JÜRGEN RENN AND JOHN STACHEL

HILBERT'S FOUNDATION OF PHYSICS: FROM A THEORY OF EVERYTHING TO A CONSTITUENT OF GENERAL RELATIVITY

1. ON THE COMING INTO BEING AND FADING AWAY OF AN ALTERNATIVE POINT OF VIEW

1.1 The Legend of a Royal Road to General Relativity

Hilbert is commonly seen as having publicly presented the derivation of the field equations of general relativity on 20 November 1915, five days before Einstein and after only half a year's work on the subject in contrast to Einstein's eight years of hardship from 1907 to 1915.[1] We thus read in Kip Thorne's fascinating account of recent developments in general relativity (Thorne 1994, 117):

> Remarkably, Einstein was not the first to discover the correct form of the law of warpage [of space-time, i.e. the gravitational field equations], the form that obeys his relativity principle. Recognition for the first discovery must go to Hilbert. In autumn 1915, even as Einstein was struggling toward the right law, making mathematical mistake after mistake, Hilbert was mulling over the things he had learned from Einstein's summer visit to Göttingen. While he was on an autumn vacation on the island of Rugen in the Baltic the key idea came to him, and within a few weeks he had the right law–derived not by the arduous trial-and-error path of Einstein, but by an elegant, succinct mathematical route. Hilbert presented his derivation and the resulting law at a meeting of the Royal Academy of Sciences in Göttingen on 20 November 1915, just five days before Einstein's presentation of the same law at the Prussian Academy meeting in Berlin.[2]

Hilbert himself emphasized that he had two separate starting points for his approach: Mie's electromagnetic theory of matter as well as Einstein's attempt to base a theory of gravitation on the metric tensor. Hilbert's superior mastery of mathematics apparently allowed him to arrive quickly and independently at combined field equa-

1 For discussions of Einstein's path to general relativity see (Norton 1984; Renn and Sauer 1999; Stachel 2002), "The First Two Acts", "Pathways out of Classical Physics …", and "Untying the Knot …", (in vols. 1 and 2 of this series). For historical reviews of Hilbert's contribution, see (Guth 1970; Mehra 1974; Earman and Glymour 1978; Pais 1982, 257–261; Corry 1997; 1999a; 1999b; 1999c; Corry, Renn, and Stachel 1997; Stachel 1989; 2002; Sauer 1999; 2002), "The Origin of Hilbert's Axiomatic Method …" and "Einstein Equations and Hilbert Action" (both in this volume).

2 For a similar account see (Fölsing 1997, 375–376).

Jürgen Renn (ed.). *The Genesis of General Relativity,* Vol. 4
Gravitation in the Twilight of Classical Physics: The Promise of Mathematics.

tions for the electromagnetic and gravitational fields. Although his use of Mie's ideas initially led Hilbert to a theory that was, from the point of view of the subsequent general theory of relativity, restricted to a particular source for the gravitational field—the electromagnetic field—he is nevertheless regarded by many historians of science and physicists as the first to have established a mathematical framework for general relativity that provides both essential results of the theory, such as the field equations, and a clarification of previously obscure conceptual issues, such as the nature of causality in generally-covariant field theories.[3] His contributions to general relativity, although initially inspired by Mie and Einstein, hence appear as a unique and independent achievement. In addition, Hilbert is seen by some historians of science as initiating the subsequent search for unified field theories of gravitation and electromagnetism.[4] In view of all these results, established within a very short time, it appears that Hilbert indeed had found an independent "royal road" to general relativity and beyond.

In a recent paper with Leo Corry, we have shown that Hilbert actually did not anticipate Einstein in presenting the field equations (Corry, Renn, and Stachel 1997).[5] Our argument is based on the analysis of a set of proofs of Hilbert's first paper,[6] hereafter referred to as the "Proofs". These Proofs not only do not include the explicit form of the field equations of general relativity, but they also show the original version of Hilbert's theory to be in many ways closer to the earlier, non-covariant versions of Einstein's theory of gravitation than to general relativity. It was only *after* the publication on 2 December 1915 of Einstein's definitive paper that Hilbert modified his theory in such a way that his results were in accord with those of Einstein.[7] The final version of his first paper, which was not published until March 1916, now includes the explicit field equations and has no restriction on general covariance (Hilbert 1916).[8] Hilbert's second paper, a sequel to his first communication, in which he first discussed causality, apparently also underwent a major revision before eventually being published in 1917 (Hilbert 1917).[9]

3 See (Howard and Norton 1993).

4 See, for example, (Vizgin 1989), who refers to "Hilbert's 1915 unified field theory, in which the attempt was first made to unite gravitation and electromagnetism on the basis of the general theory of relativity" (see p. 301).

5 See also (Stachel 1999), reprinted in (Stachel 2002).

6 A copy of the proofs of Hilbert's first paper is preserved at Göttingen, in SUB Cod. Ms. 634. They comprise 13 pages and are virtually complete, apart from the fact that roughly the upper quarter of two pages (7 and 8) is cut off. The Proofs are dated "submitted on 20 November 1915." The Göttingen copy bears a printer's stamp dated 6 December 1915 and is marked in Hilbert's own hand "First proofs of my first note." In addition, they carry several marginal notes in Hilbert's hand, which are discussed below. A complete translation of the Proofs is given in this volume.

7 The conclusive paper is (Einstein 1915e), which Hilbert lists in the references in (Hilbert 1916).

8 In the following referred to as Paper 1.

9 In the following referred to as Paper 2.

1.2 The Transformation of the Meaning of Hilbert's Work

Hilbert presented his contribution as emerging from a research program that was entirely his own—the search for an axiomatization of physics as a whole—creating a synthesis of electromagnetism and gravitation. This view of his achievement was shared by Felix Klein, who took the distinctiveness of Hilbert's approach as an argument against seeing it from the perspective of a priority competition with Einstein:

> There can be no talk of a priority question in this connection, since both authors are pursuing quite different trains of thought (and indeed, so that initially the compatibility of their results did not even seem certain). Einstein proceeds *inductively* and immediately considers arbitrary material systems. Hilbert *deduces* from previously postulated basic variational principles, while he additionally allows the restriction to electrodynamics. In this connection, Hilbert was particularly close to Mie.[10]

It is clear that, even if one disregards the non-covariant version of his theory as presented in the proofs version of his first paper, both Hilbert's original programmatic aims as well as the interpretation he gave of his own results do not fit into the framework of general relativity as we understand it today. To give one example, which we shall discuss in detail below: In the context of Hilbert's attempt at a synthesis of electromagnetism and gravitation theory, he interpreted the contracted Bianchi identities as a substitute for the fundamental equations of electromagnetism, an interpretation that was soon recognized to be problematic by Hilbert himself.

With hindsight, however, there can be little doubt that a number of important contributions to the development of general relativity do have roots in Hilbert's work: For instance, not so much the variational formulation of the gravitational field equations, an idea which had already been introduced by Einstein[11]; but the choice of the Ricci scalar as the gravitational term in this Lagrangian; and the first hints of Noether's theorem.

The intrinsic plausibility of each of these two perspectives: viewing Hilbert's work as either aiming at a theory differing from general relativity, or as a contribution to general relativity, represents a puzzle. How can Hilbert's contributions be interpreted as making sense only within an independent research program, different in essence from that of Einstein, if ultimately they came to be seen, at least by most physicists, as constituents of general relativity? This puzzle raises a profound historical question concerning the nature of scientific development: how were Hilbert's results, produced within a research program originally aiming at an electrodynamic

---

10 "Von einer Prioritätsfrage kann dabei keine Rede sein, weil beide Autoren ganz verschiedene Gedankengänge verfolgen (und zwar so, daß die Verträglichkeit der Resultate zunächst nicht einmal sicher schien). Einstein geht *induktiv* vor und denkt gleich an beliebige materielle Systeme. Hilbert *deduziert*, indem er übrigens die [...] Beschränkung auf Elektrodynamik eintreten läßt, aus voraufgestellten obersten Variationsprinzipien. Hilbert hat dabei insbesondere auch an Mie angeknüpft." (Klein 1921, 566). The text was originally published in 1917; see (Klein 1917). The quote is from a footnote to remarks added to the 1921 republication. For a recent reconstruction of Hilbert's perspective, see (Sauer 1999).

11 See "Untying the Knot …" (in vol. 2 of this series).

foundation for *all* of physics, eventually transformed into constituents of general relativity, a theory of gravitation? The pursuit of this question promises insights into the processes by which scientific results acquire and change their meaning and, in particular, into the process by which a viewpoint that is different from the one eventually accepted as mainstream emerges and eventually fades away.[12]

Hilbert's work on the foundations of physics turns out to be especially suited for such an analysis, not only because the proofs version of his first paper provides us with a previously unknown point of departure for following his development, but also because he came back time and again to these papers, rewriting them in terms of the insights he had meanwhile acquired and in the light of the developments of Einstein's "mainstream" program. In this paper we shall interpret Hilbert's revisions as indications of the conceptual transformation that his original approach underwent as a consequence of the establishment and further development of general relativity by Einstein, Schwarzschild, Klein, Weyl, and others, including Hilbert himself. We will also show that Hilbert's own understanding of scientific progress induced him to perceive this transformation as merely an elimination of errors and the introduction of improvements and elaborations of a program he had been following from the beginning.

1.3 Structure of the Paper

In the *second section* of this paper ("The origins of Hilbert's program in the 'nostrification' of two speculative physical theories"), we shall analyze the emergence of Hilbert's program for the foundations of physics from his attempt to synthesize, in the form of an axiomatic system, techniques and results of Einstein's 1913/14 non-covariant theory of gravitation and Mie's electromagnetic theory of matter. It will become clear that Hilbert's research agenda was shaped in large part by his understanding of the axiomatic formulation of physical theories, by the technical problems of achieving the synthesis of these two theories, and by open problems in Einstein's theory.

In the *third section* ("Hilbert's attempt at a theory of everything: the proofs of his first paper"), we shall interpret the proofs version of Hilbert's first paper as an attempt to realize the research program reconstructed in the second section. In particular, we shall show that, in the course of pursuing this program, he abandoned his original goal of founding all of physics on electrodynamics, now treating the gravitational field as more fundamental. We shall argue that this reversal was induced by mathematical results, to which Hilbert gave a problematic physical interpretation suggested by his research program; and that the mathematical result at the core of Hilbert's attempt to establish a connection between gravitation and electromagnetism originated in Einstein's claim of 1913/14 that generally-covariant field equations are not compatible with physical causality, a claim supported by Einstein's well-known "hole-argument." Hilbert thus turned Einstein's argument against general covariance into support for Hilbert's own attempt at a unified theory of gravitation and electro-

12 Cf. (Stachel 1994).

magnetism. Hilbert also followed Einstein's 1913/14 attempt to relate the existence of a preferred class of coordinate systems to the requirement of energy conservation. Hilbert's definition of energy, however, was not guided by Einstein's but rather by the goal of establishing a link with Mie's theory. Hilbert's unified theory thus emerges as an extension of Einstein's non-covariant theory of gravitation, in which Mie's speculative theory of matter plays the role of a touchstone, a role played for Einstein by the principle of energy-momentum conservation in classical and special relativistic physics and in Newton's theory of gravitation.

In the *fourth section* ("Hilbert's physics and Einstein's mathematics: the exchange of late 1915") we shall examine Hilbert's and Einstein's exchange of letters at the end of 1915, focussing on the ways in which they mutually influenced each other. We show that Hilbert's attempt at combining a theory of gravitation with a theory of matter had an important impact on the final phase of Einstein's work. Hilbert's vision, which Einstein temporarily adopted, provided the latter with a rather exotic perspective but allowed him to obtain a crucial result, the calculation of Mercury's perihelion precession. This, in turn, guided his completion of the general theory of relativity, but at the same time rendered obsolete its grounding in a specific theory of matter. For Hilbert's theory, on the other hand, Einstein's conclusive paper on general relativity represented a major challenge. It undermined the entire architecture; in particular, the connections Hilbert saw between energy conservation, causality, and the need for a restriction of general covariance.

In the *fifth section* ("Hilbert's adaptation of his theory to Einstein's results: the published versions of his first paper") we shall first discuss how, under the impact of Einstein's results in November 1915, Hilbert modified essential elements of his theory before its publication in March 1916. He abandoned the attempt to develop a non-covariant theory, without as yet having found a satisfactory solution to the causality problem that Einstein had previously raised for generally-covariant theories. He replaced his original, non-covariant notion of energy by a new formulation, still differing from that of Einstein and mainly intended to strengthen the link between his own theory and Mie's electrodynamics. In fact, Hilbert did not abandon his aim of providing a foundation for all of physics. He still hoped to construct a field-theoretical model of the electron and derive its laws of motion in the atom, without, however, getting far enough to include any results in his paper. His first paper was republished twice, in 1924 and 1933, each time with significant revisions. We shall show that Hilbert eventually adopted the understanding of energy-momentum conservation developed in general relativity, thus transforming his ambitious program into an application of general relativity to a special kind of source, matter as described by Mie's theory.

In the *sixth section* ("Hilbert's adoption of Einstein's program: the second paper and its revisions") we shall show that Hilbert's second paper, published in 1917, is the outcome of his attempt to tackle the unsolved problems of his theory in the light of Einstein's results, in particular the causality problem; and at the same time to keep up with the rapid progress of general relativity. In fact, instead of pursuing the consequences of his approach for microphysics, as he originally intended, he now turned to solutions of the gravitational field equations, relating them to the mathematical tradition inaugurated by Gauss and Riemann of exploring the applicability of Euclidean geometry to the physical world. In this way, he effectively worked within the program of general relativity and contributed to solving such problems as the uniqueness of the Minkowski solution and the derivation of the Schwarzschild solution; but he was less successful in dealing with the problem of causality in a generally-covariant theory. Although he followed Einstein in focussing on the invariant features of such a theory, he attempted to develop his own solution to the causality problem, different from that of Einstein. Whereas Einstein resolved the ambiguities he had earlier encountered in the hole argument by the insight that in general relativity coordinate systems have no physical significance apart from the metric, Hilbert attempted to find a purely "mathematical response" to this problem, formulating the causality condition in terms of the Cauchy or initial-value problem for the generally-covariant field equations. While it initiated an important line of research in general relativity, this first attempt not only failed to incorporate Einstein's insights into the physical interpretation of general relativity but also suffered from Hilbert's inadequate treatment of the Cauchy problem for such a theory, a treatment that was finally corrected by the editors of the revised version published in 1933.

In the *seventh section* ("The fading away of Hilbert's point of view in the physics and mathematics communities") we shall analyze the reception of Hilbert's work in contemporary literature on general relativity and unified field theories, as well as its later fate in the textbook tradition. We show that, in spite of Hilbert's emphasis on the distinctiveness of his approach, his work was perceived almost exclusively as a contribution to general relativity. It will become clear that this reception was shaped largely by the treatment of Hilbert's work in the publications of Einstein and Weyl, although, by revising his own contributions in the light of the progress of general relativity, Hilbert was not far behind in contributing to the complete disappearance of his original, distinctive point of view. This disappearance had two remarkable consequences: First, deviations of Hilbert's theory from general relativity, such as his interpretation of the contracted Bianchi identities as the coupling between gravitation and electromagnetism, went practically unremarked. Second, in spite of his attempt to depict himself as the founding father of unified field theories, the early workers in this field tended to ignore his contribution, denying him a prominent place in their intellectual ancestry. Instead, Hilbert was assigned a prominent place in the history of general relativity, even ascribing to him achievements that were not his, such as the first formulation of the field equations or the complete clarification of the question of causality. The ease with which his work could be assimilated to general relativity provides further evidence of a different kind for the tenuous and unstable character of his own framework.

In the *eighth and final section* ("At the end of a royal road") we shall compare Hilbert's and Einstein's approaches in an effort to understand Hilbert's gradual rapprochement with general relativity. Einstein had followed a double strategy in creating general relativity: trying to explore the mathematical consequences of physical principles on the one hand; and systematically checking the physical interpretation of mathematical results, on the other. Hilbert's initial approach encompassed a much narrower physical basis. Starting from a few problematic physical assumptions, Hilbert elaborated a mathematically complex framework, but never succeeded in finding any concrete physical consequences of this framework other than those that had been or could be found within Einstein's theory of general relativity. Nevertheless, Hilbert's assimilation of specific results from the mainstream tradition of general relativity into his framework eventually changed the character of this framework, transforming his results into contributions to general relativity. Thus, in a sense, Hilbert's assimilation of insights from general relativity served as a substitute for the physical component of Einstein's double strategy that was originally lacking in Hilbert's own approach. So this double strategy emerges not only as a successful heuristic characterizing Einstein's individual pathway, but as a particular aspect of the more general process by which additional knowledge was integrated into the further development of general relativity.

2. THE ORIGINS OF HILBERT'S PROGRAM IN THE "NOSTRIFICATION" OF TWO SPECULATIVE PHYSICAL THEORIES

Leo Corry has explored in depth the roots and the history of Hilbert's program of axiomatization of physics and, in particular, its impact on his 1916 paper *Foundations of Physics*.[13] We can therefore limit ourselves to recapitulating briefly some essential elements of this program. Hilbert conceived of the axiomation of physics not as a definite foundation that has to *precede* empirical research and theory formation, but as a *post-hoc* reflection on the results of such investigations with the aim of clarifying the logical and epistemological structure of the assumptions, definitions, etc., on which they are built.[14] Nevertheless, Hilbert expected that a proper axiomatic foundation of physics would not be shaken every time a new empirical fact is discovered; but rather that new, significant facts could be incorporated into the existing body of knowledge without changing its logical structure. Furthermore, Hilbert expected that, rather than emerging from the reorganization of the existing body of knowledge, the concepts used in an axiomatic foundation of physics should be those already familiar from the history of physics. Finally, Hilbert was convinced that one can distinguish sharply between the particular, empirical and the universal ingredients of a physical theory.

Accordingly, the task that Hilbert set for himself was not to find new concepts serving to integrate the existing body of physical knowledge into a coherent conceptual whole, but rather to formulate appropriate axioms involving the already-existing

---

13 See (Corry 1997; 1999a; 1999b; 1999c; see also Sauer 1999, section 1) and "The Origin of Hilbert's Axiomatic Method …" (in this volume).

14 For evidence of the following claims, see, in particular, Hilbert's lecture notes (Hilbert 1905; 1913), extensively discussed in Corry's papers.

physical concepts; axioms which allow the reconstruction of available physical knowledge by deduction from these axioms. Consequently, his interest in the axiomatization of physics was oriented toward the reductionist attempts to found all of physics on the basis of either mechanics or electrodynamics (the mechanical or electromagnetic worldview). Indeed, in his discussions of the foundations of physics before 1905, the axiomatization of mechanics was central; while, at some point after the advent of the special theory of relativity, Hilbert now placed his hopes in an axiomatization of all physics based on electrodynamics.[15] In spite of the conceptual revolution brought about by special relativity, involving not only the revision of the concepts of space and time but also the autonomy of the field concept from that of the aether, Hilbert nevertheless continued to rely on traditional concepts such as force and rigidity as the building blocks for his axiomatization program.[16]

An axiomatic synthesis of existing knowledge such as that pursued by Hilbert in physics apparently also had a strategic significance for Göttingen mathematicians making it possible for them to leave their distinctive mark on a broad array of domains, which were thus "appropriated," not only intellectually but also in the sense of professional responsibility for them. Minkowski's attempt to present his work on special relativity as a decisive mathematical synthesis of the work of his predecessors may serve as an example.[17] Discussing an accusation that Emmy Noether had neglected to acknowledge her intellectual debt to British and American algebraists, Garrett Birkhoff wrote:

> This seems like an example of German 'nostrification:' reformulating other people's best ideas with increased sharpness and generality, and from then on citing the local reformulation.[18]

2.1 Mie's Theory of Matter

By 1913, Hilbert expected that the electron theory of matter would provide the foundation for all of physics. It is therefore not surprising to find him shortly afterwards attracted to Mie's theory of matter, a non-linear generalization of Maxwell's electrodynamics that aimed at the overcoming of the dualism between "aether" and "ponderable matter." Indeed, Mie had introduced a generalized Hamiltonian formalism for electrodynamics, allowing for non-linear couplings between the field variables, in the hope of deriving the electromagnetic properties of the "aether" as well as the particulate structure of matter from one and the same variational principle.[19] Mie's theory thus not only corresponded to Hilbert's hope to found all of physics on the concepts

---

15 For a discussion of Hilbert's turn from mechanical to electromagnetic monism, see (Corry 1999a, 511–517).

16 See (Hilbert 1913, 13).

17 This attempt is extensively discussed in (Walter 1999). See also (Rowe 1989).

18 Garrett Birkhoff to Bartel Leendert van der Waerden, 1 November 1973 (Eidgenössische Technische Hochschule Zürich, Handschriftenabteilung, Hs 652:1056); quoted from (Siegmund-Schultze 1998, 270). We thank Leo Corry for drawing our attention to this letter.

of electrodynamics; but it must also have been attractive to him because it was based upon the variational calculus, a tool, with the usefulness of which for the axiomatization of physical theories Hilbert was quite familiar.[20] However, Mie's theory was far from able to provide specific results concerning the electromagnetic properties of matter, results which could be confronted with empirical data. Rather, the theory provides only a framework; a suitable "world function" (Lagrangian) must still be found, from which such concrete predictions may then be derived. Mie gave examples of such world functions that, however, were meant to be no more than illustrations of certain features of his framework. In fact, Mie could not have considered these examples as the basis of a specific physical theory since they are not even compatible with basic features of physical reality such as the existence of an elementary quantum of electricity. Concerning his principal example, later taken up by Hilbert, Mie himself remarked:

> A world that is governed by the world function
>
> $$\phi = -\frac{1}{2}\eta^2 + \frac{1}{6}a \cdot \chi^6 \tag{1}$$
>
> must ultimately agglomerate into two large lumps of electric charges, one positive and one negative, and both these lumps must continually tend to separate further and further from each other.[21]

Mie drew the obvious conclusion that the unknown world function he eventually hoped to find must be more complicated than this and the other examples he had considered.[22]

Hilbert based his work on a formulation of Mie's framework actually due to Max Born.[23] In a paper of 1914, Born showed that Mie's variational principle can be considered as a special case of a four-dimensional variational principle for the deformation of a four-dimensional continuum involving the integral:[24]

$$\int \phi(a_{11}, a_{12}, a_{13}, a_{14}; a_{21} \ldots; u_1, \ldots u_4) dx_1 dx_2 dx_3 dx_4 \ . \tag{2}$$

---

19 Mie's theory was published in three installments: (Mie 1912a; 1912b; 1913). For a concise account of Mie's theory, see (Corry 1999b), see also the Editorial Note in this volume. In the recent literature on Mie's theory, the problematic physical content of this theory (and hence of its adaptation by Hilbert) plays only a minor role; see the discussion below.

20 See, in particular, (Hilbert 1905).

21 "Eine Welt, die durch die Weltfunktion (1) regiert würde, müßte sich also schließlich zu zwei großen Klumpen elektrischer Ladungen zusammenballen, einem positiven und einem negativen, und diese beiden Klumpen müßten immer weiter und weiter voneinander wegstreben." (Mie 1912b, 38) For the meaning of Mie's formula and its ingredients in Hilbert's version, see (33) below.

22 See (Mie 1912b, 40).

23 For a discussion of Born's role as Hilbert's informant about both Mie's and Einstein's theories, see (Sauer 1999, 538–539).

24 See (Born 1914).

Here ϕ is a Lorentz scalar, and:

$$u_\alpha = u_\alpha(x_1, x_2, x_3, x_4) \qquad \alpha = 1, \ldots 4 \tag{3}$$

are the projections onto four orthogonal axes of the displacements of the points of the four-dimensional continuum from their equilibrium positions regarded as functions of the quasi-Cartesian coordinates x_1, x_2, x_3, x_4 along these axes, and

$$a_{\alpha\beta} = \frac{\partial u_\alpha}{\partial x_\beta} \tag{4}$$

are their derivatives. Furthermore, Born showed that the characteristic feature of Mie's theory lies in the ansatz that the function ϕ depends only on the antisymmetric part of $a_{\alpha\beta}$:

$$a_{\alpha\beta} - a_{\beta\alpha} = \frac{\partial u_\alpha}{\partial x_\beta} - \frac{\partial u_\beta}{\partial x_\alpha}. \tag{5}$$

Mie's four-dimensional continuum could thus be regarded as a four-dimensional spacetime generalization of MacCullagh's three-dimensional aether. MacCullagh had derived equations corresponding to Maxwell's equations for stationary electrodynamic processes from the assumption that the vortices of the aether, rather than its deformations, store its energy (Whittaker 1951, 142–145).

What role does gravitation play in Mie's theory? Mie opened the series of papers on his theory with a programmatic formulation of his goals, among them to establish a link between the existence of matter and gravitation:

> The immediate goals that I set myself are: to explain the existence of the indivisible electron and: to view the actuality of gravitation as in a necessary connection with the existence of matter. I believe one must start with this, for electric and gravitational effects are surely the most direct expression of those forces upon which rests the very existence of matter. It would be senseless to imagine matter whose smallest parts did not possess electric charges, equally senseless however matter without gravitation.[25]

Initially Mie hoped that he could explain gravitation on the basis of his non-linear electrodynamics alone, without introducing further variables. His search for a new theory of gravitation was guided by a simple model, according to which gravitation is a kind of "atmosphere," arising from the electromagnetic interactions inside the atom:

> An atom is an agglomeration of a larger number of electrons glued together by a relatively dilute charge of opposite sign. Atoms are probably surrounded by more substantial

25 "Die nächsten Ziele, die ich mir gesteckt habe, sind: die Existenz des unteilbaren Elektrons zu erklären und: die Tatsache der Gravitation mit der Existenz der Materie in einem notwendigen Zusammenhang zu sehen. Ich glaube, daß man hiermit beginnen muß, denn die elektrischen und die Gravitationswirkungen sind sicher die unmittelbarsten Äußerungen der Kräfte, auf denen die Existenz der Materie überhaupt beruht. Es wäre sinnlos, Materie zu denken, deren kleinste Teilchen nicht elektrische Ladungen haben, ebenso sinnlos aber Materie ohne Gravitation." See (Mie 1912a, 511–512).

> atmospheres, which however are still so dilute that they do not cause noticeable electric fields, but which presumably are asserted in gravitational effects.[26]

In his third and conclusive paper, however, he explicitly withdrew this model and was forced to introduce the gravitational potential as an additional variable.[27] There is thus no intrinsic connection between gravitation and the other fields in Mie's theory. By representing gravitation as an additional term in his Lagrangian giving rise to a four-vector representation of the gravitational field, he effectively returned to Abraham's gravitation theory which he had earlier rejected.[28] As a consequence, his treatment of gravitation suffers from the same objections that were raised in contemporary discussions of Abraham's theory. In summary, Mie's theory of gravitation was far from reaching the goals he had earlier set for it.

2.2 Einstein's Non-Covariant "Entwurf" Theory of Gravitation

In 1915, Hilbert became interested in Einstein's theory of gravitation after a series of talks on this topic by Einstein between 28 June and 5 July of that year in Göttingen.[29] Hilbert's attraction to Einstein's approach may have stemmed from his dissatisfaction with the contrast between Mie's programmatic statements about the need for a unification of gravitation and electromagnetism and the unsatisfactory treatment of gravitation in Mie's actual theory. This may well have motivated Hilbert to look at other theories of gravitation and perhaps even to invite Einstein. But apart from the shortcomings of Mie's theory, Hilbert's fascination with Einstein's approach to gravitation probably is rooted in the remarkable relations that Hilbert must have perceived between the structure of Mie's theory of electromagnetism and Einstein's theory of gravitation, as the latter was presented in his 1913/1914 publications and (presumably) also in the Göttingen lectures.

Like Mie's theory, Einstein's *Entwurf* theory was based on a variational principle for a Lagrangian H, here considered to be a function of the gravitational potentials (represented by the components of the metric tensor field $g_{\alpha\beta}$) and their first derivatives. In contrast to Mie, however, Einstein had specified a particular Lagrangian, from which he then derived the gravitational field equations:[30]

---

26 "Ein Atom ist eine Zusammenballung einer größeren Zahl von Elektronen, die durch eine verhältnismäßig dünne Ladung von entgegengesetztem Vorzeichen verkittet sind. Die Atome sind wahrscheinlich von kräftigeren Atmosphären umgeben, die allerdings immer noch so dünn sind, daß sie keine bemerkbaren elektrischen Felder veranlassen, die sich aber vermutlich in den Gravitationswirkungen geltend machen." See (Mie 1912a, 512–513).

27 See (Mie 1913, 5).

28 Compare (Mie 1912a, 534) with (Mie 1913, 29).

29 For notes on a part of Einstein's lectures, see "Nachschrift of Einstein's Wolfskehl Lectures" in (CPAE 6, 586–590). For a discussion of Einstein's Göttingen visit and its possible impact on Hilbert, see (Corry 1999a, 514–517).

30 Our presentation follows Einstein's major review paper, (Einstein 1914b).

$$H = \frac{1}{4}\sum_{\alpha\beta\tau\rho} g^{\alpha\beta}\frac{\partial g_{\tau\rho}}{\partial x_\alpha}\frac{\partial g^{\tau\rho}}{\partial x_\beta}. \tag{6}$$

To be more precise, Einstein was able to derive the empty-space field equations from this Lagrangian. The left-hand side of the gravitational field equations is given by the Lagrangian derivative of (6):[31]

$$\mathfrak{E}_{\mu\nu} = \frac{\partial H\sqrt{-g}}{\partial g^{\mu\nu}} - \sum_\sigma \frac{\partial}{\partial x_\sigma}\left(\frac{\partial H\sqrt{-g}}{\partial g_\sigma^{\mu\nu}}\right) \tag{7}$$

where $g_\sigma^{\mu\nu} \equiv \frac{\partial}{\partial x_\sigma} g^{\mu\nu}$. In the presence of matter, the right-hand side of the field equations is given by the energy-momentum tensor $\mathfrak{T}_{\alpha\beta}$ of matter, so that Einstein's field equations become:

$$\mathfrak{E}_{\sigma\tau} = \kappa\mathfrak{T}_{\sigma\tau}, \tag{8}$$

with the universal gravitational constant κ. In Einstein's *Entwurf* theory, the role of matter as an external source of the gravitational field is not determined by the theory, but rather to be prescribed independently. In the Lagrangian, matter thus appears simply "black-boxed," in the form of a term involving its energy-momentum tensor, rather than as an expression explicitly involving some set of variables describing the constitution of matter:

$$\int\left(\delta H - \kappa\sum_{\mu\nu}\mathfrak{T}_{\mu\nu}\delta g^{\mu\nu}\right)d\tau = 0. \tag{9}$$

Here was a possible point of contact between Mie's and Einstein's theories: Was it possible to conceive of Mie's electromagnetic matter as the source of Einstein's gravitational field? In order to answer this question, evidently one had to study how the energy-momentum tensor $\mathfrak{T}_{\alpha\beta}$ can be derived from terms of Mie's Lagrangian; in particular, what happens if Mie's matter is placed in a four-dimensional spacetime described by an arbitrary metric tensor $g_{\mu\nu}$? This naturally presupposed a reformulation of Mie's theory in generally-covariant form, with an arbitrary metric tensor $g_{\mu\nu}$ replacing the flat one of Minkowski spacetime.

Although most other expressions in his theory are generally-covariant, such as the geodesic equations of motion for a particle in the $g_{\mu\nu}$-field and the expression of energy-momentum conservation in the form of the vanishing covariant divergence of the energy tensor of matter, the field equations of Einstein's 1913/14 theory of gravitation are not. While this lack of general covariance had initially seemed to him to be a blemish on his theory, in late 1913 Einstein convinced himself that he could

31 Magnitudes in Gothic script represent tensor densities with respect to linear transformations.

even demonstrate—by means of the well-known "hole-argument"—that generally-covariant field equations are physically inadmissible because they cannot provide a unique solution for the metric tensor $g_{\mu\nu}$ describing the gravitational field produced by a given matter distribution. The hole argument involves a specific boundary value problem (whether this problem is well posed mathematically is a question that Einstein never considered) for a set of generally-covariant field equations with given sources outside of and boundary values on a "hole" (i.e. a region of spacetime without any sources in it), Einstein showed how to construct infinitely many apparently inequivalent solutions starting from any given solution. From the perspective of the hole argument, as Hilbert realized, if one considers generally-covariant field equations, then in order to pick out a unique solution these equations must be supplemented by four additional non-covariant equations. From the perspective of the 1915 theory of general relativity, however, the hole argument no longer represents an objection against generally-covariant field equations because the class of mathematically distinct solutions generated from an initial solution are not regarded as physically distinct, but merely as different mathematical representations of a single physical situation.[32]

Even in 1913/14 Einstein believed that it might be possible to formulate generally-covariant equations, from which equations (8) would follow by introducing a suitable coordinate restriction.[33] While he actually never found such equations corresponding to (8), he did find four non-covariant coordinate restrictions that he believed characteristic for his theory. He obtained these coordinate restrictions from an analysis of the behavior under coordinate transformations of the variational principle, on which his theory was based. Expressed in terms of the Lagrangian H, these four coordinate restrictions are:

$$B_\mu = \sum_{\alpha\sigma\nu} \frac{\partial^2}{\partial x_\sigma \partial x_\alpha}\left(g^{\nu\alpha}\frac{\partial H\sqrt{-g}}{\partial g_\sigma^{\mu\nu}}\right) = 0. \tag{10}$$

Einstein regarded these restrictions as making evident the non-general covariance of his theory; indeed he believed them just restrictive enough to avoid the hole-argument. Einstein also required the existence of a gravitational energy-momentum complex (non-tensorial) guaranteeing validity of four energy-momentum conservation equations for the combined matter and gravitational fields. His theory thus involved 10 field equations, 4 coordinate restrictions, and 4 conservation equations — in all 18 equations for the 10 gravitational potentials $g_{\mu\nu}$.

Einstein used the consistency of this overdetermined system as a criterion for the choice of a Lagrangian, imposing the condition that the field equations together with the energy-momentum conservation equations should yield the coordinate restric-

32 See (Stachel 1989; 71–81, sections 3 and 4).

33 See, e.g., (Einstein 1914a, 177–178). It is unclear whether Einstein expected the unknown generally-covariant equations to be of higher order than second.

tions (10). For this purpose, he assumed a general Lagrangian H depending on $g_{\mu\nu}$ and $g_{\mu\nu,\kappa}$, and then examined the four equations implied by the assumption of energy-momentum conservation for the field equations resulting from this Lagrangian. Formulating energy-momentum conservation as the requirement that the covariant divergence of the energy-momentum tensor density $\mathfrak{T}_\sigma{}^\nu$ has to vanish, and using the field equations (8), he first obtained:

$$\nabla_\nu \mathfrak{T}_\sigma{}^\nu \equiv \sum_{\nu\tau} \frac{\partial}{\partial x_\nu}(g^{\tau\nu}\mathfrak{T}_{\sigma\tau}) + \frac{1}{2}\sum_{\mu\nu}\frac{\partial g^{\mu\nu}}{\partial x_\sigma}\mathfrak{T}_{\mu\nu} = 0 \Rightarrow$$
$$\sum_{\nu\tau} \frac{\partial}{\partial x_\nu}(g^{\tau\nu}\mathfrak{E}_{\sigma\tau}) + \frac{1}{2}\sum_{\mu\nu}\frac{\partial g^{\mu\nu}}{\partial x_\sigma}\mathfrak{E}_{\mu\nu} = 0, \tag{11}$$

and then:

$$\sum_{\nu}\frac{\partial S_\sigma^\nu}{\partial x_\nu} - B_\sigma = 0, \tag{12}$$

with B_σ given by (10) and:

$$S_\sigma^\nu = \sum_{\mu\tau}\left(g^{\nu\tau}\frac{\partial H\sqrt{-g}}{\partial g^{\sigma\tau}} + g_\mu^{\nu\tau}\frac{\partial H\sqrt{-g}}{\partial g_\mu^{\sigma\tau}} + \frac{1}{2}\partial_\sigma^\nu H\sqrt{-g} - \frac{1}{2}g_\sigma^{\mu\tau}\frac{\partial H\sqrt{-g}}{\partial g_\nu^{\mu\tau}}\right). \tag{13}$$

By requiring that:

$$S_\sigma^\nu \equiv 0, \tag{14}$$

an equation that indeed is satisfied for the Lagrangian (6), it follows that (12) entails no new conditions beyond (10). In other words, for the "right" Lagrangian, the coordinate restrictions required by the hole-argument follow from energy-momentum conservation. In late 1915 Einstein found that his argument for the uniqueness of the Lagrangian, and thus for the uniqueness of the field equations, is fallacious;[34] and this insight helped to motivate him to return to generally-covariant field equations.

If one disregards the wealth of successful predictions of Newtonian gravitation theory that also buttressed Einstein's theory of 1913/14, that theory might appear almost as speculative as Mie's theory of matter. On the one hand, Einstein had been able to make several predictions based on his theory, such as the perihelion shift of Mercury, the deflection of light in a gravitational field, and gravitational redshift, that, at least in principle, could be empirically checked. On the other hand, none of these conclusions had actually received such support by the time Hilbert turned to Einstein's work: indeed, the calculated perihelion shift was in disaccord with observation.

---

34 For a historical discussion, see (Norton 1984) and "Untying the Knot …" (in vol. 2 of this series).

2.3 Hilbert's Research Program

To a mathematician of Hilbert's competence, Einstein's 1913/1914 theory must have appeared somewhat clumsy. In particular, it left several specifically mathematical questions open, such as the putative existence of the corresponding generally-covariant equations mentioned above; how the field equations (8) result from these generally-covariant equations by means of the coordinate restrictions (10); whether the hole argument for generally-covariant equations is better applied to boundary values on an open space-like hypersurface (the Cauchy problem) or a closed hypersurface (Einstein's formulation); and the closely–related question of the number of independent equations for the gravitational potentials in Einstein's system. Such questions presumably suggested to Hilbert a rather well-circumscribed research program that, taken together with his interest in Mie's theory of matter, amounted to the search for an "axiomatic synthesis" of the two speculative physical theories.

In consequence, Hilbert's initial program presumably comprised:[35]

1. a generally-covariant reformulation of both Mie's and Einstein's theories with the intention of deriving both from a single variational principle for a Lagrangian that depends on both Mie's electrodynamical and Einstein's gravitational variables;
2. an examination of the possibility of replacing Einstein's unspecified energy-momentum tensor for matter by one following from Mie's Lagrangian;
3. a further examination of the non-uniqueness of solutions to generally-covariant equations, involving a study of the question of the number of independent equations, and finally
4. the identification of coordinate restrictions appropriate to delimit a unique solution and an examination of their relation to energy-momentum conservation.

Even prior to looking at Hilbert's attempt to realize such a synthesis of Mie's and Einstein's approaches, it is clear that such a program would fit perfectly into Hilbert's axiomatic approach to physics. Indeed, the realization of this suggested initial program would: constitute a clarification of the logical and mathematical foundations of already existing physical theories in their own terms; represent the synthesis of different theories by combination of logically independent elements within one and the same formalism (in this case incorporation of Mie's variables and Einstein's variables in the same Lagrangian); replace the unspecified character of the material sources entering Einstein's theory with a daring theory of their electromagnetic nature, formulated in mathematical terms, thus shifting the boundary between experience and mathematical deduction in favor of the latter.

Unfortunately, there is no direct evidence that Hilbert developed and pursued some such research program in the course of his work in the second half of 1915 on Mie's and Einstein's theories. We have no "Göttingen notebook" that would be equivalent to Einstein's "Zurich Notebook," documenting in detail the heuristics that Hil-

35 For a similar attempt to reconstruct Hilbert's research program, see (Sauer 1999, 557–559).

bert followed.[36] However, now we have the first proofs of Hilbert's first communication that (as we have argued)[37] provide a glimpse into his thinking prior to his assimilation of Einstein's definitive paper on general relativity. In the next section we shall argue that the proofs version of Hilbert's theory can be interpreted as the result of pursuing just such a research program as that sketched above.

3. HILBERT'S ATTEMPT AT A THEORY OF EVERYTHING: THE PROOFS OF HIS FIRST PAPER

In this section we shall attempt to reconstruct Hilbert's heuristics from the Proofs and published versions of his first paper (Hilbert 1916), hereafter, Proofs and Paper 1. We will begin by reconstructing from the Proofs and other contemporary documents, the first step in the realization of Hilbert's program. This crucial step, an attempt to explore the first two points of the program, was the establishment of a relation between Mie's energy-momentum tensor and the variational derivative with respect to the metric of Mie's Lagrangian.[38] Next, we attempt to reconstruct Hilbert's calculation of Mie's energy-momentum tensor from the Born-Mie Lagrangian. We then examine the consequences of this derivation for the concept of energy, and thus for the further exploration of the second point of his program. We then discuss how these results suggest a new perspective on the relation between Mie's and Einstein's theories, from which gravitation appears more fundamental than electrodynamics. Seen from this perspective, the third point of Hilbert's program, the question of uniqueness of solutions to generally-covariant equations, took on a new significance: Hilbert turned Einstein's argument that only a non-covariant theory can make physical sense into an instrument for the synthesis of electromagnetism and gravitation. Coming to the fourth point of Hilbert's program, we show how he united his energy concept with the requirement of restricting general covariance. Finally, after examining Hilbert's attempt to derive the electromagnetic field equations from the gravitational ones, we discuss Hilbert's rearrangement of his results in the form of an axiomatically constructed theory, which he presented in the Proofs of Paper 1.

3.1 The First Result

At some point in late summer or fall of 1915, Hilbert must have discovered a relation between the energy-momentum tensor following from Mie's theory of matter, the Born-Mie Lagrangian *L*, and the metric tensor representing the gravitational poten-

36 Einstein's search for gravitational field equations in the winter of 1912/13 is documented in the so-called Zurich Notebook, partially published as Doc. 10 of (CPAE 4). Einstein's research project has been reconstructed in volumes 1 and 2 of this series. See, in particular, "Pathways out of Classical Physics …" (in vol. 1 of this series).

37 In (Corry, Renn, and Stachel 1997).

38 Henceforth, mention of the variational derivative of a Lagrangian, without further indication, always means with respect to the metric tensor.

tials in Einstein's theory of gravitation. In the Proofs and the published version of Paper 1, as well as in his contemporary correspondence, Hilbert emphasized the significance of this discovery for his understanding of the relation between Mie's and Einstein's theories. In the Proofs he wrote:

> Mie's electromagnetic energy tensor is nothing but the generally invariant tensor that results from differentiation of the invariant L with respect to the gravitational potentials $g^{\mu\nu}$ in the limit (25) [i.e. the equation $g_{\mu\nu} = \delta_{\mu\nu}$] — a circumstance that gave me the first hint of the necessary close connection between Einstein's general relativity theory and Mie's electrodynamics, and which convinced me of the correctness of the theory here developed.[39]

Hilbert expressed himself similarly in a letter of 13 November 1915 to Einstein:

> I derived most pleasure in the discovery, already discussed with Sommerfeld, that the usual electrical energy results when a certain absolute invariant is differentiated with respect to the gravitation potentials and then g is set = 0,1.[40]

On the basis of our suggested reconstruction of Hilbert's research program, it is possible to suggest what might have led him to this relation. We assume that he attempted to realize the first two steps, that is to reformulate Mie's Lagrangian in a generally-covariant setting and replace the energy-momentum tensor term in Einstein's variational principle by a term corresponding to Mie's theory. Considering (9), this would imply an expression such as $\delta H + \delta L$ under the integral, where H corresponds to Einstein's original Lagrangian and L to a generally-covariant form of Mie's Lagrangian. If the variation of Mie's Lagrangian is regarded as representing the energy-momentum tensor term, one obtains:

$$\delta L = -\kappa \sum_{\mu\nu} \mathfrak{T}_{\mu\nu} \delta g^{\mu\nu}, \tag{15}$$

where $\mathfrak{T}_{\mu\nu}$ should now be the energy-momentum tensor of Mie's theory. It may well have been an equation of this form, following from the attempt to replace the unspecified source-term in Einstein's field equations by a term depending on the generally-covariant form of Mie's Lagrangian, that first suggested to Hilbert that the energy-momentum tensor of Mie's theory could be the variational derivative of Mie's Lagrangian.

---

39 "der Mie'sche elektromagnetische Energietensor ist also nichts anderes als der durch Differentiation der Invariante L nach den Gravitationspotentialen $g^{\mu\nu}$ entstehende allgemein invariante Tensor beim Übergang zum Grenzfall (25) [i.e. the equation $g_{\mu\nu} = \delta_{\mu\nu}$] — ein Umstand, der mich zum ersten Mal auf den notwendigen engen Zusammenhang zwischen der Einsteinschen allgemeinen Relativitätstheorie und der Mie'schen Elektrodynamik hingewiesen und mir die Überzeugung von der Richtigkeit der hier entwickelten Theorie gegeben hat." (Proofs, 10)

40 "Hauptvergnügen war für mich die schon mit Sommerfeld besprochene Entdeckung, dass die gewöhnliche elektrische Energie herauskommt, wenn man eine gewisse absolute Invariante mit den Gravitationspotentialen differenziert und [d]ann $g = 0, 1$ setzt." David Hilbert to Einstein, 13 November 1915, (CPAE 8, 195). Unless otherwise noted, all translations are based on those in the companion volumes to the Einstein edition, but often modified.

If he followed the program outlined above, Hilbert would have assumed that the Lagrangian has the form:

$$H = K + L, \tag{16}$$

where K represents the gravitational part and L the electromagnetic. Indeed, this form of the Lagrangian is used both in the Proofs and the published version of Paper 1.[41]

In Paper 1, Hilbert derived a relation of the form:

$$-2\sum_{\mu}\frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}}g^{\mu m} = T_{\nu}^{m}, \tag{17}$$

where T_{ν}^{m} stands for the energy-momentum tensor density of Mie's theory.[42] This relation, which is exactly what one would expect on the basis of (15), could have suggested to Hilbert that a deep connection must exist between the nature of spacetime as represented by the metric tensor and the structure of matter as represented by Mie's theory.

3.2 Mie's Energy-Momentum Tensor as a Consequence of Generally-Covariant Field Equations

The strategy Hilbert followed to derive (17) can be reconstructed from the two versions of his paper. It consisted in following as closely as possible the standard variational techniques applied, for instance, to derive Lagrange's equations from a variational principle.[43] In Hilbert's paper, a similar variational problem forms the core of his theory. He describes his basic assumptions in two axioms:[44]

> Axiom I (Mie's axiom of the world function): *The law governing physical processes is determined through a world function* H *that contains the following arguments*:
>
> $$\begin{aligned} &g_{\mu\nu},\quad g_{\mu\nu l} = \frac{\partial g_{\mu\nu}}{\partial w_l}, \qquad g_{\mu\nu lk} = \frac{\partial^2 g_{\mu\nu}}{\partial w_l \partial w_k}, \\ &q_s\,,\quad q_{sl} = \frac{\partial q_s}{\partial w_l} \qquad\qquad (l, k = 1, 2, 3, 4), \end{aligned} \tag{18}$$
>
> *where the variation of the integral*

41 In the Proofs it was presumably introduced on the upper part of p. 8, which unfortunately is cut off.

42 See (Proofs, 10; Hilbert 1916, 404). Note that Hilbert uses an imaginary fourth coordinate, so that the minus sign emerges automatically in the determinant of the metric; he does not explicitly introduce the energy-momentum tensor T_{ν}^{m}.

43 See, for example, (Caratheodory 1935).

44 See also (Hilbert 1916, 396).

$$\int H\sqrt{g}d\tau \qquad (19)$$
$$(g = |g_{\mu\nu}|, \quad d\tau = dw_1 dw_2 dw_3 dw_4)$$

must vanish for each of the fourteen potentials $g_{\mu\nu}, q_s$.[45]

[The w_s are Hilbert's notation for an arbitrary system of coordinates.]

> Axiom II (axiom of general invariance): *The world function H is invariant with respect to an arbitrary transformation of the world parameters w_s.*[46]

Starting from an arbitrary invariant J, Hilbert formed a differential expression from it depending on $g^{\mu\nu}, g_l^{\mu\nu}, g_{lk}^{\mu\nu}, q_s, q_{sk}$, which in the published version of his paper he called PJ. He defined the operator P as follows:[47]

$$P = P_g + P_q,$$
$$P_g = \sum_{\mu, \nu, l, k}\left(p^{\mu\nu}\frac{\partial}{\partial g^{\mu\nu}} + p_l^{\mu\nu}\frac{\partial}{\partial g_l^{\mu\nu}} + p_{lk}^{\mu\nu}\frac{\partial}{\partial g_{lk}^{\mu\nu}}\right), \qquad (20)$$
$$P_q = \sum_{l, k}\left(p_l\frac{\partial}{\partial q_l} + p_{lk}\frac{\partial}{\partial q_{lk}}\right),$$

where $p^{\mu\nu}$ and p_l are arbitrary variations of the metric tensor and the electromagnetic four-potentials, respectively. Thus:

$$PJ = \sum_{\mu, \nu, l, k}\left(p^{\mu\nu}\frac{\partial J}{\partial g^{\mu\nu}} + p_l^{\mu\nu}\frac{\partial J}{\partial g_l^{\mu\nu}} + p_{lk}^{\mu\nu}\frac{\partial J}{\partial g_{lk}^{\mu\nu}} + p_l\frac{\partial J}{\partial q_l} + p_{lk}\frac{\partial J}{\partial q_{lk}}\right). \qquad (21)$$

In the mathematical terminology of the time, PJ is a "polarization" of J.[48]

As we shall see, it is possible to derive from PJ identities that realize Hilbert's goal, the derivation of (17). His procedure is described more explicitly in the published version of Paper 1, and since we assume that on this point there was no significant development of Hilbert's thinking after the Proofs, our reconstruction will make use of the published version.

In modern terminology, if $p^{\mu\nu}$ and p_l are those special variations generated by dragging the metric and the electromagnetic potentials over the manifold with some vector field p^s; i.e., if they are the Lie derivatives of the metric and the electromagnetic potentials with respect to p^s,[49] then PJ must be the Lie derivative of J with

---

45 "Axiom I (Mie's Axiom von der Weltfunktion): *Das Gesetz des physikalischen Geschehens bestimmt sich durch eine Weltfunktion H, die folgende Argumente enthält*: [(18); (1) and (2) in the original text] *und zwar muß die Variation des Integrals* [(19)] *für jedes der 14 Potentiale* $g_{\mu\nu}, q_s$ verschwinden." (Proofs, 2) The q_s are the electromagnetic four potentials.

46 "Axiom II (Axiom von der allgemeinen Invarianz): *Die Weltfunktion H ist eine Invariante gegenüber einer beliebigen Transformation der Weltparameter* w_s." (Proofs, 2)

47 See (Hilbert 1916, 398–399). Compare (Proofs, 4 and 7).

48 See, e.g., (Kerschensteiner 1887, §2).

respect to p^s. On the other hand, since J is a scalar invariant, the Lie derivative of this scalar with respect to p^s can be written directly, so that:

$$\sum_s \frac{\partial J}{\partial w_s} p^s = PJ. \tag{22}$$

With a little work,[50] equation (22) can be rewritten in the form of equation (23) below. This is the content of Hilbert's Theorem II, both in the Proofs and in Paper 1:

Theorem II. If J is an invariant depending on $g^{\mu\nu}$, $g_l^{\mu\nu}$, $g_{lk}^{\mu\nu}$, q_s, q_{sk}, then the following is always identically true in all its arguments and for every arbitrary contravariant vector p^s:

$$\begin{aligned}\sum_{\mu,\nu,l,k}\left(\frac{\partial J}{\partial g^{\mu\nu}}\Delta g^{\mu\nu}+\frac{\partial J}{\partial g_l^{\mu\nu}}\Delta g_l^{\mu\nu}+\frac{\partial J}{\partial g_{lk}^{\mu\nu}}\Delta g_{lk}^{\mu\nu}\right)\\+\sum_{s,k}\left(\frac{\partial J}{\partial q_s}\Delta q_s+\frac{\partial J}{\partial q_{sk}}\Delta q_{sk}\right) = 0;\end{aligned} \tag{23}$$

where

$$\begin{aligned}\Delta g^{\mu\nu} &= \sum_m (g^{\mu m} p_m^\nu + g^{\nu m} p_m^\mu),\\ \Delta g_l^{\mu\nu} &= -\sum_m g_m^{\mu\nu} p_l^m + \frac{\partial \Delta g^{\mu\nu}}{\partial w_l},\\ \Delta g_{lk}^{\mu\nu} &= -\sum_m (g_m^{\mu\nu} p_{lk}^m + g_{lm}^{\mu\nu} p_k^m + g_{km}^{\mu\nu} p_l^m) + \frac{\partial^2 \Delta g^{\mu\nu}}{\partial w_l \partial w_k},\\ \Delta q_s &= -\sum_m q_m p_s^m,\\ \Delta q_{sk} &= -\sum_m q_{sm} p_k^m + \frac{\partial \Delta q_s}{\partial w_k}.\end{aligned} \tag{24}$$

[51]

Hilbert next applies Theorem II to the electromagnetic part L of his Lagrangian $H = K + L$, with the assumption that L only depends on the metric $g^{\mu\nu}$, the elec-

---

49 Here $p^{\mu\nu}$ corresponds, in modern terms, to the Lie derivative of the contravariant form of the metric tensor with respect to the arbitrary vector p^j. Hilbert writes:

$$p^{\mu\nu} = \sum_s (g_s^{\mu\nu} p^s - g^{\mu s} p_s^\nu - g^{\nu s} p_s^\mu), \quad \left(p_s^j = \frac{\partial p^j}{\partial w_s}\right),$$

and similarly for the Lie derivatives of the electromagnetic potentials. While the term "Lie derivative" was only introduced in 1933 by W. Slebodzinski (see Slebodzinski 1931), it was well known in Hilbert's time that the basic idea came from Lie; see for example (Klein 1917, 471): "For this purpose one naturally determines, as Lie in particular has done in his numerous relevant publications, the formal changes that result from an arbitrary infinitesimal transformation." ("Zu diesem Zwecke bestimmt man natürlich, wie dies insbesondere Lie in seinen zahlreichen einschlägigen Veröffentlichungen getan hat, die formellen Änderungen, welche sich bei einer beliebigen infinitesimalen Transformation ... ergeben ... .") According to Schouten, the name "Lie differential" was proposed by D. Van Dantzig; see (Schouten and Struik 1935, 142).

tromagnetic potentials q_s and their derivatives q_{sk}, but *not* on the derivatives of the metric tensor. This gives the identity:[52]

$$\sum_{\mu,\nu,m} \frac{\partial L}{\partial g^{\mu\nu}}(g^{\mu m} p_m^{\nu} + g^{\nu m} p_m^{\mu}) - \sum_{s,m} \frac{\partial L}{\partial q_s} q_m p_s^m - \sum_{s,k,m} \frac{\partial L}{\partial q_{sk}}(q_{sm} p_k^m + q_{mk} p_s^m + q_m p_{sk}^m) = 0. \tag{25}$$

Since the vector field p^s is arbitrary, its coefficients as well as the coefficients of its first and second derivatives must vanish identically. Hilbert drew two conclusions, which he interpreted as strong links between a generally-covariant variational principle and Mie's theory of matter. The first concerns the form in which the electromag-

50 See (Proofs, 7–8; Hilbert 1916, 398). The equivalence of (22) and (23) is shown as follows: Since J depends on w_s through $g^{\mu\nu}$, $g_m^{\mu\nu}$, $g_{mk}^{\mu\nu}$, q_m and q_{mk} it follows that:

$$\frac{\partial J}{\partial w_s} = \frac{\partial J}{\partial g^{\mu\nu}} \cdot g_s^{\mu\nu} + \frac{\partial J}{\partial g_m^{\mu\nu}} \cdot g_{sm}^{\mu\nu} + \frac{\partial J}{\partial g_{mk}^{\mu\nu}} \cdot g_{smk}^{\mu\nu} + \frac{\partial J}{\partial q_m} \cdot q_{ms} + \frac{\partial J}{\partial q_{mk}} \cdot q_{mks}.$$

On the other hand, PJ is the Lie derivative of J through its dependence on $g^{\mu\nu}$, $g_m^{\mu\nu}$ q_m and q_{mk}, so:

$$PJ = \frac{\partial J}{\partial g^{\mu\nu}} \cdot p^{\mu\nu} + \frac{\partial J}{\partial g_m^{\mu\nu}} \cdot p_m^{\mu\nu} + \frac{\partial J}{\partial g_{mk}^{\mu\nu}} \cdot p_{mk}^{\mu\nu} + \frac{\partial J}{\partial q_m} \cdot p_m + \frac{\partial J}{\partial q_{mk}} \cdot p_{mk}$$

where $p^{\mu\nu}$ $p_m^{\mu\nu}$, $p_{mk}^{\mu\nu}$, p_m and p_{mk} stand for the Lie derivatives with respect to the vector field p^k of $g^{\mu\nu}$, $g_m^{\mu\nu}$, $g_{mk}^{\mu\nu}$, q_m and q_{mk} respectively (Hilbert's notation). Rewriting (24) in terms of the definition of the Lie derivatives of $g^{\mu\nu}$, $g_m^{\mu\nu}$, $g_{mk}^{\mu\nu}$, q_m and q_{mk}, we easily get:

$$\Delta g^{\mu\nu} = \sum_m g_m^{\mu\nu} p^m - p^{\mu\nu},$$

$$\Delta g_l^{\mu\nu} = \sum_m g_{ml}^{\mu\nu} p^m - p_l^{\mu\nu},$$

$$\Delta g_{lk}^{\mu\nu} = \sum_m g_{mlk}^{\mu\nu} p^m - p_{lk}^{\mu\nu},$$

$$\Delta q_s = \sum_m q_{sm} p^m - p_s,$$

$$\Delta q_{sk} = \sum_m q_{smk} p^m - p_{sk}.$$

Inserting these expressions into (23), and using the equations for $\frac{\partial J}{\partial w_s}$ and PJ at the beginning of this note, one sees that (23) reduces to:

$$\frac{\partial J}{\partial w_s} \cdot p^s - PJ = 0,$$

which is equivalent to (22).

netic potentials enter the Lagrangian, the second concerns the relation between this Lagrangian and Mie's energy-momentum tensor.

From Hilbert's requirements on L —that it be a generally-invariant scalar that does not depend on the derivatives of the metric tensor—he was able to show that the derivatives of the electromagnetic potentials can only enter it in the form characteristic of Mie's theory (see (5)). Setting the coefficients of p^m_{sk} in (25) equal to zero, and remembering that $p^m_{sk} = p^m_{ks}$, one obtains:

$$\left(\frac{\partial L}{\partial q_{sk}} + \frac{\partial L}{\partial q_{ks}}\right) q_m = 0. \tag{26}$$

Since q_m cannot vanish identically, it follows that:

$$\frac{\partial L}{\partial q_{sk}} + \frac{\partial L}{\partial q_{ks}} = 0, \tag{27}$$

which mean that the q_{ik} only enter L in the antisymmetric combination familiar from Mie's theory:

$$M_{ks} = q_{sk} - q_{ks}. \tag{28}$$

Thus, apart from the potentials themselves, L depends only on the components of the tensor M:

$$M = \mathrm{Rot}(q_s), \tag{29}$$

the familiar electromagnetic "six vector." Hilbert emphasized:

> *This result here derives essentially as a consequence of the general invariance, that is, on the basis of axiom II.*[53]

In order to explicitly establish the relation between his theory and Mie's, Hilbert points out that L must be a function of four invariants.[54] Hilbert only gave what he considered to be the "two simplest" of the generally-covariant generalizations of these invariants:

$$Q = \sum_{k,\,l,\,m,\,n} M_{mn} M_{lk} g^{mk} g^{nl} \tag{30}$$

---

51 "Theorem II. Wenn J eine von $g^{\mu\nu}, g^{\mu\nu}_l, g^{\mu\nu}_{lk}, q_s, q_{sk}$ abhängige Invariante ist, so gilt stets identisch in allen Argumenten und für jeden willkürlichen kontravarianten Vektor p^s [(23)] dabei ist: [(24)]."

52 See (Proofs, 9; Hilbert 1916, 403).

53 "*Dieses Resultat ergibt sich hier wesentlich als Folge der allgemeinen Invarianz, also auf Grund von Axiom II.*" (Proofs, 10) In the published version this passage reads: "*This result, which determines the character of Maxwell's equations in the first place, here derives essentially as a consequence of the general invariance, that is, on the basis of axiom II.*" ("*Dieses Resultat, durch welches erst der Charakter der Maxwellschen Gleichungen bedingt ist, ergibt sich hier wesentlich als Folge der allgemeinen Invarianz, also auf Grund von Axiom II.*") See (Hilbert 1916, 403).

54 See (Proofs, 13, and Hilbert 1916, 407). Here Hilbert followed the papers of Mie and Born; see, in particular, (Born 1914).

and:

$$q = \sum_{k,l} q_k q_l g^{kl}. \tag{31}$$

According to Hilbert, the simplest expression that can be formed by analogy to the gravitational part of the Lagrangian K is:[55]

$$L = \alpha Q + f(q), \tag{32}$$

where $f(q)$ is any function of q and α a constant. In order to recover Mie's main example (see (1)) from this more general result, Hilbert considers the following specific functional dependence:

$$L = \alpha Q + \beta q^3, \tag{33}$$

which corresponds to the Lagrangian given by Mie. In contrast to Mie, Hilbert does not even allude to the physical problems associated with this Lagrangian. And in contrast to Einstein, at no point does Hilbert introduce the Newtonian coupling constant into his equations, so that his treatment of gravitation remains as "formalistic" as that of electromagnetism.

The second consequence Hilbert drew from (25), which corresponds to what we have called above "Hilbert's first results" (see (17)), concerns Mie's energy-momentum tensor. Setting the coefficient of p_m^ν equal to zero and using (27), he obtained:[56]

$$2\sum_\mu \frac{\partial L}{\partial g^{\mu\nu}} g^{\mu m} - \frac{\partial L}{\partial q_m} q_\nu - \sum_s \frac{\partial L}{\partial M_{ms}} M_{\nu s} = 0, \quad (\mu = 1, 2, 3, 4). \tag{34}$$

Noting that:

$$2\sum_\mu \frac{\partial L}{\partial g^{\mu\nu}} g^{\mu m} = \frac{2}{\sqrt{g}} \cdot \sum_\mu \frac{\partial \sqrt{g} L}{\partial g^{\mu\nu}} g^{\mu m} + L \cdot \delta_\nu^{\ m}, \tag{35}$$

(34) can be rewritten:

$$-2\sum_\mu \frac{\partial \sqrt{g} L}{\partial g^{\mu\nu}} g^{\mu m} = \sqrt{g}\left\{ L\delta_\nu^m - \frac{\partial L}{\partial q_m} q_\nu - \sum_s \frac{\partial L}{\partial M_{ms}} M_{\nu s} \right\},$$
$$(\mu = 1, 2, 3, 4) \quad (\delta_\nu^\mu = 0, \ \mu \neq \nu, \ \delta_\mu^\mu = 1). \tag{36}$$

The right-hand side of this equation is the generally-covariant generalization of Mie's energy-momentum tensor. It is this equation that inspired Hilbert's remark about the

55 Note that Q is the term that gives rise to Maxwell's equations and that q cannot be used if the resulting theory is to be gauge invariant. See (Born and Infeld 1934).

56 See (Proofs, 10; Hilbert 1916, 404).

"Umstand, der mich zum ersten Mal auf den notwendigen engen Zusammenhang zwischen der Einsteinschen allgemeinen Relativitätstheorie und der Mie'schen Elektrodynamik hingewiesen ... hat", quoted above (p. 873). Hilbert had shown that characteristic properties of Mie's Lagrangian follow from its generally-covariant generalization, a result he interpreted as indicating that gravitation must be conceived as being more fundamental than electromagnetism, as his later work indicates.

3.3 The Definition of Energy

While (36) shows a strong link between a generally-covariant L and Mie's energy momentum tensor, it does not answer the question of how energy-momentum conservation is to be conceived in Hilbert's theory. Hilbert's theory does not allow the interpretation of an energy-momentum tensor for matter as an external source, as does that of Einstein; so Hilbert could not start from a conservation law for matter in Minkowski spacetime and simply generalize it to the case in which a gravitational field is present. Such a procedure would have conflicted with Hilbert's heuristic, according to which matter itself is conceived in terms of electromagnetic fields that, in turn, arise in conjunction with, or even as an effect of, gravitational fields.

Hilbert's heuristic for finding an appropriate definition of energy seems to be governed by a formal criterion related to his understanding of energy conservation in classical physics, as well as by a criterion with a more specific physical meaning related to the results he expected from Mie's theory. Hilbert's formal criterion is well described in a passage in his summer-semester 1916 lectures on the foundations of physics, a passage which occurs in a discussion of energy-momentum conservation in Mie's theory:

> The energy concept comes from just writing Lagrange's equations in the form of a divergence, and defining as energy what is represented as divergent.[57]

As for Hilbert's physical criterion, any definition of the energy must be compatible with his insight that the variational derivative of Mie's Lagrangian yields the electromagnetic energy-momentum tensor.

Hilbert's treatment of energy conservation in the Proofs and in Paper 1 is not easy to follow. This difficulty was felt by Hilbert's contemporaries; both Einstein and Klein had their problems with it.[58] Nevertheless, as will become clear in what follows, Hilbert's discussion was guided by the heuristic criteria mentioned above. He proceeded in three steps:

- he first identified an energy expression consisting of a sum of divergence terms (Satz 1 in the Proofs):

57 "Der Energiebegriff kommt eben daher, dass man die Lagrangeschen Gleichungen in Divergenzform schreibt, und das, was unter der Divergenz steht, als Energie definiert." Die Grundlagen der Physik I, Ms. Vorlesung SS 1916, 98 (D. Hilbert, Bibliothek des Mathematischen Seminars, Universität Göttingen); from here on "SS 1916 Lectures."

- he then formulated a divergence equation for his energy expression in analogy to classical and special-relativistic results (Satz 2 in the Proofs), and imposed this equation as a requirement implying coordinate restrictions (Axiom III):
- finally, he showed that his energy expression can be related to Mie's energy-momentum tensor (the real justification of his choice).

Here we focus on the first and last of these points, deferring the issue of coordinate restrictions to a subsequent section ("Energy-momentum conservation and coordinate restrictions").

As in his derivation of the connection between Mie's energy-momentum tensor and the variational derivative of the Lagrangian, Hilbert's starting point was his generally-covariant variational principle. However, he now proceeded somewhat differently. Instead of focussing on the electromagnetic part L, he considered the entire Lagrangian H, but now neglected the derivatives with respect to the electromagnetic potentials, i.e. the contribution of the term P_q to P (see (20)). Accordingly, Hilbert forms the expression:[59]

$$J^{(p)} = \sum_{\mu,\nu} \frac{\partial H}{\partial g^{\mu\nu}} p^{\mu\nu} + \sum_{\mu,\nu,k} \frac{\partial H}{\partial g_k^{\mu\nu}} p_k^{\mu\nu} + \sum_{\mu,\nu,k,l} \frac{\partial H}{\partial g_{kl}^{\mu\nu}} p_{kl}^{\mu\nu}, \tag{37}$$

where $p^{\mu\nu}$ corresponds, as we have seen, to the Lie derivative of the metric tensor with respect to the arbitrary vector p^j. By partial integration, Hilbert transforms this expression into:

$$\sqrt{g} J^{(p)} = -\sum_{\mu,\nu} H \frac{\partial \sqrt{g}}{\partial g^{\mu\nu}} p^{\mu\nu} + E + D^{(p)}, \tag{38}$$

with:

---

58 In (Klein 1917, 475), Klein quotes from a letter he had written to Hilbert concerning the latter's energy expression in Paper 1: "But I find your equations so complicated that I have not attempted to redo your calculations." ("Ich finde aber Ihre Formeln so kompliziert, daß ich die Nachrechnung nicht unternommen habe.") In a letter, in which Einstein asked Hilbert for a clarification of the latter's energy theorem, he wrote: "Why do you make it so hard for poor mortals by withholding the technique behind your ideas? It surely does not suffice for the thoughtful reader if, although able to verify the correctness of your equations, he cannot get a clear view of the overall plan of the analysis." ("Warum machen Sie es dem armen Sterblichen so schwer, indem Sie ihm die Technik Ihres Denkens vorenthalten? Es genügt doch dem denkenden Leser nicht, wenn er zwar die Richtigkeit Ihrer Gleichungen verifizieren aber den Plan der ganzen Untersuchung nicht überschauen kann.") See Einstein to David Hilbert, 30 May 1916, (CPAE 8, 293). In a letter to Paul Ehrenfest, Einstein expressed himself even more drastically with respect to what he perceived as the obscurity of Hilbert's heuristic: "Hilbert's description doesn't appeal to me. It is unnecessarily specialized as concerns "matter," unnecessarily complicated, and not above-board (=Gauss-like) in structure (feigning the super-human through camouflaging the methods)." ("Hilbert's Darstellung gefällt mir nicht. Sie ist unnötig speziell, was die 'Materie' anbelangt, unnötig kompliziert, nicht ehrlich (=Gaussisch) im Aufbau (Vorspiegelung des Übermenschen durch Verschleierung der Methoden).") See Einstein to Paul Ehrenfest, 24 May 1916, (CPAE 8, 288).

59 See (Proofs, 5ff.).

$$E = \sum_{\mu,\nu,s,k,l}\Big(H\frac{\partial\sqrt{g}}{\partial g^{\mu\nu}}g_s^{\mu\nu} + \sqrt{g}\frac{\partial H}{\partial g^{\mu\nu}}g_s^{\mu\nu} + \sqrt{g}\frac{\partial H}{\partial g_k^{\mu\nu}}g_{sk}^{\mu\nu} + \sqrt{g}\frac{\partial H}{\partial g_{kl}^{\mu\nu}}g_{skl}^{\mu\nu}\Big)p^s$$
$$- \sum_{\mu,\nu,s}(g^{\mu s}p_s^\nu + g^{\nu s}p_s^\mu)[\sqrt{g}H]_{\mu\nu} \tag{39}$$
$$+ \sum_{\mu,\nu,s,k,l}\Big(\frac{\partial\sqrt{g}H}{\partial g_k^{\mu\nu}}g_s^{\mu\nu} + \frac{\partial\sqrt{g}H}{\partial g_{kl}^{\mu\nu}}g_{sl}^{\mu\nu} - g_s^{\mu\nu}\frac{\partial}{\partial w_l}\frac{\partial\sqrt{g}H}{\partial g_{kl}^{\mu\nu}}\Big)p_k^s,$$

and:

$$D^{(p)} = \sum_{\mu,\nu,s,k,l}\Big\{-\frac{\partial}{\partial w_k}\Big(\sqrt{g}\frac{\partial H}{\partial g_k^{\mu\nu}}(g^{\mu s}p_s^\nu + g^{\nu s}p_s^\mu)\Big)$$
$$+\frac{\partial}{\partial w_k}\Big((p_s^\nu g^{\nu s} + p_s^\mu g^{\nu s})\frac{\partial}{\partial w_l}\Big(\sqrt{g}\frac{\partial H}{\partial g_{kl}^{\mu\nu}}\Big)\Big) \tag{40}$$
$$+\frac{\partial}{\partial w_l}\Big(\sqrt{g}\frac{\partial H}{\partial g_{kl}^{\mu\nu}}\Big(\frac{\partial p^{\mu\nu}}{\partial w_k} - g_{sk}^{\mu\nu}p^s\Big)\Big)\Big\}.$$

Hilbert had thus succeeded in splitting off a divergence term $D^{(p)}$ from the original expression $J^{(p)}$. By integrating over some region, $D^{(p)}$ could be converted into a surface term, and thus eliminated by demanding that p^s and its derivatives vanish on the boundary of that region.[60] So it would be possible to extract an energy expression from the remainder of $J^{(p)}$ if a way could be found to deal with the first term $\sum_{\mu,\nu}H\frac{\partial}{\partial g^{\mu\nu}}\sqrt{g}p^{\mu\nu}$.

Ultimately, the justification for choosing E as the energy expression depends, of course, on the possibility of a physical interpretation of this expression. As we shall see, for Hilbert this meant an interpretation in terms of Mie's theory. But, first of all, he had to show that E can be represented as a sum of divergences. For this purpose, Hilbert introduced yet another decomposition of $J^{(p)}$, derived from a generalization of (37). As we have indicated earlier, this equation may be identified as a special case of a "polarization" of the Lagrangian H with respect to the contravariant form of the metric $g^{\mu\nu}$: If one takes an arbitrary contravariant tensor $h^{\mu\nu}$, one obtains for the "first polar" of H:

$$J^{(h)} = \sum_{\mu,\nu}\frac{\partial H}{\partial g^{\mu\nu}}h^{\mu\nu} + \sum_{\mu,\nu,k}\frac{\partial H}{\partial g_k^{\mu\nu}}h_k^{\mu\nu} + \sum_{\mu,\nu,k,l}\frac{\partial H}{\partial g_{kl}^{\mu\nu}}h_{kl}^{\mu\nu}. \tag{41}$$

Applying integration by parts to this expression, Hilbert obtained:

60 Die Grundlagen der Physik II, Ms. Vorlesung WS 1916/17, 186 ff. (D. Hilbert, Bibliothek des Mathematischen Seminars, Universität Göttingen); from here on "WS 1916/17 Lectures."

$$\sqrt{g}J^{(h)} = -\sum_{\mu,\nu} H\frac{\partial\sqrt{g}}{\partial g^{\mu\nu}}h^{\mu\nu} + \sum_{\mu,\nu}[\sqrt{g}H]_{\mu\nu}h^{\mu\nu} + D^{(h)}; \tag{42}$$

here

$$[\sqrt{g}H]_{\mu\nu} = \frac{\partial\sqrt{g}H}{\partial g^{\mu\nu}} - \sum_k \frac{\partial}{\partial w_k}\frac{\partial\sqrt{g}H}{\partial g_k^{\mu\nu}} + \sum_{k,l}\frac{\partial^2}{\partial w_k \partial w_l}\frac{\partial\sqrt{g}H}{\partial g_{kl}^{\mu\nu}} \tag{43}$$

is the Lagrangian variational derivative of H, the vanishing of which is the set of gravitational field equations; and:

$$\begin{aligned} D^{(h)} = {} & \sum_{\mu,\nu,k}\frac{\partial}{\partial w_k}\Big(\frac{\partial\sqrt{g}H}{\partial g_k^{\mu\nu}}h^{\mu\nu}\Big) + \sum_{\mu,\nu,k,l}\frac{\partial}{\partial w_k}\Big(\frac{\partial\sqrt{g}H}{\partial g_{kl}^{\mu\nu}}h_l^{\mu\nu}\Big) \\ & - \sum_{\mu,\nu,k,l}\frac{\partial}{\partial w_l}\Big(h^{\mu\nu}\frac{\partial}{\partial w_k}\frac{\partial\sqrt{g}H}{\partial g_{kl}^{\mu\nu}}\Big) \end{aligned} \tag{44}$$

i.e. another divergence expression. Obviously, $J^{(h)}$ turns into $J^{(p)}$ if one sets $h^{\mu\nu}$ equal to $p^{\mu\nu}$, thus yielding the desired alternative decomposition:

$$\sqrt{g}J^{(p)} = -\sum_{\mu,\nu} H\frac{\partial\sqrt{g}}{\partial g^{\mu\nu}}p^{\mu\nu} + D^{(h)}\big|_{h=p}. \tag{45}$$

Comparing (45) with (38), it becomes clear that E indeed can be written as a divergence, and thus represents a candidate for the energy expression. In the Proofs this conclusion is presented as one of two properties justifying this designation:

> Call the expression E the energy form. To justify this designation, I prove two properties that the energy form enjoys.
>
> If we substitute the tensor $p^{\mu\nu}$ for $h^{\mu\nu}$ in identity (6) [i.e. (42)] then, taken together with (9) [i.e. (39)] it follows, provided the gravitational equations (8) [i.e. (51) below] are satisfied:
>
> $$E = (D^{(h)})_{h=p} - D^{(p)} \tag{46}$$
>
> or
>
> $$\begin{aligned} E = \sum\Big\{ & \frac{\partial}{\partial w_k}\Big(\sqrt{g}\frac{\partial H}{\partial g_k^{\mu\nu}}g_s^{\mu\nu}p^s\Big) - \frac{\partial}{\partial w_k}\Big(\frac{\partial}{\partial w_l}\Big(\sqrt{g}\frac{\partial H}{\partial g_{kl}^{\mu\nu}}\Big)g_s^{\mu\nu}p^s\Big) \\ & + \frac{\partial}{\partial w_l}\Big(\sqrt{g}\frac{\partial H}{\partial g_{kl}^{\mu\nu}}g_{sk}^{\mu\nu}p^s\Big)\Big\}, \end{aligned} \tag{47}$$
>
> that is, we have the proposition:
>
> Proposition 1: In virtue of the gravitational equations the energy form E becomes a sum of differential quotients with respect to w_s, that is, it acquires the character of a divergence.[61]

Whereas (47) for an arbitrary H involves an arbitrary combination of electromagnetic and gravitational contributions, Hilbert makes an ansatz $H = K + L$ that allows him to separate these two contributions; in particular, to relate E to his result concerning the energy-momentum tensor of Mie's theory. Accordingly, at this point, he presumably introduces in a missing part of the Proofs (as he does in the corresponding part of Paper 1) the splitting of the Lagrangian (16), and introduces the condition that L not depend on $g_s^{\mu\nu}$.[62] Finally, he writes down explicitly the electromagnetic part of the energy:

> Because K depends only on $g^{\mu\nu}, g_s^{\mu\nu}, g_{lk}^{\mu\nu}$, therefore in ansatz (17) [i.e. (16)], due to (13) [i.e. (47)], the energy E can be expressed solely as a function of the gravitational potentials $g^{\mu\nu}$ and their derivatives, provided L is assumed to depend not on $g_s^{\mu\nu}$, but only on $g^{\mu\nu}, q_s, q_{sk}$. On this assumption, which we shall always make in the following, the definition of the energy (10) [i.e. (39)] yields the expression
>
> $$E = E^{(g)} + E^{(e)}, \tag{48}$$
>
> where the "gravitational energy" $E^{(g)}$ depends only on $g^{\mu\nu}$ and their derivatives, and the "electrodynamic energy" $E^{(e)}$ takes the form
>
> $$E^{(e)} = \sum_{\mu,\nu,s} \frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}}(g_s^{\mu\nu}p^s - g^{\mu s}p_s^{\nu} - g^{\nu s}p_s^{\mu}), \tag{49}$$
>
> which proves to be *a general invariant multiplied by* $\sqrt{g}$.[63]

(The term in parentheses in equation (49) is $p^{\mu\nu}$, the Lie derivative of the contravariant metric with respect to the vector p^s.)

Hilbert's final expression (49) satisfies what we called his "physical criterion" for finding a definition of the energy since the term $\frac{\partial}{\partial g^{\mu\nu}}\sqrt{g}L$ corresponds—apart from the factor -2 —to the left-hand side of (36), and thus to Mie's energy momentum ten-

---

61 "Der Ausdruck E heiße die Energieform. Um diese Bezeichnung zu rechtfertigen, beweise ich zwei Eigenschaften, die der Energieform zukommen.
Setzen wir in der Identität (6) [i.e. (42)] für $h^{\mu\nu}$ den Tensor $p^{\mu\nu}$ ein, so folgt daraus zusammen mit (9) [(39)], sobald die Gravitationsgleichungen (8) erfüllt sind: [(46); (12) in the original text] or [(47); (13) in the original text] d. h. es gilt der Satz:
Satz 1. Die Energieform E wird vermöge der Gravitationsgleichungen einer Summe von Differentialquotienten nach w_s gleich, d. h. sie erhält Divergenzcharakter." See (Proofs, 6).

62 Compare (Hilbert 1916, 402) with (Proofs, 8), and see the discussion in "Einstein Equations and Hilbert Action …" (in this volume).

63 "Da K nur von $g^{\mu\nu}, g_s^{\mu\nu}, g_{lk}^{\mu\nu}$ abhängt, so läßt sich beim Ansatz (17) die Energie E wegen (13) lediglich als Funktion der Gravitationspotentiale $g^{\mu\nu}$ und deren Ableitungen ausdrücken, sobald wir L nicht von $g_s^{\mu\nu}$, sondern nur von $g^{\mu\nu}, q_s, q_{sk}$ abhängig annehmen. Unter dieser Annahme, die wir im Folgenden stets machen, liefert die Definition der Energie (10) den Ausdruck [(48); (18) in the original text] wo die "Gravitationsenergie" $E^{(g)}$ nur von $g^{\mu\nu}$ und deren Ableitungen abhängt und die "elektrodynamische Energie" $E^{(e)}$ die Gestalt erhält [(49); (19) in the original text] in der sie sich als *eine mit* $\sqrt{g}$ *multiplizierte allgemeine Invariante* erweist." (Proofs, 8)

sor. Hilbert's definition of energy had thus been given a "physical justification" in terms of Mie's theory. But—apart from merely formal similarities—its relation to energy-momentum conservation in classical and special-relativistic theories remains entirely unclear. In the Proofs, as we shall see below, Hilbert's energy expression served still another and even more important function, that of determining admissible coordinate systems.

3.4 Hilbert's Revision of Mie's Program and the Roots of his Leitmotiv in Einstein's Work

Apparently Hilbert was convinced that the relation he established between the variational derivative of the Lagrangian and the energy-momentum tensor (see (36)) singled out Mie's theory as having a special relation to the theory of gravitation.[64] In fact, as we have seen, this conclusion is only justified insofar as one imposes on the electrodynamic term in the Lagrangian the condition that it does not depend on $g_s^{\mu\nu}$. Nevertheless, this result apparently suggested to Hilbert that gravitation may be the more fundamental physical process and that it might be possible to conceive of electromagnetic phenomena as "effects of gravitation."[65] Such an interpretation, which was in line with the reductionist perspective implied by his understanding of the axiomatization of physics, led to a revision of Mie's original aim of basing all of physics on electromagnetism.

In the light of this possibility, the third point of Hilbert's initial research program, the question of the number of independent equations in a generally-covariant theory, must have taken on a new and increased significance. Einstein's hole argument, when applied to Hilbert's formalism, suggests that the fourteen generally-covariant field equations for the 14 gravitational and electromagnetic potentials do not have a unique solution for given boundary values. Consequently, 4 identities must exist between the 14 field equations; and 4 additional, non-covariant equations would be required in order to assure a unique solution; and if these 4 identities were somehow equivalent to the 4 equations for the electromagnetic potentials, then the latter could be considered as a consequence of the 10 gravitational equations by virtue of the unique properties of a generally-covariant variational principle, and Hilbert would indeed be entitled to claim that electromagnetism is an effect of gravitation.

As we have seen, the non-uniqueness of solutions to generally-covariant field equations and the conclusion that such field equations must obey 4 identities, are both issues raised by Einstein in his publications of 1913/14. These writings and his 1915 Göttingen lectures, which Hilbert attended, offered rich sources of information about Einstein's theory. In addition the physicist Paul Hertz, then a participant in the group

64 In fact, this relation between the special-relativistic stress-energy tensor and the variational derivative of the general-relativistic generalization of a Lagrangian giving rise to this stress-energy tensor is quite general, as was pointed out many years later in (Rosenfeld 1940, 1–30; and Belinfante 1939, 887). See also (Vizgin 1989, 304; 1994).

65 See (Proofs, 3) and (Hilbert 1916, 397).

centered around Hilbert in Göttingen, may also have kept Hilbert informed about Einstein's thinking on these issues. For example, in a letter to Hertz of August 1915, Einstein raised the problem of solving hyperbolic partial differential equations for arbitrary boundary values and discussed the necessity of introducing four additional equations to restore causality for a set of generally-covariant field equations.[66]

Einstein's treatment of these issues thus forms the background to the crucial theorem, on which Hilbert's entire approach is based, his *Leitmotiv*, labelled "Theorem I" in the Proofs:

> The guiding motive for setting up the theory is given by the following theorem, the proof of which I shall present elsewhere.
>
> Theorem I. If J is an invariant under arbitrary transformations of the four world parameters, containing n quantities and their derivatives, and if one forms from
>
> $$\delta\int J\sqrt{g}d\tau = 0 \tag{50}$$
>
> the n variational equations of Lagrange with respect to each of the n quantities, then in this invariant system of n differential equations for the n quantities there are always four that are a consequence of the remaining $n-4$ —in the sense that, among the n differential equations and their total derivatives, there are always four linear and mutually independent combinations that are satisfied identically.[67]

For a Lagrangian H depending on the gravitational and the electrodynamic potentials and their derivatives, Hilbert derived 10 field equations for the gravitational potentials $g^{\mu\nu}$ and 4 for the electrodynamic potentials q_s from such a variational principle (50):

$$\frac{\partial\sqrt{g}H}{\partial g^{\mu\nu}} = \sum_k \frac{\partial}{\partial w_k}\frac{\partial\sqrt{g}H}{\partial g_k^{\mu\nu}} - \sum_{k,l}\frac{\partial^2}{\partial w_k \partial w_l}\frac{\partial\sqrt{g}H}{\partial g_{kl}^{\mu\nu}}, \quad (\mu, \nu = 1, 2, 3, 4), \tag{51}$$

$$\frac{\partial\sqrt{g}H}{\partial q_h} = \sum_k \frac{\partial}{\partial w_k}\frac{\partial\sqrt{g}H}{\partial q_{hk}}, \quad (h = 1, 2, 3, 4). \tag{52}$$

---

66 Einstein to Paul Hertz, 22 August 1915, (CPAE 8, 163–164). See (Howard and Norton 1993) for an extensive historical discussion.

67 "Das Leitmotiv für den Aufbau der Theorie liefert der folgende mathematische Satz, dessen Beweis ich an einer anderen Stelle darlegen werde.
Theorem I. Ist J eine Invariante bei beliebiger Transformation der vier Weltparameter, welche n Größen und ihre Ableitungen enthält, und man bildet dann aus [(50)] in Bezug auf jene n Größen die n Lagrangeschen Variationsgleichungen, so sind in diesem invarianten System von n Differentialgleichungen für die n Größen stets vier eine Folge der $n-4$ übrigen — in dem Sinne, daß zwischen den n Differentialgleichungen und ihren totalen Ableitungen stets vier lineare, von einander unabhängige Kombinationen identisch erfüllt sind." (Proofs, 2–3) See (Hilbert 1916, 396–397). See (Rowe 1999) for a discussion of the debate on Hilbert's Theorem I among Göttingen mathematicians.

In both the Proofs and Paper 1, Hilbert erroneously claimed that one can consider the last four equations to be a consequence of the 4 identities that must hold, according to his Theorem I, between the 14 differential equations:

> Let us call equations (4) [i.e. (51)] the fundamental equations of gravitation, and equations (5) [i.e. (52)] the fundamental electrodynamic equations, or generalized Maxwell equations. Due to the theorem stated above, the four equations (5) [i.e. (52)] can be viewed as a consequence of equations (4) [i.e. (51)]; that is, because of that mathematical theorem we can immediately assert the claim *that in the sense explained above electrodynamic phenomena are effects of gravitation*. I regard this insight as the simple and very surprising solution of the problem of Riemann, who was the first to search for a theoretical connection between gravitation and light.[68]

We shall come back to this claim later, in connection with Hilbert's proof of a special case of Theorem I.

The fact that Hilbert did not give a proof of this theorem makes it difficult to assess its heuristic roots. No doubt, of course, some of these roots lay in Hilbert's extensive mathematical knowledge, in particular, of the theory of invariants. But the lack of a proof in Paper 1, as well as the peculiar interpretation of it in the Proofs, make it plausible that the theorem also had roots in Einstein's hole argument on the ambiguity of solutions to generally-covariant field equations.

In fact, in the Proofs, Hilbert placed the implications of Theorem I for his field theory in the context of the problem of causality, as Einstein had done for the hole argument. But while the hole argument was formulated in terms of a boundary value problem for a closed hypersurface, Hilbert posed the question of causality in terms of an initial value problem for an open one, thus adapting it to Cauchy's theory of systems of partial differential equations:

> Since our mathematical theorem shows that the axioms I and II [essentially amounting to the variational principle (50), see the discussion below] considered so far can produce only ten essentially independent equations; and since, on the other hand, if general invariance is maintained, more than ten essentially independent equations for the 14 potentials $g_{\mu\nu}$, q_s are not at all possible; therefore—provided that we want to retain the determinate character of the basic equation of physics corresponding to Cauchy's theory of differential equations— the demand for four further non-invariant equations in addition to (4) [i.e. (51)] and (5) [i.e. (52)] is imperative.[69]

Hilbert's counting of needed equations closely parallels Einstein's: the number of field equations (10 in Einstein's case and 14 in Hilbert's) plus 4 coordinate restrictions to make sure that causality is preserved. Since Hilbert, in contrast to Einstein,

---

68 "Die Gleichungen (4) mögen die Grundgleichungen der Gravitation, die Gleichungen (5) die elektrodynamischen Grundgleichungen oder die verallgemeinerten Maxwellschen Gleichungen heißen. Infolge des oben aufgestellten Theorems können die vier Gleichungen (5) als eine Folge der Gleichungen (4) angesehen werden, d. h. wir können unmittelbar wegen jenes mathematischen Satzes die Behauptung aussprechen, *daß in dem bezeichneten Sinne die elektrodynamischen Erscheinungen Wirkungen der Gravitation sind*. In dieser Erkenntnis erblicke ich die einfache und sehr überraschende Lösung des Problems von Riemann, der als der Erste theoretisch nach dem Zusammenhang zwischen Gravitation und Licht gesucht hat." (Proofs, 3; Hilbert 1916, 397–398)

had started from a generally-covariant variational principle, he obtained, in addition, 4 identities that, he claimed, imply the electrodynamic equations (52).

Additional evidence for our conjecture that Einstein's hole argument was one of the roots of Hilbert's theorem (and thus of its later elaboration by Emmy Noether) is provided by other contemporary writings of Hilbert, which will be discussed below in connection with Hilbert's second paper, in which the problem of causality is addressed explicitly.[70]

3.5 Energy-Momentum Conservation and Coordinate Restrictions

As we shall see in this section, the Proofs show that Hilbert was convinced that causality requires four supplementary non-covariant equations to fix the admissible coordinate systems. In identifying these coordinate restrictions, he again followed closely in Einstein's tracks. As did the latter, Hilbert invoked energy-momentum conservation in order to justify physically the choice of a preferred reference frame. After formulating his version of energy-momentum conservation, he introduced the following axiom:

> Axiom III (axiom of space and time). *The spacetime coordinates are those special world parameters for which the energy theorem* (15) [i.e. (57) below] *is valid*.
>
> According to this axiom, space and time in reality provide a special labeling of the world's points such that the energy theorem holds.
>
> Axiom III implies the existence of equations (16) [$d^{(g)}\sqrt{g}H \,/\, dw_s = 0$]: these four differential equations (16) complete the gravitational equations (4) [i.e. (51)] to give a system of 14 equations for the 14 potentials $g_{\mu\nu}$, q_s, *the system of fundamental equations of physics*. Because of the agreement in number between equations and potentials to be determined, the principle of causality for physical processes is also guaranteed, revealing to us the closest connection between the energy theorem and the principle of causality, since each presupposes the other.[71]

The strategy Hilbert followed to extract these coordinate restrictions from the requirement of energy conservation closely followed that of Einstein's *Entwurf* theory of 1913/14. Even before he developed the hole argument, energy-momentum conservation played a crucial role in justifying the lack of general covariance of his

---

69 "Indem unser mathematisches Theorem lehrt, daß die bisherigen Axiome I und II für die 14 Potentiale nur zehn wesentlich von einander unabhängige Gleichungen liefern können, andererseits bei Aufrechterhaltung der allgemeinen Invarianz mehr als zehn wesentlich unabhängige Gleichungen für die 14 Potentiale $g_{\mu\nu}$, q_s garnicht möglich sind, so ist, wofern wir der Cauchyschen Theorie der Differentialgleichungen entsprechend den Grundgleichungen der Physik den Charakter der Bestimmtheit bewahren wollen, die Forderung von vier weiteren zu (4) und (5) hinzutretenden nicht invarianten Gleichungen unerläßlich." (Proofs, 3–4)

70 See, e.g., his SS 1916 Lectures, in particular p. 108, as well as an undated typescript preserved at Göttingen, in SUB Cod. Ms. 642, entitled *Das Kausalitätsprinzip in der Physik*, henceforth cited as the "Causality Lecture." Page 4 of this typescript, describing a construction equivalent to Einstein's hole argument, is discussed below.

gravitational field equations. He was convinced that energy-momentum conservation actually required a restriction of the covariance group.[72] An the beginning of 1914, after having formulated the hole argument, he described the connection between coordinate restrictions and energy-momentum conservation in the *Entwurf* theory as follows:

> Once we have realized that an acceptable theory of gravitation necessarily implies a specialization of the coordinate system, it is also easily seen that the gravitational equations given by us are based upon a special coordinate system. Differentiation of equations (II) with respect to x_ν [the field equations in the form
>
> $$\sum_{\alpha\beta\mu} \frac{\partial}{\partial x_\alpha}\Big(\sqrt{-g}\gamma_{\alpha\beta} g_{\sigma\mu} \frac{\partial \gamma_{\mu\nu}}{\partial x_\beta}\Big) = \kappa(\mathfrak{T}_{\sigma\nu} + \mathfrak{t}_{\sigma\nu})]$$
>
> and summation over ν, and taking into account equations (III), [the conservations equations in the form
>
> $$\sum_\nu \frac{\partial}{\partial x_\nu}(\mathfrak{T}_{\sigma\nu} + \mathfrak{t}_{\sigma\nu}) = 0\,] \tag{53}$$
>
> yields the relations (IV)
>
> $$\Bigg[\sum_{\alpha\beta\mu\nu} \frac{\partial^2}{\partial x_\nu \partial x_\alpha}\Big(\sqrt{-g}\gamma_{\alpha\beta} g_{\sigma\mu} \frac{\partial \gamma_{\mu\nu}}{\partial x_\beta}\Big) = 0\Bigg], \tag{54}$$
>
> that is, four differential conditions for the quantities $g_{\mu\nu}$, which we write in the abbreviated form
>
> $$B_\sigma = 0. \tag{55}$$
>
> These quantities B_σ do not form a generally-covariant vector, as will be shown in §5. From this one can conclude that the equations $B_\sigma = 0$ represent a real restriction on the choice of coordinate system.[73]

In a later 1914 paper, Einstein discussed the physical significance and the transformation properties of the gravitational energy-momentum term $\mathfrak{t}_\sigma^\nu$:

> According to the considerations of §10, the equations (42 c) [i.e. (53)] represent the conservation laws of momentum and energy for matter and gravitational field combined. The $\mathfrak{t}_\sigma^\nu$ are those quantities, related to the gravitational field, which are analogies in physical

71 "Axiom III (Axiom von Raum und Zeit). *Die Raum-Zeitkoordinaten sind solche besonderen Weltparameter, für die der Energiesatz (15) gültig ist.*
Nach diesem Axiom liefern in Wirklichkeit Raum und Zeit eine solche besondere Benennung der Weltpunkte, daß der Energiesatz gültig ist.
Das Axiom III hat das Bestehen der Gleichungen (16) zur Folge: diese vier Differentialgleichungen (16) vervollständigen die Gravitationsgleichungen (4) zu einem System von 14 Gleichungen für die 14 Potentiale $g^{\mu\nu}$, q_s: *dem System der Grundgleichungen der Physik*. Wegen der Gleichzahl der Gleichungen und der zu bestimmenden Potentiale ist für das physikalische Geschehen auch das Kausalitätsprinzip gewährleistet, und es enthüllt sich uns damit der engste Zusammenhang zwischen dem Energiesatz und dem Kausalitätsprinzip, indem beide sich einander bedingen." (Proofs, 7)

72 See, e.g., (Einstein 1913, 1258).

> interpretation to the components $\mathfrak{T}_\sigma^\nu$ of the energy tensor (V-Tensor) [i.e. tensor density]. It is to be emphasized that the $\mathfrak{t}_\sigma^\nu$ do not have tensorial covariance under arbitrary admissible [coordinate] transformations but only under linear transformations. Nevertheless, we call ($\mathfrak{t}_\sigma^\nu$) the energy tensor of the gravitational field.[74]

Similarly, Hilbert notes that his energy-form is invariant with respect to linear transformations; he shows that E can be decomposed with respect to the vector p^j as follows (Proofs, 6):

$$E = \sum_s e_s p^s + \sum_{s,l} e_s^l p_l^s \tag{56}$$

where e_s and e_s^l are independent of p^j. If one compares this expression with Einstein's (53), then the analogy between the two suggests that the two-index object e_s^l should play the same role in Hilbert's theory as does the total energy-momentum tensor in Einstein's theory, satisfying a divergence equation of the form:

$$\sum_l \frac{\partial e_s^l}{\partial w_l} = 0. \tag{57}$$

Hilbert shows that this equation holds only if e_s vanishes, in which case:

$$E = \sum_{s,l} e_s^l p_l^s. \tag{58}$$

This equation can be related to energy conservation; Hilbert calls this the "normal form" of the energy. The fact that the last two equations imply each other was, for Hilbert, apparently a decisive reason for calling E the energy form. Indeed, this equivalence is the subject of his second theorem about the energy-form. Although the relevant part of the Proofs is missing,[75] Hilbert's theorem and its proof can be reconstructed:

---

73 "Nachdem wir so eingesehen haben, daß eine brauchbare Gravitationstheorie notwendig einer Spezialisierung des Koordinatensystems bedarf, erkennen wir auch leicht, daß bei den von uns angegebenen Gravitationsgleichungen ein spezielles Koordinatensystem zugrunde liegt. Aus den Gleichungen (II) folgen nämlich durch Differentiation nach x_ν und Summation über ν unter Berücksichtigung der Gleichungen (III) die Beziehungen (IV) also vier Differentialbedingungen für die Größen $g_{\mu\nu}$, welche wir abgekürzt $B_\sigma = 0$ schreiben wollen.
Diese Größen B_σ bilden, wie in §5 gezeigt ist, keinen allgemein-kovarianten Vektor. Hieraus kann geschlossen werden, daß die Gleichungen $B_\sigma = 0$ eine wirkliche Bedingung für die Wahl des Koordinatensystems darstellen." (Einstein and Grossmann 1914, 218–219)

74 "Die Gleichungen (42 c) drücken nach den in §10 gegebenen Überlegungen die Erhaltungssätze des Impulses und der Energie für Materie und Gravitationsfeld zusammen aus. $\mathfrak{t}_\sigma^\nu$ sind diejenigen auf das Gravitationsfeld bezüglichen Größen, welche den Komponenten $\mathfrak{T}_\sigma^\nu$ des Energietensors (V-Tensors) [i.e. tensor density] der physikalischen Bedeutung nach analog sind. Es sei hervorgehoben, daß die $\mathfrak{t}_\sigma^\nu$ nicht beliebigen berechtigten, sondern nur linearen Transformationen gegenüber Tensorkovarianz besitzen; trotzdem nennen wir ($\mathfrak{t}_\sigma^\nu$) den Energietensor des Gravitationsfeldes." (Einstein 1914b, 1077)

75 The top portion of the Proofs, p. 7, is missing.

Theorem 2 must have asserted that:

$$e_s = \frac{\partial e_s^l}{\partial w_l} \tag{59}$$

This assertion is easily proven by following the lines indicated in the surviving portion of Hilbert's argument. From (38) and (56) it follows that:

$$\sqrt{g}J^{(p)} + H\frac{\partial\sqrt{g}}{\partial g^{\mu\nu}}p^{\mu\nu} = e_s p^s + e_s^l p_l^s + D^{(p)}, \tag{60}$$

which can be rewritten as:

$$\sqrt{g}J^{(p)} + H\frac{\partial\sqrt{g}}{\partial g^{\mu\nu}}p^{\mu\nu} = \left(e_s - \frac{\partial e^l{}_s}{\partial w_l}\right)p^s + \overline{D^{(p)}}, \tag{61}$$

where $\overline{D^{(p)}}$ is still a divergence. If now the integral over a region Ω, on the boundary of which p^s and its first derivative vanish, is taken on both sides, then the surface terms vanish. Thus one obtains in view of (42):

$$\int_\Omega [\sqrt{g}H]_{\mu\nu}p^{\mu\nu}dx^4 = \int_\Omega \left(e_s - \frac{\partial e^l{}_s}{\partial w_l}\right)p^s(dx^4). \tag{62}$$

But the left-hand side vanishes when the gravitational field equations hold, and p^s is an arbitrary vector field, from which (59) follows.

Theorem 2 provides Hilbert with the desired coordinate restrictions:

> This theorem shows that the divergence equation corresponding to the energy theorem of the old theory
>
> $$\sum_l \frac{\partial e_s^l}{\partial w_l} = 0 \tag{63}$$
>
> holds if and only if the four quantities e_s vanish ... .[76]

After these preparations, Hilbert introduces Axiom III, quoted at the beginning of this section, which establishes a distinction between the arbitrary world parameters w_l and the restricted class of coordinates that constitute "a spacetime reference system." In fact, the latter are those world parameters satisfying the coordinate restrictions $e_s = 0$ following from Hilbert's energy condition. In analogy to the "justified coordinate transformations" of Einstein's 1913/14 theory leading from one "adapted

76 "Dieser Satz zeigt, daß die dem Energiesatz der alten Theorie entsprechende Divergenzgleichung [(63); (15) in the original text] dann und nur dann gelten kann, wenn die vier Größen e_s verschwinden ..." (Proofs, 7).

coordinate system" to another, Hilbert introduced spacetime transformations that lead from one "normal form" of the energy to another:

> To the transition from one spacetime reference system to another one corresponds the transformation of the energy form from one so-called "normal form"
>
> $$E = \sum_{s,l} e_s^l p_l^s \tag{64}$$
>
> to another normal form.[77]

The claim that Hilbert's introduction of coordinate restrictions was guided by the goal of recovering the ordinary divergence form of energy-momentum conservation is supported by his later use of this argument in a discussion with Felix Klein. In a letter to Hilbert, Klein recounted how, at a meeting of the Göttingen Academy, he had argued that, for the energy balance of a field, one should take into account only the energy tensor of matter (including that of the electromagnetic field) without ascribing a separate energy-momentum tensor to the gravitational field.[78] This suggestion was taken up by Carl Runge, who had given an expression for energy-momentum conservation that, in his letter to Hilbert, Klein called "regular" and found similar to what happens in the "elementary theory."[79] Starting from an expression for the covariant divergence of the stress-energy tensor:

$$\sum_{\mu\nu}\left(\sqrt{g}T_{\mu\nu}g_{\sigma}^{\mu\nu} + 2\frac{\partial}{\partial w_{\nu}}(\sqrt{g}T_{\mu\sigma}g^{\mu\nu})\right) = 0 \qquad \sigma = 1, 2, 3, 4 \tag{65}$$

Runge obtained his "regular" expression by imposing the four equations:

$$\sum_{\mu\nu}\sqrt{g}T_{\mu\nu}g_{\sigma}^{\mu\nu} = 0, \tag{66}$$

thus specifying a preferred class of coordinate systems. In his response, Hilbert sent Klein three pages of the Proofs to show that he had anticipated Runge's line of reasoning:

> I send you herewith my first proofs [footnote: Please kindly return these to me as I have no other record of them.] (3 pages) of my first communication, in which I also implemented Runge's ideas; in particular with theorem 1, p. 6, in which the divergence character of the energy is proven. I later omitted the whole thing as the thing did not seem to me to be fully mature. I would be very pleased if progress could now be made. For this it is necessary to retrieve the old energy conservation laws in the limiting case of Newtonian theory.[80]

---

77 "Dem Übergang von einem Raum-Zeit-Bezugssystem zu einem anderen entspricht die Transformation der Energieform von einer sogenannten "Normalform" [(64)] auf eine andere Normalform." (Proofs, 7)

78 Felix Klein to David Hilbert, 5 March 1918, (Frei 1985, 142–143).

79 For a discussion of Runge's work, see (Rowe 1999).

Hilbert's final sentence confirms that the recovery of the familiar form of energy conservation was his goal. However, at the time of the Proofs, it was clearly not his aim to eliminate the energy-momentum expression of the gravitational field from the energy balance, as the above reference to Runge might suggest. On the contrary, as we have seen above (see (48)), Hilbert followed Einstein in attempting to treat the contributions to the total energy from the electromagnetic and the gravitational parts on an equal footing.

In summary, Hilbert's first steps in the realization of his research program were the derivation of what he regarded as the unique relation between the variational derivative of Mie's Lagrangian and Mie's energy momentum tensor, and the formulation of a theorem, by means of which he hoped to show that the electromagnetic field equations follow from the gravitational ones. Albeit problematic from a modern perspective, these steps become understandable in the context of Hilbert's application of his axiomatic approach to Einstein's non-covariant theory of gravitation and Mie's theory of matter. These first steps in turn shaped Hilbert's further research. They effected a change of perspective from viewing electrodynamics and gravitation on an equal footing to his vision of deriving electromagnetism from gravitation. As a consequence, the structure of Hilbert's original, non-covariant theory, in spite of the covariance of Hilbert's gravitational equations and the different physical interpretation that he gave to his equations, is strikingly similar to that of Einstein's 1913/14 *Entwurf* theory of gravitation.

3.6 Electromagnetism as an Effect of Gravitation: The Core of Hilbert's Theory

Now we come to the part of Hilbert's program that today is often considered to contain his most important contributions to general relativity: the contracted Bianchi identities and a special case of Noether's theorem. We shall show that, in the original version of Hilbert's theory, these mathematical results actually constituted part of a different physical framework that also affected their interpretation. In a later section, we shall see how these results were transformed, primarily due to the work of Hendrik Antoon Lorentz and Felix Klein, into constituents of general relativity. In the hindsight of general relativity, it appears as if Hilbert first derived the contracted Bianchi identities, applied them to the gravitational field equations with an electromagnetic source-term, and then showed that the electrodynamic variables necessarily satisfy the Maxwell equations. This last result, however, is valid only under addi-

80 "Anbei schicke ich Ihnen meine erste Korrektur [footnote: Bitte dieselbe mir wieder freundlichst zustellen zu wollen, da ich sonst keine Aufzeichnungen habe.] (3 Blätter) meiner ersten Mitteilung, in der ich gerade die Ideen von Runge auch ausgeführt hatte; insbesondere auch mit Satz l, S. 6, in dem der Divergenzcharakter der Energie bewiesen wird. Ich habe aber die ganze Sache später unterdrückt, weil die Sache mir nicht reif erschien. Ich würde mich sehr freuen, wenn jetzt der Fortschritt gelänge. Dazu ist aber nötig im Grenzfalle zur Newtonschen Theorie die alten Energiesätze wiederzufinden." Tilman Sauer suggested that the pages sent to Klein were the three sheets of the Proofs bearing Roman numbers I, II, and III, see (Sauer 1999, 544).

tional assumptions that run counter to Mie's program. From the point of view of general relativity, Hilbert obtained Maxwell's equations as a consequence of the integrability conditions for the gravitational field equations with electromagnetic source term, as if he had treated a special case of Einstein's equations and expressed certain of their general properties in terms of this special case. From Hilbert's point of view, however, he had derived the electrodynamic equations as a consequence of the gravitational ones; his derivation was closely interwoven with other results of his theory that pointed to electromagnetism as an effect of gravitation. For him, the equation, on the basis of which he argued that electrodynamics is a consequence of gravitation, was a result of four ingredients, two of which are other links between gravitation and electrodynamics, and all of which are based on his generally-covariant variational principle:

- a general theorem corresponding to the contracted Bianchi identities,
- the field equations following from the variational principle,
- the relation between Mie's energy-momentum tensor and the variational derivative of the Lagrangian, and
- the way in which the derivatives of the electrodynamic potentials enter Mie's Lagrangian.

In the Proofs, the general theorem is:

> Theorem III. If J is an invariant depending only on the $g^{\mu\nu}$ and their derivatives and if, as above, the variational derivatives of $\sqrt{g}J$ with respect to $g^{\mu\nu}$ are denoted by $[\sqrt{g}J]_{\mu\nu}$, then the expression — in which $h^{\mu\nu}$ is understood to be any contravariant tensor —
>
> $$\frac{1}{\sqrt{g}}\sum_{\mu,\nu}[\sqrt{g}J]_{\mu\nu}h^{\mu\nu} \tag{67}$$
>
> represents an invariant; if in this sum we substitute in place of $h^{\mu\nu}$ the particular tensor $p^{\mu\nu}$ and write
>
> $$\sum_{\mu,\nu}[\sqrt{g}J]_{\mu\nu}p^{\mu\nu} = \sum_{s,l}(i_s p^s + i_s^l p_l^s), \tag{68}$$
>
> where then the expressions
>
> $$\begin{aligned} i_s &= \sum_{\mu,\nu}[\sqrt{g}J]_{\mu\nu}g_s^{\mu\nu}, \\ i_s^l &= -2\sum_{\mu}[\sqrt{g}J]_{\mu s}g^{\mu l} \end{aligned} \tag{69}$$
>
> depend only on the $g^{\mu\nu}$ and their derivatives, then we have
>
> $$i_s = \sum_l \frac{\partial i_s^l}{\partial w_l} \tag{70}$$
>
> in the sense, that this equation is identically fulfilled for all arguments, that is for the $g^{\mu\nu}$ and their derivatives.[81]

Here, (68) follows from an explicit calculation taking into account the definition of $p^{\mu\nu}$; the identity (70) follows if in analogy to (61) one rewrites (68) as:

$$[\sqrt{g}J]_{\mu\nu}p^{\mu\nu} = \left(i_s - \frac{\partial i_s^l}{\partial w_l}\right)p^s + \frac{\partial}{\partial w_l}(i_s^l p^s), \tag{71}$$

and, as in the earlier derivation, carries out a surface integration. Theorem III, in the form of (70), thus corresponds to the contracted Bianchi identities.

Hilbert next applies Theorem III to the Lagrangian $H = K + L$ using his knowledge about its electrodynamic part (see the last two "ingredients" listed above) in order to extract the electrodynamic equations from the identity for L that corresponds to (70). From a modern point of view, it is remarkable that Hilbert did not consider the physical significance of this identity for the gravitational part K of the Lagrangian, but only for the electrodynamic part. For Hilbert, however, this was natural; presumably he was convinced, on the basis of Theorem I, that generally-covariant equations for gravitation are impossible as a "stand-alone" theory. Consequently, it simply made no sense to interpret the gravitational part of these equations by itself.

Assuming the split of the Lagrangian into $K + L$, the gravitational and electrodynamic parts as in (16), he rewrites (51) as:[82]

$$[\sqrt{g}K]_{\mu\nu} + \frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}} = 0. \tag{72}$$

He next applies (69) to the invariant K:

$$i_s = \sum_{\mu,\nu}[\sqrt{g}K]_{\mu\nu}g_s^{\mu\nu}, \tag{73}$$

and

$$i_s^l = -2\sum_{\mu}[\sqrt{g}K]_{\mu s}g^{\mu l}, \quad (\mu = 1, 2, 3, 4). \tag{74}$$

From the modern point of view, it would be natural to invoke the identity (70) in order to derive its implications for the source term of the gravitational field equations, i.e., the second term of (72) in Hilbert's notation. In this way, one would obtain an integrability condition for the gravitational field equations that can be interpreted as representing energy-momentum conservation.

---

81 "Theorem III. Wenn J eine nur von den $g^{\mu\nu}$ und deren Ableitungen abhängige Invariante ist, und, wie oben, die Variationsableitungen von $\sqrt{g}J$ bezüglich $g^{\mu\nu}$ mit $[\sqrt{g}J]_{\mu\nu}$ bezeichnet werden, so stellt der Ausdruck — unter $h^{\mu\nu}$ irgend einen kontravarianten Tensor verstanden — [(67)] eine Invariante dar; setzen wir in dieser Summe an Stelle von $h^{\mu\nu}$ den besonderen Tensor $p^{\mu\nu}$ ein und schreiben [(68)] wo alsdann die Ausdrücke [(69)] lediglich von den $g^{\mu\nu}$ und deren Ableitungen abhängen, so ist [(70)] in der Weise, daß diese Gleichung identisch für alle Argumente, nämlich die $g^{\mu\nu}$ und deren Ableitungen, erfüllt ist." (Proofs, 9; Hilbert 1916, 399)

82 See (Proofs, 11; Hilbert 1916, 405).

Hilbert proceeded differently, using Theorem III to further elaborate what he considered his crucial insight into the relation between Mie's energy-momentum tensor and the variational derivative of L. Consequently he focussed on (36), from which he attempted to extract the equations for the electromagnetic field. In fact, the left-hand side of this equation can (in view of (72) and (74)) be rewritten as $-i_\nu^m$. Consequently, differentiating the right-hand side of (36) with respect to w_m and summing over m, Theorem III yields:

$$\begin{aligned} i_\nu &= \sum_m \frac{\partial}{\partial w_m}\Big(-\sqrt{g}L\delta_\nu^m + \frac{\partial\sqrt{g}L}{\partial q_m}q_\nu + \sum_s \frac{\partial\sqrt{g}L}{\partial M_{sm}}M_{s\nu}\Big) \\ &= -\frac{\partial\sqrt{g}L}{\partial w_\nu} + \sum_m \Big\{ q_\nu \frac{\partial}{\partial w_m}\Big([\sqrt{g}L]_m + \sum_s \frac{\partial}{\partial w_s}\frac{\partial\sqrt{g}L}{\partial q_{ms}}\Big) \\ &\quad + q_{\nu m}\Big([\sqrt{g}L]_m + \sum_s \frac{\partial}{\partial w_s}\frac{\partial\sqrt{g}L}{\partial q_{ms}}\Big)\Big\} \\ &\quad + \sum_s \Big([\sqrt{g}L]_s - \frac{\partial\sqrt{g}L}{\partial q_s}\Big)M_{s\nu} + \sum_{s,m}\frac{\partial\sqrt{g}L}{\partial M_{sm}}\frac{\partial M_{s\nu}}{\partial w_m}, \end{aligned} \tag{75}$$

where use has been made of:

$$\frac{\partial\sqrt{g}L}{\partial q_m} = [\sqrt{g}L]_m + \sum_s \frac{\partial}{\partial w_s}\frac{\partial\sqrt{g}L}{\partial q_{ms}} \tag{76}$$

and

$$-\sum_m \frac{\partial}{\partial w_m}\frac{\partial\sqrt{g}L}{\partial q_{sm}} = [\sqrt{g}L]_s - \frac{\partial\sqrt{g}L}{\partial q_s}. \tag{77}$$

Here $[\sqrt{g}L]_h$ denotes the Lagrangian derivative of $\sqrt{g}L$ with respect to the electrodynamic potentials q_h:

$$[\sqrt{g}L]_h = \frac{\partial\sqrt{g}L}{\partial q_h} - \sum_k \frac{\partial}{\partial w_k}\frac{\partial\sqrt{g}L}{\partial q_{hk}}; \tag{78}$$

the vanishing of which constitutes the electromagnetic field equations. At this point Hilbert makes use of the last ingredient, the special way in which the derivatives of the potentials enter Mie's Lagrangian. Taking into account (27), one obtains:

$$\sum_{m,s} \frac{\partial^2}{\partial w_m \partial w_s}\frac{\partial\sqrt{g}L}{\partial q_{ms}} = 0, \tag{79}$$

so that (75) can be rewritten as:

$$i_\nu = -\frac{\partial\sqrt{g}L}{\partial w_\nu} + \sum_m \Big(q_\nu \frac{\partial}{\partial w_m}[\sqrt{g}L]_m + M_{m\nu}[\sqrt{g}L]_m\Big) + \sum_m \frac{\partial\sqrt{g}L}{\partial q_m} q_{m\nu} + \sum_{s,m} \frac{\partial\sqrt{g}L}{\partial M_{sm}} \frac{\partial M_{s\nu}}{\partial w_m}. \tag{80}$$

While the right-hand side of this equation only involves the electrodynamic part of the Lagrangian, in view of (73) this is not the case for the left-hand side. Therefore, Hilbert once more uses the field equations, in the form of (72), for i_ν to obtain an expression entirely in terms of the electrodynamic part of the Lagrangian. For this purpose, he first writes:

$$-\frac{\partial\sqrt{g}L}{\partial w_\nu} = -\sum_{s,m} \frac{\partial\sqrt{g}L}{\partial g^{sm}} g_\nu^{sm} - \sum_m \frac{\partial\sqrt{g}L}{\partial q_m} q_{m\nu} - \sum_{m,s} \frac{\partial\sqrt{g}L}{\partial q_{ms}} \frac{\partial q_{ms}}{\partial w_\nu}, \tag{81}$$

and then uses (72) and (73) to identify the first term on the right-hand side as i_ν. Hilbert thus reaches his goal of transforming the identity following from Theorem III into an equation involving only the electromagnetic potentials. A further simplification results from noting that the last term on the right-hand side of (81) is, apart from its sign, identical to the last term of (80). (This is because:

$$\sum_{s,m} \frac{\partial\sqrt{g}L}{\partial M_{sm}} \Big(\frac{\partial M_{s\nu}}{\partial w_m} - \frac{\partial q_{ms}}{\partial w_\nu}\Big) = 0, \tag{82}$$

which follows from the definition (28) of M_{sm}.)

Finally, using (80), Hilbert obtains:

$$\sum_m \Big(M_{m\nu}[\sqrt{g}L]_m + q_\nu \frac{\partial}{\partial w_m}[\sqrt{g}L]_m\Big) = 0. \tag{83}$$

Summarizing what he had achieved, Hilbert claimed:

> ... from the gravitational equations (4) [i.e. (51)] there follow indeed the four linearly independent combinations (32) [i.e. (83)] of the basic electrodynamic equations (5) [i.e. (52)] and their first derivatives. *This is the entire mathematical expression of the general claim made above about the character of electrodynamics as an epiphenomenon of gravitation*.[83]

On closer inspection, Hilbert's claim turns out to be problematic. One might try to interpret it in either of two ways: the electromagnetic field equations follow either differentially or algebraically from (83).

---

83 "... aus den Gravitationsgleichungen (4) folgen in der Tat die vier von einander unabhängigen linearen Kombinationen (32) der elektrodynamischen Grundgleichungen (5) und ihrer ersten Ableitungen. *Dies ist der ganze mathematische Ausdruck der oben allgemein ausgesprochenen Behauptung über den Charakter der Elektrodynamik als einer Folgeerscheinung der Gravitation*." (Proofs, 12) In (Hilbert 1916, 406), "ganze" [entire] is corrected to "genaue" [exact] in the last sentence.

In the first case one would have to show that, if these equations hold on an initial hypersurface $w_4 = const$, then they hold everywhere off that hypersurface by virtue of the identities (83). Indeed it follows from these identities that, if these equations hold on $w_4 = 0$:

$$\frac{\partial[\sqrt{g}L]_4}{\partial w_4} = 0, \tag{84}$$

so that, by iteration, $[\sqrt{g}L]_4 = 0$ holds everywhere provided that it holds initially and that the other three field equations hold everywhere. But the time derivatives of the other three field equations,

$$\frac{\partial[\sqrt{g}L]_m}{\partial w_4} \qquad m = 1, 2, 3 \tag{85}$$

remain unrestricted by the identity so that one cannot simply give the electromagnetic field equations on an initial hypersurface and have them continue to hold automatically off it as a consequence of (83).

In the second case, it is clear that the field equations can only hold algebraically by virtue of (83) if the second term vanishes; this implies that the theory is gauge invariant, i.e. that the potentials themselves do not enter the field equations. In that case one indeed obtains an additional identity from gauge invariance:

$$\frac{\partial[\sqrt{g}L]_m}{\partial w_m} = 0. \tag{86}$$

(In the usual Maxwell theory this is the identity that guarantees conservation of the charge-current vector.) However, this cannot have been the argument Hilbert had in mind when stating his claim. First of all, he did not introduce the additional assumptions required—and could not have introduced them because they violated his physical assumptions;[84] and second he did not derive the identity for gauge-invariant electromagnetic Lagrangians that makes this argument work. As illustrated by Klein's later work, the derivation of these identities is closely related to a different perspective on Hilbert's results, a perspective in which electromagnetism is no longer, as in Hilbert's Proofs, treated as an epiphenomenon of gravitation, but in which both are treated in parallel.[85]

In summary, Hilbert's claim that the electromagnetic equations are a consequence of the gravitational ones turns out to be an interpretation forced upon his mathematical results by his overall program rather than being implied by them. In any case, this

84 Mie's original theory is in fact not gauge invariant, and in the version adopted by Hilbert one of the invariants involves a function of the electromagnetic potential vector, see (33).

85 Compare Klein's attempt to derive analogous equations for the gravitational and the electromagnetic potentials, from which the Maxwell equations then are derived, (Klein 1917, 472–473).

interpretation is different from that given to the corresponding results in general relativity and usually associated with Hilbert's work.

3.7 The Deductive Structure of the Proofs Version

Having attempted to reconstruct the line of reasoning Hilbert followed while developing the original version of his theory, we now summarize the way in which he presented these results in the Proofs. This serves as a review of the deductive structure of his theory, indicating which results were emphasized by Hilbert, and facilitating a comparison between the Proofs and the published versions.

We begin by recalling the elements of this deductive structure that Hilbert introduced explicitly:

- Axiom I "Mie's Axiom von der Weltfunktion," (see (19))
- Axiom II "Axiom von der allgemeinen Invarianz," (see the passage below (19))
- Axiom III "Axiom von Raum und Zeit," (see the passage above (55))
- Theorem I, Hilbert's *Leitmotiv*, (see (50))
- Theorem II, Lie derivative of the Lagrangian, (see (23))
- Theorem III, contracted Bianchi identities, (see (70))
- Proposition 1, divergence character of the energy expression, (see (47))
- Proposition 2, identity obeyed by the components of the energy expression, (see (59)).

He also used the following assumptions, introduced as part of his deductive structure without being explicitly stated:

- vanishing of the divergence of the energy expression (see (63))
- splitting of the Lagrangian into gravitational and electrodynamical terms (see (16))
- the assumption that the electrodynamical term does not depend on the derivatives of the metric tensor (see (25)).

There are, furthermore, the following physical results, not labelled as theorems:

- the field equations (see (51) and (52))
- the energy expression (see (39)) and the related coordinate restrictions (see (63))
- the form of Mie's Lagrangian (see (27))
- the relation between Mie's energy tensor and Lagrangian (see (36))
- the relation between the electromagnetic and gravitational field equations (see (83)).

The exposition of Hilbert's theory in the Proofs can be subdivided into four sections, to which we give short titles and list under each the relevant elements of his theory:

1. Basic Framework (Proofs, 1–3)

   Axioms I and II, Theorem I, and the field equations for gravitation and electromagnetism

2. Causality and the Energy Expression (Proofs, 3–8)

   the energy expression, Propositions 1 and 2, the divergence character of the energy expression, Axiom III, the coordinate restrictions, the split of the Lagrangian into gravitational and electrodynamical terms, and the structure of the electrodynamical term

3. Basic Theorems (Proofs, 8–9)

   Theorems II and III

4. Implications for Electromagnetism (Proofs, 9–13)

   the form of Mie's Lagrangian, its relation to his energy tensor, and the relation between electromagnetic and gravitational field equations.

The sequence in which Hilbert presented these elements suggests that he considered its implications for electromagnetism as the central results of his theory. Indeed, the gravitational field equations are never explicitly given and only briefly considered at the beginning as part of the general framework, whereas the presentation concludes with three results concerning Mie's theory. The centrality of these electromagnetic implications for him is also clear from his introductory and concluding remarks. Hilbert's initial discussion mentions Mie's electrodynamics first, and closes with the promise of further elaboration of the consequences of his theory for electrodynamics:

> The far reaching ideas and the formation of novel concepts by means of which Mie constructs his electrodynamics, and the prodigious problems raised by Einstein, as well as his ingeniously conceived methods of solution, have opened new paths for the investigation into the foundations of physics.
>
> In the following—in the sense of the axiomatic method—I would like to develop from three simple axioms a new system of basic equations of physics, of ideal beauty, containing, I believe, the solution of the problems presented. I reserve for later communications the detailed development and particularly the special application of my basic equations to the fundamental questions of the theory of electricity.[86]

In his conclusion, Hilbert makes clear what he had in mind here: a solution of the riddles of atomic physics:

> As one can see, the few simple assumptions expressed in axioms I, II, III suffice with appropriate interpretation to establish the theory: through it not only are our views of space, time, and motion fundamentally reshaped in the sense called for by Einstein, but I am also convinced that through the basic equations established here the most

86 "Die tiefgreifenden Gedanken und originellen Begriffsbildungen vermöge derer Mie seine Elektrodynamik aufbaut, und die gewaltigen Problemstellungen von Einstein sowie dessen scharfsinnige zu ihrer Lösung ersonnenen Methoden haben der Untersuchung über die Grundlagen der Physik neue Wege eröffnet.
Ich möchte im Folgenden—im Sinne der axiomatischen Methode—aus drei einfachen Axiomen ein neues System von Grundgleichungen der Physik aufstellen, die von idealer Schönheit sind, und in denen, wie ich glaube, die Lösung der gestellten Probleme enthalten ist. Die genauere Ausführung sowie vor allem die spezielle Anwendung meiner Grundgleichungen auf die fundamentalen Fragen der Elektrizitätslehre behalte ich späteren Mitteilungen vor." (Proofs, 1)

> intimate, hitherto hidden processes in the interior of atoms will receive an explanation; and in particular that generally a reduction of all physical constants to mathematical constants must be possible—whereby the possibility approaches that physics in principle becomes a science of the type of geometry: surely the highest glory of the axiomatic method, which, as we have seen, here takes into its service the powerful instruments of analysis, namely the calculus of variations and the theory of invariants.[87]

Hilbert's final remarks about the status of his theory *vis à vis* Einstein's work on gravitation strikingly parallel Minkowski's assessment of the relation of his four-dimensional formulation to Einstein's special theory; not just providing a mathematical framework for existing results, but developing a genuinely novel physical theory, which, properly understood, turns out to be a part of mathematics.[88]

Fig. 1 provides a graphical survey of the deductive structure of Hilbert's theory. The main elements listed above are connected by arrows; mathematical implications are represented by straight arrows and inferences based on heuristic reasoning by curved arrows. As the figure shows, apart from the field equations, Hilbert's results can be divided into two fairly distinct clusters: one comprises the implications for electromagnetism (right-hand side of the diagram); the other, the implications for the understanding of energy conservation (left-hand side of the diagram). While the assertions concerning energy conservation are not essential for deriving the other results, they depend on practically all the other parts of this theory. The main link between the two clusters is clearly Theorem I. Although no assertion of Hilbert's theory is derived directly from Theorem I, it motivates both the relation between energy conservation and coordinate restrictions and the link between electromagnetism and gravitation.

The analysis of the deductive structure of Hilbert's theory thus confirms that Theorem I is indeed the *Leitmotiv* of the theory. The two clusters of results obviously are also related to what he considered the two main physical touchstones of his theory: Mie's theory of electromagnetism and energy conservation. On the other hand, neither Newton's theory of gravitation nor any other parts of mechanics are mentioned by Hilbert. Einstein's imprint on Hilbert's theory was more of a mathematical or structural nature than a physical one.

---

87 "Wie man sieht, genügen bei sinngemäßer Deutung die wenigen einfachen in den Axiomen I, II, III ausgesprochenen Annahmen zum Aufbau der Theorie: durch dieselbe werden nicht nur unsere Vorstellungen über Raum, Zeit und Bewegung von Grund aus in dem von Einstein geforderten Sinne umgestaltet, sondern ich bin auch der Überzeugung, daß durch die hier aufgestellten Grundgleichungen die intimsten, bisher verborgenen Vorgänge innerhalb des Atoms Aufklärung erhalten werden und insbesondere allgemein eine Zurückführung aller physikalischen Konstanten auf mathematische Konstanten möglich sein muß—wie denn überhaupt damit die Möglichkeit naherückt, daß aus der Physik im Prinzip eine Wissenschaft von der Art der Geometrie werde: gewiß der herrlichste Ruhm der axiomatischen Methode, die hier wie wir sehen die mächtigen Instrumente der Analysis nämlich, Variationsrechnung und Invariantentheorie, in ihre Dienste nimmt." (Proofs, 13)

88 For Minkowski, see (Walter 1999).

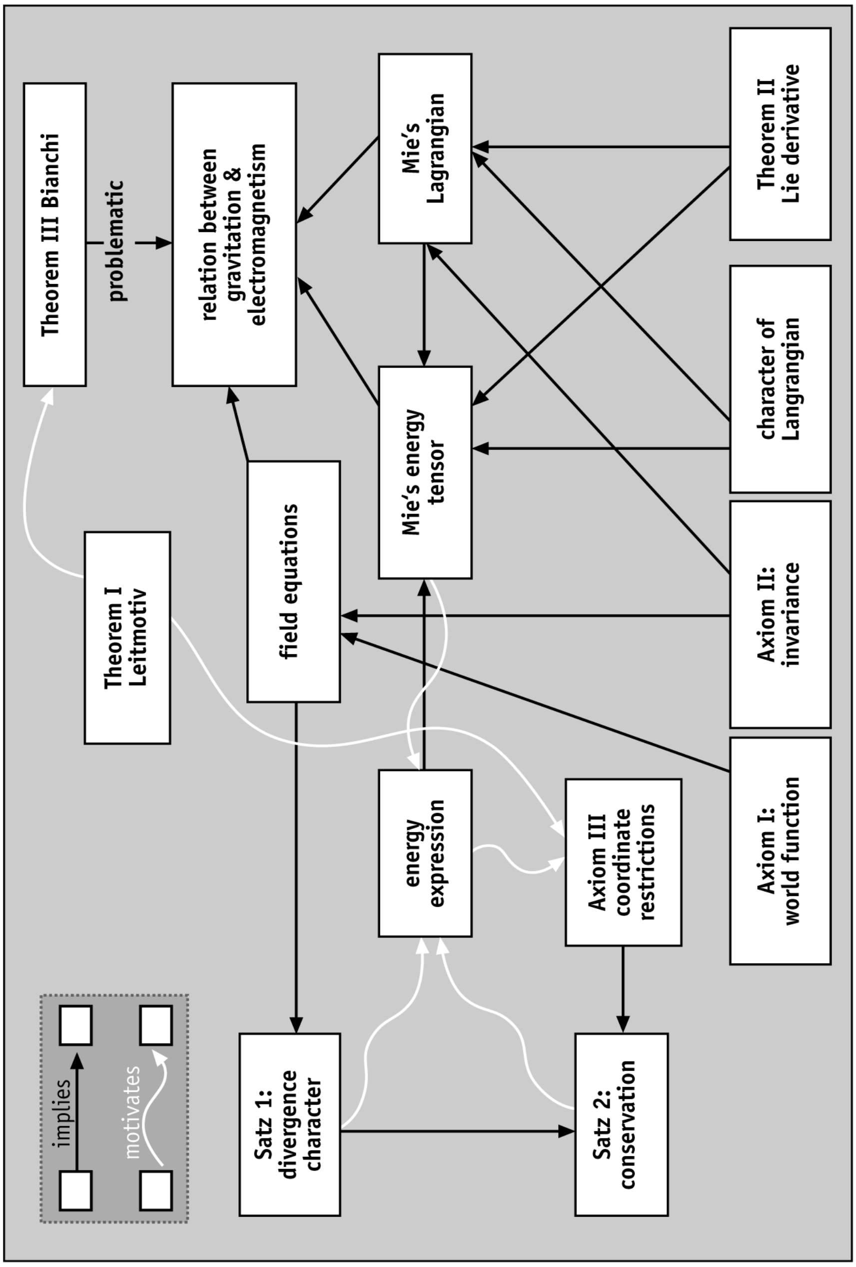

Figure 1: Deductive Structure of the Proofs (1915)

4. HILBERT'S PHYSICS AND EINSTEIN'S MATHEMATICS: THE EXCHANGE OF LATE 1915

4.1 What Einstein Could Learn From Hilbert

The Hilbert-Einstein correspondence begins with Einstein's letter of 7 November 1915.[89] That November was the month during which Einstein's theory of gravitation underwent several dramatic changes documented by four papers he presented to the Prussian Academy, culminating in the definitive version of the field equations in the paper submitted 25 November.[90] On 4 November Einstein submitted his first note, in which he abandoned the *Entwurf* field equations and replaced them with equations derived from the Riemann tensor (Einstein 1915a); he included the proofs of this paper in his letter to Hilbert. In spite of this radical modification of the field equations, the structure of Einstein's theory remained essentially unchanged from that of the non-covariant 1913 *Entwurf* theory. In both, the requirement of energy-momentum conservation is linked to a restriction to adapted coordinate systems. In Einstein's 4 November paper, this restriction implies the following equation (Einstein 1915a, 785):

$$\sum_{\alpha\beta}\frac{\partial}{\partial x_\alpha}\left(g^{\alpha\beta}\frac{\partial lg\sqrt{-g}}{\partial x_\beta}\right) = -\kappa\sum_\sigma T_\sigma^\sigma. \tag{87}$$

Einstein pointed out one immediate consequence for the choice of an adapted coordinate system:

> Equation (21a) [i.e. (87)] shows the impossibility of so choosing the coordinate system that $\sqrt{-g}$ equals 1, because the scalar of the energy tensor cannot be set to zero.[91]

That the scalar [i.e. the trace] of the energy-momentum tensor cannot vanish is obvious if one takes Einstein's standard example (a swarm of non-interacting particles or incoherent "dust") as the source of the gravitational field: the trace of its energy-momentum tensor equals the mass density of the dust. However, the physical meaning of condition (87) was entirely obscure. It was therefore incumbent upon Einstein to find a physical interpretation of it or to modify his theory once more in order to get rid of it. He soon succeeded in doing both, and formulated his new view in an addendum to the first note, published on 11 November (Einstein 1915b).

On 12 November 1915 he reported his success to Hilbert:

> For the time being, I just thank you cordially for your kind letter. Meanwhile, the problem has made new progress. Namely, it is possible to compel *general* covariance by means of the postulate $\sqrt{-g} = 1$; Riemann's tensor then furnishes the gravitational

---

89 Einstein to David Hilbert, 7 November 1915, (CPAE 8, 191).

90 See (Einstein 1915e).

91 "Aus Gleichung (21a) [i.e. (87)] geht hervor, daß es unmöglich ist, das Koordinatensystem so zu wählen, daß $\sqrt{-g}$ gleich 1 wird; denn der Skalar des Energietensors kann nicht zu null gemacht werden." (Einstein 1915a, 785)

> equations directly. If my present modification (which does not change the equations) is legitimate, then gravitation must play a fundamental role in the structure of matter. My own curiosity is impeding my work![92]

What had happened? Einstein had noticed that the condition $\sum_\sigma T^\sigma_\sigma = 0$, which follows from setting $\sqrt{-g} = 1$ in (87), can be related to an electromagnetic theory of matter: in Maxwell's theory, the vanishing of its trace is a characteristic property of the electromagnetic energy-momentum tensor. Thus, if one assumes all matter to be of electromagnetic origin, the vanishing of its trace becomes a fundamental property of the energy-momentum tensor. This has two important consequences: Condition (87) is no longer an inexplicable restriction on the admissible coordinate systems, and the 4 November field equations can be seen as a particular form of generally-covariant field equations based on the Ricci tensor. From the perspective of the 11 November revision, the condition $\sqrt{-g} = 1$ turns out to be nothing more than an arbitrary but convenient choice of coordinate systems.

The core of Einstein's new theory is strikingly simple. The left-hand side of the gravitational field equations is now simply the Ricci tensor and the right-hand side an energy-momentum tensor, the trace of which has to vanish:[93]

$$R_{\mu\nu} = -\kappa T_{\mu\nu} \qquad \sum_\sigma T^\sigma_\sigma = 0. \tag{88}$$

What distinguishes these field equations from the final equations presented on 25 November is an additional term on the right-hand side of the equations involving the trace of the energy-momentum tensor, which now need not vanish:[94]

$$R_{\mu\nu} = -\kappa\left(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T\right). \tag{89}$$

Remarkably enough, in the winter of 1912/13 Einstein had considered the linearized form of these field equations, but discarded them because they were not compatible with his expectation of how the Newtonian limit should result.[95] He had also then considered and rejected field equations of the form (88), just because they imply

---

92 "Ich danke einstweilen herzlich für Ihren freundlichen Brief. [Das] Problem hat unterdessen einen neuen Fortschritt gemacht. Es lässt sich nämlich durch das Postulat $\sqrt{-g} = 1$ die *allgemeine* Kovarianz erzwingen; der Riemann'sche Tensor liefert dann direkt die Gravitationsgleichungen. Wenn meine jetzige Modifikation (die die Gleichungen nicht ändert) berechtigt ist, dann muss die Gravitation im Aufbau der Materie eine fundamentale Rolle spielen. Die Neugier erschwert mir die Arbeit!" Einstein to David Hilbert, 12 November 1915, (CPAE 8, 194).

93 See (Einstein 1915b, 801 and 800).

94 See (Einstein 1915e, 845).

95 See Doc. 10 of (CPAE 4), "Pathways out of Classical Physics …", "Einstein's Zurich Notebook", (both in vol. 1 of this series), and the "Commentary" (in vol. 2).

the condition $\sum_{\sigma} T_{\sigma}^{\sigma} = 0$. At that time, this condition seemed unacceptable because the trace of the energy-momentum tensor of ordinary matter does *not* vanish.

The prehistory of Einstein's 11 November paper thus confronts us with a puzzle: Why did he consider it to be such a decisive advance beyond his 4 November paper and not just a possible alternative interpretation of his previous results; and why did he now so readily accept the trace-condition $\sum_{\sigma} T_{\sigma}^{\sigma} = 0$ that earlier had led him to reject this very theory? What impelled Einstein's change of perspective in November 1915?

The answer seems to lie in the changed context, within which Einstein formulated his new approach: in particular, his interaction with Hilbert. As will become evident, it would have been quite uncharacteristic of him to adopt the new approach so readily had it not been for current discussions of the electrodynamic worldview and his feeling that he was now in competition with Hilbert.[96]

In his addendum, Einstein directly referred to the supporters of the electrodynamic worldview:

> One now has to remember that, in accord with our knowledge, "matter" is not to be conceived as something primitively given, or physically simple. There even are those, and not just a few, who hope to be able to reduce matter to purely electrodynamic processes, which of course would have to be done in a theory more complete than Maxwell's electrodynamics.[97]

Only this context explains Einstein's highly speculative and fragmentary comments on an electromagnetic model of matter. That, in November 1915, Einstein conceived of a field theory of matter as a goal in its own right is also supported by his correspondence, which makes it clear that this perspective was shaped by his rivalry with Hilbert. We have already cited Einstein's letter to Hilbert, in which he wrote:

> If my present modification (which does not change the equations) is legitimate, then gravitation must play a fundamental role in the structure of matter. My own curiosity is impeding my work![98]

And when, in a letter of 14 November, Hilbert claimed to have achieved the unification of gravitation and electromagnetism, Einstein responded:

---

96 For a discussion of Hilbert's reaction to what he must have seen as an intrusion by Einstein into his domain, see (Sauer 1999, 542–543).

97 "Es ist nun daran zu erinnern, daß nach unseren Kenntnissen die "Materie" nicht als ein primitiv Gegebenes, physikalisch Einfaches aufzufassen ist. Es gibt sogar nicht wenige, die hoffen, die Materie auf rein elektromagnetische Vorgänge reduzieren zu können, die allerdings einer gegenüber Maxwells Elektrodynamik vervollständigten Theorie gemäß vor sich gehen würden." (Einstein 1915b, 799)

98 "Wenn meine jetzige Modifikation (die die Gleichungen nicht ändert) berechtigt ist, dann muss die Gravitation im Aufbau der Materie eine fundamentale Rolle spielen. Die Neugier erschwert mir die Arbeit!" Einstein to David Hilbert, 12 November 1915, (CPAE 8, 194).

> Your investigation interests me tremendously, especially since I often racked my brain to construct a bridge between gravitation and electromagnetics.[99]

A few days later (after calculating the perihelion shift on the basis of the new theory), he expressed himself similarly:

> In these last months I had great success in my work. *Generally covariant* gravitation equations. *Perihelion motions explained quantitatively.* The role of gravitation in the structure of matter. You will be astonished. I worked dreadfully hard; it is remarkable that one can sustain it.[100]

When one examines Einstein's previous writings on gravitation, published and unpublished, one finds no trace of an attempt to unify gravitation and electromagnetism. He had never advocated the electromagnetic worldview. On the contrary, he was apparently disinterested in Mie's attempt at a unification of gravitation and electrodynamics, not finding it worth mentioning in his 1913 review of contemporary gravitation theories.[101]

And soon after completion of the final version of general relativity, Einstein reverted to his earlier view that general relativity could make no assertions about the structure of matter:

> From what I know of Hilbert's theory, it makes use of an assumption about electrodynamic processes that—apart from the treatment of the gravitational field—is closely connected to Mie's. Such a specialized approach is not in accordance with the point of view of general relativity. The latter actually only provides the gravitational field law, and quite unambiguously so when general covariance is required.[102]

Einstein's mid-November 1915 pursuit of a relation between gravitation and electromagnetism was, then, merely a short-lived episode in his search for a relativistic theory of gravitation. Its novelty is confirmed by a footnote in the addendum:

> In writing the earlier paper, I had not yet realized that the hypothesis $\sum T_{\mu}^{\mu} = 0$ is, in principle, admissible.[103]

---

99 "Ihre Untersuchung interessiert mich gewaltig, zumal ich mir schon oft das Gehirn zermartert habe, um eine Brücke zwischen Gravitation und Elektromagnetik zu schlagen." Einstein to David Hilbert, 15 November 1915, (CPAE 8, 199).

100 "Ich habe mit grossem Erfolg gearbeitet in diesen Monaten. *Allgemein kovariante* Gravitationsgleichungen. *Perihelbewegungen quantitativ erklärt*. Rolle der Gravitation im Bau der Materie. Du wirst staunen. Gearbeitet habe ich schauderhaft angestrengt; sonderbar, dass man es aushält." Einstein to Michele Besso, 17 November 1915, (CPAE 8, 201).

101 See (Einstein 1913).

102 "Soviel ich von Hilbert's Theorie weiss, bedient sie sich eines Ansatzes für das elektrodynamische Geschehen, der sich [— a]bgesehen von der Behandlung des Gravitationsfeldes — eng an Mie anschliesst. Ein derartiger spezieller Ansatz lässt sich aus dem Gesichtspunkte der allgemeinen Relativität nicht begründen. Letzterer liefert eigentlich nur das Gesetz des Gravitationsfeldes, und zwar ganz eindeutig, wenn man allgemeine Kovarianz fordert." Einstein to Arnold Sommerfeld, 9 December 1915, (CPAE 8, 216).

103 "Bei Niederschrift der früheren Mitteilung war mir die prinzipielle Zulässigkeit der Hypothese $\sum T_{\mu}^{\mu} = 0$ noch nicht zu Bewußtsein gekommen." (Einstein 1915b, 800)

It thus seems quite clear that Einstein's temporary adherence to an electromagnetic theory of matter was triggered by Hilbert's work, which he attempted to use in order to solve a problem that had arisen in his own theory, and that he dropped it when he solved this problem in a different way.

So this whole episode might appear to be a bizarre and unnecessary detour. A closer analysis of the last steps of Einstein's path to general relativity shows, however, that the solution depended crucially on this detour, and hence indirectly on Hilbert's work. In fact, Einstein successfully calculated the perihelion shift of Mercury on the basis of his 11 November theory.[104] The condition $\sqrt{-g} = 1$, implied by the assumption of an electromagnetic origin of matter (see (87)), was essential for this calculation, which Einstein considered a striking confirmation of his audacious hypothesis on the constitution of matter, definitely favoring this theory over that of 4 November.[105] The 11 November theory also turned out to be the basis for a new understanding of the Newtonian limit, which allowed Einstein to accept the field equations of general relativity as the definitive solution to the problem of gravitation. Ironically, Hilbert's most important contribution to general relativity may have been enhancing the credibility of a speculative and ultimately untenable physical hypothesis that guided Einstein's final mathematical steps towards the completion of his theory.

Einstein submitted his perihelion paper on 18 November 1915. In a footnote, appended after its completion, Einstein observed that, in fact, the hypothesis of an electromagnetic origin of matter is unnecessary for the perihelion shift calculation. He announced a further modification of his field equations, finally reaching the definitive version of his theory.[106] On the same day, Einstein wrote to Hilbert, acknowledging receipt of Hilbert's work, including a system of field equations:

> The system [of field equations] you give agrees—as far as I can see—exactly with that which I found in the last few weeks and have presented to the Academy.[107]

---

104 See (Einstein 1915c).

105 See (Einstein 1915d): the abstract of this paper, probably by Einstein, summarizes the issue: "Es wird gezeigt, daß die allgemeine Relativitätstheorie die von Leverrier entdeckte Perihelbewegung des Merkurs qualitativ und quantitativ erklärt. Dadurch wird die Hypothese vom Verschwinden des Skalars des Energietensors der "Materie" bestätigt. Ferner wird gezeigt, daß die Untersuchung der Lichtstrahlenkrümmung durch das Gravitationsfeld ebenfalls eine Möglichkeit der Prüfung dieser wichtigen Hypothese bietet." ("It will be shown that the theory of general relativity explains qualitatively and quantitatively the perihelion motion of Mercury, which was discovered by Leverrier. Thus the hypothesis of the vanishing of the scalar of the energy tensor of "matter" is confirmed. Furthermore, it is shown that the analysis of the bending of light by the gravitational field also offers a way of testing this important hypothesis.")

106 See (Einstein 1915c, 831).

107 "Das von Ihnen gegebene System [of field equations] stimmt - soweit ich sehe - genau mit dem überein, was ich in den letzten Wochen gefunden und der Akademie überreicht habe." Einstein to David Hilbert, 18 November 1915, (CPAE 8, 201–202). For discussion of what Einstein may have received from Hilbert, see below.

Einstein emphasized that the real difficulty had not been the formulation of generally-covariant field equations, but in showing their agreement with a physical requirement: the existence of the Newtonian limit. Stressing his priority, he mentioned that he had considered such equations three years earlier:

> … it was hard to recognize that these equations form a generalisation, and indeed a simple and natural generalisation, of Newton's law. It has just been in the last few weeks that I succeeded in this (I sent you my first communication), whereas 3 years ago with my friend Grossmann I had already taken into consideration the only possible generally covariant equations, which have now been shown to be the correct ones. We had only heavy-heartedly distanced ourselves from it, because it seemed to me that the physical discussion had shown their incompatibility with Newton's law.[108]

Einstein's statement not only characterized his own approach, but indirectly clarified his ambivalent position with regard to Hilbert's theory. While evidently fascinated by the perspective of unifying gravitation and electromagnetism, he now recognized that, at least in Hilbert's case, this involved the risk of neglecting the sound foundation of the new theory of gravitation in the classical theory.

4.2 What Hilbert Could Learn from Einstein

Hilbert must have seen Einstein's letter of 12 November, announcing publication of new insights into a fundamental role of gravitation in the constitution of matter, as a threat to his priority.[109] At any rate, Hilbert hastened public presentation of his results. His response of 13 November gave a brief sketch of his theory and announced a 16th November seminar on it:

> Actually, I wanted first to think of a quite palpable application for physicists, namely valid relations between physical constants, before obliging with my axiomatic solution to your great problem. But since you are so interested, I would like to develop my th[eory] in very complete detail on the coming Tuesday, that is, the day after the day after tomorrow (the 16th of this mo.). I find it ideally beautiful math[ematically], and also insofar as calculations that are not completely transparent do not occur at all, and absolutely compelling in accordance with the axiom[atic] meth[od] and therefore rely on its reality. As a result of a gen. math. theorem, the (generalized Maxwellian) electrody. eqs. appear as a math. consequence of the gravitation eqs., so that gravitation and electrodynamics are actually not at all different. Furthermore, my energy concept forms the basis: $E = \Sigma(e_s t^s + e_{ih} t^{ih})$, [the t^s corresponds to p^s in Hilbert's papers, etc.] which is likewise a general invariant [see (56)], and from this then also follow from a very simple

---

108 "schwer war es, zu erkennen, dass diese Gleichungen eine Verallgemeinerung, und zwar eine einfache und natürliche Verallgemeinerung des Newton'schen Gesetzes bilden. Dies gelang mir erst in den letzten Wochen (meine erste Mitteilung habe ich Ihnen geschickt), während ich die einzig möglichen allgemein kovarianten Gleichungen, [die] sich jetzt als die richtigen erweisen, schon vor 3 Jahren mit meinem Freunde Grossmann in Erwägung gezogen hatte. Nur schweren Herzens trennten wir uns davon, weil mir die physikalische Diskussion scheinbar ihre Unvereinbarkeit mit Newtons Gesetz ergeben hatte."

109 This aspect of the Hilbert-Einstein relationship was first discussed in (Sauer 1999), where the chronology of events is carefully reconstructed.

> axiom the 4 still-missing "spacetime equations" $e_s = 0$. I derived most pleasure in the discovery already discussed with Sommerfeld that the usual electrical energy results when a certain absolute invariant is differentiated with respect to the gravitation potentials and then g are set = 0,1.[110]

This letter presents the essential elements of Hilbert's theory as presented in the Proofs. His reference to "the missing spacetime equations" suggests that he saw these equations and their relation to the energy concept as an issue common to his theory and Einstein's.

Einstein responded on 15 November 1915, declining the invitation to come to Göttingen on grounds of health.[111] Instead, he asked Hilbert for the proofs of his paper. As mentioned above, by 18 November Hilbert had fulfilled Einstein's request. He could not have sent the typeset Proofs, which are dated 6 December, so he must have sent a manuscript on 20 November, presumably corresponding to his talk. Since the Proofs are also dated 20 November, this manuscript may well have presented practically the same version of his theory. On 19 November, a day after Einstein announced his successful perihelion calculation to Hilbert, the latter sent his congratulations, making clear once more that the physical problems facing Hilbert's theory were of a rather different nature:

> Many thanks for your postcard and cordial congratulations on conquering perihelion motion. If I could calculate as rapidly as you, in my equations the electron would correspondingly have to capitulate, and simultaneously the hydrogen atom would have to produce its note of apology about why it does not radiate.
> I would be grateful if you were to continue to keep me up-to-date on your latest advances.[112]

---

110 "Ich wollte eigentlich erst nur für die Physiker eine ganz handgreifliche Anwendung nämlich treue Beziehungen zwischen den physikalischen Konstanten überlegen, ehe ich meine axiomatische Lösung ihres grossen Problems zum Besten gebe. Da Sie aber so interessiert sind, so möchte ich am kommenden Dienstag also über-über morgen (d. 16 d. M.) meine Th. ganz ausführlich entwickeln. Ich halte sie für math. ideal schön auch insofern, als Rechnungen, die nicht ganz durchsichtig sind, garnicht vorkommen. und absolut zwingend nach axiom. Meth., und baue deshalb auf ihre Wirklichkeit. In Folge eines allgem. math. Satzes erscheinen die elektrody. Gl. (verallgemeinerte Maxwellsche) als math. Folge der Gravitationsgl., so dass Gravitation u. Elektrodynamik eigentlich garnichts verschiedenes sind. Desweiteren bildet mein Energiebegriff die Grundlage: $E = \Sigma(e_s t^s + e_{ih} t^{ih})$, die ebenfalls eine allgemeine Invariante ist, und daraus folgen dann aus einem sehr einfachen Axiom die noch fehlenden 4 "Raum-Zeitgleichungen" $e_s = 0$. Hauptvergnügen war für mich die schon mit Sommerfeld besprochene Entdeckung, dass die gewöhnliche elektrische Energie herauskommt, wenn man eine gewisse absolute Invariante mit den Gravitationspotentialen differenziert und dann $g = 0, 1$ setzt." David Hilbert to Einstein, 13 November 1915, (CPAE 8, 195).

111 Einstein to David Hilbert, 15 November 1915, (CPAE 8, 199).

112 "Vielen Dank für Ihre Karte und herzlichste Gratulation zu der Ueberwältigung der Perihelbewegung. Wenn ich so rasch rechnen könnte, wie Sie, müsste bei meinen Gleichg entsprechend das Elektron kapituliren und zugleich das Wasserstoffatom sein Entschuldigungszettel aufzeigen, warum es nicht strahlt. Ich werde Ihnen auch ferner dankbar sein, wenn Sie mich über Ihre neuesten Fortschritte auf dem Laufenden halten." David Hilbert to Einstein, 19 November 1915, (CPAE 8, 202).

No doubt Einstein fulfilled this request to keep Hilbert up to date. His definitive paper on the field equations, submitted 25 November and published 2 December, must have been on Hilbert's desk within a day or two. In contrast to all earlier versions of his theory, Einstein now showed that energy-momentum conservation does not imply additional coordinate restrictions on the field equations (89). He also made clear that these field equations fulfill the requirement of having a Newtonian limit and allow derivation of the perihelion shift of Mercury.

Our analysis of the Proofs suggests that neither the astronomical implications of Einstein's theory nor the latter's treatment of the Newtonian limit directly affected Hilbert's theory since they lay outside its scope, as Hilbert then perceived it. But Einstein's insight that energy-momentum conservation does not lead to a restriction on admissible coordinate systems was of crucial significance for Hilbert. As we have seen, in Hilbert's theory the entire complex of results on energy-momentum conservation was structured by a logic paralleling that of Einstein's earlier non-covariant theory. Moreover, Theorem I, Hilbert's *Leitmotiv*, was motivated by Einstein's hole argument that generally-covariant field equations cannot have unique solutions. His definitive paper of 25 November did not explicitly mention the hole argument, but simply took it for granted that his new generally-covariant field equations avoid such difficulties.[113] Hilbert may well have checked that Einstein's definitive field equations were actually compatible[114] with the equations that follow from Hilbert's variational principle, which he had not explicitly calculated—or at least not included in the Proofs, and this compatibility would certainly have been reassuring for Hilbert. But the fact that the hole argument evidently no longer troubled Einstein must have led Hilbert to question his *Leitmotiv*, with its double role of motivating coordinate restrictions and providing the link between gravitation and electromagnetism.

Thus, Einstein's paper of 25 November 1915 represented a major challenge for Hilbert's theory. As we shall see when discussing the published version of Hilbert's paper, while Einstein temporarily took over Hilbert's physical perspective, Hilbert appears to have accepted the mathematical implications of Einstein's rejection of the hole argument.

4.3 Cooperation in the Form of Competition

In a situation such as we have described, in which the interaction between two people working on closely related problems changes the way in which each of them proceeds, it is not easy for the individuals to assess their own contributions. While Einstein was happy to have found in Hilbert one of the few colleagues, if not the only one, who appreciated and understood the nature of his work on gravitation, he also

---

113 The fact that these equations were supported by Einstein's successful calculation of the perihelion shift made it impossible for Hilbert simply to disregard them.

114 Compatible, but not the same, because of the trace term, and because of the different treatment of the stress-energy tensor, as discussed elsewhere in this paper.

resented the way in which Hilbert took over some of his results without, as Einstein saw it, giving him due credit. Einstein wrote to his friend Heinrich Zangger on 26 November 1915 with regard to his newly-completed theory:

> The theory is beautiful beyond comparison. However, only *one* colleague has really understood it, and he is seeking to "partake" [*nostrifizieren*] in it (Abraham's expression) in a clever way. In my personal experience I have hardly come to know the wretchedness of mankind better than as a result of this theory and everything connected to it. But it does not bother me.[115]

Einstein's reaction becomes particularly understandable in the light of his prior positive experience of collaboration with his friend, the mathematician Marcel Grossmann. Grossmann had restricted himself to putting his superior mathematical competence at Einstein's service.[116] What Hilbert offered was not cooperation but competition. Hilbert may well have been upset by Einstein's anticipation in print, in his paper of 11 November, of what Hilbert felt to be his idea of a close link between gravitation and the structure of matter. Even more disturbing may have been the fact that, contrary to Hilbert's assertion in the Proofs, Einstein's final formulation of his theory required no restriction on general covariance. But it is not clear exactly when Hilbert abandoned all non-covariant elements of his program, in particular his approach to the energy problem and consequent restriction to a preferred class of coordinate systems.[117]

Hilbert evidently learned of Einstein's resentment over lack of recognition by Hilbert, possibly as a result of Einstein's letter of 18 November pointing out his priority in setting up generally-covariant field equations. In any case, he began to introduce changes in his Proofs on or after 6 December, documented by handwritten marginalia, changes which not only acknowledge Einstein's priority but attempt to placate him. Hilbert's revision also provides an indication of the content of Einstein's complaints. He revised the programmatic statement in the introduction of his paper (his insertion is rendered in italics):

> In the following — in the sense of the axiomatic method — I would like to develop, *essentially* from three simple axioms a ~~new~~ system of basic equations of physics, of ideal beauty, containing, I believe, the solution of the problems presented.[118]

---

115 "Die Theorie ist von unvergleichlicher Schönheit. Aber nur *ein* Kollege hat sie wirklich verstanden und der eine sucht sie auf geschickte Weise zu "*nostrifizieren*" (Abraham'scher Ausdruck). Ich habe in meinen persönlichen Erfahrungen kaum je die Jämmerlichkeit der Menschen besser kennen gelernt wie gelegentlich dieser Theorie und was damit zusammenhängt. Es ficht mich aber nicht an." Einstein to Heinrich Zangger, 26 November 1915, (CPAE 8, 205). See the discussion of "nostrification" above.

116 See the editorial note "Einstein on Gravitation and Relativity: The Collaboration with Marcel Grossmann" in (CPAE 4, 294–301).

117 According to (Sauer 1999, 562), Hilbert had found the new energy expression by 25 January 1916.

118 "Ich möchte im Folgenden - im Sinne der axiomatischen Methode - *wesentlich* aus drei einfachen Axiomen ein ~~neues~~ System von Grundgleichungen der Physik aufstellen, die von idealer Schönheit sind, und in denen, wie ich glaube, die Lösung der gestellten Probleme enthalten ist." (Proofs, 1)

The insertion "wesentlich" was presumably motivated by Hilbert's recognition that his theory actually presupposed additional assumptions of substantial content, such as the assumption of a split of the Lagrangian into gravitational and electromagnetic parts and the assumption that the latter does not depend on derivatives of the metric (see section 3). A further assumption was the requirement that the gravitational part of the Lagrangian not involve derivatives of the metric higher than second order. Einstein had justified this requirement by the necessity for the theory to have a Newtonian limit, and it may have been Einstein's argument that drew Hilbert's attention to the fact that his theory was actually based on a much wider array of assumptions than his axiomatic presentation had indicated. More remarkably, in characterizing his system of equations, Hilbert deleted the word "neu," a clear indication that he had read Einstein's 25 November paper and recognized that the equations implied by his own variational principle are formally equivalent (because of where the trace term occurs) to Einstein's if Hilbert's electrodynamic stress-energy tensor is substituted for the unspecified one on the right-hand side of Einstein's field equations.

Hilbert's next change was presumably related to a complaint by Einstein about the lack of proper acknowledgement for what he considered to be one of his fundamental contributions, the introduction of the metric tensor as the mathematical representation of the gravitational potentials. Hilbert had indeed given the impression that Einstein's merit was confined to asking the right questions, while Hilbert provided the answers.

Hilbert's revised description of these gravitational potentials reads (his insertion is again rendered in italics):

> The quantities characterizing the events at w_s shall be:
>
> 1) The ten gravitational potentials *first introduced by <u>Einstein</u>*, $g_{\mu\nu}$ $(\mu, \nu = 1, 2, 3, 4)$ having the character of a symmetric tensor with respect to arbitrary transformation of the world parameter w_s;
>
> 2) The four electrodynamic potentials q_s having the character of a vector in the same sense.[119]

The next change represents an even more far-going recognition that Hilbert could not simply claim the results in his paper as parts of "his theory," as if it had nothing substantial in common with that of Einstein:

> The guiding motive for setting up ~~my~~ *the* theory is given by the following theorem, the proof of which I will present elsewhere.[120]

---

119 "Die das Geschehen in w_s charakterisierenden Größen seien:
1) die zehn *von <u>Einstein</u> zuerst eingeführten* Gravitationspotentiale $g_{\mu\nu}$ $(\mu, \nu = 1, 2, 3, 4)$ mit symmetrischem Tensorcharakter gegenüber einer beliebigen Transformation der Weltparameter w_s;
2) die vier elektrodynamischen Potentiale q_s mit Vektorcharakter im selben Sinne." (Proofs, 1)

120 "Das Leitmotiv für den Aufbau ~~meiner~~ *der* Theorie liefert der folgende mathematische Satz, dessen Beweis ich an einer anderen Stelle darlegen werde." (Proofs, 2)

Hilbert's final marginal notation consists of just an exclamation mark next to a minor correction of the energy expression (39)—perhaps evidence that he had identified this expression as the central problem in the Proofs. While Hilbert's first annotations were presumably intended as revisions of a text that was going to remain basically unchanged, this exclamation mark signals the abandonment of such an attempt at revision. At this point, perhaps it dawned upon Hilbert that Einstein's results forced him to rethink his entire approach.

Hilbert's recognition of the problematic character of his treatment of energy-momentum conservation appears to have been solely in reaction to Einstein's results and not as a consequence of any internal dynamics (see section 3) of the development of his theory.[121] Indeed, as our analysis of the deductive structure of Hilbert's theory showed, this treatment is well anchored in the remainder of his theory without in turn having much effect on the remainder. Hence, there was no "internal friction" that could have driven a further development of Hilbert's theory. On the contrary, since the link between energy-momentum conservation and coordinate restrictions was motivated by Hilbert's Theorem I, Einstein's abandonment of this link left Hilbert at a loss, as we have argued above. But the way in which energy-momentum conservation was connected to other results of his theory also suggested how to modify it in the direction indicated by Einstein: Hilbert had to find a new energy expression that does not imply a coordinate restriction but is still connected with Mie's energy-momentum tensor. Precisely the decoupling of his energy expression from the physical consequences of Hilbert's theory made such a modification possible. Hilbert gave up immediate publication and began to rework his theory. By early 1916 had he arrived at results that made possible this rewriting of his paper and its submission for publication; by mid-February 1916, Paper 1, which we will discuss in the following section, was in press.[122]

Meanwhile, having emerged triumphant from the exchange of November 1915, Einstein offered a reconciliation to Hilbert:

> There has been a certain ill-feeling between us, the cause of which I do not want to analyze. I have struggled against the feeling of bitterness attached to it, and this with complete success. I think of you again with unmarred friendliness and ask you to try to do the same with me. Objectively it is a shame when two real fellows who have extricated themselves somewhat from this shabby world do not afford each other mutual pleasure.[123]

121 For a different view, see (Sauer 1999, 570).

122 For a detailed chronology, see the reconstruction in (Sauer 1999, 560–565).

123 "Es ist zwischen uns eine gewisse Verstimmung gewesen, deren Ursache ich nicht analysieren will. Gegen das damit verbundene Gefühl der Bitterkeit habe ich gekämpft, und zwar mit vollständigem Erfolge. Ich gedenke Ihrer wieder in ungetrübter Freundlichkeit, und bitte Sie, dasselbe bei mir zu versuchen. Es ist objektiv schade, wenn sich zwei wirkliche Kerle, die sich aus dieser schäbigen Welt etwas herausgearbeitet haben, nicht gegenseitig zur Freude gereichen." Einstein to David Hilbert, 20 December 1915, (CPAE 8, 222). The "schäbige [.] Welt" probably refers to World War I—given Einstein and Hilbert's critical attitude to the war.

5. HILBERT'S ASSIMILATION OF EINSTEIN'S RESULTS: THE THREE PUBLISHED VERSIONS OF HIS FIRST PAPER

5.1 The New Energy Concept—An Intermediary Solution

As we have seen, modification of Hilbert's treatment of energy-momentum conservation was the most urgent step necessitated by Einstein's results of 25 November 1915. First of all, the energy-momentum conservation law should not involve coordinate restrictions but be an invariant equation. Second, the modified energy expression should still involve Mie's energy-momentum tensor; otherwise the link between gravitation and electromagnetism, fundamental to Hilbert's program, would be endangered. Third, to accord with Hilbert's understanding of energy-momentum conservation, the new energy concept must still satisfy a divergence equation. As we shall show, Hilbert's modification of his energy expression was guided by these criteria, but its relation to a physical interpretation remained as tenuous as ever.[124] The next section concerns the effect of the new energy concept on the deductive structure of Hilbert's theory.

In the introductory discussion of energy, Paper 1 emphasizes that only axioms I and II are required:

> The most important aim is now the formulation of the concept of energy, and the derivation of the energy theorem solely on the basis of the two axioms I and II.[125]

This emphasis is in contrast with the treatment in the Proofs, in which the energy concept is closely related to axiom III, which was dropped in Paper 1. Hilbert then proceeds exactly as in the Proofs, introducing a polarization of the Lagrangian with respect to the gravitational variables (see the definition of P_g, (20)):

$$P_g(\sqrt{g}H) = \sum_{\mu,\nu,k,l}\left(\frac{\partial\sqrt{g}H}{\partial g^{\mu\nu}}p^{\mu\nu} + \frac{\partial\sqrt{g}H}{\partial g_k^{\mu\nu}}p_k^{\mu\nu} + \frac{\partial\sqrt{g}H}{\partial g_{kl}^{\mu\nu}}p_{kl}^{\mu\nu}\right). \tag{90}$$

In contrast to (37), however, Hilbert polarizes $\sqrt{g}H$ instead of H. Clearly, his aim was to formulate an equation analogous to (45), but with only a divergence term on the right-hand side. Indeed, since:

$$P(\sqrt{g}H) = \sqrt{g}PH + H\sum_{\mu,\nu}\frac{\partial\sqrt{g}}{\partial g^{\mu\nu}}p^{\mu\nu}, \tag{91}$$

use of $\sqrt{g}H$ eliminates the first term of the right-hand side of (45), giving:

124 For a discussion of Hilbert's concept of energy, see also (Sauer 1999, 548–550), which stresses the mathematical roots of this concept.

125 "Das wichtigste Ziel ist nunmehr die Aufstellung des Begriffes der Energie und die Herleitung des Energiesatzes allein auf Grund der beiden Axiome I und II." (Hilbert 1916, 400)

$$P_g(\sqrt{g}H) - \sum_l \frac{\partial\sqrt{g}(a^l + b^l)}{\partial w_l} = \sum_{\mu,\nu} [\sqrt{g}H]_{\mu\nu} p^{\mu\nu}. \tag{92}$$

Since the right-hand side vanishes due to the field equations, this equation is of just the desired form.

The way in which Hilbert obtained (92) closely parallels that used in the Proofs, i.e. by splitting off divergence terms. He starts out by noting that:

$$a^l = \sum_{\mu,\nu,k} \frac{\partial H}{\partial g_{kl}^{\mu\nu}} A_k^{\mu\nu}, \tag{93}$$

where $A_k^{\mu\nu}$ is the covariant derivative of $p^{\mu\nu}$, is a contravariant vector.

Then he observes that:

$$P_g(\sqrt{g}H) - \sum_l \frac{\partial\sqrt{g}a^l}{\partial w_l} \tag{94}$$

no longer contains the second derivatives of $p^{\mu\nu}$, and hence can be written:

$$\sqrt{g}\sum_{\mu,\nu,k} (B_{\mu\nu}p^{\mu\nu} + B_{\mu\nu}^k p_k^{\mu\nu}), \tag{95}$$

where $B_{\mu\nu}^k$ is a tensor. Finally, Hilbert forms the vector:

$$b^l = \sum_{\mu,\nu} B_{\mu\nu}^l p^{\mu\nu}, \tag{96}$$

obtaining (92).

He next forms the expression for the electromagnetic variables analogous to (92) (see the definition of P_q, (20) above):

$$P_q(\sqrt{g}H) - \sum_l \frac{\partial\sqrt{g}c^l}{dw_l} = \sum_k [\sqrt{g}H]_k p_k, \tag{97}$$

with:

$$c^l = \sum_k \frac{\partial H}{\partial q_{kl}} p_k. \tag{98}$$

Adding (92) and (97), and taking account of the field equations, Hilbert could thus write:

$$P(\sqrt{g}H) = \sum_l \frac{\partial\sqrt{g}(a^l + b^l + c^l)}{\partial w_l}. \tag{99}$$

The final step consists in also rewriting the left-hand side of this equation as a divergence, using (91), which is expanded as:

$$P(\sqrt{g}H) = \sqrt{g}PH + H\sum_s\left(\frac{\partial\sqrt{g}}{\partial w_s}p^s + \sqrt{g}p_s^s\right); \tag{100}$$

using Theorem II (see (22)),[126] he then obtained:

$$P(\sqrt{g}H) = \sqrt{g}\sum_z\frac{\partial H}{\partial w_s}p^s + H\sum_s\left(\frac{\partial\sqrt{g}}{\partial w_s}p^s + \sqrt{g}p_s^s\right) = \sum_s\frac{\partial\sqrt{g}Hp^s}{\partial w_s}, \tag{101}$$

and, in view of (99),

$$\sum_l\frac{\partial}{\partial w_l}\sqrt{g}(Hp^l - a^l - b^l - c^l) = 0. \tag{102}$$

This equation could have been interpreted as giving the energy expression since, being an invariant divergence, it satisfies two of the three criteria mentioned above. But it is not related to Mie's energy-momentum tensor. So Hilbert adds yet another term $-d^l$ to the expression in the parenthesis in (102):

$$d^l = \frac{1}{2\sqrt{g}}\sum_{k,s}\frac{\partial}{\partial w_k}\left\{\left(\frac{\partial\sqrt{g}H}{\partial q_{lk}} - \frac{\partial\sqrt{g}H}{\partial q_{kl}}\right)p^s q_s\right\}, \tag{103}$$

which does not alter its character since d^l is a contravariant vector (because:

$$\frac{\partial H}{\partial q_{lk}} - \frac{\partial H}{\partial q_{kl}} \tag{104}$$

is an antisymmetric tensor) that satisfies the identity:

$$\sum_l\frac{\partial\sqrt{g}d^l}{\partial w_l} = 0. \tag{105}$$

Hilbert concluded:

> Let us now define
>
> $$e^l = Hp^l - a^l - b^l - c^l - d^l \tag{106}$$
>
> as the *energy vector, then the energy vector is a contravariant vector, which moreover depends linearly on the arbitrarily chosen vector* p^s, *and satisfies identically for that choice of this vector* p^s *the invariant energy equation*
>
> $$\sum_l\frac{\partial\sqrt{g}e^l}{\partial w_l} = 0. \tag{107}$$ [127]

126 In Paper 1, this is the only purpose for which this form of Theorem II is explicitly introduced. However, (23) presumably already had been derived from it.

While Hilbert did not explicitly introduce the condition that his energy vector be related to Mie's energy-momentum tensor, it seems to be the guiding principle of his calculation. Apparently, he wanted this connection to appear to be the result of an independently-justified definition of this vector.

In effect, starting from (106) and taking into account definitions (98) and (103), Hilbert obtained for the contribution to the energy originating from the electromagnetic term L in the Lagrangian:

$$Lp^l - \sum_k \frac{\partial L}{\partial q_{kl}} p_k - \frac{1}{2\sqrt{g}} \sum_{k,s} \frac{\partial}{\partial w_k} \left\{ \left(\frac{\partial \sqrt{g} L}{\partial q_{lk}} - \frac{\partial \sqrt{g} L}{\partial q_{kl}} \right) p^s q_s \right\}. \tag{108}$$

Using the field equations and (27), this can be rewritten as:

$$\sum_{s,k} \left(L\delta_s^l - \frac{\partial L}{\partial M_{lk}} M_{sk} - \frac{\partial L}{\partial q_l} q_s \right) p^s, \tag{109}$$

which corresponds to the right-hand side of (36), the generally-covariant generalization of Mie's electromagnetic energy-momentum tensor, contracted with p^s.

In contradistinction to the Proofs, Theorem II and (36) no longer explicitly enter this demonstration. Theorem II enters implicitly by determining the form in which the electromagnetic variables enter the Lagrangian (see (27)). Hilbert still needed Theorem II to derive his "first result," that is, to show that this energy-momentum can be written as the variational derivative of $\sqrt{g}L$ with respect to the gravitational potentials. Furthermore, (36) allows Hilbert to argue that, due to the field equations (see (72)), the electromagnetic energy and energy-vector e^l can be expressed exclusively in terms of K, the gravitational part of the Lagrangian; so that they depend only on the metric tensor and not on the electromagnetic potentials and their derivatives. Whereas, in the Proofs, this result had been an immediate consequence of the definition of the energy and of the field equations (see (49)), now it follows only with the help of Theorem II.

While Hilbert had succeeded in satisfying his heuristic criteria as well as the new challenge of deriving an invariant energy equation, the status of this equation within his theory had become more precarious. An analysis of the deductive structure of Hilbert's theory in Paper 1 (see Fig. 2) shows that it still comprises two main clusters of results: those concerning the implications of gravitation for electromagnetism and those concerning energy conservation. But the latter cluster is now even more isolated

---

127 "Definieren wir nunmehr [(106); (14) in the original text] als den *Energievektor, so ist der Energievektor ein kontravarianter Vektor, der noch von dem willkürlichen Vektor* p^s *linear abhängt und identisch für jene Wahl dieses Vektors* p^s *die invariante Energiegleichung* [(107)] *erfüllt*." (Hilbert 1916, 402)

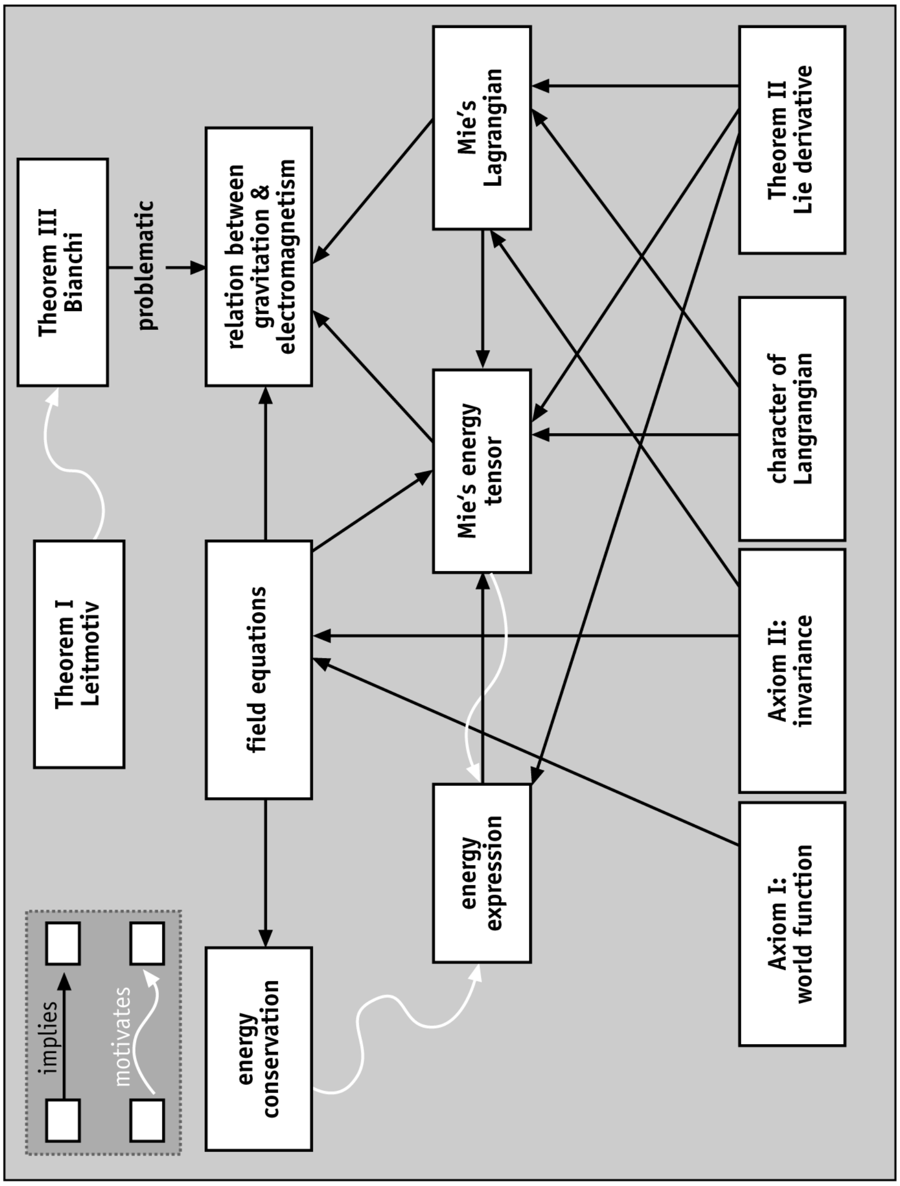

Figure 2: Deductive Structure of Paper 1 (1916)

from the rest of his theory than in the Proofs. Indeed, the new energy concept is no longer motivated by Hilbert's powerful Theorem I, but only by arguments concerning the formal properties of energy-momentum conservation and the link with Mie's energy-momentum tensor. It plays no role in deriving any other results of Hilbert's theory, nor does it serve to integrate this theory with other physical theories, a key function of the energy concept since its formulation in the 19th century. Therefore, it is not surprising that this concept only played a transitional role and was eventually replaced by the understanding of energy-momentum conservation developed by Einstein, Klein, Noether, and others.[128]

In fact, neither the physical significance nor the mathematical status of Hilbert's new energy concept was entirely clear. Physically Hilbert had failed to show that his energy equation (107) gave rise to a familiar expression for energy-momentum conservation in the special-relativistic limit, or to demonstrate that his equation was compatible with the form of energy-momentum conservation in a gravitational field that Einstein had established in 1913 (see (11)). Eventually, Felix Klein succeeded in clarifying the relation between Hilbert's and Einstein's expressions. He decomposed (107) into 140 equations and showed that 136 of these actually have nothing to do with energy-momentum conservation, while the remaining 4 correspond to those given by Einstein.[129] Mathematically, in 1917 Emmy Noether and Felix Klein found that equation (107) actually is an identity, and not a consequence of the field equations, as is the case for conservation equations in classical physics.[130] Similar identities follow for the Lagrangian of any generally-covariant variational problem. As a consequence, Hilbert's counting of equations no longer works: he assumed that his variational principle gives rise to 10 gravitational field equations plus 4 identities, which he identified with the electromagnetic equations; and that energy-momentum conservation is represented by additional equations, originally linked to coordinate restrictions. Einstein's abandonment of coordinate restrictions together with the deeper investigation of energy-momentum conservation by Noether, Klein, Einstein, and others, confronted Hilbert's approach with a severe challenge: They questioned the organization of his theory into two more-or-less independent domains, energy-momentum conservation and the implications of gravitation for electromagnetism. We shall argue that Hilbert responded to this challenge by further adapting his theory to the framework provided by general relativity.

5.2 Hilbert's Reorganization of His Theory in Paper 1

The challenge presented by Einstein's abandonment of coordinate restrictions and adoption of generally-covariant field equations forced Hilbert to reorganize his the-

---

128 For discussion, see (Rowe, 1999).

129 See (Klein 1918a, 179–185).

130 See (Klein 1917; 1918a) and also (Noether 1918). For a thorough discussion of the contemporary research on energy-momentum conservation, see (Rowe, 1999).

ory. As we have seen, he had to demonstrate the compatibility between his variational principle and Einstein's field equations (from which he had succeeded strikingly in deriving Mercury's perihelion shift), and completely rework his treatment of energy conservation. Hilbert treated both issues at the end of Paper 1. Energy conservation was no longer tied to Theorem I and its heuristic consequences as in the Proofs, but was treated along with other results of Hilbert's theory. The structure of Paper 1 is thus:[131]

1. Basic Framework (Hilbert 1916, 395–398)

   Axioms I and II, Theorem I, and the combined field equations of gravitation and electromagnetism for an arbitrary Lagrangian

2. Basic Theorems (Hilbert 1916, 398–400)

   Theorems II and III

3. New Energy Expression and Derivation of the New Energy Equation (Hilbert 1915, 400–402)

4. Implications for the Relation between Electromagnetism and Gravitation (Hilbert 1915, 402–407)

   the split of the Lagrangian into gravitational and the electrodynamical terms, the form of Mie's Lagrangian, its relation to his energy tensor, the explicit form of the gravitational field equations, and the relation between electromagnetic and gravitational field equations.

Apart from the technical and structural revisions necessitated by the new energy expression, practically all other changes concern the relation of his theory to Einstein's. Throughout Paper 1, Hilbert followed the tendency, already manifest in the marginal additions to the Proofs, to put greater emphasis on Einstein's contributions while maintaining his claim to have developed an independent approach. In the opening paragraph, Hilbert changed the order in which he mentioned Mie and Einstein. In the Proofs he wrote:

> The far reaching ideas and the formation of novel concepts by means of which Mie constructs his electrodynamics, and the prodigious problems raised by Einstein, as well as his ingeniously conceived methods of solution, have opened new paths for the investigation into the foundations of physics.[132]

In Paper 1 we read instead:

> The vast problems posed by Einstein as well as his ingeniously conceived methods of solution, and the far-reaching ideas and formation of novel concepts by means of which

131 For a sketch of Hilbert's revisions of Paper 1, see also (Corry 1999a, 517–522).

132 "Die tiefgreifenden Gedanken und originellen Begriffsbildungen vermöge derer Mie seine Elektrodynamik aufbaut, und die gewaltigen Problemstellungen von Einstein sowie dessen scharfsinnige zu ihrer Lösung ersonnenen Methoden haben der Untersuchung über die Grundlagen der Physik neue Wege eröffnet." (Proofs, 1)

> Mie constructs his electrodynamics, have opened new paths for the investigation into the foundations of physics.[133]

A footnote lists all of Einstein's publications on general relativity starting with his major 1914 review, and including the definitive paper submitted on 25 November. Although this makes clear that Hilbert must have revised his paper after that date, he failed to change the dateline of his contribution (as did Felix Klein and Emmy Noether in their contributions to the discussion of Hilbert's work in the same journal[134]). It remained "Vorgelegt in der Sitzung vom 20. November 1915," which creates the erroneous impression that there were no subsequent substantial changes in Paper 1.

The next sentence, while combining this claim with a more explicit recognition of what he considered the achievements of his predecessors, shows that Hilbert had not renounced his claim to having solved the problems posed by Mie and Einstein. In the corrected Proofs this sentence reads:

> In the following—in the sense of the axiomatic method — I would like to develop, /[*essentially*] from three simple axioms a ~~new~~ system of basic equations of physics, of ideal beauty, containing, I believe, the solution of the problems presented.[135]

In Paper 1, it reads:

> In the following — in the sense of the axiomatic method — I would like to develop, essentially from two simple axioms, a new system of basic equations of physics, of ideal beauty and containing, I believe, *simultaneously* the solution to the problems of Einstein and of Mie. I reserve for later communications the detailed development and particularly the special application of my basic equations to the fundamental questions of the theory of electricity.[136]

Although in a marginal note in the proofs version he had changed "his theory" to "the theory," he now returned to the original version:

> The guiding motive for constructing my theory is provided by the following theorem, the proof of which I shall present elsewhere.[137]

---

133 "Die gewaltigen Problemstellungen von Einstein sowie dessen scharfsinnige zu ihrer Lösung ersonnenen Methoden und die tiefgreifenden Gedanken und originellen Begriffsbildungen vermöge derer Mie seine Elektrodynamik aufbaut, haben der Untersuchung über die Grundlagen der Physik neue Wege eröffnet." (Hilbert 1916, 395)

134 See (Klein 1918a; Noether 1918).

135 "Ich möchte im Folgenden — im Sinne der axiomatischen Methode —/[*wesentlich*] aus drei einfachen Axiomen ein ~~neues~~ System von Grundgleichungen der Physik aufstellen, die von idealer Schönheit sind, und in denen, wie ich glaube, die Lösung der gestellten Probleme enthalten ist." (Proofs, 1)

136 "Ich möchte im Folgenden - im Sinne der axiomatischen Methode - wesentlich aus zwei einfachen Axiomen ein neues System von Grundgleichungen der Physik aufstellen, die von idealer Schönheit sind, und in denen, wie ich glaube, die Lösung der Probleme von Einstein und Mie gleichzeitig enthalten ist. Die genauere Ausführung sowie vor Allem die spezielle Anwendung meiner Grundgleichungen auf die fundamentalen Fragen der Elektrizitätslehre behalte ich späteren Mitteilungen vor." (Hilbert 1916, 395)

137 "Das Leitmotiv für den Aufbau meiner Theorie liefert der folgende mathematische Satz, dessen Beweis ich an einer anderen Stelle darlegen werde." (Hilbert 1916, 396)

Although Hilbert had earlier argued that his *Leitmotiv* suggested the need for four additional non-covariant equations to ensure a unique solution, he now dropped all mention of the subject of coordinate restrictions. He simply did not address the question of why, in spite of Einstein's hole argument against this possibility, it is possible to use generally-covariant field equations unsupplemented by coordinate restrictions. The only remnant in Paper 1 of the entire problem is his newly-introduced designation of the world-parameters as "allgemeinste Raum-Zeit-Koordinaten."

The significant result that Hilbert's variational principle gives rise to gravitational field equations formally equivalent to those of Einstein's 25 November theory is rather hidden in Hilbert's presentation, only appearing as an intermediate step in his demonstration that the electromagnetic field equations are a consequence of the gravitational ones. The newly-introduced passage reads:

> Using the notation introduced earlier for the variational derivatives with respect to the $g^{\mu\nu}$, the gravitational equations, because of (20) [i.e. (16)], take the form
>
> $$[\sqrt{g}K]_{\mu\nu} + \frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}} = 0. \tag{110}$$
>
> The first term on the left hand side becomes
>
> $$[\sqrt{g}K]_{\mu\nu} = \sqrt{g}\Big(K_{\mu\nu} - \frac{1}{2}Kg_{\mu\nu}\Big), \tag{111}$$
>
> as follows easily without calculation from the fact that $K_{\mu\nu}$, apart from $g_{\mu\nu}$, is the only tensor of second rank and K the only invariant, that can be formed using only the $g^{\mu\nu}$ and their first and second differential quotients, $g_k^{\mu\nu}$, $g_{kl}^{\mu\nu}$.
>
> The resulting differential equations of gravitation appear to me to be in agreement with the grand concept of the theory of general relativity established by Einstein in his later treatises.[138]

Hilbert's argument for avoiding explicit calculation of $[\sqrt{g}K]_{\mu\nu}$, which he later withdrew (see below), is indeed untenable; there are many invariants and tensors of second rank that can be constructed from the Riemann tensor. Even if one further requires such tensors and invariants to be linear in the Riemann tensor, the crucial coefficient of the trace term still remains undetermined. The explicit form of the field equations given in Paper 1 and not found in the Proofs, appears to be a direct response to Einstein's publication of 25 November; but a footnote appended to this

138 "Unter Verwendung der vorhin eingeführten Bezeichungsweise für die Variationsableitungen bezüglich der $g^{\mu\nu}$ erhalten die Gravitationsgleichungen wegen (20) [i.e. (16)] die Gestalt [(110); (21) in the original text]. Das erste Glied linker Hand wird [(111)] wie leicht ohne Rechnung aus der Tatsache folgt, daß $K_{\mu\nu}$ außer $g_{\mu\nu}$ der einzige Tensor zweiter Ordnung und K die einzige Invariante ist, die nur mit den $g^{\mu\nu}$ und deren ersten und zweiten Differentialquotienten $g_k^{\mu\nu}, g_{kl}^{\mu\nu}$ gebildet werden kann.
Die so zu Stande kommenden Differentialgleichungen der Gravitation sind, wie mir scheint, mit der von Einstein in seinen späteren Abhandlungen aufgestellten großzügigen Theorie der allgemeinen Relativität im Einklang." (Hilbert 1916, 404–405)

passage gives a generic reference to all four of Einstein's 1915 Academy publications. His cautious reference to the apparent agreement between his results and Einstein's, presumably motivated by their different frameworks, adds to the impression that Hilbert actually arrived independently at the explicit form of the gravitational field equations.

The concluding paragraph of Paper 1 acknowledges Hilbert's debt to Einstein in a more indirect way. The beginning of this paragraph of the Proofs had given the impression that Einstein posed the problems while Hilbert offered the solutions:

> As one can see, the few simple assumptions expressed in axioms I, II, III suffice with appropriate interpretation to establish the theory: through it not only are our views of space, time, and motion fundamentally reshaped in the sense called for by Einstein ...[139]

In Paper 1, Hilbert deleted the reference to axiom III and replaced "in dem von Einstein geforderten Sinne" by "in dem von Einstein dargelegten Sinne":

> As one can see, the few simple assumptions expressed in axioms I and II suffice with appropriate interpretation to establish the theory: through it not only are our views of space, time, and motion fundamentally reshaped in the sense explained by Einstein ...[140]

5.3 Einstein's Energy in Hilbert's 1924 Theory

In 1924 Hilbert published revised versions of Papers 1 and 2 (Hilbert 1924).[141] Meanwhile important developments had taken place, such as the rapid progress of quantum physics, which changed the scientific context of Hilbert's results. But it was undoubtedly the further clarifications of the significance of energy-momentum conservation in general relativity, already mentioned in the preceding sections, that affected his theory most directly. In correspondence between Hilbert and Klein (published in part in 1918),[142] this topic played a central role without, however, leading to an explicit reformulation of Hilbert's theory. Without going into detail about this important strand in the history of general relativity, we shall focus on its effect on Hilbert's 1924 revisions. In spite of the reassertion of his goal of providing foundations for all of physics, his theory was, in effect, transformed into a variation on the themes of general relativity.

---

139 "Wie man sieht, genügen bei sinngemäßer Deutung die wenigen einfachen in den Axiomen I, II, III ausgesprochenen Annahmen zum Aufbau der Theorie: durch dieselbe werden nicht nur unsere Vorstellungen über Raum, Zeit und Bewegung von Grund aus in dem von Einstein geforderten Sinne umgestaltet ..." (Proofs, 13).

140 "Wie man sieht, genügen bei sinngemäßer Deutung die wenigen einfachen in den Axiomen I und II ausgesprochenen Annahmen zum Aufbau der Theorie: durch dieselbe werden nicht nur unsere Vorstellungen über Raum, Zeit und Bewegung von Grund aus in dem von Einstein dargelegten Sinne umgestaltet ..." (Hilbert 1916, 407).

141 In the following, we will refer to the 1924 revision of Paper 1 as "Part 1" and to that of Paper 2 as "Part 2," designations which correspond to Hilbert's own division of his 1924 paper into "Teil 1" (pp. 2–11) and "Teil 2" (pp. 11–32).

142 See (Klein 1917).

On a purely technical level, Hilbert's revisions of Paper 1 appear to be rather modest; the most important one concerns Theorem III (the contracted Bianchi identities), now labelled Theorem 2. Following a suggestion by Klein (Klein 1917, 471–472), Hilbert extended this theorem to include the electromagnetic variables:

> Theorem 2. Let J, as in Theorem 1, be an invariant depending on $g^{\mu\nu}$, $g_l^{\mu\nu}$, $g_{lk}^{\mu\nu}$, q_s, q_{sk}; and as above, let $[\sqrt{g}J]_{\mu\nu}$ denote the variational derivatives of $\sqrt{g}J$ with respect to $g^{\mu\nu}$, and $[\sqrt{g}J]_\mu$, the variational derivative with respect to q_μ. Introduce, furthermore, the abbreviations [(112)]:
>
> $$\begin{aligned} i_s &= \sum_{\mu,\nu}([\sqrt{g}J]_{\mu\nu}g_s^{\mu\nu} + [\sqrt{g}J]_\mu q_{\mu s}), \\ i_s^l &= -2\sum_\mu [\sqrt{g}J]_{\mu s}g^{\mu l} + [\sqrt{g}J]_l q_s, \end{aligned} \tag{112}$$
>
> then the [following] identities hold
>
> $$i_s = \sum_l \frac{\partial i_s^l}{\partial x_l} \qquad (s = 1, 2, 3, 4).^{143} \tag{113}$$

He revised its proof accordingly.

A second, small, but significant change concerns the gravitational field equations. Hilbert now tacitly withdrew his previous claim that no derivation was needed, instead sketching a derivation and writing them, like Einstein, with the energy-momentum tensor as source. As in the earlier versions, he derived (72) but now in the form:[144]

$$[\sqrt{g}K]_{\mu\nu} = -\frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}}. \tag{114}$$

After writing down the electromagnetic field equations, Hilbert proceeded to sketch the following evaluation of the terms in (114):

> To determine the expression for $[\sqrt{g}K]_{\mu\nu}$, first specialize the coordinate system so that at the world point under consideration all the $g_s^{\mu\nu}$ vanish. In this way one finds:
>
> $$[\sqrt{g}K]_{\mu\nu} = \sqrt{g}\left(K_{\mu\nu} - \frac{1}{2}g_{\mu\nu}K\right). \tag{115}$$
>
> If, for the tensor
>
> $$-\frac{1}{\sqrt{g}}\frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}} \tag{116}$$
>
> we introduce the symbol $T_{\mu\nu}$, then the gravitational field equations can be written as

---

143 "Theorem 2. Wenn J, wie im Theorem 1, eine von $g^{\mu\nu}$, $g_l^{\mu\nu}$, $g_{lk}^{\mu\nu}$, q_s, q_{sk} abhängige Invariante ist, und, wie oben, die Variationsableitungen von $\sqrt{g}J$ bez. $g^{\mu\nu}$ mit $[\sqrt{g}J]_{\mu\nu}$, bez. q_μ mit $[\sqrt{g}J]_\mu$ bezeichnet werden, und wenn ferner zur Abkürzung: [(112)] gesetzt wird, so gelten die Identitäten [(113); (7) in the original text]." (Hilbert 1924, 5)

144 See (Hilbert 1924, 7).

$$K_{\mu\nu} - \frac{1}{2} g_{\mu\nu} K = T_{\mu\nu}. \text{ }^{145} \tag{117}$$

Although the introduction of Einstein's notation for the energy-momentum tensor may appear as no more than an adaptation of Hilbert's notation to the by-then standard usage, it actually effected a major revision in the structure of his theory. The energy-momentum tensor became the central knot binding together the physical implications of Hilbert's theory.

First of all, it served, as Hilbert's energy expressions had previously done, to relate the derivative of Mie's Lagrangian (see (34) or (36)) to Mie's energy-momentum tensor. But, in contrast to Paper 1, Mie's energy-momentum tensor no longer served as a criterion for choosing the energy-expression. The new energy expression, which Hilbert now took over from Einstein, was supported by much more than just this single result. It had emerged from the development of special-relativistic continuum physics by Minkowski, Abraham, Planck, Laue,[146] and others; and been validated by numerous applications to various areas of physics, including general relativity.

By introducing the equation:

$$T_{\mu\nu} = -\frac{\partial \sqrt{g} L}{\partial g^{\mu\nu}}, \tag{118}$$

Hilbert had returned, in a sense, to the approach of the Proofs, establishing a relation between the energy concept and the derivative of the electromagnetic Lagrangian (see (49)). He still did not make clear that this relation does not single out Mie's theory, but actually holds more generally. Introducing the notations:

$$\frac{\partial L}{\partial q_{sk}} = \frac{\partial L}{\partial M_{ks}} = H^{ks}, \tag{119}$$

and:

$$\frac{\partial L}{\partial q_k} = r^k, \tag{120}$$

As in the proofs version, Hilbert again used (35), which he now rewrites as:

$$-\frac{2}{\sqrt{g}} \sum_{\mu} \frac{\partial \sqrt{g} L}{\partial g^{\mu\nu}} g^{\mu m} = L \delta_\nu^m - \sum_s H^{ms} M_{\nu s} - r^m q_\nu, \tag{121}$$

145 "Um den Ausdruck von $[\sqrt{g}K]_{\mu\nu}$ zu bestimmen, spezialisiere man zunächst das Koordinatensystem so, daß für den betrachteten Weltpunkt die $g_s^{\mu\nu}$ sämtlich verschwinden. Man findet auf diese Weise: [(115)]. Führen wir noch für den Tensor [(116)] die Bezeichnung $T_{\mu\nu}$ ein, so lauten die Gravitationsgleichungen [(117)]." See (Hilbert 1924, 7–8).

146 For the first systematic development of relativistic continuum mechanics, see (Laue 1911a; 1911b). For further discussion, see Einstein's "Manuscript on the Special Theory of Relativity" (CPAE 4, Doc. 1, 91–98; Janssen and Mecklenburg 2006).

(see (36)). On the basis of this equation, Hilbert claims, in almost exactly the same words as in the earlier versions, that there is a necessary connection between the theories of Mie and Einstein:

> Hence the [following] representation of $T_{\mu\nu}$ results:
>
> $$T_{\mu\nu} = \sum_{\mu} g_{\mu m} T_{\nu}^{m}$$
> $$T_{\nu}^{m} = \frac{1}{2}\left\{L\delta_{\nu}^{m} - \sum_{s} H^{ms} M_{\nu s} - r^{m} q_{\nu}\right\}. \tag{122}$$
>
> The expression on the right agrees with Mie's electromagnetic energy tensor, and thus we find that Mie's electromagnetic energy tensor is nothing but the generally-invariant tensor resulting from differentiation of the invariant L with respect to the gravitational potentials $g^{\mu\nu}$ — a circumstance which gave me the first hint of the necessary close connection between Einstein's theory of general relativity and Mie's electrodynamics, and which convinced me of the correctness of the theory developed here.[147]

While Hilbert's claim remained unchanged, what he had done actually was to specialize the source term left arbitrary in Einstein's field equations. The nature of this source term can be specified on the level of the Lagrangian or of the energy-momentum tensor, and these two ways are obviously equivalent if a Lagrangian exists—but this relation is in no way peculiar to Mie's theory. The fact that the energy expression in Paper 1 was specifically chosen to produce Mie's energy-momentum tensor had obscured this circumstance, now made rather obvious by the introduction of Einstein's arbitrary energy-momentum tensor. It was no doubt difficult for Hilbert to draw this conclusion because it contradicted his program, according to which electromagnetism should arise as an effect of gravitation.

The situation was similar for Hilbert's second important application of Einstein's energy-momentum tensor, the derivation of a relation between the gravitational and electromagnetic field equations. After recognition of the close relation between the contracted Bianchi identities and energy-momentum conservation in general relativity, it was necessary for Hilbert to reconsider the link he believed he had established between the two groups of field equations. Energy-momentum conservation now played a central role in his approach, turning the link between gravitation and electromagnetism into a mere by-product. It existed, not because of any deep intrinsic connection between these two areas of physics, but due to the introduction of electromagnetic potentials into the variational principle. With the same logic, one

147 "Demnach ergibt sich für $T_{\mu\nu}$ die Darstellung: [(122)]. Der Ausdruck rechts stimmt überein mit dem Mie'schen elektromagnetischen Energietensor, und wir finden also, daß der Mie'sche elektromagnetische Energietensor ist nichts anderes als der durch Differentiation der Invariante L nach den Gravitationspotentialen $g^{\mu\nu}$ entstehende allgemein invariante Tensor—ein Umstand, der mich zum ersten Mal auf den notwendigen engen Zusammenhang zwischen der Einsteinschen allgemeinen Relativitätstheorie und der Mie'schen Elektrodynamik hingewiesen und mir die Überzeugung von der Richtigkeit der hier entwickelten Theorie gegeben hat." (Hilbert 1924, 9)

could argue that *any* form of matter giving rise to a stress-energy tensor derivable from a Lagrangian involving the metric tensor is an effect of gravitation.

This weakened link is reflected in Hilbert's new way of obtaining the desired link between gravitation and electromagnetism. Following Klein's suggestion, in Part 1 Hilbert treated the contracted Bianchi identities in parallel for both the gravitational and the electromagnetic terms in the Lagrangian:

> The application of Theorem 2 to the invariant K yields:
>
> $$\sum_{\mu\nu}[\sqrt{g}K]_{\mu\nu}g_s^{\mu\nu} + 2\sum_m \frac{\partial}{\partial x_m}\Big(\sum_\mu [\sqrt{g}K]_{\mu s}g^{\mu m}\Big) = 0. \quad (123)$$
>
> Its application to L yields:[148]
>
> $$\sum_{\mu\nu}(-\sqrt{g}T_{\mu\nu})g_s^{\mu\nu} + 2\sum_m \frac{\partial}{\partial x_m}(-\sqrt{g}T_s^{\ m}) + \sum_\mu [\sqrt{g}L]_\mu q_{\mu s} - \sum_\mu \frac{\partial}{\partial x_\mu}([\sqrt{g}L]_\mu q_s) = 0 \qquad (s = 1, 2, 3, 4). \quad (124)$$

Previously, he had derived only the first set of identities and made use of them in order to derive (83). Now Hilbert showed that both sets of identities yield the equations for energy-momentum conservation that had been central to Einstein's work since 1912. Following the work of Einstein and others, Hilbert also made clear that these equations are related to the equations of motion for the sources of the stress-energy tensor,[149] and represent a generalization of energy-momentum conservation laws in special relativity:

> As a consequence of the basic equations of electrodynamics, we obtain from this:
>
> $$\sum_{\mu\nu}\sqrt{g}T_{\mu\nu}g_s^{\mu\nu} + 2\sum_m \frac{\partial}{\partial x_m}\sqrt{g}T_s^{\ m} = 0. \quad (125)$$
>
> These equations also result as a consequence of the gravitational equations due to (15a) [i.e. (123)]. Their interpretation is that they are the basic equations of mechanics. In the case of special relativity, when the $g_{\mu\nu}$ are constants, they reduce to the equations
>
> $$\sum \frac{\partial T_s^{\ m}}{\partial x_m} = 0, \quad (126)$$
>
> which express the conservation of energy and momentum.[150]

---

148 "Die Anwendung des Theorems 2 auf die Invariante K liefert: [(123); (15a) in the original text.] Die Anwendung auf L ergibt: [(124); (15b) in the original text.]" (Hilbert 1924, 9–10)

149 See (Havas 1989, Klein 1917; 1918a; 1918b).

150 "Als Folge der elektrodynamischen Grundgleichungen erhalten wir hieraus: [(125); (16) in the original text.] Diese Gleichungen ergeben sich auch als Folge der Gravitationsgleichungen, auf Grund von (15a) [i.e. (123)]. Sie haben die Bedeutung der mechanischen Grundgleichungen. Im Falle der speziellen Relativität, wenn die $g_{\mu\nu}$ Konstante sind, gehen sie über in die Gleichungen [(126)] welche die Erhaltung von Energie und Impuls ausdrücken." (Hilbert 1924, 10)

Hilbert thus anchored his theory in the same physical foundation that had provided Einstein's search for general relativity with a stable point of reference. Only after having done this did Hilbert turn to his original goal, the link between gravitation and electromagnetism, the problematic character of which we have discussed above:

> From the identities (15b) [i.e. (124)], there follow from the equations (16) [i.e. (125)]:
>
> $$\sum_{\mu}[\sqrt{g}L]_{\mu}q_{\mu s} - \sum_{\mu}\frac{\partial}{\partial x_{\mu}}([\sqrt{g}L]_{\mu}q_s) = 0 \qquad (127)$$
>
> or
>
> $$\sum_{\mu}\left\{M_{\mu s}[\sqrt{g}L]_{\mu} + q_s\frac{\partial}{\partial x_{\mu}}[\sqrt{g}L]_{\mu}\right\} = 0; \qquad (128)$$
>
> i.e., four independent linear relations between the basic equations of electrodynamics (5) and their first derivations follow from the gravitational equations (4). This is the precise mathematical expression of the connection between gravitation and electrodynamics, which dominates the entire theory.[151]

The deductive structure of Part 1 shows the fundamental changes with respect to Paper 1 (see Fig. 3) and the central role of Einstein's energy-momentum tensor in this reorganization. In fact, this tensor suggested the particular form in which Hilbert rewrote the gravitational field equations, established the link between gravitation and electromagnetism (in terms of the choice of a specific source), and, of course, was fundamental to Hilbert's new formulation of energy-momentum conservation.

This revised deductive structure has a kernel, consisting of the variational principle, field equations, and energy-momentum conservation, that is—both from a formal and a physical perspective —fully equivalent to the kernel of Einstein's formulation of general relativity. Clearly, Hilbert's deductive presentation places greater emphasis on a variational principle than does Einstein; and the mathematically more elegant formulation of the variational principle, based on the Ricci scalar, contributes to this emphasis. Therefore, this variational formulation of general relativity is today rightly associated with Hilbert's name. On the other hand, Hilbert's original aim, the derivation of electromagnetism as an effect of gravitation, plays only a marginal role in Part 1 and still suffers from the problems indicated above. The links between the main components that had substantiated Hilbert's claim of a special relation between Mie's theory and Einstein's have been weakened, being held together only by the choice of a specific source. This link is thus no longer central to an approach presenting an alternative to that of Einstein, being little more than an attempt to supplement

151 "Aus den Gleichungen (16) [i.e. (125)] folgt auf Grund der Identitäten (15b) [i.e. (124)]: [(127)] oder [(128); (17) in the original text] d.h. aus den Gravitationsgleichungen (4) folgen vier voneinander unabhängige lineare Relationen zwischen den elektrodynamischen Grundgleichungen (5) und ihren ersten Ableitungen. Dies ist der genaue mathematische Ausdruck für den Zusammenhang zwischen Gravitation und Elektrodynamik, der die ganze Theorie beherrscht." See the comments on (83), (Hilbert 1924, 10).

Einstein's general framework with a specific physical content, Mie's electrodynamics—an attempt that is now based on the firm foundations of general relativity.

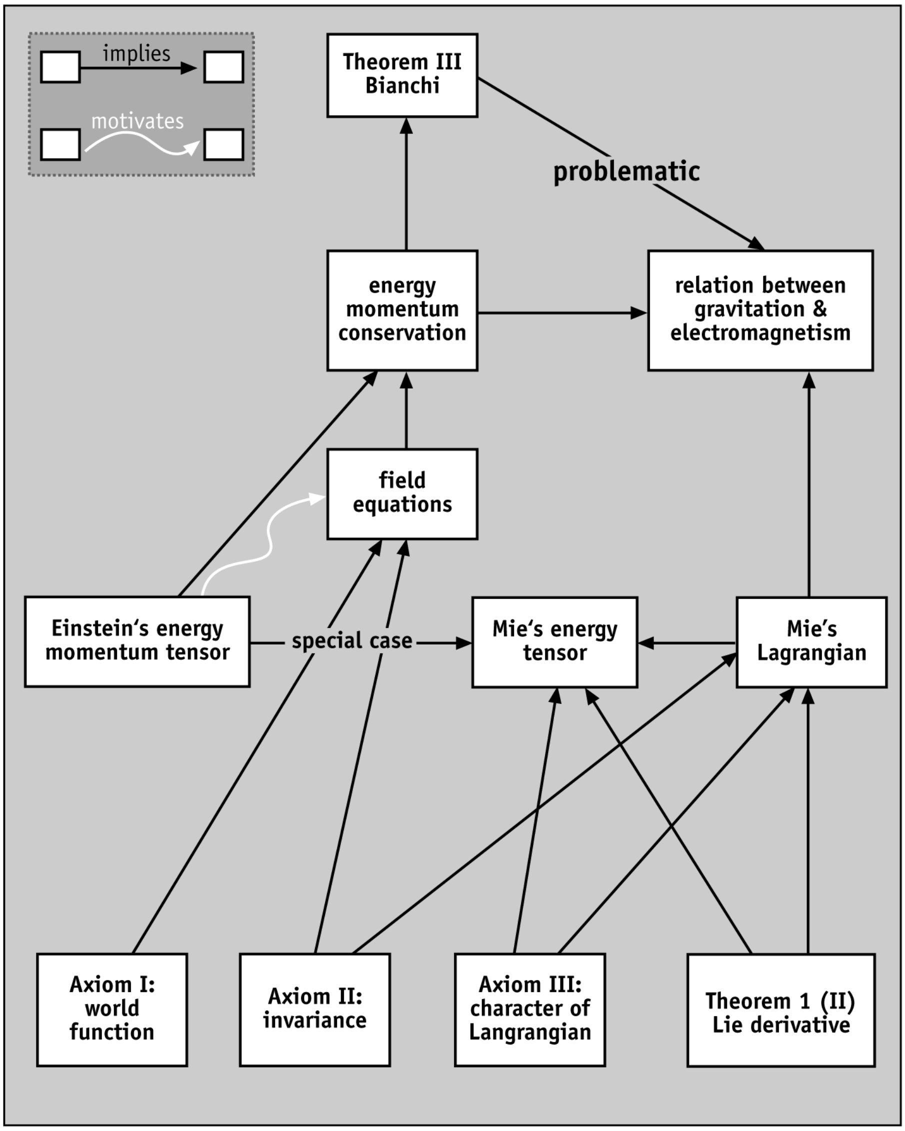

Figure 3: Deductive Structure of Part 1 (1924)

5.4 A Scientist's History

Scientists rarely investigate carefully the often only small and gradual conceptual transformations that their insights undergo in the course of historical development, often at the hands of others. Instead of undertaking such a demanding enterprise with little promise of new scientific results, they rather tend to hold onto their insights, reinterpreting them in the light of their present and prospective uses rather than in the light of past achievements, let alone failures. As we shall see, this tendency was inescapable for Hilbert, who understood the progress of physics in terms of an elaboration of the apparently universal and immutable concepts of classical physics.

Indeed, Hilbert described the 1924 Part 1 version of his theory not as a revision of his 1916 Paper 1 version, including major conceptual adjustments and a reorganization of its deductive structure, but essentially as a reprint of his earlier work:

> What follows is essentially a reprint of both of my earlier communications on the *Grundlagen der Physik*, and my comments on them, which were published by F. Klein in his communication *Zu Hilberts erster Note über die Grundlagen der Physik*, with only minor editorial differences and transpositions in order to facilitate their understanding.[152]

Indeed, the organization of Part 1 has not undergone major changes as compared to Paper 1, but seems to represent simply a tightening up; it can be subdivided into the following sections:

1. General Introduction (Hilbert 1924, 1–2)
2. Basic Setting (Hilbert 1924, 2–4)
   Axioms I and II, field equations of electromagnetism and gravitation
3. Basic Theorems (Hilbert 1924, 4–7)
   Theorems 1 (previously II) and 2 (previously III), the theorem earlier designated as Theorem I (now without numbering)
4. Implications for Electromagnetism, Gravitational Field Equations, and Energy-momentum Conservation (Hilbert 1924, 7–11)
   The character of the gravitational part of the Lagrangian, Axiom III (the split of the Lagrangian and the character of the electrodynamical part of the Lagrangian), the gravitational field equations, the form of Mie's Lagrangian, the relation between Mie's energy tensor and Mie's Lagrangian, energy-momentum conservation, and the relation between electromagnetic and gravitational field equations.

The most noteworthy changes in the order of presentation are: a new introductory section and the integration of the treatment of energy-momentum conservation with other results of Hilbert's theory towards the end. Another conspicuous change is that

152 "Das Nachfolgende ist im wesentlichen ein Abdruck der beiden älteren Mitteilungen von mir über die *Grundlagen der Physik* und meiner Bemerkungen dazu, die F. Klein in seiner Mitteilung *Zu Hilberts erster Note über die Grundlagen der Physik* veröffentlicht hat—mit nur geringfügigen redaktionellen Abweichungen und Umstellungen, die das Verständnis erleichtern sollen." (Hilbert 1924, 1)

Hilbert's *Leitmotiv*, Theorem I of Paper 1, has now lost its central place despite meanwhile having been proven by Emmy Noether. As we have seen, even in Paper 1 it no longer played the key heuristic role for Hilbert that it had originally in the Proofs. As the preceding discussion made clear, the rather unchanged form of its presentation hides major changes in the substance of his theory.

These changes are reflected in the introductory section, in a way that again downplays them.

While earlier Hilbert had introduced his own contribution as a solution to the problems raised by Mie and Einstein (Proofs) or Einstein and Mie (Paper 1), he now characterized his results as providing a simple and natural representation of Einstein's general theory of relativity, completed in formal aspects:

> The vast complex of problems and conceptual structures of Einstein's general theory of relativity now find, as I explained in my first communication, their simplest and most natural expression and, in its formal aspect, a systematic supplementation and completion by following the route trodden by Mie.[153]

In view of the overwhelming contemporary impact of Einstein's theory, Mie's role was downplayed in Hilbert's new version. Mie is no longer portrayed as posing problems of a similar profundity to those of Einstein, but as inspiring Hilbert's "simplest and most natural" presentation of general relativity, as well as "a systematic supplementation and completion in its formal aspect."

Instead of attributing a specific role in contemporary scientific discussions to Mie, Hilbert elevates him to the role of one of the founding fathers of a unified-field theoretical worldview:

> The mechanistic ideal of unity in physics, as created by the great researchers of the previous generation and still adhered to during the reign of classical electrodynamics, now must be definitively abandoned. Through the creation and development of the field concept, a new possibility for the comprehension of the physical world has gradually taken shape. Mie was the first to show a way that makes accessible to general mathematical treatment this newly risen 'field theoretical ideal of unity' as I would like to call it.[154]

Curiously neither Einstein nor Minkowski are mentioned in Hilbert's discussion of the spacetime continuum as the "foundation" of "the new field-theoretical ideal":

---

153 "Die gewaltigen Problemstellungen und Gedankenbildungen der allgemeinen Relativitätstheorie von Einstein finden nun, wie ich in meiner ersten Mitteilung ausgeführt habe, auf dem von Mie betretenen Wege ihren einfachsten und natürlichsten Ausdruck und zugleich in formaler Hinsicht eine systematische Ergänzung und Abrundung." (Hilbert 1924, 1–2) The changes in Hilbert's theory were accompanied by a change in his attitude to Einstein's achievement, by which he was increasingly impressed: see (Corry 1999a, 522–525).

154 "Das mechanistische Einheitsideal in der Physik, wie es von den großen Forschern der vorangegangenen Generation geschaffen und noch während der Herrschaft der klassischen Elektrodynamik festgehalten worden war, muß heute endgültig aufgegeben werden. Durch die Aufstellung und Entwickelung des Feldbegriffes bildete sich allmählich eine neue Möglichkeit für die Auffassung der physkalischen Welt aus. Mie zeigte als der erste einen Weg, auf dem dieses neuenstandene "feldtheoretische Einheitsideal", wie ich es nennen möchte, der allgemeinen mathematischen Behandlung zugänglich gemacht werden kann." (Hilbert 1924, 1)

> While the old mechanistic conception takes matter itself as a direct starting point and assumes it to be determined by a finite range of discrete parameters; a physical continuum, the so-called spacetime manifold, rather serves as the foundation of the new field-theoretical ideal. While previously universal laws took the form of [ordinary] differential equations with one independent variable, now partial differential equations are their necessary form of expression.[155]

Mie was exalted to the otherwise rather empty heaven of the founding fathers, leaving room for Hilbert's attempts at a unified theory of gravitation and electromagnetism. He generously mentioned other contemporary efforts as off-springs of his own contribution, a view hardly shared by his contemporaries (see below):

> Since the publication of my first communication, significant papers on this subject have appeared: I mention only Weyl's magnificent and profound investigations, and Einstein's communications, filled with ever new approaches and ideas. In the meantime, even Weyl took a turn in his development that led him too to arrive at just the equations I formulated; and on the other hand Einstein also, although starting repeatedly from divergent approaches, differing among themselves, ultimately returns, in his latest publication, to precisely the equations of my theory.[156]

This passage from Hilbert leaves unspecified to which of his equations he is referring. Given his references to Weyl and Einstein, he must mean the two sets of field equations (51) and (52), which are rather obvious ingredients of any attempted unification of gravitation and electromagnetism. The unique feature of his approach, the specific connection he introduced between these two sets of equations (see (83)) constituting the mathematical expression of electrodynamics as a phenomenon following from gravitation, had become highly problematic and was not adopted by either Weyl or Einstein.

Indeed, it was already problematic whether Weyl's and Einstein's attempts at unification were any more fortunate than Hilbert's. In his concluding paragraph, Hilbert himself expressed his doubts, which were based on the rapid progress of quantum physics, on the one hand, and the lack of any concrete physical results of such theories, on the other:

> Whether the pure field theoretical ideal of unity is indeed definitive, and what possible supplements and modifications of it are necessary to enable in particular the theoretical foundation for the existence of negative and positive electrons, as well as the consistent

155 "Während die alte mechanistische Auffassung unmittelbar die Materie selbst als Ausgang nimmt und diese durch eine endliche Auswahl diskreter Parameter bestimmt ansetzt, dient vielmehr dem neuen feldtheoretischen Ideal das physikalische Kontinuum, die sogenannte Raum-Zeit-Mannigfaltigkeit, als Fundament. Waren früher Differenzialgleichungen mit einer unabhängigen Variablen die Form der Weltgesetze, so sind jetzt notwendig partielle Differenzialgleichungen ihre Ausdrucksform." (Hilbert 1924, 1)

156 "Seit der Veröffentlichung meiner ersten Mitteilung sind bedeutsame Abhandlungen über diesen Gegenstand erschienen: ich erwähne nur die glänzenden und tiefsinnigen Untersuchungen von Weyl und die an immer neuen Ansätzen und Gedanken reichen Mitteilungen von Einstein. Indes sowohl Weyl gibt späterhin seinem Entwicklungsgange eine solche Wendung, daß er auf die von mir aufgestellten Gleichungen ebenfalls gelangt, und andererseits auch Einstein, obwohl wiederholt von abweichenden und unter sich verschiedenen Ansätzen ausgehend, kehrt schließlich in seinen letzten Publikationen geradewegs zu den Gleichungen meiner Theorie zurück." (Hilbert 1924, 2)

> development of the laws holding in the interior of the atom—to answer this is the task for the future.[157]

In spite of his doubts, Hilbert was convinced that "his theory" would endure, (see the preceding paragraph), expressing the belief that it was of programmatic significance for future developments. Even if not, at least philosophical benefit could be drawn from it:

> I am convinced that the theory I have developed here contains an enduring core and creates a framework within which there is sufficient scope for the future development of physics in the sense of a field theoretical ideal of unity. In any case, it is also of epistemological interest to see how the few, simple assumptions I put forth in Axioms I, II, III, and IV suffice for the construction of the entire theory.[158]

The fact that his theory is not based exclusively on these axioms, but also depends rather crucially on other physical concepts, such as energy, and that his theory might change in content as well structure if these concepts changes their meaning,—all of this evidently remained outside of Hilbert's epistemological scope.

6. HILBERT'S ADOPTION OF EINSTEIN'S PROGRAM: THE SECOND PAPER AND ITS REVISIONS

6.1 From Paper 1 to Paper 2

When Hilbert published his Paper 1 in early 1916, he still hoped that his unification of electromagnetism and gravitation would provide the basis for solving the riddles of microphysics. He opened his paper announcing:

> I reserve for later communications the detailed development and particularly the special application of my basic equations to the fundamental questions of the theory of electricity.[159]

and concluding:

> ... I am also convinced that through the basic equations established here the most intimate, presently hidden processes in the interior of the atom will receive an explanation,

---

157 "Ob freilich das reine feldtheoretische Einheitsideal ein definitives ist, evtl. welche Ergänzungen und Modifikationen desselben nötig sind, um insbesondere die theoretische Begründung für die Existenz des negativen und des positiven Elektrons, sowie den widerspruchsfreien Aufbau der im Atominneren geltenden Gesetze zu ermöglichen,—dies zu beantworten, ist die Aufgabe der Zukunft." (Hilbert 1924, 2)

158 "Ich glaube sicher, daß die hier von mir entwickelte Theorie einen bleibenden Kern enthält und einen Rahmen schafft, innerhalb dessen für den künftigen Aufbau der Physik im Sinne eines feldtheoretischen Einheitsideals genügender Spielraum da ist. Auch ist es auf jeden Fall von erkenntnistheoretischem Interesse, zu sehen, wie die wenigen einfachen in den Axiomen I, II, III, IV von mir ausgesprochenen Annahmen zum Aufbau der ganzen Theorie genügend sind." (Hilbert 1924, 2)

159 "Die genauere Ausführung sowie vor Allem die spezielle Anwendung meiner Grundgleichungen auf die fundamentalen Fragen der Elektrizitätslehre behalte ich späteren Mitteilungen vor." (Hilbert 1916, 395)

> and in particular that generally a reduction of all physical constants to mathematical constants must be possible ...[160]

Clearly, he intended to dedicate a second communication to the physical consequences of his theory. By March 1916 he had submitted a second installment, which was then withdrawn, no trace remaining.[161] What does remain are the notes of Hilbert's SS 1916 and WS 1916/17 Lectures, and his related Causality Lecture. The WS 1916/17 Lectures offer hints of how his theory would lead to a modification of Maxwell's equations near the sources. While this part is clearly still related to Hilbert's original project, the bulk of these notes testify to his careful study of current work by Einstein and others on general relativity, as well as containing original contributions to that project. In the second communication to the Göttingen Academy submitted at the end of December 1916 (hereafter referred to as "Paper 2"), work on general relativity occupied the entire paper (Hilbert 1917). Hilbert's lecture notes are important for understanding the transition from his original aims to Paper 2, as well as the contents of this paper.[162] One of the most remarkable features of these notes is the openness and informality with which Hilbert shares unsolved problems with his students, later explicitly stating that this was a central goal of his lectures:

> In lectures, and above all in seminars, my guiding principle was not to present material in a standard and as smooth as possible way, just to help the students to maintain ordered notebooks. Above all, I tried to illuminate the problems and difficulties and offer a bridge leading to currently open questions. It often happened that in the course of a semester the program of an advanced lecture was completely changed because I wanted to discuss issues in which I was currently involved as a researcher and which had not yet by any means attained their definite formulation.[163]

6.2 The Causality Quandary

The lecture notes make it clear that Hilbert was still in a quandary over the treatment of causality because his Proofs argument against general covariance seemed to remain valid. The bulk of the typescript notes of his SS 1916 Lectures deal with special relativity (which he calls "die kleine Relativität"): kinematics, and vector and

160 "... ich bin auch der Überzeugung, daß durch die hier aufgestellten Grundgleichungen die intimsten, bisher verborgenen Vorgänge innerhalb des Atoms Aufklärung erhalten werden und insbesondere allgemein eine Zurückführung aller physikalischen Konstanten auf mathematische Konstanten möglich sein muß ..." (Hilbert 1916, 407)

161 See the discussion in (Sauer 1999, 560 n. 129).

162 The importance of Hilbert's lectures has been emphasized by Leo Corry. See (Corry 2004).

163 "Es war mein Grundsatz, in den Vorlesungen und erst recht in den Seminaren nicht einen eingefahrenen und so glatt wie möglich polierten Wissensstoff, der den Studenten das Führen sauberer Kolleghefte erleichtert, vorzutragen. Ich habe vielmehr immer versucht, die Probleme und Schwierigkeiten zu beleuchten und die Brücke zu den aktuellen Fragen zu schlagen. Nicht selten kam es vor, daß im Verlauf eines Semesters das stoffliche Programm einer höheren Vorlesung wesentlich abgeändert wurde, weil ich Dinge behandeln wollte, die mich gerade als Forscher beschäftigten und die noch keineswegs eine endgültige Gestalt gewonnen hatten." (Reidemeister 1971, 79) Translation by Leo Corry.

tensor analysis (pp. 1–66); dynamics (pp. 66–70 and 76–82); and Maxwell's electrodynamics (pp. 70–76 and 84–89). Hilbert then discusses Mie's theory in its original, special-relativistic form (pp. 90–102), and the need to combine it with "Einstein's concept of the general relativity of events" ("des Einstein'schen Gedankens von der allgemeinen Relativität des Geschehens," p. 103). After introducing the metric tensor, he develops the field equations for gravitation and electromagnetism (pp. 103–111). Discussing these equations, he notes that the causality problem remains unsolved:

> These are 14 equations for the 14 unknown functions $g^{\mu\nu}$ and q_h $(\mu, \nu = 1\ldots4)$. The causality principle may or may not be satisfied (the theory has not yet clarified this point). In any event, unlike the case of Mie's theory, the validity of this principle cannot be inferred from simple considerations. Of these 14 equations, 4 (e.g., the 4 Maxwell equations) are a consequence of the remaining 10 (e.g., the gravitational equations). Indeed, the remarkable theorem holds that the number of equations following from Hamilton's principle always corresponds to the number of unknown functions, except in the case occurring here, that the integral is an ["a general" added by hand] invariant.[164]

He still had not resolved the causality problem when he continued the lectures during the winter semester. Among other things, the WS 1916/17 Lecture notes contain much raw material for Paper 2. For example, the discussion of causal relations between events in a given spacetime very much resembles the treatment in that paper.[165] Yet the notes do not discuss the causality question for the field equations.

The same answer to this problem presented in Paper 2 is given in the typescript (unfortunately undated) of his Causality Lecture. From its contents, it is reasonable to conjecture that this is Hilbert's first exposition of his newly-found solution. After discussing the problem for his generally-covariant system of equations and constructing an example to illustrate its nature (pp. 1–5), he comments:

> Einstein's old theory now amounts to the addition of 4 non-invariant equations. But this too is mathematically incorrect. Causality cannot be saved in this way.[166]

---

164 "Dies sind 14 Gleichungen für die 14 unbekannten Funktionen $g^{\mu\nu}$ und q_h $(\mu, \nu, h = 1\ldots4)$. 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, das unter dem Integral ["eine allgemeine" added by hand] Invariante steht." (SS 1916 Lectures, 110)

165 See Chapter XIII of the notes, *Einiges über das Kausalitätsprinzip in der Physik,* 97–103, and pp. 57–59 of Paper 2, both of which are discussed below.

166 "Die alte Theorie von Einstein läuft nun darauf hinaus, 4 nicht invariante Gleichungen hinzuzufügen. Aber auch dies ist mathematisch falsch. Auf diesem Wege kann die Kausalität nicht gerettet werden" (p. 5). As discussed above, in his *Entwurf* theory Einstein did not first set up a system of generally-covariant equations and then supplement them by non-invariant conditions; but started from non-generally-covariant field equations. But he had considered the possibility described by Hilbert that these equations have a generally-covariant counterpart, from which they could be obtained by imposing non-invariant conditions.

A similar comment appears in Paper 2:

> In his original theory, now abandoned, A. Einstein (*Sitzungsberichte der Akad. zu Berlin*, 1914, p. 1067) had indeed postulated certain 4 non-invariant equations for the $g_{\mu\nu}$, in order to save the causality principle in its old form.[167]

Neither here nor in any later publication does Hilbert repeat the claim in the lecture notes that this procedure (which he himself had followed in the Proofs) is "mathematisch falsch," which strongly suggests that the notes precede Paper 2.

This suggested temporal sequence is confirmed by another pair of passages: In his lecture, Hilbert compares the problem created by general covariance of a system of partial differential equations and that created by parameter invariance in the calculus of variations:

> The difficulty of having to distinguish between a meaningful and a meaningless assertion is also encountered in Weierstrass's calculus of variations. There the curve to be varied is assumed to be given in parametric form, and one then obtains a differential equation for two unknown functions. One then considers only those assertions that remain invariant when the parameter p is replaced by an arbitrary function of p. [168]

This comparison may well have played a significant role in his solution of the causality problem. The corresponding passage in Paper 2 generalizes this comparison:

> In the theory of curves and surfaces, where a statement in a chosen parametrization of the curve or surface has no geometrical meaning for the curve or surface itself, if this statement does not remain invariant under an arbitrary transformation of the parameters or cannot be brought to invariant form; so also in physics we must characterize a statement that does not remain invariant under any arbitrary transformation of the coordinate system as *physically meaningless*.[169]

This argument is so much more general that it is hard to believe that, once he had hit upon it, Hilbert would have reverted to its restricted application to extremalization of curves. So we shall assume the priority of the Causality Lecture notes.

In these notes, Hilbert asserts that the causality quandary can be resolved by an appropriate understanding of physically meaningful statements:

---

167 "In seiner ursprünglichen, nunmehr verlassenen Theorie hatte A. Einstein (Sitzungsberichte der Akad. zu Berlin. 1914 S. 1067) in der Tat, um das Kausalitätsprinzip in der alten Fassung zu retten, gewisse 4 nicht invariante Gleichungen für die $g_{\mu\nu}$ besonders postuliert." (Hilbert 1917, 61)

168 "Auf die Schwierigkeit, zwischen einer sinnvollen und einer sinnlosen Behauptung unterscheiden zu müssen, stösst man übrigens auch in der Weierstrass'schen Variationsrechnung. Dort wird die zu variierende Kurve als in Parametergestalt gegeben angenommen, und man erhält dann eine Differentialgleichung für zwei unbekannte Funktionen. Man betrachtet dann nur solche Aussagen, die invariant bleiben, wenn man den Parameter p durch eine willkürliche Funktion von p ersetzt." (Causality Lecture, 8)

169 "Gerade so wie in der Kurven- und Flächentheorie eine Aussage, für die die Parameterdarstellung der Kurve oder Fläche gewählt ist, für die Kurve oder Fläche selbst keinen geometrischen Sinn hat, wenn nicht die Aussage gegenüber einer beliebigen Transformation der Parameter invariant bleibt oder sich in eine invariante Form bringen läßt, so müssen wir auch in der Physik eine Aussage, die nicht gegenüber jeder beliebigen Transformation des Koordinatensystems invariant bleibt, als *physikalisch sinnlos* bezeichnen." (Hilbert 1917, 61)

> We obtain the explanation of this paradox by attempting to more rigorously grasp the concept of relativity. It does not suffice to say that the laws of the world are independent of the frame of reference, but rather every single assertion about an event or a concurrence of events only then takes on a physical meaning if it is independent of its designation, i.e. when it is invariant.[170]

In the last clause, one hears distant echoes of Einstein's assertion in his expository paper *Die Grundlage der allgemeinen Relativitätstheorie*:

> We allot to the universe four spacetime variables x_1, x_2, x_3, x_4 in such a way that for every point-event there is a corresponding system of values of the variables $x_1 \ldots x_4$. To two coincident point-events there corresponds one system of values of the variables $x_1 \ldots x_4$, i.e. coincidence is characterized by the identity of the co-ordinates. ... As all our physical experience can be ultimately reduced to such coincidences, there is no immediate reason for preferring certain systems of coordinates to others, that is to say, we arrive at the requirement of general co-variance.[171]

Perusal of this paper, published on 11 May 1916 and cited in Hilbert's WS 1916/17 Lectures,[172] may well have contributed to his new understanding of the causality problem.

However, Hilbert's interpretation of a physically meaningful statement actually differs from that of Einstein. Einstein had turned the uniqueness problem for solutions of generally-covariant field equations into an argument against the physical significance of coordinate systems. Hilbert attempted to turn the problem into its own solution by *defining* physically meaningful statements as those for which no such ambiguities arise, whether such statements employ coordinate systems or not. In his Causality Lecture, Hilbert claims to demonstrate the validity of the "causality principle," formulated in terms of physically meaningful statements:

> We would like to prove that the causality principle formulated as follows: "All meaningful assertions are a necessary consequence of the preceding ones [see the citation above]" is valid. Only this theorem is logically necessary and, for physics, also completely sufficient.[173]

To establish this principle, he considers an arbitrary set of generally-covariant field equations (which he calls "ein System invarianter Gleichungen") involving the

---

170 "Die Aufklärung dieses Paradoxons erhalten wir, wenn wir nun den Begriff der Relativität schärfer zu erfassen suchen. Man muss nämlich nicht nur sagen, dass die Weltgesetze vom Bezugssystem unabhängig sind, es hat vielmehr jede einzelne Behauptung über eine Begebenheit oder ein Zusammentreffen von Begebenheiten physikalisch nur dann einen Sinn, wenn sie von der Benennung unabhängig, d.h. wenn sie invariant ist." (Causality Lecture, 5–6)

171 "Man ordnet der Welt vier zeiträumliche Variable x_1, x_2, x_3, x_4 zu, derart, dass jedem Punktereignis ein Wertsystem der Variablen $x_1 \ldots x_4$ entspricht. Zwei koinzidierenden Punktereignissen entspricht dasselbe Wertsystem der Variablen $x_1 \ldots x_4$; d. h. die Koinzidenz ist durch die Übereinstimmung der Koordinaten charakterisiert. .... Da sich alle unsere physikalischen Erfahrungen letzten Endes auf solche Koinzidenzen zurückführen lassen, ist zunächst kein Grund vorhanden, gewisse Koordinatensysteme vor anderen zu bevorzugen, d.h. wir gelangen zu der Forderung der allgemeinen Kovarianz." (Einstein 1916a, 776–777)

172 See (WS 1916/17 Lectures, 112).

metric tensor, the electromagnetic potentials, and their derivatives.[174] He specifies the values of these fields and their derivatives on the space-like hypersurface $t = 0$, which he calls "the present" ("die Gegenwart"); and considers coordinate transformations that do not change the coordinates on this hypersurface, but are otherwise arbitrary (except for continuity and differentiability) off the hypersurface ("die Transformation soll die Gegenwart ungeändert lassen"). He then defines a physically meaningful statement as one that is uniquely determined by Cauchy data, intending to thus establish, at the same time, his principle of causality in terms of what one might call "a mathematical response" to the problem of uniqueness in a generally-covariant field theory:

> Only such a [meaningful assertion] is unequivocally determined by the initial values of $g_{\mu\nu}$, q_μ and their derivatives, and in fact these initial values are to be understood as Cauchy boundary-value conditions. It must be accepted that one can prescribe these boundary values arbitrarily, or that one can proceed to a place in the world at the moment in time when the state characterized by these values prevails. The observer of nature is also considered as standing outside these physical laws; otherwise one would arrive at the antinomies of free will.[175]

As this passage makes clear, Hilbert's proposed definition of physically meaningful statements and clarification of the problem of causality is flawed by the still-unrecognized intricacies of the Cauchy problem in general relativity. He evidently failed to realize that the classical notion of freely-choosable initial values no longer works for generally-covariant field equations since some of them function as constraints on the data that can be given on an initial hypersurface, rather than as evolution equations for that data off this surface. The next section discusses Hilbert's treatment of the problem of causality in Paper 2, including further evidence of his failure to fully grasp Einstein's insight that, in general relativity, coordinate systems have no physical significance of their own.

---

173 "Wir wollen beweisen, dass das so formulierte Kausalitätsprinzip: "Alle sinnvollen Behauptungen sind eine notwendige Folge der vorangegangenen [see the citation above]" gültig ist. Dieser Satz allein ist logisch notwendig und er ist auch für die Physik vollkommen ausreichend." (Causality Lecture, 5–6)

174 The original typescript had specified first and second derivatives of the metric and first derivatives of the electromagnetic potentials, but by hand Hilbert added "beliebig hohen" in the first case and deleted "ersten" in the second.

175 "Nur eine solche [sinnvolle Behauptung] ist durch die Anfangswerte der $g_{\mu\nu}$, q_μ und ihrer Ableitungen eindeutig festgelegt und zwar sind diese Anfangswerte als Cauchy'sche Randbedingungen zu verstehen. Dass man diese Randwerte beliebig vorgeben kann, oder dass man sich an eine Stelle der Welt hinbegeben kann, wo der durch diese Werte charakterisierte Zustand in diesem Zeitmoment herrscht, muss hingenommen werden. Der die Natur beobachtende Mensch wird eben als ausserhalb dieser physikalischen Gesetze stehend betrachtet; sonst käme man zu den Antinomien der Willensfreiheit." (Causality Lecture, 6–7)

6.3 Hilbert at Work on General Relativity

Paper 2 shows that Hilbert's original goal of developing a unified gravito-electromagnetic theory, with the aim of explaining the structure of the electron and the Bohr atom, has been modified in the light of the successes of Einstein's purely gravitational program. Hilbert's shift of emphasis in Paper 1 to the primacy of the gravitational field equations must have facilitated his shift to the consideration of the "empty-space" field equations. From Hilbert's perspective, they are just that subclass of solutions to his fourteen "unified" field equations, for which the electromagnetic potentials vanish. This makes them formally equivalent to the sub-class of solutions to Einstein's field equations with a stress-energy tensor that either vanishes everywhere, or at least outside of some finite world-tube containing the sources of the field. This formal equivalence no doubt contributed to the ease with which contemporary mathematicians and physicists assimilated Hilbert's program to Einstein's, treating Paper 2 as a contribution to the development of the general theory of relativity. This is how Hilbert's contribution came to be assimilated to the relativistic tradition, as we shall discuss in more detail below.

Let us now take a look at the six major topics Hilbert treated in Paper 2:

1. measurement of the components of the metric tensor (Hilbert 1917, 53–55);
2. characteristics and bicharacteristics of the Hamilton-Jacobi equation corresponding to the metric tensor (Hilbert 1917, 56–57);
3. causal relation between events in a spacetime with given metric (Hilbert 1917, 57–59);
4. the causality problem for the field equations determining the metric tensor (Hilbert 1917, 59–63);
5. Euclidean geometry as a solution to the field equations—in particular, the investigation of conditions that characterize it as a unique solution (Hilbert 1917, 63–66 and 70); and
6. the Schwarzschild solution, its derivation (Hilbert 1917, 67–70), and determination of the paths of (freely-falling) particles and light rays in it (Hilbert 1917, 70–76).

1) The metric tensor and its measurement: First of all, Hilbert dropped his previous use of one imaginary coordinate, perhaps influenced by Einstein's use of real coordinates, and emphasized that the $g_{\mu\nu}$, now all real, provide the "Massbestimmung einer Pseudogeometrie" (Hilbert 1917, 54). He classified the elements ("Stücke") of all curves: time-like elements measure proper time; space-like elements measure length; and null elements are segments of a light path. He introduced two ideal measuring instruments: a measuring tape ("Maßfaden") for lengths, and a light clock ("Lichtuhr") for proper times. He makes a comment that suggests, in spite of his remarks in Paper 1 and the Causality Lecture (see above), a lingering belief in some objective significance to the choice of a coordinate system, independently of the metric tensor:

> First we show that each of the two instruments suffices to compute with its aid the values of the $g_{\mu\nu}$ as functions of x_s, just as soon as a definite spacetime coordinate system x_s has been introduced.[176]

He ends with some comments on a possible axiomatic construction ("Aufbau") of the pseudogeometry, suggesting the need for two axioms:

> first an axiom should be established, from which it follows that length resp. proper time must be integrals whose integrand is only a function of the x_s and their first derivatives with respect to the parameter [p, where $x_s = x_s(p)$ is the parametric representation of a curve]; ...
>
> Secondly an axiom is needed whereby the theorems of the pseudo-Euclidean geometry, that is the old principle of relativity, shall be valid in infinitesimal regions;[177]

2) Characteristics and bicharacteristics: Hilbert defined the null cone at each point, and pointed out that the Monge differential equation (Hilbert 1917, 56):

$$g_{\mu\nu}\frac{dx_\mu}{dp}\frac{dx_\nu}{dp} = 0, \tag{129}$$

and the corresponding Hamilton-Jacobi partial differential equation:

$$g^{\mu\nu}\frac{\partial f}{\partial x_\mu}\frac{\partial f}{\partial x_\nu} = 0, \tag{130}$$

determine the resulting null cone field, the geodesic null lines being the characteristics of the first and the bicharacteristics of the second of these equations. The null geodesics emanating from any world point form the null conoid ("Zeitscheide;" many current texts apply the term "null cone" to non flat spacetimes, but we prefer the term "conoid") emanating from that point. He points out that the equation for these conoids are integral surfaces of the Hamilton-Jacobi equation; and that all timelike world lines emanating from a world point lie inside its conoid, which forms their boundary.

These topics, rather briefly discussed in Paper 2, are treated much more extensively in Hilbert's WS 1916/17 Lectures. In many ways Hilbert's discussion in

---

176 "Zunächst zeigen wir, daß jedes der beiden Instrumente ausreicht, um mit seiner Hülfe die Werte der $g_{\mu\nu}$ als Funktion von x_s zu berechnen, sobald nur ein bestimmtes Raum-Zeit-Koordinatensystem x_s eingeführt worden ist." (Hilbert 1917, 55)

177 "erstens ist ein Axiom aufzustellen, auf Grund dessen folgt, daß Länge bez. Eigenzeit Integrale sein müssen, deren Integrand lediglich eine Funktion der x_s und ihrer ersten Ableitungen nach dem Parameter ist; ...
Zweitens ist ein Axiom erforderlich, wonach die Sätze der pseudo-Euklidischen Geometrie d.h. das alte Relativitätsprinzip im Unendlichkleinen gelten soll;" (Hilbert 1917, 56)

Paper 2 reads like a précis of these notes; it becomes much more intelligible if they are consulted. Chapter IX (pp. 69–80) entitled "Die Monge'sche Differentialgleichung" also treats the Hamilton-Jacobi equation and the theory of characteristics, emphasizing their relation to the Cauchy problem, and the reciprocal relation between integral surfaces of the Hamilton-Jacobi equation (the null conoids are called "transzendentale Kegelfläche") and null curves. Chapters X (pp. 80–82, "Die vierdimensionale eigentliche u. Pseudogeometrie") and XI (pp. 82–97, "Zusammenhang mit der Wirklichkeit") cover the material in the first section of Paper 2: the measuring tape ("Massfaden") is discussed in section 38 (pp. 85–86 and pp. 91–92), and the light clock, already introduced in the context of special relativity (see the SS 1916 Lectures, 6–10), is reintroduced in section 44 (pp. 93–94, "Axiomatische Definition der Lichtuhr"). Both instruments are used to determine the components of the metric tensor as functions of the coordinates, "sobald nur ein bestimmtes Raum-Zeit Koordinatensystem x_i eingeführt worden ist" (p. 95).

3) Causal relation between events: [178] In accord with the implicit requirement that three of the coordinates be space-like and one time-like, Hilbert imposes corresponding conditions on the components of the metric tensor. But he has a unique way of motivating them:

> Up to now all coordinate systems x_s that result from any one by arbitrary transformation have been regarded as equally valid. This arbitrariness must be restricted when we want to realize the concept that two world points on the same time line can be related as cause and effect, and that it should then no longer be possible to transform such world points to be simultaneous. In declaring x_4 as the *true* time coordinate we adopt the following definition:
>
> ...
>
> So we see that the concepts of cause and effect, which underlie the principle of causality, also do not lead to any inner contradictions whatever in the new physics, if we only take the inequalities (31) always to be part of our basic equations, that is if we confine ourselves to using *true* spacetime coordinates.[179]

Again, he seems to believe that there is some residual physical significance in the choice of a coordinate system: it must reflect the relations of cause and effect between events on the same time-like world line. He defines a proper ("eigentliches") coordinate system as one, in which (in effect) the first three coordinates are space-like and the fourth time-like in nature; transformations between such proper coordinate systems are also called proper. Given Hilbert's stated goal of restricting the choice of coordinates to those that reflect the causal order on all time-like world lines, his con-

---

178 This section also includes material from Hilbert's WS 1916/17 Lectures: Chapter XII, *Einiges über das Kausalitätsprinzip in der Physik,* (pp. 97–104) covers the same ground as, but in no more detail than, the text of Paper 2.

ditions are sufficient but not necessary since they exclude retarded null coordinates, which also preserve this causal order.

4) Causality problem for the field equations: As noted, Hilbert's analysis follows his Causality Lecture. In Paper 2 he writes:

> Concerning the principle of causality, let the physical quantities and their time derivatives be known at the present in some given coordinate system: then a statement will only have physical meaning if it is invariant under all those transformations, for which the coordinates just used for the present remain unchanged; I maintain that statements of this type for the future are all uniquely determined, that is, *the principle of causality holds in this form*:
>
> *From present knowledge of the 14 physical potentials $g_{\mu\nu}$, q_s all statements about them for the future follow necessarily and uniquely provided they are physically meaningful*.[180]

A hasty reading might suggest that Hilbert is asserting the independence of all physically meaningful statements from the choice of a coordinate system, and he has often been so interpreted; but this is not what he actually says. His very definition of physically meaningful ("physikalisch Sinn haben") involves the class of coordinate systems that leave the coordinates on the initial hypersurface ("die Gegenwart") unchanged. Secondly, Hilbert uses a Gaussian coordinate system, introduced earlier,[181] in order to prove his assertion about the causality principle.[182] Finally, if his words were so interpreted, they would stand in flagrant contradiction to his earlier statements (cited above)

179 "Bisher haben wir alle Koordinatensysteme x_s die aus irgend einem durch eine willkürliche Transformation hervorgehen, als gleichberechtigt angesehen. Diese Willkür muß eingeschränkt werden, sobald wir die Auffassung zur Geltung bringen wollen, daß zwei auf der nämlichen Zeitlinie gelegene Weltpunkte im Verhältnis von Ursache und Wirkung zu einander stehen können und daß es daher nicht möglich sein soll, solche Weltpunkte auf gleichzeitig zu transformieren.

...

So sehen wir, daß die dem Kausalitätsprinzip zu Grunde liegenden Begriffe von Ursache und Wirkung auch in der neuen Physik zu keinerlei inneren Widersprüche führen, sobald wir nur stets die Ungleichungen (31) [the conditions Hilbert imposes on the metric tensor] zu unseren Grundgleichungen hinzunehmen d.h. uns auf den Gebrauch *eigentlicher* Raum-zeitkoordinaten beschränken." (Hilbert 1917, 57 and 58)

180 "Was nun das Kausalitätsprinzip betrifft, so mögen für die Gegenwart in irgend einem gegebenen Koordinatensystem die physikalischen Größen und ihre zeitlichen Ableitungen bekannt sein: dann wird eine Aussage nur physikalisch Sinn haben, wenn sie gegenüber allen denjenigen Transformationen invariant ist, bei denen eben die für die Gegenwart benutzten Koordinaten unverändert bleiben; ich behaupte, daß die Aussagen dieser Art für die Zukunft sämtlich eindeutig bestimmt sind d.h. das Kausalitätsprinzip gilt in dieser Fassung:

Aus der Kenntnis der 14 physikalischen Potentiale $g_{\mu\nu}$, q_s in der Gegenwart folgen alle Aussagen über dieselben für die Zukunft notwendig und eindeutig, sofern sie physikalischen Sinn haben." (Hilbert 1917, 61)

181 See (Hilbert 1917, 58–59).

182 See (Hilbert 1917, 61–62).

about the measurement of the metric and the causal relation between events which presuppose attaching some residual physical meaning to the choice of coordinates.

His proof consists of a brief discussion of the Cauchy problem for the field equations in a Gaussian coordinate system. One of us has discussed this aspect of his work elsewhere (Stachel 1992), so we shall be brief here. He only considers the ten gravitational field equations (51) since he interprets Theorem I of Paper 1 as showing that the other four (52) follow from them. Gaussian coordinates eliminate four of the 14 field quantities, the $g_{0\mu}$, leaving only ten (the six g_{ab}, $a, b = 1, 2, 3$, and the four q_s), so he concludes that the resulting system of equations is in Cauchy normal form. This treatment is erroneous on several counts, but we postpone discussion of this question until the next section. More relevant to the present topic is Hilbert's statement:

> Since the Gaussian coordinate system itself is uniquely determined, therefore also all statements about those potentials (34) [the ten potentials mentioned above] with respect to these coordinates are of invariant character.[183]

He never discusses the behavior of the initial data under coordinate transformations on the initial hypersurface (three-dimensional hypersurface diffeomorphisms in modern terminology), confirming that his treatment is still tied to the use of particular coordinate systems rather than being based on coordinate-invariant quantities.

Finally, his discussion of how to implement the requirement of physically meaningful assertions depends heavily on the choice of a coordinate system. He remarks:

> The forms, in which physically meaningful, i.e. invariant, statements can be expressed mathematically are of great variety.[184]

and proceeds to discuss three ways:

> *First*. This can be done by means of an invariant coordinate system. ...
>
> *Second*. The statement, *according to which a coordinate system can be found* in which the 14 potentials $g_{\mu\nu}$, q_s have certain definite values in the future, or fulfill certain definite conditions, is always an invariant and therefore a physically meaningful one. ...
>
> *Third*. A statement is also invariant and thus has physical meaning if it is supposed to be valid in any arbitrary coordinate system.[185]

---

183 "Da das Gaußische Koordinatsystem selbst eindeutig festgelegt ist, so sind auch alle auf dieses Koordinatensystem bezogenen Aussagen über jene Potentiale (34) von invariantem Charakter." (Hilbert 1917, 62)

184 "Die Formen in denen physikalisch sinnvolle d.h. invariante Aussagen mathematisch zum Ausdruck gebracht werden können, sind sehr mannigfaltig." (Hilbert 1917, 62)

185 "Erstens. Dies kann mittelst eines invarianten Koordinatensystem geschehen. ...
Zweitens. Die Aussage, wonach sich ein Koordinatensystem finden läßt, in welchem die 14 Potentiale $g_{\mu n}$, q_s für die Zukunft gewisse bestimmte Werte haben oder gewisse Beziehungen erfüllen, ist stets eine invariante und daher physikalisch sinnvoll. ...
Drittens. Auch ist eine Aussage invariant und hat daher stets physikalisch Sinn, wenn sie für jedes beliebige Koordinatensystem gültig sein soll." (Hilbert 1917, 62–63)

The first two ways explicitly depend on the choice of a coordinate system, which is not necessarily unique. As examples of the first way, he cites Gaussian and Riemannian coordinates. It is true that, discussing the second, he notes:

> The mathematically invariant expression for such a statement is obtained by eliminating the coordinates from those relations.[186]

But he does not give an example, nor does he suggest the most obvious way of realizing his goal, if indeed it was a coordinate-independent solution to the problem: the use of invariants as coordinates. As Kretschmann noted a few years later, in matter- and field-free regions the four non-vanishing invariants of the Riemann tensor may be used as coordinates. If the metric is then expressed as a function of these coordinates, its components themselves become invariants.[187] The use of such coordinates was taken up again by Arthur Komar in the 1960s, and today they are often called Kretschmann-Komar coordinates.[188]

One might think that Hilbert had in mind something like this in his third suggested way. However, the example he cites makes it clear that he meant something else:

> An example of this are Einstein's energy-momentum equations having divergence character. For, although Einstein's energy [that is, the gravitational energy-momentum pseudotensor] does not have the property of invariance, and the differential equations he put down for its components are by no means covariant as a system of equations, nevertheless the assertion contained in them, that they shall be satisfied in any coordinate system, is an invariant demand and therefore it carries physical meaning.[189]

Rather than invariant quantities, evidently he had in mind non-tensorial entities and sets of equations, which nevertheless take the same form in every coordinate system.

In summary, Hilbert's treatment in Paper 2 of the problem of causality in general relativity still suffers from many of the flaws in his original approach. In particular, physical significance is still ascribed to coordinate systems, and the claim is maintained that the identities following from Theorem I represent a coupling between the two sets of field equations. On the other hand, his efforts to explore the solutions of the gravitational field equations from the perspective of a mathematician produced significant contributions to general relativity, to be discussed later.

---

186 "Der mathematische invariante Ausdruck für eine solche Aussage wird durch Elimination der Koordinaten aus jenen Beziehungen erhalten." (Hilbert 1917, 62–63)

187 See (Kretschmann 1917).

188 See (Komar 1958).

189 "Ein Beispiel dafür sind die Einsteinschen Impuls-Energiegleichungen vom Divergenz Character. Obwohl nämlich die Einsteinsche Energie die Invarianteneigenschaft nicht besitzt und die von ihm aufgestellten Differentialgleichungen für ihre Komponenten auch als Gleichungssystem keineswegs kovariant sind, so ist doch die in ihnen enthaltene Aussage, daß sie für jedes beliebige Koordinatensystem erfüllt sein sollen, eine invariante Forderung und hat demnach einen physikalischen Sinn." (Hilbert 1917, 63)

5) Euclidean geometry: This section opens with some extremely interesting general comments contrasting the role of geometry in what Hilbert calls the old and the new physics:

> The old physics with the concept of absolute time took over the theorems of Euclidean geometry and without question put them at the basis of every physical theory. ...
>
> The new physics of Einstein's principle of general relativity takes a totally different position vis-à-vis geometry. It takes neither Euclid's nor any other particular geometry *a priori* as basic, in order to deduce from it the proper laws of physics, but, as I showed in my first communication, the new physics provides at one fell swoop through one and the same Hamilton's principle the geometrical and the physical laws, namely the basic equations (4) and (5) [the ten gravitational and four electromagnetic field equations], which tell us how the metric $g_{\mu\nu}$ —at the same time the mathematical expression of the phenomenon of gravitation—is connected with the values q_s of the electrodynamic potentials.[190]

Hilbert declares:

> With this understanding, an old geometrical question becomes ripe for solution, namely whether and in what sense Euclidean geometry—about which we know from mathematics only that it is a logical structure free from contradictions—also possesses validity in the real world.[191]

He later formulates this question more precisely:

> The geometrical question mentioned above amounts to the investigation, whether and under what conditions the four-dimensional Euclidean pseudo-geometry [i.e., the Minkowski metric] ... is a solution, or even the only regular solution, of the basic physical equations.[192]

Hilbert thus takes up a problem that emerged with the development of non-Euclidean geometry in the 19th century and considered by such eminent mathematicians as Gauss and Riemann: the question of the relation between geometry and physical real-

190 "Die alte Physik mit dem absoluten Zeitbegriff übernahm die Sätze der Euklidische Geometrie und legte sie vorweg einer jeden speziellen physikalischen Theorie zugrunde. ...
Die neue Physik des Einsteinschen allgemeinen Relativitätsprinzips nimmt gegenüber der Geometrie eine völlig andere Stellung ein. Sie legt weder die Euklidische noch irgend eine andere bestimmte Geometrie vorweg zu Grunde, um daraus die eigentlichen physikalischen Gesetze zu deduzieren, sondern die neue Theorie der Physik liefert, wie ich in meiner ersten Mitteilung gezeigt habe, mit einem Schlage durch ein und dasselbe Hamiltonsche Prinzip die geometrischen und die physikalischen Gesetze nämlich die Grundgleichungen (4) und (5), welche lehren, wie die Maßbestimmungen $g_{\mu\nu}$ — zugleich der mathematischen Ausdruck der physikalischen Erscheinung der Gravitation — mit den Werten q_s der elektrodynamischen Potentiale verkettet ist." (Hilbert 1917, 63–64)

191 "Mit dieser Erkenntnis wird nun eine alte geometrische Frage zur Lösung reif, die Frage nämlich, ob und in welchem Sinne die Euklidische Geometrie — von der wir aus der Mathematik nur wissen, daß sie ein logisch widerspruchsfreier Bau ist — auch in der Wirklichkeit Gültigkeit besitzt." (Hilbert 1917, 63)

192 "Die oben genannte geometrische Frage läuft darauf hinaus, zu untersuchen, ob und unter welchen Voraussetzungen die vierdimensionale Euklidische Pseudogeometrie ... eine Lösung der physikalischen Grundgleichungen bez. die einzige reguläre Lösung derselben ist." (Hilbert 1917, 64)

ity. For a number of reasons, this question was not central to Einstein's heuristic. He had never addressed the question posed by Hilbert: the conditions under which Minkowski spacetime is a unique solution to the gravitational field equations. To Einstein, the question of the Newtonian limit, and hence the incorporation of Newton's theory into his new theory of gravitation, was much more important than the question of the existence of matter-free solutions to his equations. Indeed, this question was a rather embarrassing one for Einstein since such solutions display inertial properties of test particles even in the absence of matter, a feature that he had difficulty in accepting because of his Machian conviction that all inertial effects must be due to interaction of masses.[193] By establishing a connection between general relativity and the mathematical tradition questioning the geometry of physical space, Hilbert made a significant contribution to the foundations of general relativity.

In attempting to answer the question of the relation between Minkowski spacetime and his equations, Hilbert first of all notes that, if the electrodynamic potentials vanish, then the Minkowski metric is a solution of the resulting equations, i.e., of the vanishing of what we now call the Einstein tensor.[194] He then poses the converse question: under what conditions is the Minkowski metric the *only* regular solution to these equations? He considers small perturbations of the Minkowski metric (a technique that Einstein had already introduced) and shows that, if these perturbations are time independent (curiously, here reverting to use of an imaginary time coordinate) and fall off sufficiently rapidly and regularly at infinity, then they must vanish everywhere. In the next section of the paper, he proves another relevant result, which we shall discuss below.

This section of Paper 2 is a condensation of material covered in his WS 1916/17 Lectures:

- in the table of contents (p. 197), pp. 104–106 are entitled: "Der Sinn der Frage: Gilt die Euklidische Geometrie?"
- pp. 109–111 are headed "Gilt die Euklidische Geometrie in der Physik?" in the typescript, with the handwritten title "Die Grundgleichungen beim Fehlen von Materie" added in the margin, and entitled "Aufstellung der Grundgleichungen beim Fehlen der Materie" in the table of contents; and
- pp. 111–112, bear the handwritten title "Zwei Sätze über die Gültigkeit der Euklidischen Geometrie" in the margin, and "Zwei noch unbewiesene Sätze über die Gültigkeit der Pseudoeuklidischen Geometrie in der Physik" in the table of contents.

The lecture notes make much clearer than Paper 2 Hilbert's motivation for a discussion of the empty-space field equations in general, and of the Schwarzschild metric in particular. In the notes, Hilbert introduces the field equations in section 51 (WS

193 For a historical discussion, see (Renn 1994).

194 "wenn alle Elektrizität entfernt ist, so ist die pseudo-Euklidische Geometrie möglich" See (Hilbert 1917, 64).

1916/17 Lectures, 106–109),[195] sandwiched between discussions of his motivation for raising the question of the validity of Euclidean geometry and his attempts to answer it. At the end of the previous section he points out:

> We would like to anticipate the results of our calculation: in general our basic physical equations have no solutions at all. In my opinion, this is a positive result of the theory: since in no way are we able to impose Euclidean geometry on nature through a different interpretation of experiments. Assuming namely that my basic physical equations to be developed are really correct, then no other physics is possible, i.e., reality cannot be understood in a different way.[196]

Hilbert evidently thought he had found a powerful argument against geometric conventionalism—presumably, he had Poincaré in mind here. He continues:

> On the other hand we shall see that under certain very specialized assumptions—perhaps the absence of matter throughout space is sufficient for this—the only solution to the differential equations are $g_{\mu\nu} = \delta_{\mu\nu}$ [the Minkowski metric].[197]

At this point, the problem of the status of geometry is broadened from three-dimensional geometry to four dimensional pseudo-geometry—and in particular the question of the status of Euclidean geometry is broadened to that of four-dimensional Minkowski pseudo-geometry. In this form, it plays a central role in Hilbert's thinking about his program. This problem, rooted as it was in a mathematical tradition going back to Gauss, led him naturally to consider what we call the empty-space Einstein field equations. He hoped that the absence of matter and non-gravitational fields might suffice to uniquely single out the Minkowski metric as a solution to his field equations (which are identical to Einstein's in this case):

> It is possible that the following theorem is correct:
> Theorem: If one removes all electricity from the world (i.e. $q_i \equiv 0$) and demands absolute regularity—i.e. the possibility of expansion in a power series—of the gravitational potentials $g_{\mu\nu}$ (a requirement that in our opinion must always be fulfilled, even in the general case), then Euclidean geometry prevails in the world, i.e. the 10 equations (3) [equation number in the original; the vanishing of the Einstein tensor] have $g_{\mu\nu} = \delta_{\mu\nu}$ as their only solution.[198]

(He explains what he means by "regular" in his discussion of the Schwarzschild metric, considered below.) Of course, Hilbert was *not* able to establish this theorem, since it is not true, as Einstein's work on gravitational waves might already have sug-

---

195 Page 107 is missing from the typescript.

196 "Wir wollen das Resultat unserer Rechnung vorwegnehmen: unsere physikalischen Grundgleichungen haben im allgemeinen keineswegs Lösungen. Dies ist meiner Meinung nach ein positives Resultat der Theorie: denn wir können der Natur die Euklidischen Geometrie durch andere Deutung der Experimente durchaus nicht aufzwingen. Vorausgesetzt nämlich, dass meine zu entwickelnden physikalischen Grundgleichungen wirklich richtig sind, so ist auch keine andere Physik möglich, d.h., die Wirklichkeit kann nicht anders aufgefasst werden." (WS 1916/17 Lectures, 106)

197 "Andererseits werden wir sehen, dass unter gewissen sehr spezialisierenden Voraussetzungen—vielleicht ist das Fehlen von Materie im ganzen Raum dazu schon hinreichend—die einzige Lösungen der Diffentialgleichungen $g_{\mu\nu} = \delta_{\mu\nu}$ [the Minkowski metric] sind."

gested (Einstein 1916c). Nor was he able to find any other set of necessary and sufficient conditions for the uniqueness of the Minkowski metric; but he did almost establish one set of sufficient conditions and proved another:

> I consider the following theorem to be very probably correct: If one removes all electricity from the world and demands for the gravitational potential, apart from the self-evident requirement of regularity, that $g_{\mu\nu}$ is independent of t, i.e. that gravitation is static, and finally [one demands] also regular behavior at infinity, then $g_{\mu\nu} = \delta_{\mu\nu}$ are the only solutions to the gravitation equations (3)[equation number in the original].
> I can now already prove this much of the theorem, that in the neighborhood of Euclidean geometry there are certainly no solutions to these equations.[199]

This is, of course, the result that he did prove in Paper 2 (see above). The proof of this result for the full, non-linear field equations hung fire for a long time with several proofs for the case of static metrics being given over the years; the proof for stationary metrics was finally given by André Lichnerowicz in 1946.[200]

6) The Schwarzschild solution: The Schwarzschild solution had already been published (Schwarzschild 1916) and Hilbert dedicates considerable space to it, both in his lecture notes and in Paper 2. He uses it in the course of his effort to exploit the new tools of general relativity for addressing the foundational questions of geometry raised in the mathematical tradition. In his lecture notes, he introduces a number of assumptions on the metric tensor in order to prove a theorem on the uniqueness of Euclidean geometry:

> 1) Let $g_{\mu\nu}$ again be independent of t.
>
> 2) Let $(g_{\nu 4} = 0)$ $(\nu = 1, 2, 3)$ [interpolated by hand: "i.e. Gaussian coordinate system, which can always be introduced by a transformation"] (Orthogonality of the t- axis to the x_1, x_2, x_3 -space, the so-called metric space.)
>
> 3) There is a distinguished point in the world, with respect to which central symmetry holds, i.e. the rotation of the coordinate system around this point is a transformation of the world onto itself.

---

198 "Es ist möglich, dass folgender Satz richtig ist:
Satz: Nimmt man alle Elektrizität aus der Welt hinweg (d.h. $q_i \equiv 0$) und verlangt man absolute Regularität—d.h. Möglichkeit der Entwicklung in eine Potenzreihe—der Gravitationspotentiale $g_{\mu\nu}$ (eine Forderung, die nach 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." (WS 1916/17 Lectures, 111–112)

199 "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 Lösung dieser Gleichungen vorhanden sind." (WS 1916/17 Lectures, 112)

200 See (Lichnerowicz 1946).

> Now the following theorem holds: If the gravitational potentials fulfill conditions 1–3, then Euclidean geometry is the only solution to the basic physical equations.[201]

The proof of this theorem leads him to consider the problem of spherically-symmetric solutions to the empty-space Einstein field equations, a problem that Hilbert notes had previously been treated by Einstein (in the linear approximation) and Schwarzschild (exactly). He claims for his own calculations only that, compared to those of others, they are "auf ein Minimum reduziert" (WS 1916/17 Lectures, 113) by working from his variational principle for the field equations (see above). Hermann Weyl gave a similar variational derivation in 1917 (Weyl 1917); the section of his book *Raum-Zeit-Materie* on the Schwarzschild metric includes a reference to Hilbert's Paper 2, which reproduces Hilbert's variational derivation, (Weyl 1918a; 1918b, 230 n.9; 1923, 250 n.19). But Pauli's magisterial survey of the theory of relativity mentions only Weyl's paper, this probably contributing to the neglect of Hilbert's contribution in most later discussions (Pauli 1921).

In Paper 2, Hilbert derives the Schwarzschild metric from the same three assumptions as in the lecture notes, emphasizing that:

> In the following I present for this case a procedure that makes no assumptions about the gravitational potentials $g_{\mu\nu}$ at infinity, and which moreover offers advantages for my later investigations.[202]

In spite of this, many later derivations of the Schwarzschild metric still continue to impose unnecessary boundary conditions. But Hilbert did not show that the assumption of time-independence is also unnecessary, as proved by Birkhoff in 1923. (The assertion that the Schwarzschild solution is the only spherically symmetric solution to the empty-space Einstein equations is known as Birkhoff's theorem.)[203]

Hilbert's discussion of the Schwarzschild solution also raises the problem of its singularities and their relation to Hilbert's theory of matter. In his lecture notes, after establishing the Schwarzschild metric, he writes:

---

201 "1) Es sei wieder $g_{\mu\nu}$ unabhängig von t.
2) Es sei $g_{\nu 4} = 0 \quad \nu = 1, 2, 3$ [interpolated by hand: "d.h. Gauss'sches Koordinatensystem, das durch Transformation immer eingeführt werden kann"] (Orthogonalität der t-Achse auf dem x_1, x_2, x_3 -Raum, dem sogenannten Streckenraum.)
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." (WS 1916/17 Lectures, 113)

202 "Ich gebe im Folgenden für diesen Fall einen Weg an, der über die Gravitationspotentiale $g_{\mu\nu}$ im Unendlichen keinerlei Voraussetzungen macht und ausserdem für meine späteren Untersuchungen Vorteile bietet." (Hilbert 1917, 67) For the derivation, see pp. 67–70.

203 See (Birkhoff 1923, 253–256).

> According to our conception of the nature of matter, we can only consider those $g_{\mu\nu}$ to be physically viable solutions to the differential equations $K_{\mu\nu} = 0$ [the Einstein equations] that are regular and singularity free.
> We call a gravitational field or a metric "regular"—this definition had to be added—when it is possible to introduce a coordinate system, such that the functions $g_{\mu\nu}$ are regular and have a non-zero determinant at every point in the world. Furthermore, we describe a single function as being regular if it and all its derivatives are finite and continuous. This is incidentally always the definition of regularity in physics, whereas in mathematics a regular function is required to be analytic.[204]

It is curious that Hilbert identifies physical regularity with *infinite* differentiability and continuity of all derivatives. Either of these requirements is much too strong: each precludes gravitational radiation carrying new information, for example gravitational shock waves.[205] But at least Hilbert attempted to define a singularity of the gravitational field. In his understanding, the Schwarzschild solution has singularities at $r = 0$ and at the Schwarzschild radius. But we now know the first singularity is real, while the second can be removed by a coordinate transformation. He remarks:

> When we consider that these singularities are due to the presence of a mass, then it also seems plausible that they cannot be eliminated by coordinate transformations. However, we will give a rigorous proof of this later by examining the behavior of geodesic lines in the vicinity of this point.[206]

Hilbert then returns to his original motif: the Schwarzschild solution as a tool for discussing foundational problems of geometry:

> In order to obtain singularity-free solutions, we must assume that *a* [i.e., the mass parameter] = 0. [This leads to the Minkowski metric.] ... This proves the ... theorem: In the absence of matter, under the stated assumptions 1–3 [see above], the pseudo-Euclidean geometry of the little relativity principle [i.e., special relativity] actually holds in physics; and for *t* = const Euclidean geometry is in fact realized in the world.[207]

---

204 "Nach unserer Auffassung vom Wesen der Materie können wir als physikalisch realisierbare Lösungen $g_{\mu\nu}$ der Differentialgleichungen $K_{\mu\nu} = 0$ [the Einstein equations] nur diejenigen ansehen, welche regulär und singularitätenfrei sind.
"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 verschiedenen Determinante haben. 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." (WS 1916/17 Lectures, 118)

205 See, e.g., (Papapetrou 1974, 169–177).

206 "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. Einen strengen Beweis dafür werden wir aber erst weiter unten geben, indem wir den Verlauf der geodätischen Linien in der Umgebung dieser Punkt untersuchen." (WS 1916/17 Lectures, 118–119)

207 "Wir müssen also, um singularitätenfreie Lösungen zu erhalten, *a* [i.e., the mass parameter] = 0 annehmen. Wir haben damit den ... Satz bewiesen: Bei Abwesenheit von Materie ($q_i = 0$) existiert unter den ... genannten Voraussetzungen 1–3 [see above] die pseudoeuklidischen Geometrie des kleinen Relativitätsprinzips in der Physik tatsächlich, und für *t* = const ist in der Welt die Euklidische Geometrie wirklich realisiert." (WS 1916/17 Lectures, 119)

In the sequel, Hilbert explores its physical significance for describing the behavior of matter in space and time. His conception of matter, based on Mie's theory, plays no significant role in this discussion, its role being taken instead by assumptions that Hilbert assimilated from Einstein's work, such as the geodesic postulate for the motion of free particles.

He then turns to the justification for considering the case $a \neq 0$:

> Then we are acting against our own prescription that we shall regard only singularity-free gravitational fields as realizable in nature. Hence we must justify the assumption $a \neq 0$. [208]

He emphasizes the extraordinary difficulty of integrating the 14 field equations, even for "the simple special case when they go over to $K_{\mu\nu} = 0$":

> Mathematical difficulties already hinder us, for example, from constructing a single neutral mass point. If we were able to construct such a neutral mass, and if its behavior in the neighborhood of this point were known, then, if we let the neutral mass degenerate increasingly to a mass point, the $g_{\mu\nu}$ at this point would display a singularity. Such a singularity we would have to regard as being allowed in the sense that the $g_{\mu\nu}$ outside the immediate neighborhood of the singularity correctly describes the course actually realized in nature. In [the Schwarzschild line element] we must now have this kind of singularity at hand. Incidentally, we can now state that the construction of a neutral mass point, even if this is possible later, will prove to be so complicated that for purposes, in which one does not look at the immediate neighborhood of the mass point, one will be able to calculate the approximately correct gravitational potentials containing a singularity with sufficient precision.
>
> We now maintain the following: If we could actually carry out the mathematical expansion leading to construction of a neutral massive particle, we would probably find laws that, for the time being, still must be formulated axiomatically; but which later will emerge as consequences of our general theory, consequences that admittedly only can be proven categorically by means of a broad-ranging theory and complex calculations. These axioms, which thus have only provisional significance, we formulate as follows:
>
> Axiom I.: The motion of a mass point in the gravitational field is represented by a geodesic line that is a time-like.
>
> Axiom II: The motion of light in the gravitational field is represented by a null geodesic curve.
>
> Axiom III.: A singular point of the metric is equivalent to a gravitational center.[209]

Hilbert calls the first two axioms, taken from Einstein's work, a "rational generalization" of the behavior of massive particles and light rays in the "old physics," in which the metric tensor takes the limiting Minkowski values. He states that the Newtonian law of gravitational attraction and the resulting Keplerian laws of planetary

---

208 "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 $a \neq 0$ rechtfertigen." (WS 1916/17 Lectures, 120)

motion follow from these axioms "in the first approximation." In this way, Hilbert integrated into his theory the essential physical elements, on which Einstein's path to general relativity was based. Even his epistemological justification for the superiority of the new theory now makes use of an argument for the integration of knowledge. Remarkably, from Hilbert's perspective, this integration not only involves knowledge of classical physics such as Newton's law of gravitation, but also of Euclidean geometry as a physical interpretation of space:

> In principle, however, this new Einsteinian law has no similarity to the Newtonian. It is infinitely more complicated than the latter. If we nevertheless prefer it to the Newtonian, this is because this law satisfies a profound philosophical principle—that of general invariance—and that it contains as special cases two such heterogeneous things as on the one hand, Newton's law and on the other, the actual validity of Euclidian geometry in physics under certain simple conditions; so that we do not have to, as was the case until now, first assume the validity of Euclidian geometry and then put together a law of attraction.[210]

Thus we see that Hilbert considers his results on the conditions of validity of Euclidean geometry on a par in importance with, and logically prior to, Einstein's and Schwarzschild's results on the Newtonian limit of general relativity.

In accord with the physical interpretation they are given in Axioms I and II, Hilbert then goes on to study the time-like and null geodesics of the Schwarzschild metric, leading to discussions of two general-relativistic effects that Einstein had already considered: the planetary perihelion precession and the deflection of light due to the Sun's gravitational field. This discussion occupies almost all of the rest of this chapter of his lecture notes (WS 1916/17 Lectures, 122–156). After a short discussion of the

209 "Die mathematischen Schwierigkeiten hindern uns z.B. schon an der Konstruktion eines einzigen neutralen Massenpunktes. Könnten wir eine solche neutrale Masse konstruieren, und würden wir den Verlauf 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 [the Schwarzschild line element] 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.

Wir behaupten nun Folgendes: Wenn wir die mathematische Entwicklung, die zur Konstruktion eines neutralen Massenteilchens führt, wirklich durchführen können, so werden wir dabei vermutlich auf Gesetze stossen, die wir einstweilen noch axiomatisch formulieren müssen, die aber später sich als Folgen unserer allgemeinen Theorie ergeben werden, als Folgen freilich, die bestimmt nur durch eine weitsichtige 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 eine 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." (WS 1916/17 Lectures, 120–121) dimensions of various physical quantities (WS 1916/17 Lectures, 156–158), he discusses the behavior of measuring threads and clocks in the Schwarzschild gravitational field (WS 1916/17 Lectures, 159–163), and concludes the chapter with a discussion of the third general-relativistic effect treated by Einstein, the gravitational redshift of spectral lines (WS 1916/17 Lectures, 163–166).

In Paper 2, these topics are treated more briefly if at all: Axioms I and II and their motivations, are discussed on pp. 70–71. The discussion of time-like geodesics occupies pp. 71–75, and the paper closes with a discussion of null geodesics on pp. 75–76. In summary, this paper must be considered a singular hybrid between the blossoming of a rich mathematical tradition that Hilbert brings to bear on the problems of general relativity, and the agony of facing the collapse of his own research program.

6.4 Revisions of Paper 2

Paper 2, like Paper 1, was republished twice: Indeed, the two were combined in the 1924 version, Paper 2 becoming Part 2 of *Die Grundlagen der Physik* (Hilbert 1924, 11–32). We shall refer to this version as "Part 2." The reprint of Hilbert 1924 in the *Gesammelte Abhandlungen* was edited by others, presumably under Hilbert's supervision (Hilbert 1935, 268–289). We shall refer to this version as "Part 2–GA." Compared to Paper 1, Hilbert's additions and corrections to Paper 2 are less substantial, as is to be expected since Paper 2 was written largely within the context of general relativity. Most changes are minor improvements, e.g. in connection with recent literature on the theory. There are three significant changes however. One, introduced by Hilbert at the beginning of Part 2, concerns Hilbert's view of the relation between Papers 1 and 2, the other two by the editors of the *Gesammelte Abhandlungen* in Part 2–GA. The second concerns the Cauchy problem, and the third concerns his understanding of invariant assertions. We shall discuss these revisions, both major and minor.

The first significant change concerns the paper's goal: Paper 2 states that "it seems necessary to discuss some more general questions of a logical as well as physical nature" ("erscheint es nötig, einige allgemeinere Fragen sowohl logischer wie physikalischer Natur zu erörtern" Hilbert 1917, 53). Part 2 states: "now the relation of the theory with experience shall be discussed more closely" ("Es soll nun der Zusammenhang der Theorie mit der Erfahrung näher erörtert werden" Hilbert 1924, 11). This revision confirms our interpretation of Paper 2 as resulting, in its original

210 "Prinzipiell aber hat dieses neue Einsteinsche Gesetz gar keine Ähnlichkeit mit dem Newtonschen. Es ist unmöglich 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, sodass wir also nicht, wie dies bis jetzt der Fall war, zuerst die Gültigkeit der Euklidischen Geometrie voraussetzen, und dann ein Attraktionsgesetz anflicken müssen." (WS 1916/17 Lectures, 122)

version, from the tension between Hilbert's concern about the unsolved problems of his theory, in particular the problem of causality, and his immersion in the challenging applications of general relativity, in particular to astronomy. Since Hilbert's revision of Paper 1 had effectively transformed his theory into a version of general relativity, the revision of Paper 2 could now be presented as relating this theory to its empirical basis, the astronomical problems being addressed by contemporary general relativity.

We shall now discuss the changes, which occur in four of the six topics discussed (see above):

1. The metric tensor and its measurement: Part 2 drops all reference to "Messfaden." The discussion of measurement is based entirely on the "Lichtuhr," but is otherwise parallel to that in Paper 2 (Hilbert 1924, 11–13).
2. The causality problem for the field equations (Hilbert 1924, 16–19): There are several changes in the discussion. The wording, with which Hilbert introduces the problem now reads:

   > Our basic equations of physics [the gravitational and the electromagnetic field equations] in no way take the form characterized above [Cauchy normal form]: rather four of them are, as I have shown, a consequence of the rest ...[211]

   Note that "wie ich gezeigt habe" replaces "nach Theorem I" (see p. 59 of Paper 2). Hilbert says that, if there were 4 additional invariant equations, then the system of equations in Gaussian normal coordinates "ein überbestimmtes System bilden würde" (see p. 16 of Part 2) replacing "untereinander in Widerspruch ständen" (see p. 60 of Paper 2).

   In the discussion of the first way, in which "physically meaningful, i.e., invariant assertions can be expressed mathematically" (Hilbert 1917, 62; 1924, 18), he corrects a number of the equations in his example. His discussion of the third way is shortened considerably, now reading:

   > An assertion is also invariant and is therefore always physically meaningful if it is valid for any arbitrary coordinate system, without the need for the expressions occurring in it to possess a formally invariant character.[212]

   In Paper 2, this sentence had ended with "...gültig sein soll," and the paragraph had given the example of Einstein's gravitational energy-momentum complex.
3. Euclidean geometry: His discussion is the same, except that the discussion of gravitational perturbations drops the use of an imaginary time coordinate and Euclidean metric (Hilbert 1924, 19–23, 26).

---

211 "Unsere Grundgleichungen der Physik sind nun keineswegs von der oben charakterisierten Art; vielmehr sind, wie ich gezeigt habe, vier von ihnen eine Folge der übrigen ..." (Hilbert 1924, 16)

212 "Auch ist eine Aussage invariant und hat daher stets physikalischen Sinn, wenn sie für jedes beliebige Koordinatensystem gültig ist, ohne daß dabei die auftretenden Ausdrücke formal invarianten Charakter zu besitzen brauchen." (Hilbert 1924, 19)

4. The Schwarzschild solution (Hilbert 1924, 23–32): He adds a footnote to the light ray axiom:

> Laue has shown for the special case $L = \alpha Q$ [i.e., for the usual Maxwell Lagrangian] how this theorem can be derived from the electrodynamic equations by considering the limiting case of zero wavelength.[213]

followed by a reference to Laue's 1920 paper (Laue 1920) showing that Hilbert kept up with the relativity literature. He also dropped a rather trivial footnote to Axiom I (massive particles follow time-like world lines):

> This last restrictive addition [i.e., "Zeitlinie"] is to be found neither in Einstein nor in Schwarzschild.[214]

He adds a more careful discussion of circular geodesics, the radius of which equals the Schwarzschild radius (Hilbert 1924, 30, compared to 1917, 75), but otherwise the discussion of geodesics remains the same.

When the 1924 version of his two papers was republished in 1935 in his *Gesammelte Abhandlungen*, the editors introduced two extremely significant changes, as well as more trivial ones that we shall not discuss, that retract the last elements of Hilbert's attempt to provide a solution to the causality problem for his theory. These changes in Part 2–GA are footnotes marked "Anm[erkung] d[er] H[erausgeber]". The first occurs in the discussion of the causality principle for generally-covariant field equations (Hilbert 1924, 18–19; 1935, 275–277). The sentence:

> Since the Gaussian coordinate system itself is uniquely determined, therefore also all assertions with respect to these coordinates about those potentials (24) [equation number in the original] are of invariant character.[215]

is dropped; and a lengthy footnote is added (Hilbert 1935, 275–277). This footnote shows that the editors[216] correctly understood the nature of the fourteen field equations. Six of the ten gravitational and three of the four electromagnetic equations contain second time derivatives of the six spatial components of the metric tensor and three spatial components of the electromagnetic potentials. Thus, their values together with those of their first time derivatives on the initial hypersurface determine their evolution off that hypersurface. But these initial values are subject to constraints, set by the remaining four gravitational and one electromagnetic equation, which contain no second time derivative. Due to the differential identities satisfied

---

213 "Laue hat für den Spezialfall $L = \alpha Q$ [i.e., for the usual Maxwell Lagrangian] gezeigt, wie man diesen Satz aus den elektrodynamischen Gleichungen durch Grenzübergang zur Wellenlänge Null ableiten kann." (Hilbert 1924, 27).

214 "Dieser letzte einschränkende Zusatz findet sich weder bei Einstein noch bei Schwarzschild." (Hilbert 1917, 71)

215 "Da das Gaußische Koordinatensystem selbst eindeutig festgelegt ist, so sind auch alle auf dieses Koordinatensystem bezogenen Aussagen über jene Potentiale (24) von invariantem Charakter." (Hilbert 1924, 18)

216 Paul Bernays, Otto Blumenthal, Ernst Hellinger, Adolf Kratzer, Arnold Schmidt, and Helmut Ulm.

by the field equations, if these constraint equations hold initially, they will continue to hold by virtue of the remaining field equations. This footnote culminates in the statement:

> Thus causal lawfulness does not express the full content of the basic equations; rather, in addition to this lawfulness, these equations also yield *restrictive conditions on the respective initial state*.[217]

The editors also explain that, in the gauge-invariant electromagnetic case, it is only the fields and not the potentials that are determined by the field equations. The editors' addition thus presents a lucid account of the Cauchy problem in general relativity, and shows that Hilbert's attempt to formulate a principle of causality for his theory in terms of the classical notion of initial data (i.e. values that can be freely chosen at any given moment in time, which then determine their future evolution) had not taken into account the existence of constraints on the initial data.

The second footnote occurs in the discussion of how to satisfy the requirement that physically meaningful assertions be invariant by use of an invariant coordinate system (Hilbert 1924, 18–19). The footnote, which actually undermines claims in Hilbert's paper, reads:

> In the case of each of the three types of preferred coordinate systems named here, there is only a partial fixation of the coordinates. The Gaussian nature of a coordinate system is preserved by arbitrary transformations of the space coordinates and by Lorentz transformations, and a coordinate system in which the vector r^k has the components $(0, 0, 0, 1)$, is transformed into another such system by an arbitrary transformation of the spatial coordinates together with a spatially varying shift of the temporal origin.
>
> The characterization of a Gaussian coordinate system by conditions (23) [equation number in the original] and likewise that of the third-named preferred coordinate system through the conditions for r^k is in fact not completely invariant insofar as the specification of the fourth coordinate—introduced through conditions (21) [equation number in the original; the conditions for a "proper" coordinate system]—plays a role in it.[218]

217 "Somit bringt die kausale Gesetzlichkeit nicht den vollen Inhalt der Grundgleichungen zum Ausdruck, diese liefern vielmehr außer jener Gesetzlichkeit noch *einschränkende Bedingungen für den jeweiligen Anfangszustand*." (Hilbert 1935, 277)

218 "Bei den drei hier genannten Arten von ausgezeichneten Koordinatensystemen handelt es sich jedesmal nur um eine partielle Festlegung der Koordinaten. Die Eigenschaft des Gaußischen Koordinatensystems bleibt erhalten bei beliebigen Transformationen der Raumkoordinaten und bei Lorentztransformationen, und ein Koordinatensystem, in welchem der Vektor r^k die Komponenten $(0, 0, 0, 1)$ hat, geht wieder in ein solches über bei einer beliebigen Transformation der Raumkoordinaten nebst einer örtlich variablen Verlegung des zeitlichen Nullpunktes.
Die Charakterisierung des Gaußischen Koordinatensystems durch die Bedingungen (23) und ebenso die des drittgenannten ausgezeichneten Koordinatensystems durch die Bedingungen für r^k ist übrigens insofern nicht völlig invariant, als darin die Auszeichnung der vierten Koordinate zur Geltung kommt, die mit der Aufstellung der Bedingungen (21) eingeführt wurde." (Hilbert 1935, 277)

The editors of Hilbert's papers corrected two major mathematical errors that survived his own revision of Paper 2, and since he was still active when this edition of his papers was published, it can be assumed that these changes were made with his consent, if not participation.

7. THE FADING AWAY OF HILBERT'S POINT OF VIEW AND ITS SUBSUMPTION BY EINSTEIN'S PROGRAM

Early on, Einstein and Weyl set the tone for the way in which Hilbert's papers on the *Foundations of Physics* were integrated into the mainstream of research in physics and mathematics. Not only did the articles by Einstein and Weyl receive immediate attention when first published in the *Sitzungsberichte* of the Prussian Academy of Sciences, but they were soon incorporated into successive editions of *Das Relativitätsprinzip*, then the standard collection of original works on the development of relativity.[219] Three out of four of Einstein's works added to the third edition mention Hilbert, as does Weyl's contribution to the fourth edition—although, as we shall see, the latter's omissions are as significant as his attributions. Translated into French, English and other languages, and in print to this day, countless scholars had their impression of the scope and history of relativity shaped by this book.

First we shall discuss Einstein's two mentions of Hilbert in 1916. (His third in 1919 is related to Weyl's 1918 paper, so we shall discuss it afterwards.) In contrast with Hilbert's need to reorganize his theory in reaction to Einstein's work, Einstein could assimilate Hilbert's results into the framework of general relativity without being bothered by the latter's differing interpretation of them. This assimilation, in turn, assigned Hilbert a place in the history of general relativity.

Einstein's 1916 review paper on general relativity mentions Hilbert in a discussion of the relation between the conservation identities for the gravitational field equations and the field equations for matter:

> Thus the field equations of gravitation contain four conditions [the conservation equations for the energy-momentum tensor of matter] which govern the course of material phenomena. They give the equations of material processes completely of the latter are capable of being characterized by four independent differential equations.[220]

---

219 See (Blumenthal 1913; 1919; 1923; 1974). All editions were edited by the mathematician Otto Blumenthal. The first edition appeared as the second volume of his series *Fortschritte der Mathematischen Wissenschaften in Monographien* (the first being a collection of Minkowski's papers on electrodynamics), "als eine Sammlung von Urkunden zur Geschichte des Relativitätsprinzips" ("Vorwort" [n.p.]). The third edition in 1919 included additional papers by Einstein on general relativity, the fourth edition added Weyl's first paper on his unified theory of gravitation and electromagnetism. The fifth edition in 1923 is the basis of the editions currently in print, and of the translations into other languages. It would be interesting to know how Blumenthal chose the papers to include in what became the canonical source book on relativity.

A footnote adds a reference to Paper 1.[221] Thus, Einstein subsumed into the general theory of relativity, as a particular case of an important general result, what Hilbert regarded as an outstanding achievement of his theory. Hilbert's interpretation of this result as embodying a unique coupling between gravitation and electromagnetism, is not even mentioned.

In the same year, Einstein published his own derivation of the generally-covariant gravitational field equations from a variational principle. While in the 1916 review paper he had given a non-invariant "Hamiltonian" (= Lagrangian) for the field equations modulo the coordinate condition $\sqrt{-g} = 1$, he now proceeded in a manner reminiscent of Hilbert's in Paper 1. He uses the same gravitational variables (the $g_{\mu\nu}$ and their first and second derivatives), but Einstein's $q_{(\rho)}$ "describe matter (including the electromagnetic field" ("beschreiben die Materie (inklusive elektromagnetisches Feld)") and hence are arbitrary in number and have unspecified tensorial transformation properties. By his straightforward generalization, Einstein transformed Hilbert's variational derivation into a contribution to general relativity, without adopting the latter's perspective on this derivation as providing a synthesis between gravitation and a specific theory of matter. Rather, Einstein's generalization made it possible to regard Hilbert's theory as no more than a special case.

Einstein prefaced his calculations with some observations placing his work in context:

> H. A. Lorentz and D. Hilbert have recently succeeded [footnoted references to Lorentz's four papers of 1915–1916 and Hilbert's Paper 1] in presenting the theory of general relativity in a particularly comprehensive form by deriving its equations from a single variational principle. The same shall be done in this paper. My aim here is to present the fundamental connections as transparently and comprehensively as the principle of general relativity allows. In contrast to Hilbert's presentation, I shall make as few assumptions about the constitution of matter as possible.[222]

---

220 "The Foundation of the General Theory of Relativity" p. 810, in (CPAE 6E, Doc. 30, 187). "Die Feldgleichungen der Gravitation enthalten also gleichzeitig vier Bedingungen [the conservation equations for the energy-momentum tensor of matter], welchen der materielle Vorgang zu genügen hat. Sie liefern die Gleichungen des materiellen Vorganges vollständig, wenn letzterer durch vier voneinander unabhängige Differentialgleichungen charakterisierbar ist." (Einstein 1916a, 810)

221 The reference to "p. 3," is probably to a separately paginated off-print; see the discussion in (Sauer 1999).

222 Einstein, "Hamilton's Principle and the General Theory of Relativity" Sitzungsberichte 1916, 1111–1116, citation from p. 1111, in (CPAE 6E, Doc. 41, 240). "In letzter Zeit ist es H. A. Lorentz und D. Hilbert gelungen [footnoted references to Lorentz's four papers of 1915–1916 and Hilbert's Paper 1], der allgemeinen Relativitätstheorie dadurch eine besonders übersichtliche Gestalt zu geben, daß sie deren Gleichungen aus einem einzigen Variationsprinzipe ableiteten. Dies soll auch in der nachfolgenden Abhandlung geschehen. Dabei ist es mein Ziel, die fundamentalen Zusammenhänge möglichst durchsichtig und so allgemein darzustellen, als es der Gesichtspunkt der allgemeinen Relativität zuläßt. Insbesondere sollen über die Konstitution der Materie möglichst wenig spezialisierende Annahmen gemacht werden, im Gegensatz besonders zur Hilbertschen Darstellung." (Einstein 1916b, 1111)

Thus Einstein both gave Hilbert credit for his accomplishments and circumscribed their nature: Like Lorentz, Hilbert was supposedly looking for a variational derivation of the general-relativistic field equations, but included assumptions about the constitution of matter that were too special. In an earlier, unpublished draft, Einstein's tone was even sharper:

> Hilbert, following the assumption introduced by Mie that the H function depends on the components of a four-vector and their first derivatives, I do not consider very promising.[223]

In private correspondence, he was still more harsh, but also gave his reasons for disregarding Hilbert's point of view:

> Hilbert's assumption about matter appears childish to me, in the sense of a child who knows none of the perfidy of the world outside. [...] At all events, mixing the solid considerations originating from the relativity postulate with such bold, unfounded hypotheses about the structure of the electron or matter cannot be sanctioned. I gladly admit that the search for a suitable hypothesis, or Hamilton function, for the construction of the electron, is one of the most important tasks of theory today. The "axiomatic method" can be of little use here, though.[224]

Evidently, Einstein clearly perceived the diverse status of the physical assumptions underlying general relativity, on the one hand, and Hilbert's theory, on the other. From Einstein's point of view, Hilbert's detailed results, such as his variational derivation of the Schwarschild metric could be—and were—acknowledged as contributions to the development of general relativity, without any need to refer to the grandiose program, within which Hilbert had originally placed them.

In view of his own claims in this regard, one might expect Hilbert's work to have played a prominent role in the developing search for a unified field theory.[225] But his fate was that of a transitional figure, eclipsed by both his predecessors and his successors. His achievements were perceived as individual contributions to general relativity rather than as genuine milestones on the way towards a unified field theory. Evidently, this "mixed score" was the price Hilbert had to pay for being made one of the founding fathers of general relativity.

In his first contribution to unified field theory, Weyl assigned a definite place to Hilbert, if largely by omission. After presenting his generalization of Riemannian

---

223 "Die von Hilbert im Anschluss an Mie eingeführte Voraussetzung, dass sich die Funktion H durch die Komponenten eines Vierervektors q_ρ und dessen erste Ableitungen darstellen lasse, halte ich für wenig aussichtsvoll." See note 3 to Doc. 31 in (CPAE 6, 346).

224 "Der Hilbertsche Ansatz für die Materie erscheint mir kindlich, im Sinne des Kindes, das keine Tükken der Aussenwelt kennt. [...] Jedenfalls ist es nicht zu billigen, wenn die soliden Überlegungen, die aus dem Relativitätspostulat stammen, mit so gewagten, unbegründeten Hypothesen über den Bau des Elektrons bezw. der Materie verquickt werden. Gerne gestehe ich, dass das Aufsuchen der *geeigneten* Hypothese bezw. Hamilton'schen Funktion für die Konstruktion des Elektrons eine der wichtigsten heutigen Aufgaben der Theorie bildet. Aber die "axiomatische Methode" kann dabei wenig nützen." Einstein to Hermann Weyl, 23 November 1916, (CPAE 8, 365–366).

225 For a historical discussion, see (Majer and Sauer 2005; Goenner 2004).

geometry to include what he called "gauge invariance" (Eichinvarianz),[226] Weyl turned to unified field theory:

> Making the transition from geometry to physics, we must assume, in accord with the example of Mie's theory [references to Mie's papers of 1912/13 and Weyl's recently-published *Raum-Zeit-Materie*], that the entire lawfulness of nature is based upon a certain integral invariant, the action
>
> $$\int W d\omega = \int \mathfrak{W} dx \qquad (\mathfrak{W} = W\sqrt{g}),$$
>
> in such a way that the actual world is distinguished from all possible four-dimensional metric spaces, by the fact that the action contained in every region of the world takes an extremal value with respect to those variations of the potentials g_{ij}, ϕ_i that vanish at the boundaries of the region in question.[227]

In spite of its obvious relevance, there is no mention here of Hilbert. The sole mention comes in what we shall refer to as "the litany" since this or a similar list occurs so frequently in the subsequent literature:

> We shall show in fact, in the same way that, according to the investigations of Hilbert, Lorentz, Einstein, Klein and the author [reference follows to Paper 1 for Hilbert], the four conservation laws of matter (of the energy-momentum-tensor) are connected with the invariance of the action under coordinate transformations containing four arbitrary functions; the charge conservation law is linked to a newly introduced "scale-invariance" depending on a fifth arbitrary function.[228]

This passage, (incorrectly) attributing to Hilbert a clarification of energy-momentum conservation in general relativity and disregarding his attempt to create a unified field theory, makes his "mixed score" particularly evident. In a footnote added to the republication of his paper in *Das Relativitätsprinzip,* Weyl notes that:

> The problem of defining all invariants W admissible as actions, while requiring that they contain the derivatives of g_{ij} up to second order at most, and those of ϕ_i only up to first order, was solved by R. Weitzenböck [Weitzenböck 1920], [229]

---

226 This generalization was named a Weyl space by J.A. Schouten (see Schouten 1924).

227 "Von der Geometrie zur Physik übergehend, haben wir nach dem Vorbild der Mieschen Theorie anzunehmen, daß die gesamte Gesetzmäßigkeit der Natur auf einer bestimmten Integralinvariante, der Wirkungsgröße $\int W d\omega = \int \mathfrak{W} dx \qquad (\mathfrak{W} = W\sqrt{g})$ beruht, derart, daß die wirkliche Welt unter allen möglichen vierdimensionalen metrischen Räumen dadurch ausgezeichnet ist, daß für sie die in jedem Weltgebiet enthaltene Wirkungsgröße einen extremalen Wert annimmt gegenüber solchen Variationen der Potentiale g_{ik}, ϕ_i, welche an den Grenzen des betreffenden Weltgebiets verschwinden." (Weyl 1918c, 475)

228 "Wir werden nämlich zeigen: in der gleichen Weise, wie nach Untersuchungen von Hilbert, Lorentz, Einstein, Klein und dem Verf. [reference follows to Paper 1 for Hilbert] die vier Erhaltungsätze der Materie (des Energie-Impuls-Tensors) mit der, vier willkürliche Funktionen enthaltenden Invarianz der Wirkungsgröße gegen Koordinatentransformationen zusammenhängen, ist mit der hier neu hinzutretenden, eine fünfte willkürliche Funktion hereinbringenden "Maßstab-Invarianz" [...] das Gesetz von der Erhaltung der Elektrizität verbunden." (Weyl 1918c, 475)

without mentioning that this is the solution to the problem raised by Hilbert's ansatz for the invariant Lagrangian, first introduced in Paper 1. Little wonder that those whose knowledge of the history of relativity came from *Das Relativitätsprinzip* had no idea of Hilbert's original aims and little more of his achievements.

Hilbert fared a little better in Weyl's *Raum-Zeit-Materie,* the first treatise on general relativity (Weyl 1918a; 1918b; 1919; 1921; 1923).[230] The discussion of the energy-momentum tensor in the first edition (section 27) credits Hilbert with having shown that (Weyl 1918a; 1918b, 184):

> [...] Mie's electrodynamics can be generalized from the assumptions of the special to those of the general theory of relativity. This was done by Hilbert.[231]

Footnote 5 cites Paper 1 and adds (Weyl 1918a; 1918b, 230):

> The connection between Hamilton's function and the energy-momentum tensor is established here, and the gravitational equations articulated almost simultaneously with Einstein, if only within the confines of Mie's theory,[232]

Hilbert's work has already been subsumed under general relativity. Curiously, both textual reference and footnote disappear from all later editions (but see the discussion below of the fifth edition). Presumably because Weyl had already mentioned Hilbert, the latter's name does not appear in the litany in the first edition (footnote 6), listing those who had worked on the derivation of the energy-momentum conservation laws. By the third edition, Hilbert has been added to the litany (Weyl 1919, 266 n. 8), and remained there. In his discussion of causality for generally-covariant field equations in the first edition, Weyl credits Papers I and II (Weyl 1918a; 1918b, 190 and 230, n. 9); again, this note disappears from all later editions. Paper 2 is also cited in the first edition in connection with the Schwarzschild solution (Weyl 1918a; 1918b, 230, n. 15), and the introduction of geodesic normal coordinates (Weyl 1918a; 1918b, 230, n. 21).

The third edition carries over these references to Paper 2 and adds one in connection with linearized gravitational waves (Weyl 1919, 266, n. 14); and the fourth edition includes all these footnotes. Perhaps questions had been raised concerning

---

229 "Die Aufgabe, alle als Wirkungsgrößen zulässigen invarianten *W* zu bestimmen, wenn gefordert ist, daß sie die Ableitungen der g_{ik} höchstens bis zur 2., die der ϕ_i nur bis zur 1. Ordnung enthalten dürfen, wurde von R. Weitzenböck [Weitzenböck 1920] gelöst." (Blumenthal 1974, 159; translation from Lorentz et al. 1923.) This seventh edition from 1974 is an unchanged reprint of the fifth edition of 1923, 159, n. 2. Weitzenböck has his own version of the litany: "Die obersten physikalischen Gesetze: Feldgesetze und Erhaltungsätze werden nach den klassischen Arbeiten von Mie, Hilbert, Einstein, Klein und Weyl aus einem Variationsprinzip [...] hergeleitet"(p. 683). It is not clear why Lorentz is omitted from the litany; perhaps he was too much of a physicist for Weitzenböck.

230 The second edition of 1918 was unchanged, the fourth of 1921 was translated into English and French; the fifth of 1923, being thereafter reprinted without change.

231 "[...] die Miesche Elektrodynamik von den Voraussetzungen der speziellen auf die der allgemeinen Relativitätstheorie übertragen werden [kann]. Dies ist von Hilbert durchgeführt worden."

232 "Hier ist auch der Zusammenhang zwischen Hamiltonscher Funktion and Energie-Impuls-Tensor aufgestellt und wurden, etwa gleichzeitig mit Einstein, wenn auch nur im Rahmen der Mieschen Theorie, die Gravitationsgleichungen ausgesprochen."

Weyl's treatment of Hilbert in the book; at any rate, the footnote to the litany citing Hilbert in the fifth edition again credits him with a contribution to general relativity, rather than to unified field theories:

> In the first communication, Hilbert established the invariant field equations simultaneously with and independently of Einstein, but within the framework of Mie's hypothetical theory of matter.[233]

In short, in none of the editions is Hilbert mentioned in connection with unified field theories.

Pauli's standard 1921 review article on relativity is another major source, still consulted mainly in the English translation of 1958 (with additional notes) by physicists and mathematicians for historical and technical information about relativity and unified field theories (Pauli 1921; 1958). Pauli adopted what we may call the Einstein-Weyl line on Hilbert, considering him a somewhat unfortunate founding father of general relativity. After describing Einstein's work on general relativity culminating in the November 1915 breakthrough, Pauli adds in a footnote (Pauli 1921):[234]

> At the same time as Einstein, and independently, Hilbert formulated the generally covariant field equations [reference to Paper 1]. His presentation, though, would not seem to be acceptable to physicists, for two reasons. First, the existence of a variational principle is introduced as an axiom. Secondly, of more importance, the field equations are not derived for an arbitrary system of matter, but are specifically based on Mie's theory of matter ... .

His discussion of invariant variational principles in section 23 cites the litany: "investigations by Lorentz, Hilbert, Einstein, Weyl and Klein[235] on the role of Hamilton's Principle in the general theory of relativity" (Pauli 1921).[236]

Later (section 56), he discusses the question of causality in "a generally relativistic [i.e, generally-covariant] theory," arguing from general covariance to the existence of 4 identities between the 10 field equations, and concluding (Pauli 1921):[237]

> The contradiction with the causality principle is only apparent, since the many possible solutions of the field equations are only formally different. Physically they are completely equivalent. The situation described here was first recognized by Hilbert.

This passage represents a striking example of erroneously crediting Hilbert with a contribution to general relativity while neglecting his actual achievements. To make matters worse, Pauli's footnote cites Paper 1, rather than Paper 2; after also crediting Mach with a version of this insight, he adds (Pauli 1921):[238]

---

233 "In der 1. Mitteilung stellte Hilbert gleichzeitig und unabhängig von Einstein die invarianten Feldgleichungen auf, aber im Rahmen der hypothetischen Mieschen Theorie der Materie." (Weyl 1923, 329, n. 10)

234 Section 50, cited from translation in (Pauli 1958, 145 n. 277).

235 See Felix Klein to Wolfgang Pauli, 8 May 1921 in (Pauli 1979, 31).

236 Cited from translation in (Pauli 1958, 68).

237 Cited from translation in (Pauli 1958, 160).

238 Cited from translation in (Pauli 1958, 160, n. 315).

> Furthermore it deserves mentioning that Einstein had, for a time, held the erroneous view that one could deduce from the non-uniqueness of the solution that the gravitational equations could not be generally-covariant [reference to (Einstein 1914b)].

Pauli does acknowledge various contributions to general relativity in Paper 2.[239] But his discussion of unified field theories (Part V), like Weyl's, jumps from Mie (section 64) to Weyl (section 65) without mention of Hilbert.

By examining a couple of early treatises on relativity by non-German authors, we can get some idea of the propagation of the Einstein-Weyl line as canonized by Pauli. Jean Becquerel's *Le Principe de la Relativité et la Théorie de la Gravitation* was the first French treatise on general relativity. In Chapter 16 on "Le Principe d'Action Stationnaire," Becquerel asserts:

> Lorentz and Hilbert [references to Papers 1 and 2], and then Einstein succeeded in presenting the general equations of the theory of gravitation as consequences of a unique stationary action principle, …[240]

followed by section 103 on "Méthode de Lorentz et d'Hilbert" (Becquerel 1922, 257–262). Paper 2 is cited in connection with linearized gravitational waves (Becquerel 1922, 216), but there is no mention of Hilbert in Chapter 18 on "Union du Champ de Gravitation et du Champ Électromagnétique. Géometries de Weyl et d'Eddington" (Becquerel 1922, 309–335).

Until recently Eddington's treatise, *The Mathematical Theory of Relativity,* was widely read, cited and studied by students; and was translated into French and German (Eddington 1923; 1924). The two English editions cite Papers 1 and 2 in the bibliography, with a reference to section 61 on "A Property of Invariants,"[241] which demonstrates the theorem:[242]

> The Hamiltonian [i.e, Lagrangian] derivative of any fundamental invariant is a tensor whose divergence vanishes.

Outside the Bibliography, few references are given in the English editions; but Eddington added material to the German translation, including several references to Hilbert (Eddington 1925). On p. 114, footnote 1 credits Hilbert (Paper 2) with realizing that the assumption of asymptotic flatness is not needed in the derivation of the Schwarzschild metric. On p. 116, he credits Paper 2 for an "elegante Methode" for deducing the Christoffel symbols from the geodesic equation; and on p. 183, he credits the same paper for the first strict proof that one can always satisfy the linearized

239 See (Pauli 1921), section 13 for Axiom II; section 22 for discussion of the restrictions on coordinate systems if three coordinates are to be space-like and one time-like; and section 60 for the proof that linearized harmonic coordinate conditions may always be imposed.

240 "Lorentz et Hilbert [references to Papers 1 and 2], puis Einstein, ont réussi à presenter les équations générales de la theorie de la gravitation comme des conséquences d'un unique principe d'action stationnaire," (Becquerel 1922, 256).

241 See (Eddington 1924, 264): "wherever possible the subject matter is indicated by references to the sections in this book chiefly concerned."

242 See (Eddington 1924, 140–141).

harmonic coordinate conditions by an infinitesimal coordinate transformation. And that is it.

We see that, by the mid-1920s, and with minor variations within the accepted limits, the Einstein-Weyl line on Hilbert's role was already becoming standard in the literature on relativity.

8. AT THE END OF A ROYAL ROAD

The preceding discussion has shown that Hilbert did not discover a royal road to the field equations of general relativity. In fact, he did not formulate these equations at all but, at the end of 1915, developed a theory of gravitation and electromagnetism that is incompatible with Einstein's general relativity. Nevertheless, this theory can hardly be considered an achievement parallel to that of Einstein's creation of general relativity, to be judged by criteria independent of it. Not only is the dependence of Hilbert's theory on and similarity to Einstein's earlier, non-covariant Entwurf theory of gravitation too striking; but its contemporary reception as a contribution to general relativity and regardless of the extent to which Hilbert accepted the transformation of his theory into such a contribution, this is evidence of the theory's evanescent and heteronomous character. It could thus appear as if our account, in the end, describes a race for the formulation of a relativistic theory of gravitation with a clear winner—Einstein—and a clear loser—Hilbert. In contrast to the legend of Hilbert's royal road, such an account would bring us essentially back to Pauli's sober assessment of Hilbert's work as coming close to the formulation of general relativity but being faulted by its dependence on a specific theory of matter. However, as we have shown, this interpretation ascribes to Hilbert results in general relativity that he neither intended nor achieved, and ignores contributions that lay outside the scope of general relativity but were nevertheless crucial for its development. In view of such conundrums, we therefore propose not to consider the Einstein-Hilbert race as the competition between two individuals and their theories but as an event within a larger, collective process of knowledge integration.

As formulated by Einstein in 1915, general relativity incorporates elements of classical mechanics, electrodynamics, the special theory of relativity, and planetary astronomy, as well as such mathematical traditions as non-Euclidean geometry and the absolute differential calculus. It integrates these elements into a single, coherent conceptual framework centered around new concepts of space, time, inertia and gravitation. Without this enormous body of knowledge as its underpinning, it would be hard to explain the theory's impressive stability and powerful role even in today's physics. This integration was the result of an extended and conflict-laden process, to which not only Einstein but many other scientists contributed. From the point of view of historical epistemology, it was a collective process in an even deeper sense.[243] It involved a substantial, shared knowledge base, structured by fundamental concepts, models, heuristic etc., which were transmitted by social institutions, utilizing material representations, such as textbooks, and appropriated by individual learning processes. While individual thinking is governed to a large degree by these shared resources, it also affects and amplifies them, occasionally even changing these epistemic structures. On the basis of such an epistemology, which takes into account the interplay between shared knowledge resources and individual thinking, the emergence and fading away of a theory such as Hilbert's can be understood as an aspect of the process of integration of knowledge that produced general relativity.

To answer the question of from where alternative solutions (or attempted solutions) to the same problem come, we shall look at some of the shared knowledge of the time available for formulating theories such as those of Einstein and Hilbert. To explain the fading-away of Hilbert's theory, we then discuss the interplay between individual thinking and the knowledge resources that led to the formulation of general relativity and the transformation of Hilbert's theory into a contribution to it. It will become clear that, in both cases, the same mechanism was at work. In the case of general relativity, it integrated the various components of shared knowledge and resulted in the creation of a stable epistemic structure, which represents that integrated knowledge. In the case of Hilbert's theory, the same process disaggregated the various components of shared knowledge that had been brought together in a temporary structure, and rearranged and integrated them into a more stable structure.

The available knowledge offered a limited number of approaches to the problem that occupied both Einstein and Hilbert in late 1915: the formulation of differential equations governing the inertio-gravitational potential represented by the metric tensor. Two fundamentally different models underlying contemporary field theories of electrodynamics embodied the principal alternatives. One, the "monistic model," conceived all physical phenomena, including matter, in terms of fields. The other "fields-with-matter-as-source model" (or "Lorentz model") was based on a dualism of fields and matter. The first model was the basis for attempts to formulate an "electromagnetic world picture," which remained fragmentary and never succeeded in accounting for most contemporary physical knowledge. The second model was the basis for Lorentz's formulation of electron theory, the epitome of classical electrodynamics, in which matter acts as source for electrodynamic fields that, in turn, affect the motion of material bodies. Rather than attempting to reduce classical mechanical concepts to electrodynamic field concepts, the task associated with the electrodynamic world picture, Lorentz's electron theory successfully integrated electromag-

---

243 See (Csikszentmihalyi 1988): "All of the definitions [of creativity] ... of which I am aware assume that the phenomenon exists... either inside the person or in the work produced... After studying creativity for almost a quarter of a century, I have come to the reluctant conclusion that this is not the case. We cannot study creativity by isolating individuals and their works from the social and historical milieu in which their actions are carried out. This is because what we call creative is never the result of individual actions alone; it is the product of three main shaping forces: a set of social institutions or *field*, that selects from the variations produced by individuals those that are worth preserving; a stable cultural *domain* that will preserve and transmit the selected new ideas or forms to the following generations; and finally the *individual*, who brings about some change in the domain, a change that the field will consider to be creative." This concept is further discussed in (Stachel 1994).

netic and classical mechanical phenomena. The first model became the core of Hilbert's approach in an attempt to create a unified field theory, while Einstein's search for gravitational field equations was guided by the second. To a large extent, the difference between the two models accounts for the differences between Hilbert's and Einstein's approaches, including their differing capacity to incorporate available physical knowledge into their theories. The information about matter compatible with Hilbert's theory was essentially only Mie's speculative theory: The source-term in Einstein's gravitational field equations could embody the vast amount of information contained in special-relativistic continuum theory, including energy-momentum conservation, as well as Maxwell's theory.

The information available for solving the problem of gravitation was not exhausted by the two different physical models of the interaction between fields and matter. Contemporary mathematics also provided a reservoir of useful tools. The series of attempts between 1912 and 1915 to formulate a theory of gravitation, including contributions by Abraham, Nordström, and Mie, as well as Einstein and Hilbert, illustrates the range of mathematical formalisms available, from partial differential equations for a scalar field to the absolute differential calculus applied to the metric tensor. As did the physical models, different mathematical formalisms showed varying capacities for integrating the available knowledge about matter and gravitation, such as that embodied in Newtonian gravitation theory or in the observational results on Mercury's perihelion shift. To explore its capacity to integrate knowledge, a formalism needs to be elaborated and its consequences interpreted, if possible, as representations of that knowledge. The degree of such successful elaboration and interpretation, the "exploration depth" of a given formalism, determines its acceptability as a possible solution to the physical problem at hand. In early 1913, believing that the Newtonian limit could not be recovered from generally-covariant field equations, Einstein proposed the non-covariant *Entwurf* theory, from which it could be. At the end of 1915, on the basis of an increased "exploration depth" of the formalism, he decided in favor of generally-covariant equations.

Which physical models and mathematical formalisms are favored in a given historical situation depends on many factors, among them their accessibility and specific epistemological preferences that make some of them appear more attractive to certain groups than others. It was natural for a mathematician of Hilbert's caliber to start from a generally-covariant variational principle based on the metric tensor, while Einstein, ignorant of the appropriate mathematical resources, initially tried to develop his own, "pedestrian" calculus for dealing with the metric tensor.[244] It is clear that the monistic field theory model must have appealed more to Hilbert, a mathematician in search for an axiomatic foundation for all of physics, than the conceptually more clumsy dualistic model. The latter, on the other hand, was a more natural starting point for physicists such as Abraham, Einstein, and Nordström, who were familiar

244 See his calculations in "Einstein's Zurich Notebook," e.g. on p. 08L (in vol. 1 of this series). See also the "Commentary" (in vol. 2).

with the extraordinary successes of this model in the domain of electromagnetism. Images of knowledge also determine decisions on the depth and direction of exploration of a given formalism. While the question of the Newtonian limit was crucial to the physicist Einstein, Hilbert did not deal at all with this problem.

Constructs formulated by individual scientists, such as Hilbert's proposal for an axiomatic foundation of physics, are largely contingent; but their building blocks (concepts, models, techniques) are taken from the reservoir of the socially available knowledge characteristic of a given historical situation. This reservoir of shared background knowledge accounts for more than just the intercommunicability of individual contributions such as those of Hilbert and Einstein. Given that such contributions are integrated into already-shared knowledge by various processes of intellectual communication and assimilation, an equilibration process must take place between the individual constructs and the shared knowledge-reservoir. It is the outcome of this process that decides on whether a research program is progressive or degenerating in the sense of Lakatos but also the fate of an individual contribution, its longevity (the case of general relativity), its mutation, or its rapid fading-away (the case of Hilbert's contribution).

Whatever is individually constructed will be brought into contact with other elements of the shared knowledge-base, and thus integrated into it in multiple ways that, of course, are shaped by the social structures of scientific communication. The fate of an individual construct depends on the establishment of such connections. If individual constructs are not embedded, for whatever reasons, within the structures of socially available knowledge, they effectively disappear; if they are so embedded, they will be transmitted as part of shared knowledge. Usually, individual contributions are not assimilated wholesale to shared knowledge but only in a piecemeal fashion. One finds Hilbert's name associated, for instance, with the variational derivation of the field-equations but not with the program of an axiomatic foundation of physics. The "packaging" of individual contributions as they are eventually transmitted and received by a scientific community is not governed by the individual perspectives of their authors but by the more stable cognitive structures of the shared knowledge. The reception of Hilbert's contribution is thus not different from that of most scientific contributions that become assimilated into the great banquet of shared knowledge. It rarely happens that its basic epistemic structures, such as the concepts of space and time in classical physics, are themselves challenged by the growth of knowledge. Usually, these fundamental structures simply overpower any impact of individual contributions by the sheer mass of integrated knowledge they reflect. Only when individual constructs come with their own power of integrating large chunks of shared knowledge do they have a chance of altering these structures. This, in turn, only happens when the individual contributions themselves result from a process of knowledge integration and its reflection in terms of new epistemic structures.

Einstein's theory of general relativity is the result of such an integration process. Over a period of several years, he had attempted not only to reconcile classical physical knowledge about gravitation with the special-relativistic requirement of the finite propagation speed of physical interactions; but also with insights into the inseparability of gravitation and inertia, and with the special-relativistic generalization of energy-momentum conservation. Each of these building blocks: Newtonian theory, metric structure of space and time, the equivalence principle, and energy-momentum conservation, was associated with a set of possible mathematical representations, more or less well defined by physical requirements. In the case of energy-momentum conservation, for instance, Einstein had quickly arrived at an appropriate mathematical formulation, which stayed fixed throughout his search for the gravitational field equations. The inseparability of gravitation and inertia as expressed by the equivalence principle, on the other hand, could be given various mathematical representations; for Einstein the most natural at the time seemed to be the role of the metric tensor as the potentials for the inertio-gravitational field. The available mathematical representations of Einstein's building blocks were not obviously compatible with each other. In order to develop a theory comprising as much as possible of the knowledge incorporated in these building blocks, Einstein followed a double strategy.[245] On the one hand, he started from those physical principles that embody the vast store of knowledge in classical and special-relativistic physics and explored the consequences of their mathematical representations in terms of the direction of his other building blocks (his "physical strategy"). On the other hand, he started from those building blocks that had not yet been integrated into a physical theory, such as his equivalence principle, chose a mathematical representation, and explored its consequences, in the hope of being able to find a physical interpretation that also would integrate his other building blocks (his "mathematical strategy"). Eventually, he succeeded in formulating a theory that complies with these heterogeneous requirements; but only at the price of having to modify, in a process of reflection on his own premises, some of the original building blocks themselves, with far-going consequences for the structuring of the physical knowledge embodied in these building blocks, e.g. about the meaning of coordinate systems in a physical theory. That such modifications eventually became more than just personal idiosyncrasies and have had a lasting effect on the epistemic structures of physical knowledge is due to the fact that they were stabilized by the knowledge they helped to integrate into general relativity.

Hilbert's theory was clearly not based on a comparable process of knowledge integration and hence shared the fate of most scientific contribution: dissolution and assimilation to the structures of shared knowledge. Even if, in 1915, he had derived the field equations of general relativity, his theory would not have had the same "exploration depth" as that of Einstein's 1915 version, and hence not covered a similarly large domain of knowledge. Hilbert's theory is rather comparable to one of Einstein's early intermediate versions, for instance to that involving the (linearized) Einstein tensor, briefly considered in the Zurich Notebook in the winter of 1912/13. Einstein quickly rejected this candidate because it appeared to him impossible to derive the Newtonian limit from it, while Hilbert intended to publish his version in late 1915, although he had not checked its compatibility with the Newtonian limit.

245 See "Pathways out of Classical Physics …" (in vol. 1 of this series).

This difference in reacting to a similar candidates for solving the problem of the gravitational field equations obviously does not reveal any difference in the epistemic status of Hilbert's theory compared to Einstein's intermediate version but only by a different attitude with regard to a given exploration depth, motivated by the different image of knowledge that Hilbert associated with his endeavor. Such motivations make little difference to the fate of a theory in the life of the scientific community. In fact, the subsequent elaborations, revisions, and transformations of Hilbert's result testify to an equilibration process similar to that also undergone by Einstein's intermediate versions, in which ever new elements of shared knowledge found their way into Hilbert's construct. In the end, as we have seen, his theory comprises the same major building blocks of physical knowledge as those, on which general relativity is based. The exchange with Einstein and others had effectively compensated for Hilbert's original neglect of the need to consider his results in the light of physical knowledge, and thus substituted, in a way, for the "physical strategy" of Einstein's heuristics, constituting a "collective process of reflection." The fact that the equilibration process leading to general relativity essentially went on in private exchanges between Einstein and a few collaborators, while the equilibration process transforming Hilbert's theory of everything into a constituent of general relativity went on in public, as a contest between Einstein and Hilbert, Berlin and Göttingen, physics and mathematics communities, plays an astonishingly small role in the history of knowledge.

9. ACKNOWLEDGEMENTS

The research for this paper was conducted as part of a collaborative research project on the history of general relativity at the Max Planck Institute for the History of Science. We would like to warmly thank our colleagues in this project, in particular Leo Corry, Tilman Sauer, and Matthias Schemmel, for the discussions, support, and criticism, all of it intensive, which have made this collaboration a memorable and fruitful experience for both of us. Without the careful editorial work by Stefan Hajduk, who not only checked references, consistency, and language but also coordinated the contributions of the authors when they were not in one place, the paper would not have reached its present form. The authors are furthermore indebted to Giuseppe Castagnetti, Peter Damerow, Hubert Goenner, Michel Janssen, Ulrich Majer, John Norton, and David Rowe for helpful comments and conversations on the subject of this paper. We are particularly grateful to the library staff of the Max Planck Institute for the History of Science and its head, Urs Schoepflin, for their creative and indefatigable support in all library-related activities of our work. We thank the Niedersächsische Staats- und Universitätsbibliothek Göttingen (*Handschriftenabteilung*) and the library of the Mathematisches Institut, Universität Göttingen, for making unpublished material available to our project. Finally, we thank Ze'ev Rosenkranz and the Einstein Archives, The Hebrew University of Jerusalem, for permission to quote from Einstein's letters.

REFERENCES

Belinfante, F. J. 1939. "Spin of Mesons." *Physica* 6:887–898.

Becquerel, Henri. 1922. *Le Principe de la Relativité et la Théorie de la Gravitation*. Paris: Gauthier-Villars.

Birkhoff, George D., and Rudolph E. Langer. 1923. *Relativity and Modern Physics*. Cambridge, Ma.: Harvard University Press.

Blumenthal, Otto, ed. 1913. *Das Relativitätsprinzip*. 1st ed. Leipzig, Berlin: Teubner.

———. 1919. *Das Relativitätsprinzip*. 3rd ed. Leipzig, Berlin: Teubner.

———. 1923. *Das Relativitätsprinzip*. 5th ed. Leipzig, Berlin: Teubner.

———. 1974. *Das Relativitätsprinzip*. 7th ed. Darmstadt: Wissenschaftliche Buchgesellschaft.

Born, Max. 1914. "Der Impuls-Energie-Satz in der Elektrodynamik von Gustav Mie." *Königliche Gesellschaft der Wissenschaften zu Göttingen. Nachrichten* (1914):23–36.

Born, Max, and Leopold Infeld. 1934. "Foundations of the New Field Theory." *Royal Society of London. Proceedings* A 144:425–451.

Caratheodory, Constantin. 1935. *Variationsrechnung und partielle Differentialgleichungen erster Ordnung*. Leipzig, Berlin: B. G. Teubner.

Corry, Leo. 1997. "David Hilbert and the Axiomatization of Physics (1894–1905)." *Archive for History of Exact Sciences* 51:83–198.

———. 1999a. "David Hilbert between Mechanical and Electromagnetic Reductionism (1910–1915)." *Archive for History of Exact Sciences* 53:489–527.

———. 1999b. "From Mie's Electromagnetic Theory of Matter to Hilbert's Unified Foundations of Physics." *Studies in History and Philosophy of Modern Physics* 30 B (2):159–183.

———. 1999c. "David Hilbert: Geometry and Physics (1900–1915)." In J. J. Gray (ed.), *The Symbolic Universe: Geometry and Physics (1890–1930)*, Oxford: Oxford University Press, 145–188.

———. 2004. *David Hilbert and the Axiomatization of Physics, 1898–1918: From "Grundlagen der Geometrie" to "Grundlagen der Physik"*. Dordrecht: Kluwer.

Corry, Leo, Jürgen Renn, and John Stachel (eds.). 1997. *Belated Decision in the Hilbert-Einstein Priority Dispute*. Vol. 278, Science.

CPAE 4: Martin J. Klein, A. J. Kox, Jürgen Renn, and Robert Schulmann (eds.). 1995. *The Collected Papers of Albert Einstein*. Vol. 4. *The Swiss Years: Writings, 1912–1914*. Princeton: Princeton University Press.

CPAE 6: A. J. Kox, Martin J. Klein, and Robert Schulmann (eds.). 1996. *The Collected Papers of Albert Einstein*. Vol. 6. *The Berlin Years: Writings, 1914–1917*. Princeton: Princeton University Press.

CPAE 6E: *The Collected Papers of Albert Einstein*. Vol. 6. *The Berlin Years: Writings, 1914–1917*. English edition translated by Alfred Engel, consultant Engelbert Schucking. Princeton: Princeton University Press, 1996.

CPAE 8: Robert Schulmann, A. J. Kox, Michel Janssen, and József Illy (eds.). 1998. *The Collected Papers of Albert Einstein*. Vol. 8. *The Berlin Years: Correspondence, 1914–1918*. Princeton: Princeton University Press.

CPAE 8E: *The Collected Papers of Albert Einstein*. Vol. 8. *The Berlin Years: Correspondence, 1914–1918*. English edition translated by Ann M. Hentschel, consultant Klaus Hentschel. Princeton: Princeton University Press, 1998.

Csikszentmihalyi, Mihaly. 1988. "Society, Culture and Person: a Systems View of Creativity." In R. J. Sternberg (ed.), *The Nature of Creativity*. Cambridge: Cambridge University Press.

Earman, John, and Clark Glymour. 1978. "Einstein and Hilbert: Two Months in the History of General Relativity." *Archive for History of Exact Sciences* 19:291–308.

Eddington, Arthur Stanley. 1923. *The Mathematical Theory of Relativity*. Cambridge: The University Press.

———. 1924. *The Mathematical Theory of Relativity*. 2nd ed. Cambridge: The University Press.

———. 1925. *Relativitätstheorie in Mathematischer Behandlung*. Translated by Alexander Ostrowski Harry Schmidt. Berlin: Springer.

Einstein, Albert. 1913. "Zum gegenwärtigen Stande des Gravitationsproblems." *Physikalische Zeitschrift* 14 (25):1249–1262. (English translation in volume 3 of this series.)

———. 1914a. "Prinzipielles zur verallgemeinerten Relativitätstheorie und Gravitationstheorie." *Physikalische Zeitschrift* 15:176–180.

———. 1914b. "Die formale Grundlage der allgemeinen Relativitätstheorie." *Königlich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (1914) (XLI):1030–1085.

———. 1915a. "Zur allgemeinen Relativitätstheorie." *Königlich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (1915) (XLIV):778–786.

———.. 1915b. "Zur allgemeinen Relativitätstheorie (Nachtrag)." *Königlich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (1915) (XLVI):799–801.

———. 1915c. "Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie." *Königlich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (1915) (XLVII):831–839.

———. 1915d. [Zusammenfassung der Mitteilung "Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie."] *Königlich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (1915) (XLVII):803.

———. 1915e. "Die Feldgleichungen der Gravitation." *Königlich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (1915) (XLVIII–XLIX):844–847.

———. 1916a. "Die Grundlage der allgemeinen Relativitätstheorie." *Annalen der Physik* 49 (7):769–822.

———. 1916b. "Hamiltonsches Prinzip und allgemeine Relativitätstheorie." *Königlich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (1916) (XLII):1111–1116.

———. 1916c. "Näherungsweise Integration der Feldgleichungen der Gravitation." *Sitzung der physikalisch-mathematischen Klasse* 668–96. (CPAE 6, Doc. 32, 348–57)

Einstein, Albert, and Marcel Grossmann. 1914. "Kovarianzeigenschaften der Feldgleichungen der auf die verallgemeinerte Relativitätstheorie gegründeten Gravitationstheorie." *Zeitschrift für Mathematik und Physik* 63 (1 / 2):215–225.

Fölsing, Albrecht. 1997. *Albert Einstein: a biography*. New York: Viking.

Frei, Günther, ed. 1985. *Der Briefwechsel David Hilbert-Felix Klein (1886–1918)*. Vol. 19, *Arbeiten aus der Niedersächsischen Staats- und Universitätsbibliothek Göttingen*. Göttingen: Vandenhoeck & Ruprecht.

Goenner, Hubert. 2004. "On the History of Unified Field Theories." *Living Reviews of Relativity* 7 <http:// /www.livingreviews.org>.

Goenner, Hubert, Jürgen Renn, Jim Ritter, and Tilman Sauer (eds.). 1999. *The Expanding Worlds of General Relativity*. (*Einstein Studies* vol. 7.) Boston: Birkhäuser.

Guth, E. 1970. "Contribution to the History of Einstein's Geometry as a Branch of Physics." In *Relativity,* edited by M. Carmeli et al. New York, London: Plenum Press, 161–207.

Havas, Peter. 1989. "The Early History of the 'Problem of Motion' in General Relativity." In *Einstein and the History of General Relativity*, edited by Don Howard and John Stachel. (*Einstein Studies* vol. 1.) Boston: Birkhäuser, 234–276.

Hilbert, David. 1905. "Logische Prinzipien des mathematischen Denkens." Ms. Vorlesung SS 1905, annotated by E. Hellinger, Bibliothek des Mathematischen Seminars, Göttingen.

———. 1912–13. "Molekulartheorie der Materie." Ms. Vorlesung WS 1912–13, annotated by M. Born, Nachlass Max Born #1817, Stadtbibliothek Berlin.

———. 1913. "Elektronentheorie." Ms. Vorlesung SS 1913, Bibliothek des Mathematischen Seminars, Göttingen.

———. 1916. "Die Grundlagen der Physik. (Erste Mitteilung)." *Königliche Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Nachrichten* (1915):395–407. (English translation in this volume.)

———. 1917. "Die Grundlagen der Physik (Zweite Mitteilung)." *Königliche Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Nachrichten* (1917):53–76. (English translation in this volume.)

———. 1924. "Die Grundlagen der Physik." *Mathematische Annalen* 92:1–32.

———, ed. 1935. *Gesammelte Abhandlungen, Band III: Analysis, Grundlagen der Mathematik, Physik, Verschiedenes, Lebensgeschichte*. [1932–35, 3 vols.]. Berlin: Springer.

———. 1971. "Über meine Tätigkeit in Göttingen." In *Hilbert-Gedenkenband*, ed. K. Reidemeister. Berlin, Heidelberg, New York: Springer, 79–82.

Howard, Don, and John D. Norton. 1993. "Out of the Labyrinth? Einstein, Hertz, and the Göttingen Answer to the Hole Argument." In *The Attraction of Gravitation: New Studies in the History of General Relativity,* edited by John Earman, Michel Janssen and John D. Norton. Boston/Basel/Berlin: Birkhäuser, 30–62.

Janssen, Michel and Matthew Mecklenburg. 2006. "Electromagnetic Models of the Electron and the Transition from Classical to Relativistic Mechanics." In *Interactions: Mathematics, Physics and Philosophy, 1860–1930*, edited by V. F. Hendricks et al. *Boston Studies in the Philosophy of Science*, Vol. 251. Dordrecht: Springer, 65–134.

Kerschensteiner, Georg, ed. 1887. *Paul Gordan's Vorlesungen über Invariantentheorie. Zweiter Band: Binäre Formen*. Leipzig: Teubner.

Klein, Felix. 1917. "Zu Hilberts erster Note über die Grundlagen der Physik." *Königliche Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Nachrichten* (1917):469–482.

———. 1918a. "Über die Differentialgesetze für die Erhaltung von Impuls und Energie in der Einsteinschen Gravitationstheorie." *Königliche Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Nachrichten* (1918):171–189.

———. 1918b. "Über die Integralform der Erhaltungsätze und die Theorie der räumlich-geschlossenen Welt." *Königliche Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Nachrichten* (1918):394–423.

———. 1921. "Zu Hilberts erster Note über die Grundlagen der Physik." In *Gesammelte Mathematische Abhandlungen,* edited by R. Fricke and A. Ostrowski. Berlin: Julius Springer, 553–567.

Komar, Arthur. 1958. "Construction of a Complete Set of Independent Observables in the General Theory of Relativity." *Physical Review* 111:1182–1187.

Kretschmann, Erich. 1917. "Über den physikalischen Sinn der Relativitätspostulate, A. Einsteins neue und seine ursprüngliche Relativitätstheorie." *Annalen der Physik* 53 (16):575–614.

Laue, Max. 1911a. "Zur Dynamik der Relativitätstheorie." *Annalen der Physik* 35: 524–542.

———. 1911b. *Das Relativitätsprinzip*. Braunschweig: Friedrich Vieweg und Sohn.

Laue, Max von. 1920. "Theoretisches über neuere optische Beobachtungen zur Relativitätstheorie." *Physikalische Zeitschrift* 21:659–662.

Lichnerowicz, André. 1946. "Sur le caractère euclidien d'espaces-temps extérieurs statiques partout réguliers." *Academie des Sciences* (Paris). *Comptes Rendus* 222:432–436.

Lorentz, Hendrik A., et al. 1923. *The Principle of Relativity.* London: Methuen & Co.

Majer, Ulrich and Tilman Sauer. 2005. "'Hilbert's World Equations' and His Vision of a Unified Science." In *The Universe of General Relativity*, edited by A. Kox and J. Eisenstaedt. (*Einstein Studies*, vol. 11.) Boston: Birkhäuser, 259–276.

Mehra, Jagdish. 1974. *Einstein, Hilbert, and the Theory of Gravitation. Historical Origins of General Relativity Theory.* Dordrecht, Boston: D. Reidel Publishing Company.

Mie, Gustav. 1912a. "Grundlagen einer Theorie der Materie. Erste Mitteilung." *Annalen der Physik* 37:511–534. (English translation of excerpts in this volume.)

———. 1912b. "Grundlagen einer Theorie der Materie. Zweite Mitteilung." *Annalen der Physik* 39:1–40.

———. 1913. "Grundlagen einer Theorie der Materie. Dritte Mitteilung." *Annalen der Physik* 40:1–66. (English translation of excerpts in this volume.)

Noether, Emmy. 1918. "Invariante Variationsprobleme." *Königliche Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Nachrichten* (1918):235–257.

Norton, John D. 1984. "How Einstein Found His Field Equations, 1912–1915." *Historical Studies in the Physical Sciences* 14:253–316.

Pais, Abraham. 1982. *'Subtle is the Lord ...' The Science and the Life of Albert Einstein*. Oxford, New York, Toronto, Melbourne: Oxford University Press.

Papapetrou, Achille. 1974. *Lectures on General Relativity.* Dordrecht/Boston: D. Reidel.

Pauli, Wolfgang. 1921. "Relativitätstheorie." In *Encyklopädie der mathematischen Wissenschaften, mit Einschluss ihrer Anwendungen*, edited by Arnold Sommerfeld. Leipzig: B. G. Teubner, 539–775.

———. 1958. *Theory of Relativity.* Translated by G. Field. London: Pergamon.

———. 1979. *Scientific Correspondence with Bohr, Einstein, Heisenberg, a.o. Volume 1: 1919–1929*. New York: Springer.

Reidemeister, Kurt, ed. 1971. *Hilbert-Gedenkenband*. Berlin, Heidelberg, New York: Springer,

Renn, Jürgen. 1994. "The Third Way to General Relativity." Preprint n° 9, *Max Planck Institute for the History of Science,* Berlin (http://www.mpiwg-berlin.mpg.de/Preprints/P9.PDF). (Revised edition in vol. 3 of this series.)

Renn, Jürgen, and Tilman Sauer. 1996. "Einsteins Züricher Notizbuch." *Physikalische Blätter* 52:865–872.

———. 1999. "Heuristics and Mathematical Representation in Einstein's Search for a Gravitational Field Equation." In (Goenner et al. 1999, 87–125).

Rosenfeld, Leon. 1940. "Sur le tenseur d'impulsion-énergie." *Mémoires de l'Academie royale de Belgique* 18 (16):1–30.

Rowe, David. 1989. "Klein, Hilbert, and the Göttingen Mathematical Tradition." *Osiris* 5:186–213.

———. 1999. "The Göttingen Response to General Relativity and Emmy Noether's Theorems." In *The Visual World: Geometry and Physics (1890–1930)*, ed., J. J. Gray. Oxford: Oxford University Press.

Sauer, Tilman. 1999. "The Relativity of Discovery: Hilbert's First Note on the Foundations of Physics." *Archive for History of Exact Sciences* 53:529–575.

———. 2002. "Hopes and Disappointments in Hilbert's Axiomatic 'Foundations of Physics.'" In *History of Philosophy and Science: new trends and perspectives*, ed. M. Heidelberger and F. Stadler. Dordrecht: Kluwer, 225–237.

Schouten, Jan A. 1924. *Der Ricci-Kalkül*, 1st ed. Berlin: Springer-Verlag.

Schouten, Jan A., and Dirk J. Struik. 1935. *Algebra und Übertragungslehre*. Vol. 1, *Einführung in die neueren Methoden der Differentialgeometrie*. Groningen, Batavia: P. Noordhoff.

Schwarzschild, Karl. 1916. “Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie.” *Königlich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (1916) (VII):189–196.

Siegmund-Schultze, Reinhard. 1998. *Mathematiker auf der Flucht vor Hitler: Quellen und Studien zur Emigration einer Wissenschaft*. Vol. 10, *Dokumente zur Geschichte der Mathematik*. Braunschweig Wiesbaden: Vieweg.

Slebodzinski, Wladyslaw. 1931. “Sur les equations de Hamilton.” *Bulletin de l'Academie royale de Belgique* (5) (17):864–870.

Stachel, John. 1989. “Einstein's Search for General Covariance, 1912–1915.” In *Einstein and the History of General Relativity,* edited by Don Howard and John Stachel. Boston/Basel/Berlin: Birkhäuser, 63–100.

———. 1992. “The Cauchy Problem in General Relativity - The Early Years.” In *Studies in the History of General Relativity,* edited by Jean Eisenstaedt and A. J. Kox. Boston/Basel/Berlin: Birkhäuser, 407–418.

———. 1994. “Scientific Discoveries as Historical Artifacts.” In *Current Trends in the Historiography of Science,* edited by Kostas Gavroglu. Dordrecht, Boston: Reidel, 139–148.

———. 1999. “New Light on the Einstein-Hilbert Priority Question.” *Journal of Astrophysics and Astronomy* 20:91–101. Reprinted in (Stachel 2002).

———. 2002. *Einstein from ‘B’ to ‘Z’*. (*Einstein Studies* vol. 9.) Boston: Birkhäuser.

Thorne, Kip S. 1994. *Black Holes and Time Warps: Einstein's Outrageous Legacy.* New York, London: Norton.

Vizgin, Vladimir P. 1989. “Einstein, Hilbert, and Weyl: The Genesis of the Geometrical Unified Field Theory Program.” In *Einstein and the History of General Relativity,* edited by Don Howard and John Stachel. Boston/Basel/Berlin: Birkhäuser, 300–314.

———. 1994. *Unified Field Theories in the First Third of the 20th Century.* Translated by Barbour, Julian B. Edited by E. Hiebert and H. Wussing. Vol. 13, *Science Networks, Historical Studies*. Basel, Boston, Berlin: Birkhäuser.

Walter, Scott. 1999. “Minkowski, Mathematicians, and the Mathematical Theory of Relativity.” In (Goenner et al. 1999, 45–86).

Weitzenböck, Roland. 1920. “Über die Wirkungsfunktion in der Weyl'schen Physik.” *Akademie der Wissenschaften (Vienna). Mathematisch-naturwissenschaftliche Klasse. Sitzungsberichte* 129:683–696.

Weyl, Hermann. 1917. “Zur Gravitationstheorie.” *Annalen der Physik* 54:117–145.

———. 1918a. *Raum-Zeit-Materie. Vorlesungen über allgemeine Relativitätstheorie*. 1st ed. Berlin: Julius Springer.

———. 1918b. *Raum-Zeit-Materie. Vorlesungen über allgemeine Relativitätstheorie*. 2nd ed. Berlin: Julius Springer.

———. 1918c. "Gravitation und Elektrizität.” *Königlich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (1918):465–480.

———. 1919. *Raum-Zeit-Materie. Vorlesungen über allgemeine Relativitätstheorie*. 3rd, revised ed. Berlin: Julius Springer.

———. 1921. *Raum-Zeit-Materie. Vorlesungen über allgemeine Relativitätstheorie*. 4th ed. Berlin: Julius Springer.

———. 1923. *Raum-Zeit-Materie. Vorlesungen über allgemeine Relativitätstheorie*. 5th, revised ed. Berlin: Julius Springer.

Whittaker, Edmund Taylor. 1951. *A History of the Theories of Aether and Electricity.* Vol. 1: *The Classical Theories*. London: Nelson.