $3 + 1$ und

[118]See p. 239 above.

dieser kann, wenn die drei ersten Ungleichungen erfüllt sind, nur dann vorhanden sein, wenn $g' < 0$ ist. Dass endlich $g'_{44} < 0$ wird, sieht man leicht ein: Da die Bahnkurve des Massenpunktes eine Zeitlinie ist, so ist für dieselbe $G(\dot{x}_i(p)) < 0$. Im transformierten Bezugssystem sind $\dot{x}'_1 = 0$, $\dot{x}'_2 = 0$, $\dot{x}'_3 = 0$, also folgt, weil $G < 0$ eine invariante Gleichung ist, $g'_{44}\dot{x}'^2_4 < 0$ oder $g'_{44} < 0$, und zwar unabhängig davon, ob die Transformation auf Ruhe eine erlaubte ist oder nicht.

Es kann also jeder Massenpunkt, der sich auf einer Zeitlinie bewegt, auf Ruhe transformiert werden, und es kann sich jeder noch so kleine Planet, der sich um die Sonne bewegt, als Mittelpunkt der Welt betrachten, ohne dass das Kausalitätsprinzip verletzt wird, Dies müssen wir ihm auch gönnen, ja, dieses Resultat ist sogar ein sehr befriedigendes.

142 § 72. Geradlinige Bewegung des Massenpunktes

Nun wenden wir uns einer anderen Aufgabe zu und behandeln das Gegenstück zur Kreisbewegung, nämlich

die Bewegung in gerader Richtung [119]

auf das Massenzentrum. Statt $r =$ const. nehmen wir also jetzt $\varphi =$ const. $= 0$ an. Die Lagrangeschen Gleichungen der Bewegungen lauten jetzt

$$\frac{d}{dp}\left(2\frac{r}{r-\alpha}\dot{r}\right) - \left(\frac{r}{r-\alpha}\right)_r \dot{r}^2 + \frac{\alpha}{r^2}\dot{t}^2 = 0, \tag{19}$$

$$\frac{r-\alpha}{r}\dot{t} = 1, \tag{12'}$$

(aus (11') folgt jetzt $B = 0$). Aus diesen beiden Gleichungen eliminieren wir den Parameter p, und zwar folgt wegen $\dot{t} = \frac{r}{r-\alpha}$ und $\frac{d}{dp} = \frac{r-\alpha}{r}\frac{d}{dt}$ aus (19) die Gleichung

$$\frac{r}{r-\alpha}\frac{d}{dt}\left(2\frac{r}{r-\alpha}\frac{r}{r-\alpha}\frac{dr}{dt}\right) + \frac{\alpha}{(r-\alpha)^2}\frac{r^2}{(r-\alpha)^2}\left(\frac{dr}{dt}\right)^2 + \frac{\alpha}{r^2}\frac{r^2}{(r-\alpha)^2} = 0$$

oder

$$\frac{d^2r}{dt^2} = \frac{3\alpha}{2r(r-\alpha)}\left(\frac{dr}{dt}\right)^2 - \frac{\alpha(r-\alpha)}{2r^3}, \tag{20}$$

als eine Differentialgleichung 2. Ordnung. Die entsprechende Gleichung in der Newtonschen Theorie heisst $\frac{d^2r}{dt^2} = -\kappa\frac{M}{r^2}$.[120] *Die neue Differentialgleichung*

[119] Added by Hilbert in pencil: “Hier besser Note S. 24.” Hilbert’s comment is probably a reference to *Hilbert 1917*, p. 76, (this Volume, p. 71). P. 76 is the 24th page of that paper, and Hilbert probably referred to the pagination of an offprint version. Offprints were often paginated starting with page 1 in contrast to the version that appeared in the journal issue, see *Sauer 1999*, note 74.

[120] Deleted: “Für grosses r wird in derselben die Beschleunigung konstant (auf der Erdoberfläche g).”

(20) ist infolge des Auftretens eines Gliedes mit $\left(\frac{dr}{dt}\right)^2$ *erheblich komplizierter als die Newtonsche*, aber sie hat noch denselben Charakter. Insbesondere kommt in ihr als einzige Konstante, die aus den Tatsachen zu entnehmen ist, wieder nur die Sonnenmasse α vor.

Wir stellen nun die Frage, die wir für allgemeines α und für beliebige Geschwindigkeit noch gar nicht berührt haben: *Wirkt die Gravitation anziehend oder abstossend*, d. h. | ist $\frac{d^2r}{dt^2} < 0$ oder $\frac{d^2r}{dt^2} > 0$? Dies hängt vom Vorzei- 143
chen der rechten Seite der Gleichung (20) ab. Aus derselben erhalten wir als Bedingung für die

$$\begin{array}{l} \textit{Abstossung:} \\ \textit{Anziehung:} \end{array} \quad \frac{3\alpha}{2r(r-\alpha)}\left(\frac{dr}{dt}\right)^2 \begin{array}{c} > \\ < \end{array} \frac{\alpha(r-\alpha)}{2r^3}.$$

Also muss $\frac{dr}{dt} = |v| < \frac{1}{\sqrt{3}}\frac{r-\alpha}{r}$ sein, damit Anziehung und $|v| > \frac{1}{\sqrt{3}}\frac{r-\alpha}{r}$ damit Abstossung eintritt. In einer beliebigen, festen Entfernung vom Massenzentrum wirkt bei hinreichend kleiner Geschwindigkeit die Gravitation, wie man sieht, unabhängig vom Vorzeichen von α immer anziehend. Dieser Fall ist in der Himmelsmechanik realisiert.

Für $\lim \alpha = 0$ muss $|v| < \frac{1}{\sqrt{3}}$ sein, damit die Gravitation anziehend wird; die Geschwindigkeit kann also noch ausserordentlich gross sein, d.h. *die Anziehung gilt*, wie man sieht, *weit über die Gültigkeitsgrenze des Newtonschen Gesetzes hinaus*. Im Newtonschen Falle, nämlich für $\lim \alpha = 0$ und $\lim \frac{dr}{dt} = 0$ erhalten wir aus (20) auch das Newtonsche Gesetz: $\frac{d^2r}{dt^2} = -\frac{\frac{\alpha}{2}}{r^2}$. Wiederum finden wir $\frac{\alpha}{2} = \kappa M$.

Ersetzen wir in der Newtonschen Gleichung r für grosse Werte durch $x + a$, wobei $\frac{x}{a}$ sehr klein sein soll und entwickeln nach Potenzen von $\frac{x}{a}$, so erhalten wir

$$\frac{d^2x}{dt^2} = -\frac{\frac{\alpha}{2}}{a^2}\left\{1 - \frac{2x}{a} + \cdots\right\}.$$

Also ist die Beschleunigung in grosser Entfernung vom Massenzentrum angenähert konstant (auf der Erdoberfläche $= g$).

Für kleine Gravitation, aber *endliche* Geschwindigkeit erhalten wir aus (20) $\frac{d^2r}{dt^2} = -\frac{\frac{\alpha}{2}}{r^2}\left(1 - 3\left(\frac{dr}{dt}\right)^2\right)$. | Es tritt dann also *zur Newtonschen Kraft* $-\frac{\frac{\alpha}{2}}{r^2}$ 144
noch eine ihr *entgegengesetzt wirkende Kraft* $\frac{\frac{\alpha}{2}}{r^2}3\left(\frac{dr}{dt}\right)^2$, die auch proportional der Masse und ausserdem proportional dem Quadrat der Geschwindigkeit ist, hinzu.

Wir können die Gleichung aber noch *anders deuten*, wenn wir sie in der Form

$$\frac{1}{1 - 3\left(\frac{dr}{dt}\right)^2}\frac{d^2r}{dt^2} = -\frac{\frac{\alpha}{2}}{r^2}$$

schreiben. Dann wirkt auf den Massenpunkt nur die Newtonsche Kraft $-\frac{\frac{\alpha}{2}}{r^2}$, dagegen *vergrössert sich* jetzt unter dem Einfluss der Geschwindigkeit *die*

Masse im Verhältnis $1 : \frac{1}{1-3\left(\frac{dr}{dt}\right)^2}$. Man beachte die merkwürdige *Analogie zur Elektrodynamik*, wo die auf das Elektron wirkende Kraft

$$\mathfrak{k} = \frac{1}{1-\frac{3}{2}\left(\frac{dr}{dt}\right)^2}\frac{d^2r}{dt^2} \text{ ist, d. h. Masse} = \frac{1}{1-\frac{3}{2}\left(\frac{dr}{dt}\right)^2}.$$

Da sich der Massenpunkt auf einer Zeitlinie bewegt, ist in (9')

$$A = \frac{r}{r-\alpha}\left(\frac{dr}{dt}\right)^2 - \frac{r-\alpha}{r} < 0.$$

Es wird also immer $\frac{dr}{dt} = |v| < \frac{r-\alpha}{r}$ sein.

§ 73. Geradlinige Bewegung des Lichtstrahls

Für $A = 0$ haben wir das auf das Zentrum gerichtete oder das davon wegstrebende Licht. Dann wird aus der Ungleichung eine Gleichung: $|v| = \frac{dr}{dt} = \frac{r-\alpha}{r}$. Ein solcher *Lichtstrahl wird also immer abgestossen*, da seine Geschwindigkeit stets zu gross ist, als dass die Gravitation noch anziehend wirken könnte. Die Lichtgeschwindigkeit im Unendlichen in beliebiger Richtung erhält man, wenn man in $\frac{r}{r-\alpha}\left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{d\varphi}{dt}\right)^2 - \frac{r-\alpha}{r} = 0$ $r^2\left(\frac{d\varphi}{dt}\right)^2$ ersetzt durch $v^2 - \left(\frac{dr}{dt}\right)^2$.
145 Dann folgt aus $\frac{r}{r-\alpha}\left(\frac{dr}{dt}\right)^2 + v^2 - \left(\frac{dr}{dt}\right)^2 - \frac{r-\alpha}{r} = 0$ für grosses r *unabhängig von der Richtung* $|v| = 1$.

Der geradlinig auf das Zentrum zueilende Lichtstrahl, dessen Geschwindigkeit also immer abnimmt, *hat* wegen $\frac{dr}{dt} = \frac{r-\alpha}{r}$ *im Punkt* $r = \alpha$ *die Geschwindigkeit Null*. Er erreicht diesen Punkt übrigens erst nach unendlich langer Zeit. Man kann diese Gleichung nämlich integrieren: es wird $t = r + \alpha \lg(r-\alpha)$ und man erhält $t = \infty$ für $r = \alpha$ Da unsere Gleichung eigentlich $\left|\frac{dr}{dt}\right| = \frac{r-\alpha}{r}$ lautet, ist auch der umgekehrte Strahlengang möglich und es[121] wird ein Lichtstrahl, der in sehr geringer Entfernung vom Punkt $r = \alpha$ die zugehörige sehr kleine, radial nach aussen gerichtete Beschleunigung besitzt, geradlinig ins Unendliche laufen, aber erst nach sehr langer Zeit (mit der Geschwindigkeit 1) dort ankommen. *Die Lichtgeschwindigkeit* im Gravitationsfeld ist natürlich nur im Unendlichen von der Richtung des Lichtstrahls unabhängig, im allgemeinen *hängt* sie *von der Richtung ab*. Im zentrisch symmetrischen Gravitationsfeld z. B. ist auf dem Kreis $r = \frac{3}{2}\alpha$ die Geschwindigkeit in der Normalen $\frac{dr}{dt} = \frac{r-\alpha}{r} = \frac{\frac{3}{2}\alpha-\alpha}{\frac{3}{2}\alpha} = \frac{1}{3}$. In der Tangente dagegen ist sie, wie wir auf S. 137 sahen, gleich $\frac{1}{\sqrt{3}}$. Den topologischen Verlauf der Lichtstrahlen werden wir weiter unten diskutieren.[122]

[121]"ist auch der umgekehrte Strahlengang möglich und es" was interlineated with ink in an unknown hand.

[122]See §§ 76, 77 below.

§ 74. Untersuchung der Bahnkurven in zweiter Näherung 146

Nachdem wir erkannt haben, dass die Bahnkurve eines nach der Einsteinschen Theorie sich bewegenden Massenpunktes, d. h. die Lösungen der Differentialgleichung (13), für $\lim \alpha = 0$ und $\lim B = 0$ in *erster Näherung* die Kegelschnitte der Newtonschen Theorie sind, machen wir uns nun an

die Untersuchung der Bahnkurven in zweiter Näherung.

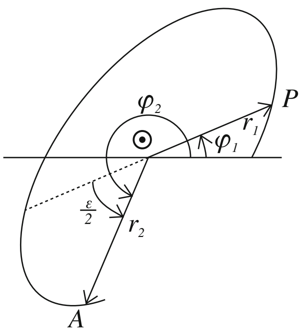
Bekanntlich findet der Astronom das Newtonsche Attraktionsgesetz in all den mannigfachen Erscheinungen der ganzen Himmelsmechanik, *mit Ausnahme einer kleinen*, unbedeutenden *Störung beim Merkur, auf Glänzendste bestätigt*,[123] d. h. er kann alle Bewegungen der Planeten, Kometen und Monde, sowie die der Doppelsterne unter der Annahme dieses Kraftgesetzes zwanglos erklären und genau vorausberechnen. Um diese Leistung der Newtonschen Theorie gebührend einzuschätzen, muss man wissen, dass die astronomischen Messungen sich sehr exakt durchführen lassen und z. B. die physikalischen Messungen — mit alleiniger Ausnahme der Spektralanalyse — an Genauigkeit weit übertreffen, so dass das Newtonsche Gesetz einer viel feineren Prüfung unterzogen werden konnte als die meisten anderen physikalischen Gesetze. Bedenkt man noch, dass dieses Attraktionsgesetz ausserdem die ganze Mechanik erklärt, die wir auf der Erdoberfläche vor uns haben, so begreift man, dass das Newtonsche Gesetz als eines der best fundierten Naturgesetze galt, und dass jene kleine Abweichung bei der Merkurbahn durch ad hoc eingeführte Annahmen zu erklären gesucht wurde.[124]

An eine *Abänderung des Newtonschen Gesetzes* zu denken, | wagte man gar 147
nicht. So wurde es ein *Prüfstein der Einsteinschen Theorie*, ob aus ihr die merkwürdige Bahnkurve des Merkurs ohne weiteres folge, was nun in der Tat der Fall ist.

Wir werden sehen, dass in der Einsteinschen Theorie alle Bahnkurven, sobald sie nur beliebig wenig von der Kreisbahn abweichen, nicht mehr geschlossen sind, ganz im Gegensatz zur Newtonschen Theorie, wo die dem Kreis benachbarten Kurven Ellipsen sind. Dass die aus jener Theorie folgende Abweichung von der Ellipsenbahn nun nur beim Merkur beobachtet werden konnte, hat den Grund darin, dass für alle Planeten die Exzentrizität der Ellipsenbahn so klein und — mit Ausnahme des Merkur — die Entfernung von der Sonne so gross ist, dass die Sonnenmasse α selbst, die ja auch die Dimension einer Länge hat, gegenüber jener Entfernung eine so kleine Grösse ist, dass in der Gleichung (13) schon die *erste Näherung* die astronomischen Tatsachen mit hinreichender Genauigkeit beschreibt. Um so verblüffender ist es aber auch, dass gerade die *einzige* mit Sicherheit *feststellbare Abweichung vom Newtonschen Gesetz durch die Einsteinsche Theorie vollkommen erklärt* werden kann.

[123] For contemporary reviews, see *Zenneck 1901*, *Newcomb 1895*.

[124] For a historical discussion of attempts to explain the anomaly, see *Earman and Janssen 1993*.



Man bezeichnet in der Astronomie den Punkt der Planetenbahn, in dem die Entfernung von der Sonne ein Minimum ist, als *Perihel*, denjenigen, in dem sie ein Maximum ist, als *Aphel*. Die Newtonsche
148 | Theorie verlangt nun, dass der Winkel $\varphi_2 - \varphi_1$, den der Radius-Vektor: Sonne-Planet zwischen diesen beiden Bahnpunkten beschreibt, genau gleich π ist. Als

Perihelbewegung ϵ

d. h. als Winkel, um den das Perihel in Abweichung von der Newtonschen Theorie bei einem Umlauf des Planeten in der Bahnebene vorrückt oder zurückbleibt, haben wir also $\varepsilon = 2(\varphi_2 - \varphi_1 - \pi)$ zu nehmen, und *diesen Winkel wollen wir jetzt* aus der Einsteinschen Theorie *berechnen*. Zuerst leiten wir die Formel ab, die das Vorrücken des Perihels ganz *allgemein* durch die Sonnenmasse, die halbe grosse Achse und die Exzentrizität der Ellipse ausdrückt, und erst nachher setzen wir die speziellen Werte für den Merkur ein.

Wir müssen von der Differentialgleichung (13)

$$\left(\frac{d\rho}{d\varphi}\right)^2 = \frac{1+A}{B^2} - \frac{\alpha A}{B^2}\rho - \rho^2 + \alpha\rho^3 = a + b\rho - \rho^2 + \alpha\rho^3$$

ausgehen. Wir bezeichnen noch die Entfernungen Sonne-Perihel und Sonne-Aphel mit bezw. r_1 und r_2 und setzen $\rho_1 = \frac{1}{r_1}$, $\rho_2 = \frac{1}{r_2}$. Diese beiden Werte finden wir als Wurzeln der kubischen Gleichung, die man durch Nullsetzen der rechten Seite von (13) erhält,

$$a + b\rho - \rho^2 + \alpha\rho^3 = 0,$$

und zwar müssen wir jene beiden Wurzeln nehmen, die für $\alpha = 0$ in die Wurzeln der entsprechenden quadratischen Gleichung der Newtonschen Theorie

$$a + b\rho - \rho^2 = 0$$

übergehen. Wir erhalten aus (13)

$$\frac{d\varphi}{d\rho} = \frac{1}{\sqrt{a + b\rho - \rho^2 + \alpha\rho^3}}$$

149 und finden für den Winkel

$$\varphi_2 - \varphi_1 = \int_{\rho_1}^{\rho_2} \frac{d\rho}{\sqrt{a + b\rho - \rho^2 + \alpha\rho^3}}. \tag{21}$$

Rechts steht ein ganz bestimmter Ausdruck von a, b und α. Zu integrieren ist zwischen jenen beiden Wurzeln der kubischen Gleichung. Für $\alpha = 0$ lässt sich das Integral sofort auswerten und man findet $\varphi_2 - \varphi_1 = \pi$, so dass das Vorrücken des Perihels $\varepsilon = 0$ wird, wie es auch sein muss. Wir entwickeln nun den Integrand (21) nach Potenzen von α. Es ist

$$\begin{aligned} a + b\rho - \rho^2 + \alpha\rho^3 &= \alpha(\rho - \rho_1)(\rho - \rho_2)(\rho - \rho_3) \\ &= -(\rho - \rho_1)(\rho - \rho_2) \cdot \alpha\rho_3 \left(1 - \frac{\alpha\rho}{\alpha\rho_3}\right), \end{aligned}$$

wobei ρ_3 diejenige Wurzel der kubischen Gleichung ist, die für $\alpha = 0$ unendlich wird. Jetzt muss noch ρ_3 berechnet werden. Bekanntlich ist $\rho_1 + \rho_2 + \rho_3 = \frac{1}{\alpha}$ (die Summe der Wurzeln ist gleich dem negativen Koeffizienten von ρ^2). Also wird

$$a + b\rho - \rho^2 + \alpha\rho^3 = -(\rho - \rho_1)(\rho - \rho_2)(1 - \alpha\rho_1 - \alpha\rho_2) \left(1 - \frac{\alpha\rho}{1 - \alpha\rho_1 - \alpha\rho_2}\right)$$

und

$$\frac{1}{\sqrt{a + b\rho - \rho^2 + \alpha\rho^3}} = + \left(1 + \frac{1}{2}\alpha(\rho_1 + \rho_2)\right) \left(1 + \frac{1}{2}\alpha\rho\right) \frac{1}{\sqrt{(\rho_1 - \rho)(\rho - \rho_2)}}.$$

Diesen Ausdruck setzen wir unter dem Integralzeichen ein und erhalten

$$\varphi_2 - \varphi_1 = \left(1 + \frac{1}{2}\alpha(\rho_1 + \rho_2)\right) \int_{\rho_1}^{\rho_2} \frac{1 + \frac{1}{2}\alpha\rho}{\sqrt{(\rho_1 - \rho)(\rho - \rho_2)}} d\rho.$$

Nun können wir die Integration ausführen:

$$\begin{aligned} \varphi_2 - \varphi_1 = \left\{1 + \frac{1}{2}\alpha(\rho_1 + \rho_2)\right\} \Bigg[& \arcsin \frac{\rho - \frac{\rho_1 + \rho_2}{2}}{\frac{\rho_2 - \rho_1}{2}} \\ & + \frac{1}{2}\alpha \left\{ \sqrt{(\rho_1 - \rho)(\rho - \rho_2)} + \frac{\rho_1 + \rho_2}{2} \arcsin \frac{\rho - \frac{\rho_1 + \rho_2}{2}}{\frac{\rho_2 - \rho_1}{2}} \right\} \Bigg]_{\rho_1}^{\rho_2} \end{aligned}$$

. 150

§ 75. Die Perihelbewegung des Merkur

Wir erhalten also

$$\begin{aligned} \varphi_2 - \varphi_1 &= \left\{1 + \frac{1}{2}\alpha(\rho_1 + \rho_2)\right\} \Bigg(\overbrace{\arcsin 1 - \arcsin(-1)}^{\pi} \\ &\qquad + \frac{\alpha}{2} \left\{ 0 + \frac{\rho_1 + \rho_2}{2} (\arcsin 1 - \arcsin(-1)) \right\} \Bigg) \\ &= \left\{1 + \frac{1}{2}\alpha(\rho_1 + \rho_2)\right\} \left\{\pi + \frac{\pi}{4}\alpha(\rho_1 + \rho_2)\right\} = \left(1 + \frac{3}{4}\alpha(\rho_1 + \rho_2)\right)\pi. \end{aligned}$$

Somit wird der Winkel ε, der das *Vorrücken des Perihels* angibt

$$\varepsilon = \frac{3}{2}\pi\alpha(\rho_1 + \rho_2).$$

Nun ist aber die Summe der beiden Wurzeln $\rho_1 + \rho_2$ der kubischen Gleichung angenähert gleich der Summe der beiden Wurzeln $\sigma_1 + \sigma_2$ der quadratischen Gleichung $a+b\rho-\rho^2 = 0$, d.h. es muss $\rho_1+\rho_2 = \sigma_1+\sigma_2+\alpha f(\sigma_1, \sigma_2)$ sein, weil für $\alpha = 0$ $\rho_1+\rho_2 = \sigma_1+\sigma_2$ wird.[125] Sei $S_i = \frac{1}{\sigma_i}$ die grösste bezw. kleinste Entfernung des Planeten von der Sonne, wenn derselbe sich genau in einer Ellipse bewegen würde, so gilt die Beziehung $\sigma_1 + \sigma_2 = \frac{S_1+S_2}{S_1 \cdot S_2} = \frac{2a}{a^2(1-e^2)}$ Hierin bedeutet a die grosse Halbachse und e die numerische Exzentrizität der Ellipse. Setzt man diesen Ausdruck oben ein, so erhält man als *Perihelbewegung*

$$\varepsilon = \frac{3\alpha\pi}{a(1-e^2)}.$$

In diese Formel hat man nun in jedem Spezialfall für alle Konstanten ihre numerischen Werte einzusetzen. Um *die Perihelbewegung des Merkur* zu finden, hat man also für α die Sonnenmasse einzusetzen und a und e aus der beobachteten Merkurbahn zu entnehmen. Wir behalten uns alle Dimensionsüberlegungen auf später vor, um hier die theoretischen Untersuchungen nicht unterbrechen zu müssen. Vor der Hand genügt es zu wissen, dass α aus der experimentell bestimmten Sonnenmasse berechnet werden kann und angenä-
151 hert gleich 3 km gefunden wird.[126] Setzt | man alle beobachteten Daten ein, so findet man für den Merkur

$$\varepsilon = 43'' \qquad \textit{in 100 Jahren}$$

in wundervoller Uebereinstimmung mit den Beobachtungen des Astronomen Leverrier: $\varepsilon = 45'' \pm 5''$.[127]

Aus der allgemeinen Formel *ersieht man von neuem, dass* $\alpha > 0$ sein muss, weil $\varepsilon > 0$ ist. Dies allein ist schon eine qualitative Bestätigung der Theorie, besonders auch, da in allen Fällen, wo eine Perihelbewegung einigermassen bestimmt hat nachgewiesen werden können, *immer* $\varepsilon > 0$ gefunden wurde. (Literatur: A. Einstein, Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie. Sitzungsber. d. Preuss. Akad. d. Wiss. (47) S. 831 1915).[128]

[125] On the left hand page, Hilbert wrote in pencil: "Statt σ_1, σ_2 schreibe ρ_1^0, ρ_2^0!"

[126] See p. 280 below.

[127] This is the value quoted in *Einstein 1915c*, p. 839. Einstein cites *Newcomb 1895* as his source, but see the discussion in *Earman and Janssen 1993*, pp. 130–132. See also *Roseveare 1982* for a historical discussion.

[128] *Einstein 1915c*.

§ 76. Diskussion der Differentialgleichung der Bahnkurven für die Lichtbewegung

Wir wollen nun nochmals auf *die Bewegung der Lichtstrahlen im Gravitationsfeld* eines Massenpunktes zurückkommen, und zwar wollen wir

die Gestalt der Bahnkurven

beschreiben, ohne uns dabei auf Beweise einzulassen. Alles Nähere wird man in einer demnächst in den Göttinger Nachrichten erscheinenden Mitteilung von V. Fréedericksz finden.[129] Da für das Licht $A = 0$ wird, nimmt die Differentialgleichung (13) jetzt die Form an

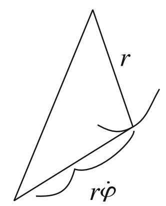
$$\left(\frac{d\rho}{d\varphi}\right)^2 = \frac{1}{B^2} - \rho^2 + \alpha\rho^3. \tag{22}$$

Die Konstante B hat hier eine einfache Bedeutung. Wir erinnern nämlich daran, dass B die Konstante des Flächensatzes ist, wie man aus der Lagrangeschen Differentialgleichung | $r^2\dot{\varphi} = B$ ersieht.[130] B ist also gleich dem 152
doppelten Inhalt des Dreiecks, dessen Grundlinie die in einem Bahnpunkt in der Richtung der Tangente an die Bahnkurve aufgetragene Länge der Lichtgeschwindigkeit und dessen Höhe der kürzeste Abstand dieser Tangente vom Gravitationszentrum ist. Da die Lichtgeschwindigkeit nun im Unendlichen den Wert 1 hat, so ist B selbst der kürzeste Abstand, den die Tangente an die Bahnkurve im unendlichen fernen Punkt derselben (Asymptote der Bahnkurve) vom Gravitationszentrum hat. *B ist sozusagen die Distanz, um welche der Lichtstrahl im Unendlichen am Zentrum vorbeizielt.* (L. Flamm, Beiträge zur Einsteinschen Gravitationstheorie, Physikal. Zeitschr. 1916 S. 452)[131]

Das allgemeine Integral der Differentialgleichung (22) ist von der Form $\rho = f(B^2, \varphi - \varphi_0)$. Es entsteht also dadurch, dass man diejenige Bahnkurve, die ein Lichtstrahl beschreibt, der von einem bestimmten unendlich fernen Punkt ($\rho = 0$) z. B. dem Punkt $\varphi_0 = 0$ in einer durch die Konstante B bestimmten Richtung ausgeht, um den Punkt $r = 0$ sich drehen lässt. Wir setzen in

[129] This paper by Fréedericksz, in fact, never appeared, see *Hilbert 1917*, p. 77, (this Volume, p. 71 and its note 63).

[130] At the top of the page, Hilbert sketched the following figure in pencil:

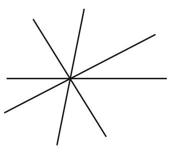


[131] *Flamm 1916*.

der weiteren Diskussion immer $\varphi_0 = 0$, betrachten also nur die einparametrige Kurvenschar, die der Lichtstrahl beschreibt, wenn man der Konstante B immer andere Werte erteilt. Es ist dies die *Gesamtheit der Bahnkurven des Lichtes, das von einem unendlich fernen Punkt* nach allen möglichen Richtungen *ausgeht*. Die Asymptoten an diesen Bahnkurven im unendlich fernen Punkt bilden also eine Schar paralleler Geraden. Wir können uns nun ein all-
153 gemeines Bild | dieser Kurven machen, weil wir wissen, dass das Licht sich auf einem Kreis vom Radius $r = \frac{3}{2}\alpha$, also auf einer *geschlossenen Kurve* bewegen kann. *Dieser Kreis muss daher ein Integral unserer Differentialrechnung sein*, wenn in dieselbe für B ein gewisser Wert eingesetzt wird, den man leicht berechnen kann; und zwar folgt aus $r^2\dot{\varphi} = B$, $r^2\dot{\varphi}^2 - \frac{r-\alpha}{r}\dot{t}^2 = 0$, $\frac{r-\alpha}{r}\dot{t} = 1$ — drei Gleichungen, die wir auf S. ⟨126⟩ aufgestellt hatten — für $r = \frac{3}{2}\alpha$ der gesuchte Wert $B^2 = \frac{27}{4}\alpha^2$.

§ 77. Die Poincarésche Zykeltheorie und die Krümmung der Lichtstrahlen

Dieser *merkwürdige Fall, dass eine Differentialgleichung eine einzige geschlossene Kurve als Lösung hat*, ist von *Poincaré* in seiner Himmelsmechanik zuerst untersucht worden. Seine Grundidee war, dass diese *geschlossene Kurve den Verlauf aller anderen Bahnkurven*, d. h. aller anderen Lösungen der Differentialgleichungen für andere Werte von B^2 *charakterisieren muss*. Man nennt diese ausgezeichnete Lösung ein *Zykel*, und die in dem erwähnten Werk ausgebaute Theorie heisst die *Poincarésche Zykeltheorie*. Da alle Lösungen der Differentialgleichung, die zu verschiedenen Werten des Parameters B^2 gehören, durch den unendlichen fernen Punkt $\rho = 0$ hindurchgehen, so ist auch diese eine singuläre Stelle, allerdings eine sehr einfache nämlich in der Poincaréschen Terminologie ein sogenannter *Strahlenpunkt*.

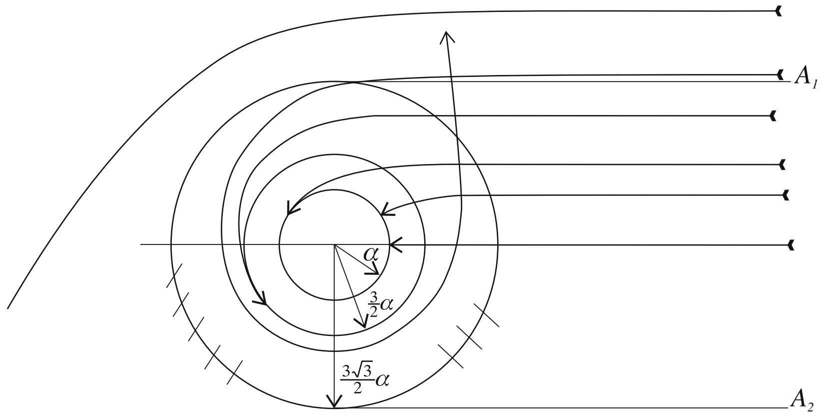


Für grosse Werte von r, d. h. für kleine Werte von ρ wird nämlich die Differentialgleichung (22) zu $\left(\frac{d\rho}{d\varphi}\right)^2 = \frac{1}{B^2} - \rho^2$. Ihre Lösung ist die gerade Linie $\rho = \frac{\sin\varphi}{B}$, d. h. die Bahnkurven
154 sind in grosser Entfernung vom Gravitationszentrum | Gerade, und zwar bedeutet B diesen Abstand der Asymptote von der x-Achse.[132] Setzen wir für $B = \frac{3\sqrt{3}}{2}\alpha$, so erhalten wir die Asymptoten der zu diesem Wert von B gehörigen Bahnkurve. Da für dieses B auch der Kreis $r = \frac{3}{2}\alpha$ ein Integral[133] ist, so können wir schon vermuten, dass sich *jene Bahnkurve asymptotisch um diesen Kreis herumschlingen wird*, ohne ihn je zu erreichen. Die nähere Diskussion der

[132] The preceding half sentence was interlineated by Hilbert in pencil.

[133] "Integral" was corrected from "singuläres Integral" in pencil.

Differentialgleichung zeigt, dass dies tatsächlich der Fall ist. Den Verlauf der übrigen Lösungen der Gleichung ersieht man aus nebenstehender Figur.[134]



In derselben sind die drei[135] Kreise $r = \alpha$, $r = \frac{3}{2}\alpha$ und $r = \frac{3\sqrt{3}}{2}\alpha$ gezeichnet. Jene beiden Kurven also, deren Asymptoten A_1 und A_2 sind,[136] schlingen sich asymptotisch um den Kreis $r = \frac{3\alpha}{2}$[137] herum. Die Kurven, deren Asymptoten zwischen A_1 und A_2 liegen, schneiden den Kreis $r = \frac{3}{2}\alpha$ und gelangen bis zum innersten Kreis, | wo sie aufhören. Diejenigen Bahnkurven endlich, 155
deren Asymptoten ausserhalb von A_1 und A_2 liegen, schlingen sich, wenn ihre Asymptoten nahe genug an einer der beiden ausgezeichneten Asymptoten liegen, beliebig oft um den Kreis $r = \frac{3}{2}$ herum, entfernen sich dann wieder und verlaufen asymptotisch geradlinig in einer anderen Richtung ins Unendliche. Hat dagegen die Asymptote an die Kurve eine grosse Entfernung $B\delta$[138] von A_1 und A_2, so schlingt sich die Kurve überhaupt nicht um den mittleren Kreis herum, sondern wird in seiner Nähe nur etwas gekrümmt und geht dann wieder ins Unendliche. Ist der Abstand δ äusserst gross, d. h. ist $\lim \alpha = 0$, *so hat die Bahnkurve die Gestalt eines Hyperbelastes* $r = \frac{B}{\frac{\alpha}{2B} + \cos\varphi}$, der gegen die Sonne zu konkav ist und die Sonne selbst im Brennpunkt stehen hat; wird schliesslich δ unendlich gross, d. h. ist $\alpha = 0$, so wird die Bahnkurve zur Geraden. Für $\lim \alpha = 0$ erhält man also die Ablenkung, die der Lichtstrahl im Gravitationsfeld erfährt, wenn man den Winkel, den die beiden Hyperbelasymptoten einschliessen, berechnet. Dieser Winkel hat die Grösse $\frac{\alpha}{B}$.

Setzt man für $\frac{\alpha}{B}$ die Sonnenmasse ein und wählt B so, dass der Lichtstrahl gerade den Sonnenrand tangiert, d. h. so, dass *die Ablenkung durch die Sonne*

[134]Laue included virtually the same figure into his *Laue 1921*, p. 226, see also *Eisenstaedt 1987*, p. 306, for historical discussion.

[135]"drei" was deleted in pencil and corrected to "2".

[136]"sind" was corrected by Hilbert in pencil from "den äussersten Kreis berühren".

[137]"den Kreis $r = \frac{3\alpha}{2}$" was corrected in pencil from "den mittleren Kreis".

[138]"B was interlineated with pencil."

ein Maximum wird, so findet man

$$\frac{\alpha}{B} = 1.7''.$$

Dieser Effekt, müsste festzustellen sein, wenn man einen Fixstern, dessen Strahlen nahe am Sonnenrande vorbeigehen, bei einer Sonnenfinsternis beob-
156 achtet. Doch ist er experimentell noch |[139] nicht bestätigt worden, da bis heute noch keine dahinzielenden Untersuchungen angestellt werden konnten.[140] Auch durch den Planeten Jupiter sollte eine Ablenkung der Lichtstrahlen verursacht werden, die — obgleich viel kleiner als die durch die Sonne bewirkte — doch eben noch feststellbar sein muss. Leider liegen auch hierüber noch keine Beobachtungen vor.[141]

---

[139] On the left hand page, Hilbert wrote in pencil the following notes and then later deleted them again:
Will man für grosse Entfern⟨un⟩g⟨en⟩ und zugleich grosse B die Differentialgl. int⟨e⟩g⟨rieren,⟩ so setze $\rho B = \sigma$ so wird (S. 151)

$$\left(\frac{d\sigma}{d\varphi}\right)^2 \frac{1}{B^2} = \frac{1}{B^2} - \frac{\sigma^2}{B^2} + \frac{\alpha\sigma^3}{B^3}$$

$$\left(\frac{d\sigma}{d\varphi}\right)^2 = 1 - \sigma^2 + \frac{\alpha\sigma^3}{B}$$

$$\sigma = \cos\varphi + \tau \qquad \frac{\alpha}{B} = \delta$$

$$\left(-\sin\varphi + \frac{d\tau}{d\varphi}\right)^2 = 1 - (\cos\varphi + \tau)^2 + \frac{\alpha}{B}(\cos\varphi + \tau)^3$$

$$-2\sin\varphi\frac{d\tau}{d\varphi} = -2\cos\varphi\cdot\tau + \frac{\alpha}{\cdot}$$

$$\left(\frac{d\sigma}{d\varphi}\right)^2 = 1 - \sigma^2 + \delta\sigma^3 \qquad \frac{1}{\sigma} = \tau$$

$$+\frac{1}{\tau^4}\left(\frac{d\tau}{d\varphi}\right)^2 = 1 - \frac{1}{\tau} + \frac{\delta}{\tau^3}$$

$$\left(\frac{d\tau}{d\varphi}\right)^2 = \tau^4 - \tau^2 + \delta\tau$$

[140] Einstein's prediction of light deflection in the gravitational field of the sun was confirmed by the results of two English expeditions to measure the deflection of star light during the solar eclipse of 29 May 1919, see *Dyson et al. 1920*. The results were announced on 6 November 1919, and made Einstein world famous almost instantaneously, see *CPAE9 2004*, pp. xxxi–xxxvii, and references cited therein, for a historical discussion.

[141] In *Einstein 1916a*, p. 822, the value of gravitational deflection for a light ray grazing the limb of Jupiter is given as $0,02''$. In correspondence, Einstein had claimed the feasibility of an observational confirmation of light deflection in the gravitational field of Jupiter (see Einstein to Otto Naumann, 7 December 1915, and Einstein to Karl Schwarzschild, 9 January 1916 *CPAE8-A 1998*, Docs. 160, 181). He had also suggested that Erwin Freundlich should try to undertake the corresponding measurements at the Göttingen observatory under his supervision, and had communicated with Hilbert, both in correspondence and in personal conversation, about Hilbert's support for such a plan (see Einstein to Hilbert, 18 February 1916, 30 March 1916, and Hilbert to Einstein, 27 May 1916 *CPAE8-A 1998*, Docs. 193, 207, 222, see also note 152 on p. 285 below). For the first observational confirmation of light deflection in the gravitational field of Jupiter, see *Treuhaft and Lowe 1991*.

§ 78. Dimensionsbetrachtungen. Berechnung von α für die Sonne und für ein Wasserstoffmolekül

Wir sind bis jetzt der Frage, welches

die Dimensionen der physikalischen Grössen

sind, absichtlich aus dem Wege gegangen. Wir wollen uns jetzt kurz mit dieser Frage beschäftigen, die deswegen von grosser Bedeutung ist, weil man bei allen numerischen Rechnungen gerade die Dimensionen kennen muss, um bekannte physikalische Daten, die der Physiker natürlich immer in seinem Masssystem, dem sog. cm, gr, sec (c. g. s)-Masssystem, angibt, auf unser Masssystem umrechnen zu können.

Die Grundgleichung der Newtonschen Mechanik ist $\frac{d^2x}{dt^2} = -\kappa \frac{M}{x^2}$. Dabei muss die Konstante κ eine solche Dimension besitzen, dass die rechte Seite der Gleichung dieselbe Dimension wie die linke, nämlich $\frac{\text{Länge}}{(\text{Zeit})^2}$ hat. Bezeichnen wir die Längen-, Massen-, Zeit-Dimension mit bezw. l, m, t, so sind die Dimensionen der Newtonschen Gleichung $\frac{l}{t^2} = \kappa \frac{m}{l^2}$. *Also ist κ von der Dimension $l^3 m^{-1} t^{-2}$* und zwar ist der experimentell gefundene Wert $\kappa = 6.7 \cdot 10^{-8} \text{cm}^3 \text{gr}^{-1} \text{sec}^{-2}$. *Um die Grösse α unserer Theorie* für irgend eine Masse *ausrechnen zu können*, muss man, wie wir wissen, die Gleichung $\left(\frac{d\varphi}{dt}\right)^2 = -\frac{\frac{\alpha}{2}}{r^3}$ | vergleichen mit der 157
entsprechenden Gleichung der Newtonschen Theorie $\left(\frac{d\varphi}{dt}\right)^2 = -\kappa \frac{M}{r^3}$. Diese letztere Gleichung kann man übrigens aus den beiden Differentialgleichungen

$$\frac{d^2x}{dt^2} = -\kappa \frac{Mx}{(x^2+y^2)^{\frac{3}{2}}}, \qquad \frac{d^2y}{dt^2} = -\kappa \frac{My}{(x^2+y^2)^{\frac{3}{2}}}$$

sofort ableiten. Wir suchen dieselben nämlich durch den eine Kreisbewegung darstellenden Ansatz

$$x = \sqrt{x^2+y^2} \cos \omega t, \qquad y = \sqrt{x^2+y^2} \sin \omega t$$

zu integrieren, wobei wir also $x^2 + y^2 = r^2 = \text{const.}$ lassen müssen. ω hat dann die Bedeutung der Winkelgeschwindigkeit und soll ebenfalls eine Konstante sein. Aus den beiden Differentialgleichungen finden wir dann als Gleichung, der die beiden Konstanten r und w genügen müssen, die bekannte Beziehung $\omega^2 = -\kappa \frac{M}{r^3}$, d. h. eben jene gesuchte Gleichung $\left(\frac{d\varphi}{dt}\right)^2 = -\kappa \frac{M}{r^3}$. Diese Gleichung können wir *nicht* ohne weiteres mit der entsprechenden Einsteinschen $\left(\frac{d\varphi}{dt}\right)^2 = -\frac{\frac{\alpha}{2}}{r^3}$ vergleichen, weil *die Zeit* in beiden *in verschiedenen Einheiten gemessen* wird; in der letzteren ist die Zeitdimension nämlich so gewählt, dass die Lichtgeschwindigkeit im Unendlichen gleich 1 wird, während der Physiker als Lichtgeschwindigkeit $c = 3 \cdot 10^{10} \text{cm}^1 \text{sec}^{-1}$ annimmt. Damit auch in der Einsteinschen Gleichung dieselbe Zeiteinheit wie in der Newtonschen steht, muss in der ersteren t durch ct ersetzt werden. In der Tat wird dann aus $\frac{dr}{dt} = 1$ für die neue Zeitvariable $\frac{dr}{dt} = c$. Wir haben also einerseits $\left(\frac{d\varphi}{dt}\right)^2 = -c^2\frac{\frac{\alpha}{2}}{r^3}$ und andererseits $\left(\frac{d\varphi}{dt}\right)^2 = -\kappa\frac{M}{r^3}$ oder

$$\frac{\alpha}{2} = \kappa\frac{M}{c^2}.$$

158 Rechter Hand muss eine Länge stehen, da α die Dimension einer Länge hat. Dies ist in der Tat der Fall, denn es ist $\frac{\text{cm}^3}{\text{grsec}^2}\frac{\text{gr}}{\frac{\text{cm}^2}{\text{sec}^2}} = \text{cm}$. Um den *Wert der Konstanten* $\frac{\alpha}{2}$ *für die Sonne* zu berechnen, hat man nun für M die bekannte Sonnenmasse $1.95 \cdot 10^{33} gr$ einzusetzen. Dann wird $\frac{\alpha}{2} = 1.45 \cdot 10^5$cm (der wirkliche Sonnendurchmesser ist $6.9 \cdot 10^{10}$cm).[142] Das Licht kann, wie wir wissen, auf einem Kreis vom Radius $r = \frac{3}{2}\alpha$ kreisen; dieser Radius hat also für die Sonne eine Länge von ungefähr 4.5km.

Als Gegenstück dazu berechnen wir noch $\frac{\alpha}{2}$ *für ein* H_2*-Molekül*. Hier hat man für M den experimentell gefundenen Wert von $1\,.\,64 \cdot 10^{-24} gr$ einzusetzen und erhält $\frac{\alpha}{2} = 1\,.\,23 \cdot 10^{-53}$cm. Dies ist ein sehr wichtiges Datum; einerseits gibt es einen Begriff von der Dimension, in der das Licht, also auch ein sehr schnell sich bewegendes Elektron um den positiven Kern eines Atoms kreisen müsste. Da andererseits der Durchmesser des Elektrons von der Grössenordnung 10^{-13}cm, also im Verhältnis zur Masse des H_2-Moleküls[143] schon unendlich gross ist, so wird man nicht erwarten dürfen, dass infolge der Grössenordnung von $\frac{\alpha}{2}$ für die Elektronenbewegung ein von der Newtonschen Mechanik abweichendes Bewegungsgesetz gelte. Unsere auf S. 129 geäusserte dahin gehende Vermutung erweist sich also als unrichtig.[144]

§ 79. Verhalten des Massfadens im zentrischen Gravitationsfeld bei tangentialer und radialer Lage

Wir haben schon zwei Folgerungen aus unserer Theorie gezogen, die einer Prüfung durch das Experiment zugänglich sind, nämlich erstens die Perihel-
159 bewegung der Planeten und | zweitens die Krümmung der Lichtstrahlen im Gravitationsfeld. Nun wollen wir noch einen dritten Effekt besprechen, der ebenfalls an der Erfahrung prüfbar ist. Wir befassen uns nämlich mit dem

Verhalten von Massfaden und Lichtuhr im Gravitationsfeld,

weil daraus ein ungemein wichtiger Schluss, nämlich die Rotverschiebung der Spektrallinien einer sich im Gravitationsfeld befindlichen Lichtquelle gegenüber einer ausserhalb desselben sich befindenden gezogen werden kann. Um

[142] "Sonnendurchmesser" should be "Sonnenradius".

[143] "Moleküls" was corrected from "Atoms".

[144] See p. 259 above.

diese komplizierte Erscheinung zu verstehen, machen wir uns am zentrisch-symmetrischen Gravitationsfeld nochmals klar, dass in demselben *die Euklidische Geometrie durchaus nicht gilt*. Wir denken uns in der Lage, alle nun anzustellenden Messungen gleichzeitig ausführen zu können, d. h. wir setzen $t = \text{const}$ Dann nimmt unser Linienelement die Form an

$$ds^2 = \frac{r}{r-\alpha}dr^2 + r^2 d\varphi^2.$$

Der Apparat, mit dem wir jetzt zu messen haben, ist der Massfaden, da alle Weltlinien nun Strecken sind. Derselbe gestattet uns also, die vom gewählten Bezugssystem unabhängige Länge einer endlichen — und nicht nur einer infinitesimalen — Strecke zu messen; denn *endliche* Strecken müssen wir messen können, um zu zeigen, dass die Euklidische Geometrie nicht mehr gültig ist. Im Infinitesimalen gilt dieselbe ja auch jetzt noch. Es kommt uns nun darauf an, ein *einfaches Experiment* ausfindig zu machen, *das die Nichtexistenz der Euklidischen Geometrie beweist*. Dazu müssen uns die Koordinaten r u. φ jedes Punktes bekannt sein.[145]

Falls wir *die Koordinaten* r und φ jedes Punktes der $r\varphi$-Ebene noch nicht 160
kennen, so *können wir* sie *folgendermassen bestimmen*: Man kann durch drei Messungen mit dem Massfaden in jedem Punkt der $r\varphi$-Ebene die $g_{\mu\nu}$ in Bezug auf irgend ein willkürliches Koordinatensystem experimentell bestimmen (siehe S. 96).[146] Diese Grösssen $g_{\mu\nu}$ denken wir uns in jedem Punkte der Geometrie angeschrieben. Dann können wir — wie unsere Theorie zeigt — ein solches Koordinatensystem einführen, dass $g_{11} = \frac{r}{r-\alpha}$, $g_{12} = 0$, $g_{22} = r^2$ wird. In demselben hat dann jeder Punkt der Ebene zwei experimentell zu findende Zahlen r und φ als Namen.

Die hier beschriebene allgemein anwendbare Methode, um die Funktionen $g_{\mu\nu}$ zu finden, lässt sich sicher in diesem besonders durchsichtigen Fall der zentrischen Symmetrie erheblich vereinfachen. Es ist klar, dass man selbst bei Hinzunahme der Zeit in unserem Gravitationsfeld

$$ds^2 = \frac{r}{r-\alpha}dr^2 + r^2 d\varphi^2 - \frac{r-\alpha}{r}dt^2$$

r mit dem Massfaden und auch mit der Lichtuhr durch irgend welche Messungen finden können muss; denn r ist eine Invariante. *Alle Invarianten* aber *lassen sich durch Massfaden und Lichtuhr bestimmen*. Ebenso ist in diesem Linienelement dt eine Invariante und kann also mit der Lichtuhr gemessen werden. Diese Bemerkung ist prinzipiell von Wichtigkeit für die unten zu besprechende Rotverschiebung der Spektrallinien.

Nachdem nun r und φ gefunden sind, können wir die Länge S einer geschlossenen Kurve $r = \text{const}$ mit dem Massfaden messen. Diese Kurve nennen wir
einen Kreis. Längs | desselben ist $ds = rd\varphi$ und daher wird $s = 2r\pi$. Auf dem 161

[145] The preceding sentence was added.

[146] See p. 236 above.

Kreise, auf welchem *der Massfaden überall tangential zum Gravitationsfeld angelegt* wird, *gelten noch* die Massverhältnisse der *Euklidischen Geometrie*. Betrachten wir also die Oberfläche der im dreidimensionalen r, ϑ, φ-Raum liegenden Kugel $r = \text{const}$, konstruieren auf ihr ein rechtwinkliges Dreieck, dessen Seiten als geodätische Linien grösste Kreise der Kugel sein müssen, und machen die mit dem Massfaden zu messende Länge der Katheten $= 1$, so finden wir als Länge der Hypothenuse $\sqrt{2}$. Wir bemerken noch, dass die geodätischen Linien auf der Kugeloberfläche keineswegs geodätische Linien sind, wenn wir die Geometrie im dreidimensionalen Raum

$$ds^2 = \frac{r}{r-\alpha}dr^2 + r^2 d\vartheta^2 + r^2 \sin^2 \vartheta d\varphi^2$$

betrachten, d. h. wenn wir auch Kurven als Verbindungslinien zweier Punkte der Kugeloberfläche zulassen, die *nicht* auf der Kugeloberfläche selbst verlaufen.

Legen wir nun den Massfaden im Gravitationsfeld

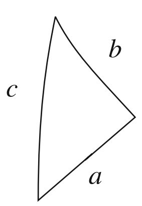
$$ds^2 = \frac{r}{r-\alpha}dr^2 + r^2 d\varphi^2$$

in radialer Richtung an, d. h. halten wir die φ konstant und lassen r und dr wachsen, so finden wir als Zunahme der Länge, die wir auf dem Massfaden ablesen können,

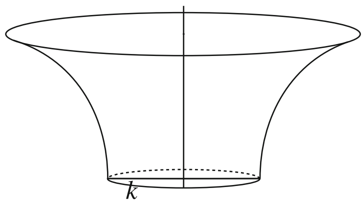
$$ds = \sqrt{\frac{r}{r-\alpha}}dr = \left(1 + \frac{\alpha}{2r}\right) dr,$$

während wir, wenn die Euklidische Geometrie Geltung hätte, $ds = dr$ finden müssten. Die am Massfaden abgelesene Zahl hat sich eben beim Fortschreiten um dr um mehr als dr vermehrt, nämlich um $\left(1 + \frac{\alpha}{2r}\right) dr$. *In Bezug auf dieses Koordinatensystem scheint es daher, als ob die am Massfaden angebrach-*
162 *ten | Zahlen näher zusammengerückt wären.* Wir können unser Experiment deswegen dahin interpretieren, *dass sich der in radialer Richtung angelegte Massfaden unter dem Einfluss des Gravitationsfeldes zusammengezogen hat.*

§ 80. Geometrische Interpretation auf der Rotationsfläche einer Parabel



Denken wir uns nun ein gleichschenkliges rechtwinkliges Dreieck mit der Kathetenlänge $s = 1$; die eine Kathete a möge eine radial gerichtete geodätische Linie ($\varphi = \text{const.}$) sein. Die andere sei eine geodätische Linie b, die auf der ersten senkrecht steht. Die Hypothenuse c ist dann diejenige geodätische Linie, welche die beiden Endpunkte der Katheten verbindet. *Legen wir* jetzt *den Massfaden*, der auf den beiden Katheten angelegt, die Länge 1 anzeigt, *längs der Hypothenuse an, so finden wir*



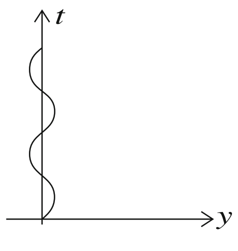
keineswegs als Länge derselben $\sqrt{2}$. Dies kann man sich leicht plausibel machen. Die durch das Linienelement $ds^2 = \frac{r}{r-\alpha}dr^2 + r^2 d\varphi^2$ definierte Geometrie lässt sich nämlich als *Geometrie auf einer Fläche deuten*, die durch *Rotation einer Parabel* um ihre Leitlinie entsteht. In der Umgebung des vom rotierenden Scheitelpunkt beschriebenen Kreises κ haben wir also von der Euklidischen Geometrie durchaus verschiedene Massverhältnisse. Da die Parabel aber, wenn man sich von dieser Stelle entfernt, sehr bald zum Durchmesser angenähert parallel wird, so nähert man sich dann auch schnell den Massverhältnissen der Euklidischen Geometrie (siehe L. Flamm, Physikal.⟨Zeitschrift⟩)[147]

§ 81. Die Rotverschiebung der Spektrallinien 163

Nun *betrachten wir die Pseudogeometrie* in einem x, y, t-Raum ($\vartheta = \frac{\pi}{2}$) unseres Gravitationsfeldes. Hier werden wir einen entsprechenden Effekt erhalten, und zwar einen, der der experimentellen Beobachtung zugänglich ist, nämlich

die Rotverschiebung der Spektrallinien auf den Fixsternen.

Das Linienelement unserer Pseudogeometrie kann man gemäss den auf S. 115 angegebenen Formeln leicht auf rechtwinklige Koordinaten umrechnen.[148] Es kommen dann Glieder mit dx^2, $dxdy$, dy^2 und dt^2 vor. Für einen in der x, t-Ebene[149] ($y = 0$) sich befindenden Punkt wird der Faktor von dt^2, der nachher allein von Wichtigkeit ist, zu $\frac{x-\alpha}{x}$: das Linienelement hat dann die Form



$$ds^2 = -\frac{x-\alpha}{x}dt^2 + \text{quadratische Form in } dx \text{ und } dy.$$

Um die Zeit ablesen zu können, müssen wir eine *Uhr* haben. Eine solche finden wir in denkbarer Vollkommenheit in der Natur realisiert durch *ein schwingendes und Licht emittierendes Molekül*, etwa das eine helle gelbe Spektrallinie a⟨u⟩ssendende Natriummolekül. Damit die Lichtschwingungen für einen in der x-Richtung blickenden Beobachter sichtbar sind, muss das Elektron im Molekül um den auf der x-Achse

[147] *Flamm 1916*.

[148] See p. 248 above.

[149] “in der x, t-Ebene” was corrected from “auf der Achse”.

ruhenden Kern in der dazu senkrechten y-Richtung[150] Schwingungen ausführen. Wir machen nun die *axiomatische Annahme dass die auf der Zeitachse*
164 *abgelesene Zeit t einer Schwingung immer die nämliche sei*. Befindet | sich das Molekül in dem zentrisch-symmetrischen Gravitationsfeld, so wird diese Schwingungsdauer aber, wie wir nun zeigen wollen, von der Entfernung des Moleküls vom Gravitationszentrum abhängen.

Wir machen ferner die Annahmen, dass

1) *die Eigenzeit* τ einer Schwingung, — die man also auf der Uhr ablesen müsste, welche die Bewegung des schwingenden Elektrons mitmacht — *unendlich klein* sei,

2) *die Amplitude dy* der Schwingung noch unendlich klein sei im Verhältnis zur Eigenzeit τ, d. h. dass auch die Geschwindigkeit des Elektrons *unendlich klein* sei.

Dann ist folgendes *Axiom* einleuchtend:

Die Eigenzeit τ einer Schwingung des Moleküls ist *unabhängig von der Stelle des Gravitationsfeldes*, an welcher sich das Molekül befindet.

In der Tat ist unter obigen Voraussetzungen $g_{\mu\nu} = \text{const.}$ Man kann es also immer durch eine Transformation — bei welcher ja die Invariante τ ungeändert bleibt — erreichen, dass die neuen $g'_{\mu\nu} = \delta_{\mu\nu}$ werden. Dann herrschen aber — wo sich auch das Molekül befindet — immer die nämlichen Massverhältnisse der Pseudoeuklidischen Geometrie.

Dieses *Axiom hat* natürlich ebenso wie das weiter oben genannte, *nur provisorischen Charakter*. Wenn die neue Physik einmal vollständig ausgebaut sein wird, so muss dasselbe eine Folge der allgemeinen Theorie sein.

Nun fassen wir zwei Moleküle ins Auge, von denen eines auf der Sonneno-
165 berfläche in einer Entfernung x_0 vom | Zentrum, das andere auf der Erde in der praktisch unendlich grossen Entfernung X_0 sich befinden möge. Für das erstgenannte Molekül berechnen wir nun die Eigenzeit τ einer Periode. Weil der positive Kern ruht, wird $dx = 0$; weil die Geschwindigkeit des negativen Elektrons unendlich klein ist, wird auch $dy = 0$. Also erhalten wir, wenn P eine Periode auf der Sonnenoberfläche bezeichnet,

$$\tau = \int_P ds = \int_P \sqrt{\frac{x_0 - \alpha}{x_0}} dt = \sqrt{\frac{x_0 - \alpha}{x_0}} t = \left\{ 1 - \frac{1}{2} \frac{\alpha}{x_0} \right\} t.$$

Ebenso erhalten wir für das auf der Erde schwingende Molekül, wenn wir über eine Periode P^* integrieren und berücksichtigen, dass $\frac{X_0 - \alpha}{X_0}$ angenähert gleich 1 ist,

$$\tau = \int_{P^*} \sqrt{\frac{X_0 - \alpha}{X_0}} dt = T.$$

[150]"in der dazu senkrechten y-Richtung" was corrected from "in einer dazu senkrechten Richtung, etwa in der y-Richtung".

Nun sind *die Schwingungszeiten t* und *T proportional den Wellenlängen* bezw. λ und Λ. Also wird, weil die Eigenzeit τ einer Periode beidemal dieselbe war,

$$\lambda : \Lambda = 1 : \left(1 - \frac{1}{2}\frac{\alpha}{x_0}\right) \qquad \text{oder}$$

$$\frac{\lambda - \Lambda}{\lambda} = \frac{\Delta\lambda}{\lambda} = \frac{\alpha}{2x_0}.$$

Setzt man für α *die Sonnenmasse* und für x_0 den *Sonnenradius* ein so findet man[151]

$$\frac{\Delta\lambda}{\lambda} = \frac{3}{7}10^{-5}.$$

Nehmen wir noch an, dass die Lichtgeschwindigkeit an den einzelnen Stellen des Raumes sich zeitlich nicht ändert, dass wir also die Schwingung eines auf der Sonne schwingenden Moleküls unverändert von der Erde aus beobachten können, so folgt hieraus:

Die Spektrallinien eines lichtaussehenden Gases sind auf der Sonne nach Rot verschoben gegenüber den Linien, die dasselbe Gas auf der Erde emittiert, und zwar findet diese | *Verzögerung* der Schwingung als unmittelbare Folge 166
der Gravitation *am Ort der Schwingung* selbst statt und wird nicht erst dadurch erzeugt, dass das Licht der emittierenden Quelle von der Sonne zur Erde strahlt. Für das Bestehen dieses Effektes sprechen nach E. Freundlich spektrale Beobachtungen an Fixsternen bestimmter Typen; eine endgültige Prüfung aber steht noch aus.[152]

III. Kapitel

§ 82. Die elektrodynamischen Erscheinungen als Wirkungen der Gravitation

Der Fall der Anwesenheit von Materie in unserer vierdimensionalen Pseudogeometrie.

---

[151] On p. [158] (see p. 280 above), Hilbert had computed the value of $\alpha/2$ to be $1.45 \cdot 10^5$ cm, and had given the value of the solar diameter as $6.9 \cdot 10^{10}$ cm which is actually the value for the solar radius. The value for $\Delta\lambda/\lambda$ given by Einstein is $2 \cdot 10^{-6}$, see *Einstein 1907b*, p. 459, *Einstein 1911*, p. 905, and *Einstein 1914*, p. 1084. The discrepancy of a factor of 2 arises from the earlier confusion of solar diameter and radius.

[152] See *Freundlich 1915a* and *Freundlich 1915b*. Freundlich's claim that the observational data provided evidence for Einstein's gravitational red shift was disputed by Hugo von Seeliger and Hans Ludendorff, a dispute that contributed to Freundlich's difficulty in obtaining an academic position. When Einstein and Hilbert discussed plans to provide Freundlich a position at the Göttingen observatory in early 1916 (see note 141 on p. 278 above), it was suggested that Freundlich should also work on the observation of solar red shift. See *Hentschel 1997*, ch. 4, for a historical discussion of early attempts to find evidence for gravitational redshift and Freundlich's role in it. Astronomical evidence of gravitational red shift remained controversial for a long time. For the first unambiguous, terrestial confirmation of the gravitational red shift of spectral lines see *Pound and Rebka 1960*, and for a comprehensive discussion of the history of gravitational redshift, see *Hentschel 1998*.

ist der letzte Schritt, den wir tun müssen, um unsere Theorie so auszugestalten, dass sie keine Geometrie mehr ist, sondern eine theoretische Physik, die sich experimentell prüfen lässt. Dazu knüpfen wir wieder an die Untersuchungen an, die wir auf S. 104ff angestellt hatten.[153] Wir lassen also die drei auf S. 121 genannten Axiome,[154] insbesondere das ad hoc eingeführte Axiom, dass Unstetigkeiten der Massbestimmung mit dem Vorhandensein von Materie äquivalent sein sollen, wieder fallen und treiben *reine Kontinuumsphysik*. Wir sind damals auf Grund allgemeiner Betrachtungen über das Wesen der Materie zu dem Schluss gekommen, dass die physikalischen Grundgleichungen aus einem Variationsprinzip entspringen müssen, in welchem als Funktion unter dem Integralzeichen eine allgemeine Invariante H steht, deren Wahl wir uns vorbehielten. Wir sagten nur, dass sie als Argumente die 10 Gravitati-
167 onspotentiale $g_{\mu\nu}$ samt ihren ersten und | zweiten Ableitungen — wovon die letzteren nur linear auftreten sollen — und die 4 Komponenten q_1, q_2, q_3, q_4 des elektrodynamischen Vektors samt ihren ersten Ableitungen enthält. Wir haben oben in der Folge nur den Spezialfall, dass die ganze Welt frei von Materie ist, behandelt; nun tun wir also den noch fehlenden *letzten Schritt* und nehmen an, das die *Welt mit der Materie erfüllt* ist.

Schon bevor wir uns über die Wahl der Hamiltonschen Funktion H schlüssig machen, können wir eine *merkwürdige Folgerung* aus unserer Theorie ziehen: Aus unserem Variationsproblem folgen 14 Differentialgleichungen, nämlich durch Lagrangesche Differentiation nach den $g^{\mu\nu}$ die 10 Gravitationsgleichungen

$$\frac{\partial\sqrt{g}H}{\partial g^{\mu\nu}} - \sum_k \frac{\partial}{\partial x_k}\frac{\partial\sqrt{g}H}{\partial g_k^{\mu\nu}} + \sum_{kl}\frac{\partial^2}{\partial x_k \partial x_l}\frac{\partial\sqrt{g}H}{\partial g_{kl}^{\mu\nu}} = 0, \quad (\mu,\nu = 1,2,3,4), \quad (23)$$

die wir auch abgekürzt in der Form

$$[\sqrt{g}H]_{\mu\nu} = 0 \quad (23')$$

schreiben und durch Lagrangesche Differentiation nach den q_h die 4 verallgemeinerten Maxwellschen Gleichungen

$$\frac{\partial\sqrt{g}H}{\partial q_k} - \sum_k \frac{\partial}{\partial x_k}\frac{\partial\sqrt{g}H}{\partial q_{hk}} = 0, \qquad (h = 1,2,3,4), \quad (24)$$

die wir abgekürzt als

$$[\sqrt{g}H]_h = 0. \quad (24')$$

schreiben. *Von diesen 14 Gleichungen sind*, wie wir schon bemerkten (siehe S. 108),[155] *4 eine Folge der 10 übrigen*, d. h. es bestehen zwischen diesen 14 Differentialgleichungen 4 wirkliche Identitäten. Betrachtet man also die Gesamtheit der Lösungen von irgend 10 jener 14 Gleichungen, z. B. alle Lösungen

168 der Gleichungen (23), so enthalten dieselben noch 4 willkürliche Funk|tionen der 4 Veränderlichen x_1, x_2, x_3, x_4, da wir ja 14 unbekannte Funktionen und nur 10 Gleichungen zu ihrer Bestimmung haben. Diese 4 willkürlichen Funktionen werden nun durch die 4 noch nicht herangezogenen elektrodynamischen Gleichungen (24) *nicht* bestimmt, sondern es werden nur gewisse Lösungen der 10 Gravitationsgleichungen, die mit dem Problem selber nichts zu tun haben, durch die 4 letzten Gleichungen als ungültig ausgeschieden.

In diesem Sinne kann man also die 4 elektrodynamischen Gleichungen als Folge der Gravitationsgleichungen ansehen. Für den Fall, dass die Invariante H jene spezielle Gestalt hat, die wir ihr unten geben werden, habe ich diesen Satz in meiner ersten Mitteilung über die Grundlagen der Physik in den Göttinger Nachrichten (1915) bewiesen.[156] *Damit ist ein bekanntes, von Riemann stammendes Problem gelöst*, nämlich die Frage, inwiefern die elektrischen Erscheinungen mit der Gravitation zusammenhängen.

§ 83. Die vier Invarianten, aus denen die Funktion L zusammengesetzt werden muss

Unsere Aufgabe ist es nun, *die Hamiltonsche Funktion H* so *zu wählen*, dass die resultierenden physikalischen Grundgleichungen alle physikalischen Beobachtungen richtig beschreiben. Hierzu sind weitere Axiome erforderlich. Soll in der Grenze d. h. bei Abwesenheit von Materie, die Euklidische Geometrie herauskommen, so muss notwendig

$$H = K + L$$

gesetzt werden, wo K die Riemannsche Krümmungsinvariante ist, so dass wir nur über die Invariante L noch frei verfügen können. Die *einfachste Annahme* nun, die wir *über L* machen können, ist | die, dass dieselbe ausser den q_h und 169
deren ersten Ableitungen nur die $g_{\mu\nu}$ selbst, nicht aber deren Ableitungen enthält. Die $g_{\mu\nu}$ selber müssen allerdings sicher in L vorkommen, weil sich aus dem Vektor q_h und dessen ersten Ableitungen allein keine allgemeinen Invarianten bilden lassen. Diese Annahme ist nun aber eine ganz *gewaltige Vereinfachung*; denn aus diesen Argumenten q_h, q_{hk}, $g_{\mu\nu}$ kann man *nur 4 allgemeine Invarianten* zusammensetzen, also kann L selbst auch nur eine noch zu bestimmende Funktion dieser 4 Invarianten sein. Diese letzteren wollen wir nun angeben.[157]

[153] See p. 241 above.

[154] See p. 253 above.

[155] See p. 243 above.

[156] *Hilbert 1915*, p. 397, (this Volume, p. 30).

[157] Hilbert, in fact, missed a fifth invariant:

$$Q_3 \equiv \sum_{klmnop} M_{kl} M_{no} q_m q_p g^{kn} g^{mo} g^{lp}.$$

In den Bezeichnungen, die ich in meiner ersten Mitteilung verwendet habe, haben wir[158]

$$Q = \sum_{klmn} M_{mn} M_{lk} g^{mk} g^{nl}. \tag{1}$$

Dabei bezeichnet $M_{lk} = q_{kl} - q_{lk}$ eine Komponente eines schief symmetrischen Tensors, des sogenannten elektromagnetischen Sechservektors. Seinen Komponenten M_{23}, M_{31}, M_{12}, M_{14}, M_{24}, M_{34} entsprechen in der dreidimensionalen Schreibweise bezw. $\mathfrak{H}_x$, $\mathfrak{H}_y$, $\mathfrak{H}_z$, -i$\mathfrak{E}_x$, -i$\mathfrak{E}_y$, -i$\mathfrak{E}_z$, wobei mit $\mathfrak{H}$ und $\mathfrak{E}$ die magnetische und elektrische Feldstärke bezeichnet wurde. In dieser Invariante — und ebenso in den folgenden — kommen *die Ableitungen* der elektrodynamischen Potentiale *nur in den Verbindungen* $q_{kl} - q_{lk}$ vor; dies ist ein ganz *merkwürdiger Umstand*. Die Vektorkomponenten q_h selbst treten in dieser Invariante nicht auf.

$$q = \sum_{kl} q_k q_l g^{kl} \tag{2}$$

170 Dies ist die einzige Invariante, in der keine Ableitungen der| $g_{\mu\nu}$ vorkommen.[159]

$$Q_1 = \sum_{mnkl} M_{mn} M^*_{lk} g^{mk} g^{nl}, \tag{3}$$

wobei mit M^* der zu M *duale Sechservektor* bezeichnet wird, der entsteht, wenn man in M die drei ersten Komponenten der Reihe nach an die Stelle der drei zweiten setzt und umgekehrt die drei zweiten Komponenten an erster Stelle schreibt. Man kann diese Regel dadurch zum Ausdruck bringen, dass man für

$$M^*_{lm} = \frac{1}{2} \sum_{\alpha\beta} G^{\alpha\beta}_{lm} M_{\alpha\beta}$$

schreibt, wobei

$$G^{\alpha\beta}_{lm} = \sum_{\mu\nu} \frac{1}{\sqrt{g}} \delta_{\alpha\beta\mu\nu} g_{\mu l} g_{\mu m}$$

bedeutet. Dabei hat $\delta_{\alpha\beta\mu\nu}$ den Wert $+1$ bezw. -1, je nachdem man von 1234 zu $\alpha\beta\mu\nu$ durch eine gerade oder ungerade Anzahl von Vertauschungen gelangt. Die Invariante 3) gleicht der ersten insofern, als hier wieder nur Ableitungen der q_h, aber nicht diese Grössen selbst auftreten.

$$Q_2 = \sum_{lk} B_k B_l g^{kl}, \qquad \text{wobei der Vektor} \tag{4}$$

$$B_k = \sum_{mn} M^*_{nk} q_m g^{mn}$$

In *Pauli 1921*, §. 64, it is also claimed that there exist only four invariants but, in addition to Q_3, only Hilbert's Q, q, and Q_1 were listed. In the supplementary notes to the English edition of 1958, Pauli pointed out that there exist, in fact, five invariants, and that he had missed the one that Hilbert here calls Q_2 (*Pauli 1958*, p. 223)

[158] Cp. *Hilbert 1915*, p. 407, (this Volume, p. 44 above).

[159] "$g_{\mu\nu}$" should be "q_h".

ist. Dies ist die *einzige Invariante, die sowohl* q_h *als auch* q_{hk} *enthält*. Beide Grössen treten ausserdem in ihr — wie auch in den vorher genannten — quadratisch auf.

§ 84. Wahl der Hamiltonschen Funktion L

Die *Invariante* L muss nun eine Funktion sein, die Q, q, Q_1 und Q_2 als Argumente enthält.[160] Die einfachste Annahme ist nun die, Q_1 und Q_2 überhaupt wegzulassen und L nur aus Q und q zusammenzusetzen. Dann erhält man nämlich aus der Theorie die bekannten Maxwellschen Gleichungen. Diesen Fall werden wir auch unten allein betrachten; doch stellt er meiner Ansicht nach kein Definitivum dar. Vielmehr wird man | wohl, um die durch die 171
Quantentheorie bedingten *an der heutigen Elektrodynamik anzubringenden Abänderungen* aus der Theorie zu erhalten, noch mindestens eine der beiden Invarianten Q_1und Q_2 hinzuziehen müssen. Dazu scheint mir gerade Q_2 besonders gut geeignet, weil in dieser Invariante *gleichzeitig* q_k *und* q_{hk} *auftreten*. Diese Invariante liefert nämlich[161] nur an den Stellen, wo q_h von Null verschieden ist — d. h. in der Nähe von elektrischen Massen — einen merklichen Beitrag zu den Feldgleichungen. Da die Maxwellschen Gleichungen aber gerade in der Nähe solcher elektrischer Massen versagen, so kann man durch die Hinzunahme von Q_2 möglicherweise die gesuchte Verbesserung dieser Gleichungen erzielen.
Wenn man *axiomatisch* festsetzt, dass in L die Ableitungen q_{hk} nur quadratisch vorkommen sollen, so dürfen in L die *Invarianten* Q, Q_1, Q_2[162] nur *linear* auftreten. Einzig die Wahl der Funktion, welche q als Argument enthält, wollen wir noch offen lassen, weil in q keine Ableitungen der q_h vorkommen. Diese Funktion muss späterhin durch Axiome festgelegt werden. Einstweilen schreiben wir also

$$L = aQ + a_1Q_1 + a_2Q_2 + f(q).$$

Die aus dieser Form der Invarianten L sich ergebende Theorie wollen wir nicht weiter verfolgen, sondern machen nun die *vereinfachende Annahme* — die vielleicht *nur für grosse Entfernungen von elektrischen Massen zulässig* ist — $L = aQ + f(q)$. Setzen wir noch $a = -\frac{1}{2}$, so wird die Hamiltonsche Funktion zu

$$H = K - \frac{1}{2}Q + f(q). \tag{25}$$

Unter Benutzung unserer abkürzenden Bezeichnungsweise für Lagrangesche 172

[160] As pointed out in *Pauli 1921*, § 64, the lack of gauge invariance poses a severe problem for Mie's theory. Of the five possible invariants, only Q and Q_1 are gauge invariant. Moreover, Q_1 is actually a pseudo-scalar and changes sign under parity transformations. Hence, only it's square would be admissible, see *Pauli 1958*, p. 223.

[161] "nämlich" was corrected from "also".

[162] "Q, Q_1, Q_2" was corrected from Q, q, Q_1, Q_2".

Ableitungen erhalten wir dann *die 10 Gravitationsgleichungen*

$$[\sqrt{g}K]_{\mu\nu} + \frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}} = 0, \qquad (\mu, \nu = 1, 2, 3, 4) \tag{26}$$

und wenn wir die auf S. 110 erwähnte Umformung[163]

$$[\sqrt{g}K]_{\mu\nu} = \sqrt{g}(K_{\mu\nu} - \frac{1}{2}Kg_{\mu\nu})$$

verwenden, so können wir diese Gleichung auch schreiben

$$K_{\mu\nu} - \frac{1}{2}Kg_{\mu\nu} = -\frac{1}{\sqrt{g}}\frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}}, \qquad (\mu, \nu = 1, 2, 3, 4), \tag{26'}$$

wobei die rechten Seiten noch auszurechnen sind. Dies werden wir erst weiter unten tun. Hierzu kommen nun noch *die 4 verallgemeinerten Maxwellschen Gleichungen*

$$[\sqrt{g}L]_h = 0, \qquad (h = 1, 2, 3, 4). \tag{27}$$

Da zwischen den 14 Gleichungen (26) und (27) 4 Identitäten bestehen, so lassen sich also die 4 Maxwellschen Gleichungen als Folge der 10 Gravitationsgleichungen darstellen, und zwar sind *bei unserer speziellen Wahl der Hamiltonschen Funktion H* die Gleichungen (26) in Bezug auf die 4 Funktionen q_h *Integrale der Gleichungen (27)*. In der Tat enthalten die Gravitationsgleichungen nur die Grössen M_{lk} — d.h. $\mathfrak{E}$ und $\mathfrak{H}$ — nicht aber deren Ableitungen, während in den Maxwellschen Gleichungen auch noch diese 4 Ableitungen vorkommen. Man gelangt also von den 4 letzten Gleichungen durch Integration zu den 10 ersten.

§ 85. Die Maxwellschen Gleichungen als Folge für $g_{\mu\nu} = \delta_{\mu\nu}$

Wir wollen nun zuerst zeigen, dass *für $g_{\mu\nu} = \delta_{\mu\nu}$ die Gleichungen (27) in*
173 *die gewöhnlichen Maxwellschen Gleichun|gen übergehen*. Hierauf müssen wir untersuchen, was aus den 10 Gravitationsgleichungen wird, wenn man $g_{\mu\nu} = \delta_{\mu\nu}$ einsetzt. Da in den Gleichungen (27) nicht nach den $g^{\mu\nu}$ differenziert wird, so dürfen wir von vornherein $g_{\mu\nu} = \delta_{\mu\nu}$ in dieselben einsetzen. Dann wird

$$Q = -\sum_{kl} M_{kl}^2 \qquad \text{und} \qquad q = \sum_h q_h^2.$$

Lassen wir k nur die Werte $k > l$ durchlaufen, so schreibt sich $L = \sum_{k>l} M_{kl}^2 + f(q)$. Durch Lagrangesche Differentiation nach q_h erhalten wir hieraus

$$\sum_k \frac{\partial}{\partial x_k}\frac{\partial L}{\partial q_{hk}} - \frac{\partial L}{\partial q_h} = 0, \qquad (h = 1, 2, 3, 4).$$

[163]See p. 245 above.

Nun ist $\frac{\partial L}{\partial q_{hk}} = \frac{\partial}{\partial q_{hk}} \sum_{m>n} M_{mn}^2 = -2M_{hk}$, also wird $\frac{\partial L}{\partial q_{hk}} = 2M_{kh}$. Hieraus folgt

$$\sum_k \frac{\partial}{\partial x_k} \frac{\partial L}{\partial q_{hk}} = 2 \sum_k \frac{\partial}{\partial x_k} M_{kh} = 2(\operatorname{Div} M)_h.$$

Ferner ist $\frac{\partial L}{\partial q_h} = \frac{\partial f}{\partial q_h}$. Setzt man also $\frac{\partial f}{\partial q_h} = -2r_h$, wobei die r_h die *Vererdichte der Elektrizität* sein soll, so werden die Lagrangeschen Gleichungen zu

$$(\operatorname{Div} M)_h = -r_h, \qquad (h = 1, 2, 3, 4). \tag{28}$$

Dies ist aber die *eine Gruppe von vier Maxwellschen Gleichungen*, die man so in Worte fassen kann:

„*die Weltdivergenz des elektromagnetischen Sechservektors ist gleich der negativen elektrischen Dichte*.“

Nun lautet ein bekannter Satz der Vektoranalysis, den man leicht verifizieren kann, $\operatorname{Div} \operatorname{Div} M = 0$, also folgt aus (28)

$$\operatorname{Div} r = 0. \tag{29}$$

Dies ist die sogenannte *Kontinuitätsgleichung*, d. h. die Bedingung dafür, dass 174
der Vektor r_h eine Dichte sein kann. Diese Bedingung besagt, dass in jedem Raumpunkt so viel Elektrizität hinein- wie hinausströmt.

Bezeichnen wir ferner mit $(\operatorname{curl} q)^*$ den zu $\operatorname{curl} q$ dualen Vektor, so lautet eine andere Gleichung der Vektoranalysis $\operatorname{Div}(\operatorname{curl} q)^* = 0$, also folgt aus (28) noch

$$\operatorname{Div} M^* = 0. \tag{30}$$

Dies ist die *die zweite Gruppe von Maxwellschen Gleichungen*.

§ 86. Die Maxwellschen Gleichungen als Ausgangspunkt der Relativitätstheorie

Die Maxwellschen Gleichungen sind bekanntlich der Ausgangspunkt und die *Quelle aller Relativitätstheorie* gewesen; und es ist interessant, sich dieser merkwürdigen historischen Entwicklung zu erinnern, um einzusehen, welch ungeheuren Umweg die Wissenschaft machen musste, um uns den Gedanken der allgemeinen Relativität aufzuzwingen. Zuerst stoppelte Maxwell aus den bekannten elektrodynamischen Gesetzen jene merkwürdigen nach ihm benannten Gleichungen zusammen. Dann brachte man sie auf eine einfachere und durchsichtigere Gestalt, indem man die Zeit als gleichberechtigte Veränderliche neben die drei Raumkoordinaten setzte. Hierauf zeigten Lorentz und Einstein die orthogonale Invarianz dieser Gleichungen, d. h. die Invarianz gegenüber der Lorentztransformation ($g_{\mu\nu} = \delta_{\mu\nu}$). Der nächste Schritt vorwärts war Einsteins *kleine* Relativitätstheorie; und schliesslich sagte sich Einstein

„wenn schon, denn schon“, und schuf so seine *grosse* Relativitätstheorie. Hand in Hand damit gingen Experimente, deren Bedeutung von den | Experimenta- 175
toren selbst gar nicht erkannt wurde; denn sie sahen nicht, dass diese Experimente dazu *zwingen*, das allgemeine Relativitätsprinzip anzuerkennen. — Die experimentellen Untersuchungen wiederum wurden erst dadurch ermöglicht, dass man lernte, bessere Glasplatten zu schleifen.

§ 87. Der Energietensor $T_{\mu\nu}$

Was wird nun aus den 10 Gravitationsgleichungen (26), *wenn man in dieselben* $g_{\mu\nu} = \delta_{\mu\nu}$ *einsetzt?* Da wir jetzt nach den $g^{\mu\nu}$ differenzieren, so müssen wir zuerst $\frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}}$ berechnen und dürfen erst danach den $g_{\mu\nu}$ die speziellen Werte $\delta_{\mu\nu}$ erteilen. Durch einfaches Ausrechnen erhalten wir die Identität[164]

$$\sum_\mu \frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}} g^{\mu\sigma} = L\delta_\nu^\sigma - \sum_k \frac{\partial L}{\partial M_{\sigma k}} M_{\nu k} - \frac{\partial L}{\partial q_\sigma} q_\nu.$$

Nun setzen wir $g_{\mu\nu} = \delta_{\mu\nu}$, dann geht dieselbe[165] über in

$$\left(\frac{\partial\sqrt{g}L}{\partial g^{\nu\sigma}}\right)_{g_{\mu\nu}=\delta_{\mu\nu}} = \left(-\frac{1}{2}Q + f\right)\delta_{\nu\sigma} - 2\sum_k M_{\sigma k} M_{\nu k} - \frac{\partial f}{\partial q_\sigma} q_\nu,$$

also erhalten wir für die rechte Seite von (26) auch[166]

$$\left(\frac{1}{\sqrt{g}}\frac{\partial\sqrt{g}L}{\partial g^{\nu\sigma}}\right)_{g_{\mu\nu}=\delta_{\mu\nu}} = \left(-\frac{1}{2}Q + f\right)\delta_{\nu\sigma} - 2\sum_k M_{\sigma k} M_{\nu k} - \frac{\partial f}{\partial q_\sigma} q_\nu. \qquad (31)$$

Rechter Hand steht aber nichts anderes als *der bekannte Sechzehnertensor der alten Elektrodynamik*, dessen Komponenten die *Maxwellschen Spannungen* und *der Poyntingsche Vektor* bilden.
Wir werden nun den *Tensor*

$$\frac{1}{\sqrt{g}}\frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}} = T_{\mu\nu} \qquad (32)$$

immer als Energie ansprechen, wenn auch $g_{\mu\nu} \neq \delta_{\mu\nu}$ ist; doch werden wir die Berechtigung dieser Annahme erst weiter unten bei der Besprechung des Impulsenergiesatzes einsehen können.
176 Unsere 10 Gravitationsgleichungen (26) haben mithin | für allgemeine $g_{\mu\nu}$ folgende *merkwürdige Gestalt und Bedeutung*

$$K_{\mu\nu} - \frac{1}{2}K g_{\mu\nu} = T_{\mu\nu}, \qquad (33)$$

in Worten:

[164]“Identität” was corrected from “Intensität”.
[165]“dieselbe” was corrected from “die linke Seite”.
[166]“(26)” should be “(26’)”; “auch” was added.

$$\textit{Krümmungstensor} \; -\tfrac{1}{2}Kg_{\mu\nu} = \textit{Energietensor}.$$

Einstein drückt dies auch so aus: Die Gleichungen (33) besagen, dass die drei Begriffe *Energie*, *träge Masse* und *schwere Masse ein und dasselbe* sind.

Es ist durchaus bemerkenswert, dass sich der Energietensor so verblüffend einfach aus dem zweiten Term L der Hamiltonschen Funktion ergibt als $\frac{1}{\sqrt{g}}\frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}}$. Man ersieht aus dieser Form sofort, dass die Energie ein symmetrischer Tensor ist — es ist ja $g^{\mu\nu} = g^{\nu\mu}$ —, während dies in der alten Elektrodynamik, in der man keine $g^{\mu\nu}$ kennt und also auch nicht danach differenzieren kann, mühsam bewiesen wird und als ein merkwürdiger Satz gilt, aus dem dann folgt, dass die Energie träge Masse besitzt. Hier erscheint die Energie direkt als Gegenstück zum Gravitationspotential: *Die Ableitungen von $\sqrt{g}L$ nach den Gravitationspotentialen geben den Energietensor $T_{\mu\nu}$*, gewiss eine wunderbar sich ineinander kettende Tatsachenreihe.

Multiplizieren wir die Gleichungen (33) mit bez. $g^{\mu\nu}$ und summieren über μ und ν, so erhalten wir linker Hand (siehe S. 110)[167] $-K$ und rechter Hand $\sum_{\mu\nu} g^{\mu\nu}T_{\mu\nu} = T$. Also kann man diese Gleichungen auch so schreiben, dass linker Hand nur der Krümmungstensor steht[168]

$$K_{\mu\nu} = T_{\mu\nu} - \frac{1}{2}Tg_{\mu\nu}. \qquad (33')$$

177

§ 88. Die alte Elektrodynamik folgt nicht aus der neuen Theorie für $g_{\mu\nu} = \delta_{\mu\nu}$

Nun müssen wir auf unsere Frage zurückkommen: Was wird aus den Gravitationsgleichungen, wenn man in dieselben für $g_{\mu\nu}$ die Werte $\delta_{\mu\nu}$ und für q_h irgend ein Lösungssystem der Maxwellschen Gleichungen einsetzt. Wir ersehen sofort: Es wird

$$T_{\mu\nu} = \frac{1}{\sqrt{g}}\frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}} = 0,$$

d. h. *es muss der Energietensor verschwinden, damit die Gravitationsgleichungen erfüllt sind*. Dann dürfen aber die q_h nicht irgend welche Lösungen der Maxwellschen Gleichungen sein, wir müssen vielmehr — damit die Welt von Energie frei ist — die trivialen Lösungen der Maxwellschen Gleichungen $q_h = 0$ wählen. Mit diesen müssen wir in die Gravitationsgleichungen eingehen, damit dieselben durch $g_{\mu\nu} = \delta_{\mu\nu}$ zu erfüllen sind.

Wir erhalten also das *scheinbar paradoxe Resultat*, das *die alte Elektrodynamik für $g_{\mu\nu} = \delta_{\mu\nu}$ nicht als Grenzfall aus der neuen* entspringt. Tatsächlich

[167] See p. 245 above.

[168] In the following form, the gravitational field equations were presented by Einstein when he published them for the first time, see *Einstein 1915d*.

ist dieses Ergebnis aber im besten Einklang mit unserer Theorie; denn wir ersehen daraus, dass die Anwesenheit von Materie — d. h. $q_h \neq 0$ — immer ein Abweichen der Massbestimmung von derjenigen der Pseudoeuklidischen Geometrie bedingt. Dies ist aber gerade eine notwendige Ergänzung der Antwort, die wir auf die auf S. 109 aufgeworfene Frage:[169] Gilt die Euklidische Geometrie in der Physik? gaben. Wir konnten zeigen, dass diese Geometrie unter der Voraussetzung $q_h = 0$ und unter gewissen einschränkenden Voraussetzungen für die $g_{\mu\nu}$ (siehe S. 113)[170] tatsächlich gilt. Nunmehr sehen wir, dass die Annahme $q_h = 0$ durchaus notwendig war.

178 § 89. Näherungsweise Integration

Um ein Lösungssystem der 14 physikalischen Grundgleichungen zu erhalten, in welchen die q_h von Null verschiedene Werte haben, dürfen wir den $g_{\mu\nu}$ nicht *genau* die Werte $\delta_{\mu\nu}$ erteilen; doch können die $g_{\mu\nu}$ noch immer *angenähert* diese Werte annehmen, wie wir ersehen, wenn wir nun

die angenäherte Integration der Feldgleichungen

behandeln. Dazu haben wir nur eine *kleine Modifikation in unserem Ansatz* der Hamiltonschen Funktion anzubringen oder besser: Wir haben eine andere Auffassung zum Ausdruck zu bringen. Wir ersetzen nämlich L durch ϵL, so dass

$$H = K + \epsilon L \tag{34}$$

wird. Den Parameter ϵ denken wir uns klein, damit wir die Lösungen der Differentialgleichungen nach steigenden Potenzen desselben entwickeln können. Dieser Ansatz ist durchaus *keine Einschränkung der Allgemeinheit*; durch denselben bringen wir vielmehr zum Ausdruck, dass wir eine bestimmte Integrationsmethode anwenden wollen. Dass wir damit das Richtige getroffen haben, wird der Erfolg lehren.

Wir setzen nun die Lösungen $g_{\mu\nu}$ und q_h in der Form an

$$g_{\mu\nu} = g_{\mu\nu}^{(0)} + \epsilon g_{\mu\nu}^{(1)} + \epsilon^2 g_{\mu\nu}^{(2)} + \cdots , \qquad q_h = q_h^{(0)} + \epsilon q_h^{(1)} + \epsilon^2 q_h^{(2)} + \cdots .$$

Für $\epsilon = 0$ haben wir nur die 10 Gleichungen

$$[\sqrt{g}K]_{\mu\nu} = 0$$

zu erfüllen. Ein Lösungssystem derselben — vielleicht sogar das einzige, wenn wir absolute Regularität verlangen — ist $\delta_{\mu\nu}$. Wir haben daher für $g_{\mu\nu}$ die Entwicklung

$$g_{\mu\nu} = \delta_{\mu\nu} + \epsilon h_{\mu\nu} + \epsilon^2 g_{\mu\nu}^{(2)} + \cdots ,$$

179 in welcher der Koeffizient von ϵ, den wir auch noch berechnen wollen, zur Abkürzung mit $h_{\mu\nu}$ bezeichnet wurde. Jedes physikalische Problem läuft jetzt darauf hinaus, die weiteren Glieder in dieser Reihe sowie in derjenigen für q_h zu finden. *Setzen wir also den Koeffizienten von ϵ gleich Null*, so erhalten wir aus den 4 elektrodynamischen Gleichungen für das Verschwinden dieses Koeffizienten die 4 Bedingungsgleichungen

$$([\sqrt{g}L]_h)_{g_{\mu\nu}=\delta_{\mu\nu}} = 0, \qquad (h = 1, 2, 3, 4). \tag{35}$$

Dies sind aber, wie wir eben gesehen haben, *genau die alten Maxwellschen Gleichungen*.[171]

Wir bezeichnen ein Lösungssystem derselben mit $q_h^{(0)}$. Dieses haben wir nun einzusetzen in die *erste Näherung der Gravitationsgleichungen*, die durch Nullsetzen des Koeffizienten von ϵ entsteht, dann lauten dieselben

$$\left\{([\sqrt{g}K]_{\mu\nu})_{g_{\mu\nu}=\delta_{\mu\nu}+\epsilon h_{\mu\nu}}\right\}_{(\epsilon^1)} = -\left(\frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}}\right)_{g_{\mu\nu}=\delta_{\mu\nu}}. \tag{36}$$

Hierbei ist linker Hand, weil K — im Gegensatz zu L — nicht mit ϵ multipliziert ist, für $g_{\mu\nu}$ die *erste Annäherung* $\delta_{\mu\nu} + \epsilon h_{\mu\nu}$ einzusetzen und dann sind die Glieder mit ϵ zu sammeln, während rechter Hand die *nullte Annäherung* $\delta_{\mu\nu}$ schon den Koeffizienten von ϵ liefert. Rechts steht jetzt schon wieder der Energietensor der alten Elektrodynamik. Aber da derselbe nicht verschwinden muss, ist alles in bester Ordnung; und die 14 Gleichungen haben durch diese Entwicklung nach steigenden Potenzen von ϵ ein ganz anderes Aussehen bekommen.

Unsere Gravitationstheorie steht also | durchaus *nicht im Widerspruch zur* 180
alten Physik. Sie ist vielmehr ein *Schlüssel zum Verständnis* der letzteren; insbesondere wird die Bedeutung der Energie erst jetzt durchsichtig. Die 10 Gravitationsgleichungen[172] und die 4 elektrodynamischen Gleichungen geben in nullter[173] Näherung genau die Erfahrungstatsachen wieder, nämlich: Die *Pseudoeuklidische Geometrie und die gewöhnlichen Maxwellschen Gleichungen*. Es erscheint also hier die Pseudoeuklidische Geometrie als eine Annäherung an die wirkliche Physik, und dies ist es gerade, was die Entwicklung nach ϵ so schön und elegant macht. Man kann diese Integrationsmethode recht eigentlich als die „*Erschaffung der Welt aus dem Nichts*" bezeichnen; denn das Nichts — d.h. $\epsilon = 0$ — ist nach unserer Auffassung eben die Pseudoeuklidische Geometrie.

---

[169] See p. 244 above.

[170] See p. 247 above.

[171] Strictly speaking, Hilbert's assertion is true only for $f(q) \equiv 0$.

[172] Deleted: "geben in nullter".

[173] "nullter" was corrected from "erster".

§ 90. Die erste Annäherung

Um die Gleichungen (36) weiter zu behandeln, stützen wir uns auf Rechnungen, die Einstein (Näherungsweise Integration der Feldgleichungen der Gravitation, Sitzungsber. d. Preuss. Akad. d. Wiss. (32) 1916 S. 688ff)[174] durchgeführt hat. Wir setzen mit Einstein

$$h_{\mu\nu} = k_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}\sum_s k_{ss}, \tag{37}$$

wobei $k_{\mu\nu} = k_{\nu\mu}$ sein soll. Einstein nimmt weiter an, dass die $k_{\mu\nu}$ den vier Bedingungen

$$\sum_s \frac{\partial}{\partial x_s} k_{\mu s} \qquad (\mu = 1, 2, 3, 4) \tag{38}$$

genügen mögen. Dies involviert natürlich auch *vier Bedingungen für die Funktionen $h_{\mu\nu}$*. Man kann aber den Beweis erbringen, dass diese Bedingungen
181 keine Einschränkung der Allgemein|heit bedeuten (Hilbert, Die Grundlagen der Physik, 2 Mitteil. Gött.Nachr. 1916).[175] Es lässt sich nämlich durch eine *infinitesimale Transformation*

$$x'_s = x_s + \varepsilon\varphi_s(x_1, x_2, x_3, x_4)$$

— wobei das ε wieder jenes ist, nach welchem wir entwickeln — immer erzwingen, dass bei geeigneter Wahl der vier willkürlichen Funktionen φ_s die neuen $k'_{\mu\nu}$ die vier Bedingungen (38) erfüllen. Sind aber diese vier Gleichungen erfüllt, so erhält man nach Einstein (a.a.O.)

$$(K_{\mu\nu})_{g_{\mu\nu}=\delta_{\mu\nu}+\varepsilon h_{\mu\nu}} \equiv \varepsilon\Box(k_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}\sum_s k_{ss}), \quad (\varepsilon^2)$$

wobei $\Box = \sum_\iota \frac{\partial^2}{\partial x_\iota^2}$ bedeutet und wobei nach der in der Zahlentheorie üblichen Schreibweise mit $\equiv$, (ε^2) gemeint ist, dass nur Glieder mit ε^1 berücksichtigt sind. Diesen Ausdruck setzen wir in die Gravitationsgleichungen in der Gestalt (33') ein. In denselben schreiben wir noch für

$$T_{\mu\nu} = \varepsilon t_{\mu\nu} + \varepsilon^2 t^{(2)}_{\mu\nu} + \dots .$$

Hierin ist also, wie man aus (31) und (32) ersieht, $t_{\mu\nu}$ der Energietensor der alten Elektrodynamik. Wir erhalten nun als Koeffizient von ε

$$\Box\left(k_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}\sum_s k_{ss}\right) = t_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}\sum_s t_{ss}. \tag{39}$$

[174] *Einstein 1916b.*

[175] *Hilbert 1917*, pp. 65–66, (this Volume, pp. 59–61 above).

Setzt man hierin $\mu = \nu$ und summiert über μ, so wird

$$\Box\left(\sum_\mu k_{\mu\mu} - 2\sum_s k_{ss}\right) = \sum_\mu t_{\mu\mu} - 2\sum_s t_{ss}$$

oder

$$\Box\sum_\mu k_{\mu\mu} = \sum_\mu t_{\mu\mu}.$$

Unter Benutzung dieser Relation *schreibt sich (39) in der Form*

$$\Box k_{\mu\nu} = t_{\mu\nu}, \qquad (\mu, \nu = 1, 2, 3, 4). \tag{40}$$

Sind die $q_s^{(0)}$ aus (35) gefunden, so kann man aus | (31) die rechten Seiten von 182
(40) berechnen. Aus diesen Gleichungen entnimmt man dann die $k_{\mu\nu}\langle,\rangle$ hierauf aus (37) die $h_{\mu\nu}$ und erhält so *in erster Näherung die Gravitationspotentiale*

$$g_{\mu\nu} = \delta_{\mu\nu} + \varepsilon h_{\mu\nu}.$$

§ 91. Das Wesen der Gravitation

Die Gleichungen (40) liefern uns *wichtige Aufschlüsse über das Wesen der Gravitation*. In der Bezeichnungsweise der Potentialtheorie schreiben sich dieselben als

$$\Delta k_{\mu\nu} - \frac{\partial^2 k_{\mu\nu}}{\partial t^2} = t_{\mu\nu}. \tag{41}$$

Dies ist aber nichts anderes als die bekannte *Schwingungsgleichung* der Elektrodynamik. Als Integral derselben erhalten wir das *retardierte Potential*

$$k_{\mu\nu}(x, y, z, t) = \frac{1}{4\pi}\iiint \frac{t_{\mu\nu}(\xi, \eta, \zeta, t-r)}{r} d\xi d\eta d\zeta, \tag{42}$$

wobei $r = \sqrt{(x-\xi)^2 + (y-\eta)^2 + (x-\zeta)^2}$ bedeutet und über alle Stellen des Raumes zu integrieren ist, an denen $t_{\mu\nu} \neq 0$ ist. Nun kann man aber *alle Schlüsse, die in der Optik* an diese Differentialgleichung *geknüpft werden*, in dieser Näherung auch *auf die Gravitation übertragen*, insbesondere: Die *Gravitation pflanzt sich in Wellen fort*, die sich in erster Näherung geradlinig und gleichförmig mit Lichtgeschwindigkeit bewegen.

Dies ist aber die Beantwortung einer alten Frage von fundamentaler Bedeutung. Früher hatte man die Vorstellung, dass sich die Gravitation unendlich rasch fortpflanze. Als dann das kleine Relativitätsprinzip aufkam, geriet man mit dieser Ansicht auf Widersprüche gegen das Kausalitätsprinzip. Dann
nahm man an, dass sich die Gravitation mit Lichtgeschwindigkeit | fortpflan- 183
ze, aber dies war nichts weiter als eine leere Festsetzung, bis die neue Theorie über den Zusammenhang zwischen Licht und Gravitation Aufschluss gab.

Die der Gleichung (40) entsprechende Gleichung der Elektrodynamik ist bekanntlich

$$\Box q_s = -r_s$$

mit dem Integral

$$q_s = -\frac{1}{4\pi}\iiint \frac{r_s(\xi,\eta,\zeta,t-r)}{r} d\xi d\eta d\zeta,$$

wobei r_s die elektrische Viererdichte ist. *In der Einsteinschen Gravitationsgleichung* (40) steht also die elektromagnetische Energie an Stelle der elektrischen Dichte.

§ 92. Elektronentheorie

Um die *Elektronentheorie* aus unserer allgemeinen Theorie der Materie zu erschliessen, muss man in den Näherungen der Differentialgleichungen einen Schritt weitergehen. Um z. B. eine statische, um einen Punkt *symmetrische Elektrizitätsverteilung* zu erhalten, hat man

$$r_1 = 0, \quad r_2 = 0, \quad r_3 = 0, \quad r_4 = \varphi(r)$$

zu setzen. Nun hatten wir aber[176]

$$\frac{\partial f}{\partial q_s} = -2r_s,$$

also folgt für $g_{\mu\nu} = \delta_{\mu\nu}$

$$q_1 = 0, \quad q_2 = 0, \quad q_3 = 0.$$

Eine solche Lösung der elektrodynamischen Gleichung bedeutet, *dass die ganze Welt aus einem einzigen, dauernd ruhenden Elektron besteht*. Diese Untersuchung führt in erster Näherung auf die sehr schönen und interessanten Resultate von Mie (Ann.d.Phys 1912 (37) S.511, (39) S.1, 1913 (40) S.1).[177]

184 § 93. Hamiltonsches Prinzip in der Mechanik

Als Abschluss dieses Kollegs wollen wir

einige Betrachtungen über den Impuls-Energiesatz

[176] See p. 291 above.

[177] *Mie 1912a*; *Mie 1912b*; *Mie 1913*. For English translations of these papers, see *Renn and Schemmel 2007a*; for historical discussion of Mie's theory and Hilbert's reception of it, see *Kohl 2000*, *Corry 2004*, ch. 6, and *Smeenk and Martin 2007*.

anstellen und zwar möchte ich Ihnen eine besondere einfache Ableitung dieses mit dem Hamiltonschen Prinzip aufs engste zusammenhängenden Satzes geben.

Um die leitende Idee, die unseren Betrachtungen zu Grunde liegt, zu erfassen, machen wir uns die Verhältnisse erst an einem *besonders einfachen Hamiltonschen Prinzip*, nämlich an demjenigen, aus welchem in der Mechanik die *Gesetze der Bewegung eines Massenpunktes* in einem Kraftfeld abgeleitet werden. Bezeichnet man mit T die kinetische Energie und mit U die potentielle Energie des Massenpunktes und setzt $H = T - U$, so lautet dieses Hamiltonsche Prinzip

$$\delta \int H dt = 0, \tag{43}$$

wobei t die Zeit bedeutet. Hierin ist U eine beliebige Funktion von x, y, z allein und T eine Funktion von x, y, z ausserdem eine homogene quadratische Funktion von $\frac{dx}{dt}$, $\frac{dy}{dt}$, $\frac{dz}{dt}$, sodass H eine Funktion der 6 Argumente x, y, z, $\frac{dx}{dt}$, $\frac{dy}{dt}$, $\frac{dz}{dt}$ wird. *Der Parameter t tritt also unter dem Integral nicht explizit auf*, sondern nur die Ableitungen von x, y, z nach diesem Parameter. Eine solche nicht explizit vorkommende Variable heisst nach Helmholtz eine *zyklische Koordinate*.[178]

Bekanntlich ist nun *ein Integral* dieses Hamiltonschen Prinzip (43) *die Gesamtenergie*

$$T + U = \text{const.}$$

Dies liegt aber gerade an dem Auftreten einer zyklischen Koor|dinate. Allgemein gilt folgender leicht zu verifizierende 185

Satz : Enthält die Funktion H unter dem Integralzeichen des Hamiltonschen Prinzips die angegebenen Argumente, so bilde man die sogenannte *Legendresche Funktion*

$$F = H - \dot{x}\frac{\partial H}{\partial \dot{x}} - \dot{y}\frac{\partial H}{\partial \dot{y}} - \dot{z}\frac{\partial H}{\partial \dot{z}}; \qquad (\dot{x} = \frac{dx}{dt} \quad \text{u. s. w.})$$

dann ist die Gleichung

$$\frac{d}{dt}F = 0$$

eine Folge der drei aus dem Variationsprinzip fliessenden *Lagrangeschen Differentialgleichungen*. Hieraus folgt nun $F = \text{const.}$ Da T eine homogene quadratische Funktion von $\frac{dx}{dt}$, $\frac{dy}{dt}$, $\frac{dz}{dt}$ ist und U diese Variabelen überhaupt nicht enthält, so ist

$$\dot{x}\frac{\partial H}{\partial \dot{x}} + \dot{y}\frac{\partial H}{\partial \dot{y}} + \dot{z}\frac{\partial H}{\partial \dot{z}} = 2T, \quad \text{also wird} \quad F = T - U - 2T$$

oder

$$T + U = \text{const.} \qquad \text{w. z. b. w.}$$

[178] See *Helmholtz 1884a*, *Helmholtz 1884b*, see also *Hertz 1894*, §§ 546ff, and *Boltzmann 1904*, ch. IV. For a historical discussion of the notion of cyclic coordinates, see *Lützen 2005*, ch. 18.

§ 94. Die 4 Impulsenergiesätze der Physik

In der Physik liegen nun die Verhältnisse viel komplizierter; denn hier steht im Hamiltonschen Prinzip ein vierfaches Integral und die Funktion unter dem Integralzeichen enthält *die 4 zyklischen Koordinaten* x, y, z, t bez. x_1, x_2, x_3, x_4. Diese 4 Parameter treten in der Hamiltonschen Funktion nicht explizit auf, weil

die physikalischen Gesetze von Ort und Zeit unabhängig sind. In der Tat würde im entgegengesetzten Fall die Hamiltonsche Funktion in jedem Weltpunkt eine andere sein und infolgedessen wären die 14 das physikalische Geschehen bestimmenden Lagrangeschen Differentialgleichungen an jedem Ort und zu jeder Zeit verschiedene. Entsprechend den 4 zyklischen Koordinaten erhalten
186 wir auch 4 Differentialgleichungen an Stelle | der einen Gleichung (44). *Diese 4 Gleichungen können wir als unmittelbaren Ausdruck dafür ansehen, dass das physikalische Geschehen in jedem Weltpunkt denselben Gesetzen zu gehorchen hat.* Die Gleichungen haben — wie wir sehen werden — *Divergenzform* und lassen sich infolgedessen auch nicht so einfach integrieren wie (44).

§ 95. Abspalten eines Divergenzausdrucks

Um diese 4 Differentialgleichungen abzuleiten, hatte man früher komplizierte Rechnungen nötig. Auch wir werden jetzt bei der Ableitung derselben nicht ohne ein gewisses Mass von Rechnung auskommen; doch ist der Weg, den wir einschlagen klar und durchsichtig und entbehrt nie eines bestimmten leitenden Gedankens.

Wir gehen aus vom *Variationsprinzip*

$$\delta \iiiint (K+L)\sqrt{g}\, dW = 0. \tag{45}$$

Dasselbe *können wir* nun — freilich auf Kosten der allgemeinen Invarianz — *beträchtlich vereinfachen*. Da wir nämlich wissen, dass die zweiten Ableitungen $g^{\mu\nu}_{hk}$ der $g^{\mu\nu}$ in L gar nicht und in H nur *linear* auftreten, so können wir folgende *Umformung* vornehmen: Wie man aus Formel (2) S. 110 entnimmt,[179] stammen die Glieder $g^{\mu\nu}_{hk}$ in $K\sqrt{g}$ aus den beiden Ausdrücken

$$\Phi = \sum_{\mu\nu\lambda} \sqrt{g}\, g^{\mu\nu} \frac{\partial}{\partial x_\lambda} \left\{ \begin{matrix} \mu\nu \\ \lambda \end{matrix} \right\} \qquad \text{und} \quad \psi = \sum_{\mu\nu\lambda} \sqrt{g}\, g^{\mu\nu} \frac{\partial}{\partial x_\nu} \left\{ \begin{matrix} \mu\lambda \\ \lambda \end{matrix} \right\}.$$

Durch partielle Integration erhalten wir z. B. für Φ folgende Umformung

$$\Phi = \sum_{\lambda} \frac{\partial}{\partial x_\lambda} \left(\sqrt{g} \sum_{\mu\nu} g^{\mu\nu} \left\{ \begin{matrix} \mu\nu \\ \lambda \end{matrix} \right\} \right) - \sum_{\mu\nu\lambda} \left\{ \begin{matrix} \mu\nu \\ \lambda \end{matrix} \right\} \frac{\partial}{\partial x_\lambda} (\sqrt{g}\, g^{\mu\nu}).$$

[179]See p. 245 above.

Das erste Glied rechter Hand ist ein Ausdruck von *Divergenzcharakter*; einen solchen bezeichnen wir abgekürzt mit Div. | Dieselbe Umformung führen wir 187
auch mit ψ durch. Den beiden zweiten Gliedern in Φ und ψ kann man noch unter Benutzung von Formeln, die sich bei Einstein (Die Grundlage der allgemeinen Relativitätstheorie, Sonderabdr⟨uck⟩ a. d. Ann. d. Phys. 1916, S. 35)[180] finden, eine andere Gestalt geben und zwar erhält man wiederum nach Einstein (Hamiltonsches Prinzip und allgemeine Relativitätstheorie, Sitzungsber. d. Preuss. Akad. d. Wiss. 1916, (42) S. 1113)[181]

$$K\sqrt{g} = \sqrt{g} \sum_{\mu\nu\alpha\beta} g^{\mu\nu} \left[\left\{ \begin{matrix} \mu\alpha \\ \beta \end{matrix} \right\} \left\{ \begin{matrix} \beta\nu \\ \alpha \end{matrix} \right\} - \left\{ \begin{matrix} \mu\nu \\ \alpha \end{matrix} \right\} \left\{ \begin{matrix} \alpha\beta \\ \beta \end{matrix} \right\} \right] + \mathrm{Div}\,.$$

Wir bezeichnen

$$\mathfrak{K} = \sqrt{g} \sum_{\mu\nu\alpha\beta} g^{\mu\nu} \left[\left\{ \begin{matrix} \mu\alpha \\ \beta \end{matrix} \right\} \left\{ \begin{matrix} \beta\nu \\ \alpha \end{matrix} \right\} - \left\{ \begin{matrix} \mu\nu \\ \alpha \end{matrix} \right\} \left\{ \begin{matrix} \alpha\beta \\ \beta \end{matrix} \right\} \right]. \tag{46}$$

$\mathfrak{K}$ ist also ein homogener quadratischer Ausdruck in den $g_h^{\mu\nu}$ in gewisser Analogie z. B. zum Dirichletschen Integral oder zum Hamiltonschen Prinzip (43) der Mechanik, in welchen beiden die Grössen $\frac{dx_i}{dt}$ quadratisch auftreten.
Da wir von $K\sqrt{g}$ einen Ausdruck von Divergenzcharakter angespalten[182] haben, um $\mathfrak{K}$ zu erhalten, so ist $\frac{\mathfrak{K}}{\sqrt{g}}$ *keine allgemeine Invariante*. Dagegen kann man zeigen, dass $\frac{\mathfrak{K}}{\sqrt{g}}$ eine *projektive Invariante* der Koordinaten x_s ist; d. h. es ist zwar durch Weglassen des Ausdrucks Div die allgemeine, aber nicht die projektive Invarianz von K zerstört worden. Wir bezeichnen noch

$$\mathfrak{L} = L\sqrt{g}; \tag{46'}$$

dann schreibt sich also das Hamiltonsche Prinzip (45) in der Form

$$\delta \iiiint (\mathfrak{K} + \mathfrak{L} + Div)\, dW. \tag{47}$$

§ 96. Infinitesimale Transformation 188

Nun machen wir Gebrauch von der Tatsache, dass die Variablen x_s unter dem Integral (47) nicht explizit auftreten und zwar führen wir

eine infinitesimale Transformation

[180] *Einstein 1916a*.
[181] *Einstein 1916c*, p. 1113.
[182] "angespalten" should be "abgespalten".

an dem Integral

$$\iiiint K\sqrt{g}\,dW = \iiiint \mathfrak{K}\,dW + \iiiint Div\,dW \tag{48}$$

durch, indem wir

$$x_s \quad \text{durch} \quad x_s + \Delta x_s$$

ersetzen. Hierin bedeutet $\Delta x_s = \epsilon\varphi_s(x_1, x_2, x_3, x_4)$ eine Funktion von x_1, x_2, x_3, x_4, die nur sehr kleine Werte annimmt. Diese Transformation wird uns fast ohne Rechnung einige *sehr merkwürdige identische Gleichungen* erschliessen. Berechnen wir jetzt das Integral (48) für die transformierten Variablen und subtrahieren davon dieselbe Gleichung, gebildet für die ursprünglichen Veränderlichen, so wird die linke Seite gleich Null, weil hier die allgemeine Invariante $K\sqrt{g}\,dW$ steht, deren Wert ja bei der Transformation ungeändert bleibt. Auf der rechten Seite kann das vierfache Integral

$$\iiiint \text{Div}\ dW$$

mit Hilfe des *Gauss'schen Satzes* in ein dreifaches, genommen über die dreidimensionale Begrenzung des vierdimensionalen Weltstückes umgeformt werden. Betrachtet man nun nur solche Funktionen Δx_s, die auf dieser Begrenzung verschwinden, so ändert auch $\iiiint \text{Div}\ dW$ durch die Transformation seinen Wert nicht und hebt sich nach Subtraktion der beiden Gleichungen (48) von-
189 einander weg. Also folgt |

$$\Delta \iiiint \mathfrak{K}\,dW = \Delta \iiiint \left(\frac{\mathfrak{K}}{\sqrt{g}}\right)\sqrt{g}\,dW = 0, \tag{49}$$

wobei das Δ vor dem Integralzeichen bedeutet, dass die Differenz der Werte des Integrals $\iiiint \mathfrak{K}\,dW$ nach und vor der Transformation zu nehmen ist. Man bemerke übrigens, dass die Gleichung (49) gefunden wurde, *ohne dass vom Hamiltonschen Prinzip irgend wie Gebrauch gemacht wurde*. Nun ist $\sqrt{g}\,dW$ eine allgemeine Invariante, also wird $\Delta\sqrt{g}\,dW = 0$ und daher kann man (49) in der Form

$$\iiiint \left(\Delta\frac{\mathfrak{K}}{\sqrt{g}}\right)\sqrt{g}\,dW = 0 \tag{49'}$$

schreiben.

Um diese Gleichung weiter zu behandeln, müssen wir untersuchen, wie sich die Funktionen $g^{\mu\nu}$ ändern, wenn x_s durch $x_s + \Delta x_s$ ersetzt wird. Aus dem Transformationsgesetz der $g^{\mu\nu}$ bei allgemeinen Transformationen findet man in diesem Spezialfall (Einstein, Hamiltonsches Prinzip usw. a.a.O. S. 1114)[183]

$$\Delta g^{\mu\nu} = \sum_\alpha \left(g^{\mu\alpha}\frac{\partial\Delta x_\nu}{\partial x_\alpha} + g^{\nu\alpha}\frac{\partial\Delta x_\mu}{\partial x_\alpha}\right).$$

[183] The following equations are the same as *Einstein 1916c*, eqs. (11), (12).

Ebenso ergibt sich

$$\Delta g_{\sigma}^{\mu\nu} = \frac{\partial}{\partial x_\sigma}(\Delta g^{\mu\nu}) - \sum_{\alpha} g_{\alpha}^{\mu\nu} \frac{\partial \Delta x_\alpha}{\partial x_\sigma}.$$

Nun können wir $\Delta \frac{\mathfrak{K}}{\sqrt{g}}$ berechnen und zwar erhalten wir, wenn wir die Glieder zusammenfassen, je nachdem sie $\frac{\partial \Delta x_\sigma}{\partial x_\nu}$ oder $\frac{\partial^2 \Delta x_\sigma}{\partial x_\nu \partial x_\alpha}$ als Faktor enthalten, für alle Variationen und nicht nur für solche, die auf der Begrenzung verschwinden,

$$\Delta\left(\frac{\mathfrak{K}}{\sqrt{g}}\right) = \frac{1}{\sqrt{g}}\left\{\sum_{\sigma\nu} T_{\sigma}^{\nu} \frac{\partial \Delta x_\sigma}{\partial x_\nu} + 2\sum_{\mu\nu\alpha\sigma} \frac{\partial \mathfrak{K}}{\partial g_{\alpha}^{\mu\sigma}} g^{\mu\nu} \frac{\partial^2 \Delta x_\sigma}{\partial x_\nu \partial x_\alpha}\right\}, \tag{50}$$

wobei zur Abkürzung

$$T_{\sigma}^{\nu} = 2\sum_{\mu} \frac{\partial \mathfrak{K}}{\partial g^{\mu\sigma}} g^{\mu\nu} + 2\sum_{\mu\alpha} \frac{\partial \mathfrak{K}}{\partial g_{\alpha}^{\mu\sigma}} g_{\alpha}^{\mu\nu} + \mathfrak{K}\delta_{\sigma}^{\nu} - \sum_{\mu\alpha} \frac{\partial \mathfrak{K}}{\partial g_{\nu}^{\mu\alpha}} g_{\sigma}^{\mu\alpha} \tag{51}$$

gesetzt ist. Man sieht leicht, dass T_{σ}^{ν} in den $g_{\alpha}^{\mu\nu}$ homogen quadratisch ist. 190
Jetzt machen wir von der Eigenschaft Gebrauch, dass $\frac{\mathfrak{K}}{\sqrt{g}}$ eine projektive Invariante ist. Bei einer *linearen* infinitesimalen Transformation wird also die linke Seite von (50) zu Null. Auf der rechten Seite fallen die Glieder mit $\frac{\partial^2 \Delta x_\sigma}{\partial x_\nu \partial x_\alpha}$ weg und *es folgt*

$$T_{\sigma}^{\nu} = 0, \qquad (\sigma, \nu = 1, 2, 3, 4). \tag{52}$$

Diesen *16 Identitäten* muss also der Ausdruck $\mathfrak{K}$ genügen. Wir haben dieselben durch rein gedankliche Überlegungen mit einem Minimum von Rechnung gefunden. Vom Hamiltonschen Prinzip ist auch dabei noch nicht die Rede gewesen.
Infolge dieser Gleichungen kommen für beliebige Δx_s in (50) nur die Glieder mit $\frac{\partial^2 \Delta x_\sigma}{\partial x_\nu \partial x_\alpha}$ vor; und für alle Variationen, die auf dem Rande verschwinden, wird aus (49')

$$\iiiint \sum_{\mu\nu\alpha\sigma} g^{\mu\nu} \frac{\partial \mathfrak{K}}{\partial g_{\alpha}^{\mu\sigma}} \frac{\partial^2 \Delta x_\sigma}{\partial x_\nu \partial x_\alpha} dW = 0. \tag{53}$$

Nun formen wir den Ausdruck unter dem Integralzeichen, – wie immer in der Variationsrechnung in solchen Fällen – durch *zweimalige partielle Integration* um und erhalten, weil Δx_s im übrigen noch beliebig wählbar ist

$$\sum_{\mu\nu\alpha} \frac{\partial^2}{\partial x_\nu \partial x_\alpha}\left(g^{\mu\nu} \frac{\partial \mathfrak{K}}{\partial g_{\alpha}^{\mu\sigma}}\right) = 0, \qquad (\sigma = 1, 2, 3, 4). \tag{54}$$

Dies sind *4 weitere Identitäten*, welche $\mathfrak{K}$ erfüllen muss und die man auch durch blosses Ausrechnen hätte finden können. Dieselben sind *eine direkte Folge davon, dass die Variablen x_s in $\mathfrak{K}$ nicht explizit auftreten*. Wir hätten die Rechnung übrigens noch kürzer gestalten können, wenn wir nur diese 4 Gleichungen, nicht aber die 16 Gleichungen (52) hätten finden wollen, von 191
welchen wir unten noch Gebrauch zu machen haben.

§ 97. Der Energietensor S_σ^ν

Nun führen wir den folgenden Tensor

$$S_\sigma^\nu = \sum_{\mu\alpha} \frac{\partial}{\partial x_\alpha} \left(g^{\mu\nu} \frac{\partial \mathfrak{K}}{\partial g_\alpha^{\mu\sigma}} \right) \tag{55}$$

ein. Mit dessen Hilfe lassen sich die Gleichungen (54) schreiben als

$$\sum_\nu \frac{\partial}{\partial x_\nu} S_\sigma^\nu = 0, \qquad (\sigma = 1, 2, 3, 4). \tag{56}$$

In der gewöhnlichen Elektrodynamik ist bekanntlich *die Energie ein Sechzehnertensor, dessen Divergenz verschwindet*. Hier haben wir aber gerade 4 Divergenzgleichungen vor uns, *also müssen wir S_σ^ν als Energie ansprechen*.

Dies ist das eine wichtige Resultat unserer Rechnung. Das andere ist, dass *jede einzelne Komponente der Energie S_σ^ν nach (55) selbst wieder die Form einer Divergenz* hat. Dies ist natürlich die Folge der Gleichung (54) d. h. davon, dass eine Summe von zweiten Ableitungen zu Null wird. Diese merkwürdige *Divergenzform jeder Komponente des Energietensors* hat in der gewöhnlichen Elektrodynamik kein Gegenstück und wird erst durch die Gravitation verständlich. Es folgt aus dieser Gestalt der Energie, dass sich jedes Integral einer einzelnen Energiekomponente, erstreckt über ein Weltstück, in ein Oberflächenintegral verwandeln lässt.

§ 98. Einsetzen in das Hamiltonsche Prinzip

Bis hierher war vom Hamiltonschen Prinzip nicht die Rede. Deswegen haben wir auch noch keine

192 *Beziehungen zu unseren Weltgleichungen*

gefunden, das soll nun nachgeholt werden. Für alle an der Begrenzung verschwindenden Variationen hat das Hamiltonsche Prinzip (47) die Form

$$\delta \iiiint (\mathfrak{K} + \mathfrak{L})\, dW = 0. \tag{57}$$

Durch Bildung der Lagrangeschen Ableitungen nach den $g^{\mu\nu}$ erhalten wir die 10 Gravitationsgleichungen

$$\sum_\mu g^{\mu\sigma} \left| \sum_\alpha \frac{\partial}{\partial x_\alpha} \frac{\partial \mathfrak{K}}{\partial g_\alpha^{\mu\nu}} - \frac{\partial \mathfrak{K}}{\partial g^{\mu\nu}} = \frac{\partial \mathfrak{L}}{\partial g^{\mu\nu}} \right. , \qquad (\mu, \nu = 1, 2, 3, 4).$$

Wir multiplizieren dieselben bez. mit $g^{\mu\sigma}$ und summieren über μ. Hierauf formen wir das erste Glied durch Produktintegration um und erhalten wegen (55)

$$S_\nu^\sigma = \sum_{\alpha\mu} g_\alpha^{\mu\sigma} \frac{\partial \mathfrak{K}}{\partial g_\alpha^{\mu\nu}} - \sum_\mu g^{\mu\sigma} \frac{\partial \mathfrak{K}}{\partial g^{\mu\nu}} = \sum_\mu g^{\mu\sigma} \frac{\partial \mathfrak{L}}{\partial g^{\mu\nu}}. \tag{58}$$

Unter Benutzung der Lagrangeschen Gleichungen hat sich also die Energie so umformen lassen, dass die Ableitungen nach den Variablen x_s nicht mehr explizit auftreten. Die *Energie* S_ν^σ *zerfällt in die beiden Teile*

$$\mathfrak{G}_\nu^\sigma = \sum_\mu g^{\mu\sigma} \frac{\partial \mathfrak{K}}{\partial g^{\mu\nu}} + \sum_{\alpha\mu} g_\alpha^{\mu\sigma} \frac{\partial \mathfrak{K}}{\partial g_\alpha^{\mu\nu}}, \langle 184 \rangle$$

$$\mathfrak{E}_\nu^\sigma = \sum_\mu g^{\mu\sigma} \frac{\partial \mathfrak{L}}{\partial g^{\mu\nu}} + \sum_\mu g^{\mu\sigma} \frac{\partial \sqrt{g}\,\mathfrak{L}}{\partial g^{\mu\nu}},$$

von denen wir den ersteren als die *Energie des Gravitationsfeldes*, den letzteren als die *Energie des elektromagnetischen Feldes* ansprechen. Diese geht — wie aus Gleichung (31) folgt[185] —, für $g_{\mu\nu} = \delta_{\mu\nu}$ in den bekannten Energietensor der alten Elektrodynamik über. Der kovariante Tensor (32) folgt aus dem gemischten Tensor $\mathfrak{E}_\nu^\sigma$ als $\tau_{\mu\nu} = \frac{1}{\sqrt{g}} \sum_p \mathfrak{E}^p g_{p\nu}$.[186] Der Gravitationsenergie kann man noch eine *einfachere Form* geben, wenn man (52) benutzt. Dann wird nämlich

$$\mathfrak{G}_\nu^\sigma = \frac{1}{2}\left(\sum_{\alpha\mu} g_\nu^{\mu\alpha} \frac{\partial \mathfrak{K}}{\partial g_\sigma^{\mu\alpha}} - \mathfrak{K}\delta_\nu^\sigma\right).$$

Die Energie hat nur im projektiven Sinn Tensorcharakter, da wir den streng 193
allgemeinen Charakter der Invarianz durch Abspaltung des Gliedes mit Divergenzcharakter im Hamiltonschen Prinzip haben fallen lassen. Wir sagten auf S. 108, dass nur invariante Aussagen einen physikalischen Sinn haben.[187] Die hier abgeleiteten Gleichungen sind nun nicht invariant. Aber sie haben trotzdem physikalischen Sinn; denn *sie gelten für jedes Bezugssystem*, d. h. die entsprechenden Gleichungen für ein anderes Koordinatensystem gelten auch und zwar als Folge der Gravitationsgleichungen. Diese Eigenschaft der Gleichungen ist aber offenbar etwas von Invarianz durchaus verschiedenes und die dahingehenden Behauptungen von Einstein sind daher unrichtig.[188]

§ 99. Ausbau der Theorie

Wir sind in diesem Kolleg soweit gekommen, dass wir die ganze allgemeine mathematische Grundlage der neuen Physik in allen wesentlichen Teilen

[184] After the equal sign, Hilbert added in pencil: "Vorzeichen?"

[185] See p. 292 above.

[186] The preceding sentence was interlineated.

[187] See p. 243 above.

[188] Probably a reference to Einstein's comments in the last paragraph of *Einstein 1916c*.

vollständig hergestellt haben. Der *weitere Ausbau der Theorie* hat nun darin zu bestehen, die Integration der physikalischen Grundgleichungen anzustreben. Dies wird in der Weise geschehen, wie ich es auf S. 183 angedeutet habe: Durch eine Entwicklung nach ε baut man die Elektrodynamik in die Pseudogeometrie hinein. Die zweite Annäherung der Grundgleichungen liefert die Elektronentheorie. Die experimentelle Physik ist heute viel weiter vorangeschritten als die theoretische und ich hoffe, dass gerade die vielen unaufgeklärten Fragen der Elektronentheorie nun ihre Lösung finden werden.

Description of the Text

Collection: Georg-August-Universität Göttingen, Mathematisches Institut, Inv. Nr. 16205k.

Size: Cover size 22.5 cm ×28.2 cm; page size approx. 22.0 cm ×27.7 cm

Cover Annotations: On the spine, in gold lettering, is the notation, 'Hilbert, Grundlagen // der // Physik // II. // 1916/17'

Composition: 7 signatures of 9 to 20 sheets each (double pages); in all 203 pages, inclusive front- and end-sheets.

Pagination: The title page is not numbered. The following pages are continuously numbered from 1 to 193. Between page 33 and 34, a page numbered 33a was inserted. Page 107 is missing. The table of contents follows after page 193 and is paginated from 194 to 200.

Original Title: On the title page: DIE GRUNDLAGEN DER PHYSIK / II. / Vorlesung / von / D. HILBERT / Göttingen / Wintersemester 1916/17.//Ausgearbeitet von / R. BÄR.

Text: Typewritten text with equations, mathematical signs and some corrections added in black ink by R. Bär who arranged the script. Drawings have been added in pencil. Verso sides of pages are left blank. Paragraph headings are written, often in abbreviated form, in the margin of the page. Hilbert added occasional corrections and additions in pencil. On the botton of page 21 a paper strip, which is about 4 centimeters high, has been glued over the page, and page 74 was stuck together using two different sheets of paper.

http://www.springer.com/978-3-540-20606-4

David Hilbert's Lectures on the Foundations of Physics
1915-1927
Relativity, Quantum Theory and Epistemology
Sauer, T.; Majer, U. (Eds.)
2009, XII, 690 p., Hardcover
ISBN: 978-3-540-20606-4