MAX-PLANCK-INSTITUT FÜR WISSENSCHAFTSGESCHICHTE

Max Planck Institute for the History of Science

PREPRINT 118 (1999)

Jürgen Renn and John Stachel

Hilbert's Foundation of Physics: From a Theory of Everything to a Constituent of General Relativity

ISSN 0948-9444

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

JÜRGEN RENN AND JOHN STACHEL
EDITED BY STEFAN HAJDUK

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

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 five days before Einstein on 20 November 1915 – 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:[2]

> 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.

Hilbert himself emphasized that he actually 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 equations 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 later general 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

---

1 For discussions of Einstein's path to general relativity see *Norton 1984* and *Renn and Sauer 1998*. For historical reviews of Hilbert's contribution, see *Guth 1970*; *Mehra 1974*; *Earman and Glymour 1978*; *Pais 1982*, pp. 257-261; *Corry 1996*; *Corry, Renn, and Stachel 1997*; *Corry 1997*; *Corry 1999a*; *Sauer 1999, Corry (Forthcoming)*; and *Corry 1999b.*

2 *Thorne 1994*, p. 117; for a similar account see *Fölsing 1997*, pp. 375-376.

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 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 we have shown that Hilbert actually did not anticipate Einstein in presenting the field equations.[5] Our argument is based on the analysis of a set of proofs of Hilbert's first paper; in the following they are referred to as the "Proofs."[6] These Proofs not only do not include the explicit form of the field equations of general relativity, but they also show Hilbert's original 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.[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.[9]

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

---

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 *Corry, Renn, and Stachel 1997.*

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. 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, the Proofs carry several marginal notes in Hilbert's hand which are discussed below.

7 The conclusive paper is *Einstein 1915e*.

8 *Hilbert 1915*. In the following referred to as Paper 1.

9 *Hilbert 1917*. In the following referred to as Paper 2.

Klein who took the distinctiveness of Hilbert’s approach also as an argument against seeing it from the perspective of a priority competition with Einstein:[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.

It is indeed clear that 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, even if one disregards the non-covariant version of his theory as presented in the proofs version of his first paper. As we shall discuss in detail below, in the context of Hilbert’s attempt at a synthesis of electromagnetism and gravitation theory, he interpreted, for instance, the contracted Bianchi identities as substituting 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, not so much the mere fact of a variational formulation of gravitational field equations, which had already been introduced by Einstein, but, for instance, the relation of the gravitational Lagrangian to the Ricci scalar, and first hints of Noether’s theorem.

The intrinsic plausibility of both of these two perspectives, seeing Hilbert’s work as aiming at a theory different from general relativity or as a contribution to general relativity, respectively, represents a puzzle. In fact, how can Hilbert’s contributions be interpreted as making sense only within an independent research program, different in essence from that of Einstein, if they ultimately 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 foundation of 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 emerges and eventually fades away that is different from the one eventually accepted as mainstream.

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

10 *Klein 1921*, p. 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*.

came back time and again to his original 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 merely as an elimination of errors and the introduction of improvements and elaborations of a program he had been following from the beginning.

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 results and techniques of Einstein's 1913/14 non-covariant theory of gravitation with Mie's electrodynamic theory of matter in the form of an axiomatic system. It will become clear that Hilbert's research agenda was largely shaped by his understanding of such an axiomatic synthesis of physical theories, by the technical problems of achieving this synthesis, and by open problems in Einstein's own 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 agenda reconstructed in the second section. We shall show, in particular, that, in the course of pursuing this agenda, he reversed his original idea of founding all of physics on electrodynamics, now treating the gravitational field equations 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 agenda. The mathematical statement at the core of Hilbert's attempt to establish a connection between gravitation and electromagnetism in turn, as we shall also see, had its origin in Einstein's erroneous claim of 1913/14 that generally covariant field equations are not compatible with physical causality, a claim supported by Einstein's ill-fated "hole-argument." Hilbert thus effectively turned Einstein's argument against general covariance into support for his unified theory of gravitation and electromagnetism. Hilbert followed Einstein also in relating the choice of a preferred class of coordinate systems to the requirement of energy conservation. His definition of energy was, however, not guided by Einstein's example but rather by the goal of establishing a link with Mie's theory. Hilbert's theory thus emerges an extension of Einstein's non-covariant theory of gravitation, in which now Mie's speculative theory of matter plays the role of a touch stone, a role that for Einstein was played by the knowledge of classical and special relativistic physics as it is incorporated in the principle of energy-momentum conservation 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 at the end of 1915, focussing on the ways in which they mutually influenced each other. It is shown 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 on general relativity. Hilbert’s vision, which Einstein temporarily adopted, provided the latter with a perspective that was rather exotic but allowed him to attain a crucial result, the calculation of Mercury’s perihelion precession, which, in turn, guided his completion of general relativity, at the same time rendering 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, undermining its entire architecture and, 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, Hilbert modified essential elements of his theory before its publication in March 1916. He abandoned, in particular, the attempt to develop a non-covariant theory, without however, as yet having found a satisfactory solution to the causality problem that Einstein had raised for generally covariant theories. He replaced his original, non-covariant notion of energy by a new formulation, still different from that of general relativity and mainly intended to tighten the link between his own theory and Mie’s electrodynamics. In fact, Hilbert did not abandon his ambition to provide a foundation for all of physics. He evidently hoped to construct a field-theoretical model of the electron and derive its laws of motion, without, however, getting far enough to include any results on these issues in his paper. His first paper was republished twice, in 1924 and 1933, with significant revisions. We will show that Hilbert eventually adopted the understanding of energy-momentum conservation as developed in general relativity, and thus effectively transformed 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 the revision 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. While, 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, 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 by the insight that coordinate systems have no physical significance in general relativity, Hilbert attempted to find a purely "mathematical response" to this problem, formulating causality in terms of a Cauchy initial-value problem. This 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 a generally covariant theory, corrected only 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 the contemporary standard 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 almost exclusively perceived as a contribution to general relativity. It will become clear that this reception was largely shaped 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 vanishing of his original, distinct point of view. This vanishing had two remarkable consequences: First, actual deviations of Hilbert's theory from general relativity, such as his interpretation of the contracted Bianchi identies as a coupling between gravitation and electromagnetism, went practically unnoticed. Second, in spite of his attempt to fashion himself as the founding father of unified field theories, the early workers in this field tended to ignore his contribution and refused him a prominent place in their intellectual ancestry. Instead, Hilbert was assigned such a place in the history of general relativity, even at the price of 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 in general relativity. The ease with which his work could be assimilated to general relativity provides evidence of a different kind for the tenuous and unstable character of his own framework.

In the *eigth and final section* ("At the end of a royal road") we shall compare Hilbert's and Einstein's approaches in order to explain Hilbert's gradual rapprochment with general relativity. While 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 comparatively much narrower physical base. Starting from only a few problematic physical assumptions, Hilbert elaborated a mathematically complex framework and 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 and, in turn, transformed his own results into contributions to general relativity. Thus, in a way, Hilbert's assimilation of insights from general relativity served as a substitute for the physical component of Einstein's double strategy that was originally missing from Hilbert's own approach. This double strategy hence emerges not only as a successful heuristics characterizing Einstein's individual pathway but as a particular aspect of the more general process by which knowledge was integrated into the emerging theory of general relativity.

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

In a series of papers 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 *Foundations of Physics*.[11] 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 and firm foundation that has to *precede* empirical research and theory formation, but as a *post-hoc* reflection on their results with the aim of clarifying the logical and epistemological structure of the assumptions, definitions, etc., on which they is built.[12] Nevertheless, Hilbert expected that a proper axiomatic foundation of physics could 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 the concepts in an axiomatic foundation of physics to be those already familiar from the history of physics, rather than new ones emerging from the reorganization of the existing body of knowledge. Finally, Hilbert was convinced that one can strictly distinguish between the empirical and the universal ingredients of the body of knowledge of physics.

In accordance with this understanding, the task that Hilbert set for himself was not to find new concepts that could serve to integrate the existing body of physical knowledge into a coherent conceptual whole, but rather to formulate appropriate axioms involving the already existing physical concepts that 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 (Mechanical or electromagnetic world view). Indeed, in his early 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 set his hopes on an axiomatization of all of physics based on electrodynamics.[13] In spite of the conceptual revolution brought about by special relativity, not only concerning the revision of the concepts of space and time but also the con-

11 See *Corry 1996*; *Corry 1997*; *Corry 1999a*; *Corry 1999b*; *Corry (Forthcoming)*; see also *Sauer 1999*, section 1.

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

13 For a discussion of Hilbert's turn from a mechanical to an electromagnetic reductionism, see *Corry 1999a*, pp. 511-517.

ceptual autonomy of the field concept from that of the ether, Hilbert nevertheless continued to count on traditional concepts such as force and rigidity as the building blocks for his axiomatization program.[14]

An axiomatic synthesis of existing knowledge such as that pursued by Hilbert in physics apparently also had a strategic significance for Göttingen mathematicians since it made possible to leave their distinctive mark on a broad array of domains which were thus “appropriated,” not only intellectually but also in the sense of the professional responsibility for these domains. Minkowski’s attempt to present his work on special relativity as the decisive mathematical synthesis of the work of his predecessors may serve as an example.[15]In the context of an accusation that Emmy Noether had neglected to acknowledge her intellectual debt to British and American algebraists, Garrett Birkhoff wrote:[16]

> 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.

Mie’s theory of matter

By 1913, Hilbert expected that the electron theory 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 with the aim of overcoming the inherent dualism between “ether” 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 “ether” as well as the particulate structure of matter from one and the same variational principle.[17] Mie’s theory thus not only corresponded to Hilbert’s expectations with regard to founding all of physics on the concepts of electrodynamics, but it must also have been attractive to him because it was based upon variational calculus, a tool the usefulness of which for the axiomatization of physical theories Hilbert was quite familiar with.[18] However, Mie’s theory was far from yielding specific results concerning the electromagnetic properties of matter that could be confronted with the results of empirical research. Rather it correponds to a framework for postulating a suitable “world function” (Lagrangian), from which such concrete predictions may then perhaps

---

14 See *Hilbert 1913*, p. 13.

15 This attempt is extensively discussed in *Walter 1998;* see also *Rowe 1989*.

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

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

18 See, in particular, *Hilbert 1905*.

be derived. Mie gave examples for his world function which, however, were merely meant to be illustrations of certain features of his framework. In fact, Mie could not consider these examples to be the basis of a specific physical theory since they are not even compatible with elementary aspects of physical reality such as the existence of an elementary quantum of electricity. Concerning his principal example, later taken up by Hilbert, Mie remarked:[19]

> Eine Welt, die durch die Weltfunktion
>
> $$\varphi = -\frac{1}{2}\eta^2 + \frac{1}{6}a \quad \chi^6 \tag{1}$$
>
> 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.

On a more general level, Mie drew the conclusion that the unknown world function he eventually hoped to find must be more complicated than this and other examples he had considered.[20]

Hilbert based his research on a formulation of Mie’s framework that is due to Max Born.[21] 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:[22]

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

where φ is a Lorentz scalar and:

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

are the projections of the displacements of the points of the four-dimensional continuum from their equilibrium positions onto four orthogonal axes 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. Born showed furthermore that the characteristic feature of Mie’s theory lies in the Ansatz that the function φ depends only on the antisymmetric part of $a_{\alpha\beta}$:

19 *Mie 1912b*, p. 38. For the meaning of Mie’s formula and its ingredients, compare Hilbert’s version of this formula; see (33) below.

20 See *Mie 1912b*, p. 40.

21 For a discussion of Born’s role as Hilbert’s informant on both Mie’s and Einstein’s theories, see *Sauer 1999*, pp. 538-539.

22 See *Born 1914*.

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

He could thus regard Mie's four-dimensional continuum as a generalization of MacCullagh's three-dimensional ether to a four-dimensional space-time continuum. MacCullagh had derived equations corresponding to Maxwell's equations for stationary electrodynamic processes from the assumption that the vortices of the ether, rather than its deformations, store its energy.

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:[23]

> 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.

Mie initially 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:[24]

> 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.

In his third and conclusive paper, however, he explicitly withdrew this model and saw himself forced to introduce the gravitational potential as an additional variable into his theory.[25] 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 excluded.[26] As a consequence, his treatment of gravitation suffers from the same objections that were raised in the contemporary discussion of Abraham's theory. In summary, Mie's theory of gravitation was far from reaching the goals he had earlier formulated.

---

23 See *Mie 1912a*, pp. 511-512.

24 See *Mie 1912a*, pp. 512-513.

25 See *Mie 1913*, p. 5.

26 Compare *Mie 1912a*, p. 534, with *Mie 1913*, p. 29.

Einstein's non-covariant theory of gravitation

In 1915, Hilbert also became interested in Einstein's theory of gravitation, originally after a series of talks the latter gave between 28 June and 5 July of this year in Göttingen on this topic.[27] Hilbert's interest in Einstein's approach may have its roots in the contrast between Mie's programmatic statements about the need for a unification of gravitation and electromagnetism and the unsatisfactory treatment of gravitation actually given in Mie's theory. This may well have raised Hilbert's expectations and provided him with a motivation for looking at other theories of gravitation and perhaps even for the invitation to 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 and that of Einstein's theory of gravitation as presented in his 1913/1914 publications and presumably also in the 1915 Göttingen lectures.

Like Mie's theory, Einstein's was built upon 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's electrodynamics, however, Einstein had specified a particular Lagrangian, from which he then derived the gravitational field equations:[28]

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

To be precise, Einstein was only 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):[29]

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

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

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

27 For notes on a part of Einstein's lectures, see "Nachschrift of Einstein's Wolfskehl Lectures" in *Kox et al. 1996*, pp. 586-590. For a discussion of Einstein's Göttingen visit and its potential impact on Hilbert, see *Corry 1999a*, pp. 514-517.

28 Our presentation follows Einstein's major review paper *Einstein 1914b.*

29 The magnitudes in Gothic script represent tensor densities with respect to linear transformations.

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

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

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

The field equations of Einstein's 1913/1914 theory of gravitation are not generally covariant, although most other expressions in his theory are, 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 stress-energy tensor of matter. While this lack of general covariance of the field equations had initially seemed to be a blemish on his theory, in late 1913 Einstein convinced himself that he could even demonstrate – by means of the notorious "hole"-argument – that generally covariant field equations are 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, for which Einstein was able to construct infinitely many apparently inequivalent solutions for a set of generally covariant field equations with given sources and boundary values. This boundary value problem is formulated for a region of space-time devoid of matter - hence the name "hole"-argument. From the perspective of this argument, if one considers generally covariant field equations at all, then they must be supplemented by four additional non-covariant equations necessary to pick out a unique solution. From the perspective of the mature theory of general relativity of 1915, however, the hole argument is no longer seen as an objection against generally covariant field equations because the mathematically distinct solutions are not regarded as representing physically distinct situations, but merely as different representations of the same physical situation.[30]

Even in 1913/14 Einstein believed that one could formulate generally covariant equations, from which then the equations (8) would follow by introducing a suitable coordi-

30 See *Stachel 1989*; sections 3 and 4, pp. 71-81.

nate restriction.[31] While he actually failed to find those generally covariant field equations corresponding to (8), he did identify four non-covariant coordinate restrictions, which he believed to be characteristic for his theory. He obtained these coordinate restrictions from an analysis of the transformational behavior of the variational principle on which his theory was based. Expressed in terms of the Lagrangian H, these four coordinate restrictions are:

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

Einstein regarded these conditions as making it evident that his theory is not generally covariant, indeed so restrictive as to avoid the hole-argument. In addition, Einstein required the existence of a gravitational energy-momentum complex (non-tensorial) that guarantees the validity of energy-momentum conservation for the combined matter and gravitational fields, resulting in another four equations. He thus had 10 field equations, 4 coordinate restrictions, and 4 conservation equations, all in all 18 equations for the 10 gravitational potentials $g_{\mu\nu}$.

Einstein used this overdetermination as a criterion for the choice of the right Lagrangian, imposing the condition that the field equations together with the energy-momentum conservation equations should yield the coordinate restrictions (10). For this purpose, he first assumed a general Lagrangian H depending on $g_{\mu\nu}$, $g_{\mu\nu,\kappa}$ and then examined the four equations implied by the assumption of energy-momentum conservation for the resulting field equations. Formulating energy-momentum conservation in terms of the requirement that the covariant divergence of the energy-momentum tensor density $\mathbf{T}_\sigma{}^\nu$ has to vanish and using the field equations (8), he first obtained:

$$\begin{aligned} {}_\nu\mathbf{T}_\sigma{}^\nu \equiv \sum_{\nu\tau}\frac{\partial}{\partial x_\nu}(g^{\tau\nu}\mathbf{T}_{\sigma\tau}) + \frac{1}{2}\sum_{\mu\nu}\frac{\partial g^{\mu\nu}}{\partial x_\sigma}\mathbf{T}_{\mu\nu} &= 0 \\ \sum_{\nu\tau}\frac{\partial}{\partial x_\nu}(g^{\tau\nu}\mathbf{E}_{\sigma\tau}) + \frac{1}{2}\sum_{\mu\nu}\frac{\partial g^{\mu\nu}}{\partial x_\sigma}\mathbf{E}_{\mu\nu} &= 0, \end{aligned} \tag{11}$$

and then:

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

with B_σ given by (10) and:

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

31 See, e.g., *Einstein 1914a*, pp. 177-178. It is, however, unclear whether Einstein expected the unknown generally covariant equations to be of higher order.

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 given by (6), and thus for the uniqueness of the field equations, is fallacious.[32] This insight helped to motivate him to return to generally covariant field equations.

If one disregards all richness of Newtonian gravitation theory that buttressed even Einstein's gravitational theory of 1913/14, that theory may appear almost as speculative as Mie's theory of matter. On the one hand, Einstein had been able to draw a number of conclusions from his theory that, at least in principle, could be empirically checked, such as the perihelion shift of Mercury, the deflection of light in a gravitational field, and gravitational redshift. But on the other hand, none of these conclusions had actually received such support by the time Hilbert turned to Einstein's work, and the calculated perihelion shift even turned out to be in disaccord with observation in both sign and value.

Hilbert's research program

To a mathematician of Hilbert's competence, Einstein's 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 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. Such questions presumably suggested to Hilbert a rather well circumscribed program of work, which, taken together with his interest in Mie's theory of matter, amounted to an "axiomatic synthesis" of the two speculative physical theories.

In consequence, Hilbert's initial program presumably comprised:[33]

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 depending on both Mie's electrodynamical and Einstein's gravitational variables;
2. an examination of the possibility of replacing Einstein's unspecified energy-momentum tensor by one following from Mie's Lagrangian;

32 For historical discussion, see *Norton 1984*.

33 For a similar attempt to reconstruct Hilbert's research program, see *Sauer 1999*, pp. 557-559.

3. a further examination of the non-uniqueness of generally-covariant equations, involving a study of the question of the number of independent equations, and finally
4. the identification of appropriate coordinate restrictions and an examination of their relation to energy-momentum conservation.

Even prior to looking at Hilbert's attempts to realize such a synthesis of Mie's and Einstein's approaches, it thus becomes clear that such a program would fit perfectly Hilbert's understanding of his axiomatic approach to physics. Indeed, the realization of what we suggest was Hilbert's initial program would correspond to a clarification of the logical and mathematical foundations of already existing physical theories, if possible in their own terms; it would represent the synthesis of different theories by the combination of logically independent elements within one and the same formalism (for example incorporation of Mie's variables and Einstein's variables in the same Lagrangian), and it would replace the unspecified characterization of the material sources entering Einstein's theory with a daring theory of their constitution formulated in mathematical terms, thus shifting the boundary between experience and mathematical deduction in favor of the latter.

Unfortunately, there is no direct evidence for our claim 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 Hilbert followed.[34] However, we do have the first proofs of Hilbert's first communication which (as we have argued)[35] provides a glimpse into an earlier stage of 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 a research program such 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 (hereafter, Paper 1).[36] We will begin by reconstructing from the Proofs of Hilbert's first paper, as well as from other contemporary documents, the first step in the realization of Hilbert's research program. This crucial step, resulting from an attempt to explore the first two points of Hilbert's initial program, consisted of the establishment of a relation between Mie's energy momentum-tensor and the

34 Einstein's search for a gravitational field equation in the winter of 1912/13 is documented in the so-called Zurich Notebook, partially published as Doc. 10 of *Klein et al. 1995*. Einstein's research notes have been reconstructed in the course of a joint project by Michel Janssen, John Norton, Jürgen Renn, John Stachel, and Tilman Sauer; for a recent publication based on this work, see *Renn and Sauer 1998*.

35 *Leo Corry, Jürgen Renn, and John Stachel 1997.*

36 *Hilbert 1915*.

variational derivative of Mie's Lagrangian with respect to the metric.[37] Next, we attempt to reconstruct the precise calculations by which Hilbert obtained Mie's energy momentum-tensor from the Born-Mie Lagrangian. We then examine the consequences of Hilbert's derivation of Mie's energy-momentum tensor for his concept of energy, and thus for the further exploration of the second point of his program. We then discuss how Hilbert's results suggested a new perspective on the relation between Mie's and Einstein's theories, in which gravitation appears to be more fundamental than electrodynamics. Seen from this perspective, the third point of Hilbert's program, dealing with the question of uniqueness of solutions to generally-covariant equations, took on a new significance. In particular, 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 research program, we show how he united his concept of energy 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.

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 potential in Einstein's theory of gravitation. In the Proofs and published versions 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. He thus wrote in the proofs version:[38]

> 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) [equation in the original, i.e. $g_{\mu\nu} = \delta_{\mu\nu}$] - 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.

Hilbert expressed himself similarly in a letter to Einstein of 13 November 1915:[39]

> 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.

37 Unless indicated to the contrary, when we speak of the variational derivative of the Lagrangian, we shall always mean with respect to the metric tensor.

38 Proofs, p. 10.

39 David Hilbert to Einstein, 13 November 1915, *Schulmann et al. 1998*, p. 195.

On the basis of our reconstruction of Hilbert's initial research program, it is possible to suggest how Hilbert's heuristics might have led him to this relation. We only have to assume that he attempted to realize the first two steps listed above, 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 in Einstein's original version, one obtains:

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

where $\mathbf{T}_{\mu\nu}$ should now be the energy-momentum tensor of Mie's theory. It may well have been an equation of this form, following quite naturally 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 can be obtained as the variational derivative of Mie's Lagrangian.

Following a program such as that outlined above, Hilbert would, first of all, have assumed that the Lagrangian has the form:

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

where K represents the gravitational part and L the electromagnetic part. Indeed, this form of the Lagrangian is used both in the Proofs and the published version of Paper 1.[40]

In his paper, Hilbert derived a relation for this Lagrangian of the form:

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

where $T_{\nu}^{\ m}$ stands for the energy-momentum tensor density of Mie's theory.[41] 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 space and time as represented by the metric tensor and the structure of matter as represented by Mie's theory.

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

41 Compare Proofs, p. 10; *Hilbert 1915*, p. 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 $T_{\nu}^{\ m}$.

Mie's energy-momentum tensor as a consequence of generally covariant field equations

The strategy that Hilbert followed in order to derive (17) can be reconstructed from the two versions of his paper. It consisted in following as closely as possible the usual variational techniques applied, for instance, to derive Lagrange's equations from a variational principle.[42] In Hilbert's paper, a similar variational problem forms the core of his theory. He describes his basic setting in terms of two axioms. The first of these reads:[43]

> Axiom I (Mie's Axiom von der Weltfunktion): *Das Gesetz des physikalischen Geschehens bestimmt sich durch eine Weltfunktion H, die folgende Argumente enthält:*
>
> $$g_{\mu\nu}, \quad g_{\mu\nu l} = \frac{\partial g_{\mu\nu}}{\partial w_l}, \qquad g_{\mu\nu lk} = \frac{\partial^2 g_{\mu\nu}}{\partial w_l \partial w_k}, \qquad \qquad (18)$$
> $$q_s, \quad q_{sl} = \frac{\partial q_s}{\partial w_l} \qquad (l, k = 1, 2, 3, 4)$$
>
> *und zwar muß die Variation des Integrals*
>
> $$\int H \sqrt{g} d\tau \qquad \qquad (19)$$
> $$(g = |g_{\mu\nu}|, \quad d\tau = dw_1 dw_2 dw_3 dw_4)$$
>
> für jedes der 14 Potentiale $g_{\mu\nu}, q_s$ verschwinden.

The w_s are Hilbert's notation for an arbitrary system of coordinates. Hilbert's second axiom establishes the generally covariant character of this variational problem:

> Axiom II (Axiom von der allgemeinen Invarianz): *Die Weltfunktion H ist eine Invariante gegenüber einer beliebigen Transformation der Weltparameter* w_s.

Hilbert formed a differential expression for an arbitrary invariant J 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,* where he defined the operator *P* as follows:[44]

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

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

42 See, for example, *Caratheodory 1935*.

43 Proofs, p. 2; see also *Hilbert 1915*, p. 396.

44 See *Hilbert 1915*, pp. 398-399; compare Proofs, p. 4 and p. 7.

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

In the mathematical terminology of the time, PJ is a "polarization" of J.[45]

As we shall see in the following, it is possible to derive from the expression for PJ identities that realize Hilbert's goal, the derivation of (17). Since 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 in Hilbert's thinking from the Proofs to the published version, our reconstruction will make use of the latter.

If $p^{\mu\nu}$ and p_l are these variations generated by dragging the metric and the electromagnetic potentials over the manifold by some vector field p^s, i.e., in modern terminology, if they are the Lie derivatives of the metric and the electromagnetic potentials with respect to p^s,[46] then PJ must be the Lie derivative of J with respect to p^s. On the other hand, since J is a scalar invariant, PJ can be also expressed as the Lie derivative of this scalar with respect to p^s so that:

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

The last equation can be rewritten in the form (23) below, which is the statement of Hilbert's Theorem II, both in the Proofs and in the published version:[47]

> Theorem II. Wenn J eine von $g^{\mu\nu}, g_l^{\mu\nu}, g_{lk}^{\mu\nu}, q_s, q_{sk}$ abhängige Invariante ist, so gilt stets identisch in allen Argumenten und für jeden willkürlichen kontravarianten Vektor p^s

45 See, e.g., *Kerschensteiner 1887*, § 2.

46 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*, p. 471, where he writes: "Zu diesem Zweck 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*, p. 142.

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

dabei ist:

---

47 See Proofs, pp. 7-8, and *Hilbert 1915*, p. 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:

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

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

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

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:

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

which is clearly equivalent to (22).

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

Hilbert now focusses on the electromagnetic part of L his Lagrangian $H = K + L$, to which he applies Theorem II under the assumption that L only depends on the metric $g^{\mu\nu}$, the electromagnetic potentials q_s and their derivatives q_{sk}, but *not* on the derivatives of the metric tensor. This gives the identity:[48]

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

Since the vector field p^s is arbitrary, its coefficients as well as the coefficients of its derivatives must each vanish separately. Hilbert drew two conclusions from this fact, 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 electromagnetic variables appear in the Lagrangian, the second concerns the relation between this Lagrangian and Mie's energy-momentum tensor.

Although Hilbert had only specified certain general properties of L, such as the requirement that it be a generally invariant scalar that does not depend on the derivatives of the metric tensor, he was nevertheless able to show that the derivatives of the electromagnetic potentials can only appear in the specific form characteristic of Mie's theory (compare (5)). In fact, setting the coefficients of p_{sk}^m in (25) equal to zero, and remembering that $p_{sk}^m = p_{ks}^m$, one obtains:

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

or, since q_m cannot vanish identically:

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

48 See Proofs, p. 9, and *Hilbert 1915*, p. 403.

which implies that the derivatives of the electrodynamic potentials only appear in the antisymmetric combination familiar from Mie's theory:

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

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

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

corresponding to the familiar electromagnetic "six vector." Hilbert emphasized:[49]

> Dieses Resultat ergibt sich hier wesentlich als Folge der allgemeinen Invarianz, also auf Grund von Axiom II.

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.[50] Hilbert only gave what he considered to be the "two simplest" of the generally-covariant generalization of these invariants:

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

and:

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

According to Hilbert, the simplest expression that can be formed in analogy to the gravitational part of the Lagrangian K is:[51]

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

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

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

where β is another constant. This now corresponds to the Lagrangian given by Mie. Remarkably, in contrast to Mie, Hilbert does not even allude to the physical problems associated with this Lagrangian. And in contrast to Einstein, Hilbert at no point intro-

49 Proofs, p. 10. This passage reads in the published version: "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 1915*, p. 403.

50 See Proofs, p. 13, and *Hilbert 1915*, p. 407; Hilbert followed here the papers of Mie and Born; compare, in particular, *Born 1914*.

51 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*.

duces 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) concerns Mie's energy-momentum tensor; it corresponds to what we have called above "Hilbert's first results" (compare (17)). Setting the coefficient of p_m^ν equal to zero and using (27), he obtained:[52]

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

Noting that:

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

(34) can then be rewritten as:

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

The right-hand side of this expression can be identified as the generally-covariant generalization of Mie's energy-momentum tensor. It is this equation that inspired the above-quoted remark by Hilbert about the "Umstand, der mich zum ersten Mal auf den notwendigen engen Zusammenhang zwischen der Einsteinschen allgemeinen Relativitätstheorie und der Mie'schen Elektrodynamik hingewiesen ... hat." Hilbert had thus arrived at a result of the form (17) and, in addition, shown that characteristic properties of Mie's Lagrangian follow from its generally-covariant generalization. The latter result was interpreted by Hilbert an indication that gravitation must be conceived as being more fundamental than electromagnetism, as his later works indicate.

The definition of energy

While (36) indicates a strong link between a generally-covariant term in the Lagrangian and Mie's energy momentum tensor, taken by itself, it does not answer the question of how energy-momentum conservation is to be conceived in Hilbert's theory. Because of the different structure of his theory, Hilbert could not simply follow Einstein's precedent. In particular, Hilbert's theory does not allow for the interpretation of an energy-momentum tensor for matter as an external source, as does that of Einstein; so Hilbert could not, as does the latter, start with a conservation law for matter in Minkowski space-time and then generalize it to the case in which a gravitational field is present. Such a procedure

52 See Proofs, p. 10; *Hilbert 1915*, p. 404.

would have been in conflict with Hilbert's heuristics, according to which, as we have seen, matter itself is conceived in terms of electromagnetic fields that, in turn, arise in conjunction with, or even as an effect of, gravitational fields.

It seems that Hilbert's heuristics for finding an appropriate definition of energy is governed by formal criteria related to his understanding of energy conservation in classical physics, as well as by criteria with a more specific physical meaning related to the results he expected concerning Mie's theory. Hilbert's formal criteria are 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:[53]

> Der Energiebegriff kommt eben daher, dass man die Lagrangeschen Gleichungen in Divergenzform schreibt, und das, was unter der Divergenz steht, als Energie definiert.

As concerns Hilbert's physical criteria, any definition of the energy should 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, both in the proofs version 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.[54] Nevertheless, as will become clear in what follows, Hilbert's discussion of his energy concept can be construed as 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),
- he then formulated a divergence equation for his energy expression in analogy to classical and special-relativistic physics (Satz 2 in the Proofs) and imposed this equation as a requirement implying coordinate restrictions for his theory (Axiom III),
- and finally, he showed that his energy expression can be related to Mie's energy-momentum tensor (the true justification of his choice of an energy expression).

---

53 Die Grundlagen der Physik I, Ms. Vorlesung SS 1916, p. 98 (D. Hilbert, Bibliothek des Mathematischen Seminars, Universität Göttingen); from here on "SS 1916 Lectures."

54 In *Klein 1917*, p. 475, Klein quotes from a letter he had written to Hilbert concerning the latter's energy expression in Paper 1: "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: "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, *Schulmann et al. 1998*, p. 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 heuristics: "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, *Schulmann et al. 1998*, p. 288.

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," see below).

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 and its treatment according to the usual variational techniques. However, he now proceeded somewhat differently. Instead of focussing on the electromagnetic part *L* of the Lagrangian, he took 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* (compare (20)). Accordingly, in the Proofs, Hilbert forms the expression:[55]

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

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:

$$\begin{aligned} E = {} & \sum \left(H \frac{\partial \sqrt{g}}{\partial g^{\mu\nu}} g_s^{\mu\nu} + \sqrt{g} \frac{\partial H}{\partial g^{\mu\nu}} g_s^{\mu\nu} + \sqrt{g} \frac{\partial H}{\partial g_k^{\mu\nu}} g_{sk}^{\mu\nu} + \sqrt{g} \frac{\partial H}{\partial g_{kl}^{\mu\nu}} g_{skl}^{\mu\nu} \right) p^s \\ & - \sum (g^{\mu s} p_s^\nu + g^{\nu s} p_s^\mu) [\sqrt{g} H]_{\mu\nu} \\ & + \sum \left(\frac{\partial \sqrt{g} H}{\partial g_k^{\mu\nu}} g_s^{\mu\nu} + \frac{\partial \sqrt{g} H}{\partial g_{kl}^{\mu\nu}} g_{sl}^{\mu\nu} - g_s^{\mu\nu} \frac{\partial}{\partial w_l} \frac{\partial \sqrt{g} H}{\partial g_{kl}^{\mu\nu}} \right) p_k^s, \end{aligned} \tag{39}$$

and:

$$\begin{aligned} D^{(p)} = \sum \Bigg\{ & -\frac{\partial}{\partial w_k} \left(\sqrt{g} \frac{\partial H}{\partial g_k^{\mu\nu}} (g^{\mu s} p_s^\nu + g^{\nu s} p_s^\mu) \right) \\ & + \frac{\partial}{\partial w_k} \left((p_s^\nu g^{\nu s} + p_s^\mu g^{\nu s}) \frac{\partial}{\partial w_l} \left(\sqrt{g} \frac{\partial H}{\partial g_{kl}^{\mu\nu}} \right) \right) \\ & + \frac{\partial}{\partial w_l} \left(\sqrt{g} \frac{\partial H}{\partial g_{kl}^{\mu\nu}} \left(\frac{\partial p^{\mu\nu}}{\partial w_k} - g_{sk}^{\mu\nu} p^s \right) \right) \Bigg\}. \end{aligned} \tag{40}$$

55 See Proofs, p. 5ff.

In other words, Hilbert succeeded in splitting off a divergence term $D^{(p)}$ from the original expression $J^{(p)}$. By integrating over some region, $D^{(p)}$ could, in principle, be converted into a surface term and thus eliminated by demanding that p^s and its derivatives vanish on the boundary of that region.[56] So it was plausible to extract an energy expression from the remainder of $J^{(p)}$, if a way could be found to deal with the first term $\sum H\frac{\partial}{\partial g^{\mu\nu}}\sqrt{g}p^{\mu\nu}$.

Ultimately, the justification of choosing E as the energy expression depends, of course, on the possibility of a physical interpretation for this expression – which, as we shall see, for Hilbert 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 $J^{(p)}$. This alternative decomposition is derived from a generalization of (37). This equation may be identified, as we have indicated earlier, as a special case of a "polarization" of the Lagrangian H with respect to the contravariant form of the metric $g^{\mu\nu}$. In fact, 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:

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

where now:

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

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

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

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:

---

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

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

Comparing (45) with (38), it becomes clear that E indeed satisfies a divergence equation and thus represents a candidate for the energy expression. In the Proofs this conclusion is presented as one of two properties justifying this designation:[57]

> D e r A u s d r u c k E h e i ß e d i e E n e r g i e f o r m. 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) [i.e. (51) below] erfüllt sind:
>
> $$E = (D^{(h)})_{h=p} - D^{(p)} \tag{46}$$
>
> oder
>
> $$E = \sum\left\{\frac{\partial}{\partial w_k}\left(\sqrt{g}\frac{\partial H}{\partial g_k^{\mu\nu}}g_s^{\mu\nu}p^s\right) - \frac{\partial}{\partial w_k}\left(\frac{\partial}{\partial w_l}\left(\sqrt{g}\frac{\partial H}{\partial g_{kl}^{\mu\nu}}\right)g_s^{\mu\nu}p^s\right) + \frac{\partial}{\partial w_l}\left(\sqrt{g}\frac{\partial H}{\partial g_{kl}^{\mu\nu}}g_{sk}^{\mu\nu}p^s\right)\right\}, \tag{47}$$
>
> d. h. es gilt der Satz:
>
> S a t z 1. Die Energieform E wird v e r m ö g e d e r G r a v i t a t i o n s g l e i c h u n g e n einer Summe von Differentialquotienten nach w_s gleich, d. h. sie erhält Divergenzcharakter.

Whereas (47) for an arbitrary H involves electromagnetic and gravitational contributions in an undifferentiated way, Hilbert makes an Ansatz $H = K + L$ that allows him to separate these two contributions, in particular in order to relate E to his result concerning the energy-momentum tensor of Mie’s theory. Accordingly, at this point, he introduces (in a part of the Proofs that is now missing but corresponding to a part of Paper 1) the splitting of the Lagrangian (16), and introduces the condition that electromagnetic part not depend on $g_s^{\mu\nu}$.[58] Finally, he writes down explicitly the electromagnetic part of the energy:[59]

> Da K nur von $g^{\mu\nu}, g_s^{\mu\nu}, g_{lk}^{\mu\nu}$ abhängt, so läßt sich beim Ansatz (17) [i.e. (16)] die Energie E wegen (13) [i.e. (47)] 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) [i.e. (39)] den Ausdruck
>
> $$E = E^{(g)} + E^{(e)} \tag{48}$$

57 See Proofs, p. 6.

58 Compare *Hilbert 1915*, p. 402, with Proofs, p. 8.

59 Proofs, p. 8.

> wo die "Gravitationsenergie" $E^{(g)}$ nur von $g^{\mu\nu}$ und deren Ableitungen abhängt und die "elektrodynamische Energie" $E^{(e)}$ die Gestalt erhält
>
> $$E^{(e)} = \sum_{\mu, \nu, s} \frac{\partial \sqrt{g} L}{\partial g^{\mu\nu}} (g_s^{\mu\nu} p^s - g^{\mu s} p_s^{\nu} - g^{\nu s} p_s^{\mu}), \qquad (49)$$
>
> in der sie sich als *eine mit* $\sqrt{g}$ *multiplizierte allgemeine Invariante* erweist.

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

Hilbert's final expression (49) satisfies, in fact, what we have called above his "physical criteria" 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 tensor. In this way, Hilbert's definition of energy had been given its "physical justification" in terms of Mie's theory, while its relation to energy-momentum conservation in classical and special-relativistic physics remains – apart from merely formal similiarities – entirely unclear. As we shall see below, Hilbert's energy expression served still another and even more important function in the Proofs, that of determining admissible coordinate systems.

Hilbert's reversal of Mie's program and the roots of his *Leitmotiv* in Einstein's work

Hilbert was apparently convinced that the relation he had established between the variational derivative of the Lagrangian and the energy-momentum tensor (compare (36)) singled out Mie's theory as having a special relation to the theory of gravitation.[60] In fact, as we have seen, this conclusion is only justified insofar as one imposes on the electrodynamic part of 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 even be possible to conceive of electromagnetic phenomena as "effects of gravitation."[61] Such an interpretation was in line with the reductionist perspective implied by his understanding of the axiomatization of physics and led to a reversal of Mie's original aim of basing all of physics on electromagnetism.

In the light of the possibility of considering electromagnetism as an effect of gravitation, 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 for him. If it was indeed true, as suggested by Einstein's hole argu-

---

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

61 See Proofs, p. 3 amd *Hilbert 1915*, p. 397.

ment applied to Hilbert's formalism, that the 14 generally covariant field equations for the 14 gravitational and electromagnetic potentials do not have a unique solution for given boundary values so that 4 additional, non-covariant equations were required in order to assure a unique solution, then 4 identities must exist between the 14 field equations. If these 4 identities are somehow equivalent to the 4 equations for the electromagnetic potentials, then the latter can be considered as a consequence of the 10 gravitational equations by virtue of the unique properties of a generally-covariant variational principle. Then Hilbert would be indeed 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 the field equations must involve 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 knowledge about Einstein's theory. In addition the physicist Paul Hertz, then participating in the activities of the group 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.[62]

Einstein's treatment of these issues thus formed the background to the crucial theorem on which Hilbert's entire approach is based, his "Leitmotiv," labelled "Theorem I" in the proofs version of his paper:[63]

> Das Leitmotiv für den Aufbau meiner 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
>
> $$\delta\int J\sqrt{g}\,d\tau = 0 \qquad (50)$$
>
> 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.

62 Einstein to Paul Hertz, 22 August 1915, *Schulmann et al. 1998*, pp. 163-164. See *Norton and Howard 1993* for an extensive historical discussion.

63 Proofs, pp. 2-3; compare *Hilbert 1915*, pp. 396-397. See *Rowe (Forthcoming)* for a discussion of the debate on Hilbert's Theorem I among Göttingen mathematicians.

From such a variational principle (50) 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:

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

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

In both the Proofs and the published versions of Paper 1, Hilbert erroneously claimed that one can consider the last four equations as a consequence of the 4 identities that must hold, according to his Theorem I, between the 14 differential equations:[64]

> Die Gleichungen (4) [i.e. (51)] mögen die Grundgleichungen der Gravitation, die Gleichungen (5) [i.e. (52)] die elektrodynamischen Grundgleichungen oder die verallgemeinerten Maxwellschen Gleichungen heißen. Infolge des oben aufgestellten Theorems können die vier Gleichungen (5) [i.e. (52)] als eine Folge der Gleichungen (4) [i.e. (51)] 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.

We shall come back to this claim later, in connection with Hilbert's proof of a special case of Theorem 1.

The fact that Hilbert did not give a proof of the theorem makes it difficult to assess its heuristic roots. It is possible, of course, that these roots lay somewhere in Hilbert's extended mathematical knowledge, in particular of the theory of invariants. But the lack of a proof in Paper 1, as well as the peculiar interpretation it is given in the proofs version of this paper, make it plausible that the theorem had its ultimate origin in Einstein's hole argument on the ambiguity of solutions to generally-covariant field equations.

In fact, in the proofs version of Paper 1, Hilbert drew the implications of Theorem I for his field theory in the context of the problem of causality, just as Einstein had done for the hole argument. But while the latter 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:[65]

> Indem unser mathematisches Theorem lehrt, daß die bisherigen Axiome I und II [essentially amounting to the variational principle (50), see the discussion below] für die 14 Potentiale nur zehn wesentlich von einander unabhängige Gleichungen liefern können,

64 Proofs, pp. 3-4, and Hilbert 1915, pp. 397-398.

65 Proofs, pp. 3-4.

> 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) [i.e. (51)] und (5) [i.e. (52)] hinzutretenden nicht invarianten Gleichungen unerläßlich.

Hilbert's counting of equations thus closely parallels that of Einstein: the 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, 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 at the root 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 where the problem of causality is explicitly addressed.[66]

Energy-momentum conservation and coordinate restrictions

In the Proofs, Hilbert showed himself convinced that causality requires four supplementary non-covariant equations to fix the admissible coordinate systems. In identifying these coordinate restrictions, supposedly necessary in order to make the generally covariant field equations physically acceptable, he again followed, as we shall see in this section, closely in Einstein's tracks. As did the latter, Hilbert invoked energy-momentum conservation in order to justify the choice of a preferred reference frame. After formulating his version of energy-momentum conservation, he introduced the following axiom:[67]

> Axiom III (Axiom von Raum und Zeit). *Die Raum-Zeitkoordinaten sind solche besonderen Weltparameter, für die der Energiesatz (15)* [i.e. (57) below] *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) [$d^{(g)}\sqrt{g}H / dw_s = 0$] zur Folge: diese vier Differentialgleichungen (16) vervollständigen die Gravitationsgleichungen (4) [i.e. (51)] 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.

---

66 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 typrescript describes a construction equivalent to Einstein's hole argument discussed below.

67 Proofs, p. 7.

The strategy that Hilbert followed to extract these coordinate restrictions from the requirement of energy conservation indeed closely followed that of Einstein's theory of 1913/14. Even before Einstein developed the hole argument, energy-momentum conservation had played a crucial role justifying the lack of general covariance of his gravitational field equations. He was convinced that energy-momentum conservation actually required a restriction of the covariance group.[68] In the beginning of 1914, after having formulated the hole argument, he established the connection between coordinate restrictions and energy-momentum conservation in the "Entwurf" theory as follows:[69]

> 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) [the field equations in the form
>
> $$\sum_{\alpha\beta\mu}\frac{\partial}{\partial x_\alpha}\left(\sqrt{-g}\,\gamma_{\alpha\beta}\,g_{\sigma\mu}\frac{\partial\gamma_{\mu\nu}}{\partial x_\beta}\right) = \kappa(\mathbf{T}_{\sigma\nu}+\mathbf{t}_{\sigma\nu})\;]$$
>
> folgen nämlich durch Differentiation nach x_ν und Summation über ν unter Berücksichtigung der Gleichungen (III) [the conservations equations in the form
>
> $$\sum_{\nu}\frac{\partial}{\partial x_\nu}(\mathbf{T}_{\sigma\nu}+\mathbf{t}_{\sigma\nu}) = 0\;] \tag{53}$$
>
> die Beziehungen
>
> $$\sum_{\alpha\beta\mu\nu}\frac{\partial^2}{\partial x_\nu\partial x_\alpha}\left(\sqrt{-g}\,\gamma_{\alpha\beta}\,g_{\sigma\mu}\frac{\partial\gamma_{\mu\nu}}{\partial x_\beta}\right) = 0, \tag{54}$$
>
> also vier Differentialbedingungen für die Größen $g_{\mu\nu}$, welche wir abgekürzt
>
> $$B_\sigma = 0 \tag{55}$$
>
> 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.

In a later 1914 paper, Einstein discussed the physical significance and the transformation properties of the gravitational energy-momentum term $\mathbf{t}_{\sigma\nu}$:[70]

> Die Gleichungen (42 c) [i.e. (53))] drücken nach den in § 10 gegebenen Überlegungen die Erhaltungssätze des Impulses und der Energie für Materie und Gravitationsfeld zusammen aus. $\mathbf{t}_\sigma^\nu$ sind diejenigen auf das Gravitationsfeld bezüglichen Größen, welche den Komponenten $\mathbf{T}_\sigma^\nu$ des Energietensors (*V*-Tensors) [i.e. tensor density] der physikalischen Bedeutung nach analog sind. Es sei hervorgehoben, daß die $\mathbf{t}_\sigma^\nu$ nicht beliebigen

68 See, e.g., *Einstein 1913*, p. 1258.

69 *Einstein 1914*, pp. 218-219.

70 *Einstein 1914b*, p. 1077.

berechtigten, sondern nur linearen Transformationen gegenüber Tensorkovarianz besitzen; trotzdem nennen wir ($\mathfrak{t}_\sigma^\nu$) den Energietensor des Gravitationsfeldes.

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:[71]

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

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 of 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 only holds if e_s vanishes, which implies:

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

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 version of Paper 1 is incomplete,[72] Hilbert's theorem and its proof can be reconstructed:

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 along the line indicated in the surviving portion of Hilbert's argument. From (38) and (56) it follows:

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

which can be rewritten as:

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

71 Proofs, p. 6.

72 The top of Proofs, p. 7, is missing

where $\overline{D^{(p)}}$ is still a divergence expression. 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):

$$0 = \int_\Omega [\sqrt{g}H]_{\mu\nu} p^{\mu\nu} dx^4 = \int_\Omega \left(e_s - \frac{\partial e^l_s}{\partial w_l}\right) p^s dx^4. \tag{62}$$

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 for which he was looking:[73]

> Dieser Satz zeigt, daß die dem Energiesatz der alten Theorie entsprechende Divergenzgleichung
>
> $$\sum_l \frac{\partial e^l_s}{\partial w_l} = 0 \tag{63}$$
>
> dann und nur dann gelten kann, wenn die vier Größen e_s verschwinden ...

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 space-time coordinates. In fact, the latter are those world parameters that satisfy 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 coordinate system" to another, Hilbert introduced space-time transformations that lead from one "normal form" of the energy to another other:[74]

> Dem Übergang von einem Raum-Zeit-Bezugssystem zu einem anderen entspricht die Transformation der Energieform von einer sogenannten "Normalform"
>
> $$E = \sum_{s,l} e^l_s p^s_l \tag{64}$$
>
> auf eine andere Normalform.

The claim that Hilbert's introduction of coordinate restrictions was guided by his goal to recover the ordinary divergence form of energy-momentum conservation in analogy to Einstein's (53) 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.[75] This suggestion was taken

---

73 Proofs, p. 7.

74 Proofs, p. 7.

75 Felix Klein to David Hilbert, 5 March 1918, *Frei 1985*, pp. 142-143.

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.”[76] Starting from an expression for the covariant divergence of the stress-energy tensor:

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

Runge obtained his “regular” expression by imposing the four equations:

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

thus specifying a preferred class of coordinate systems. In his response, Hilbert sent Klein three pages of the proofs version of Paper 1, to show that he had anticipated Runge’s line of reasoning:

> 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 Mittei-ung, 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.[77]

Hilbert’s final sentence indeed confirms that the recovery of the familiar form of energy conservation was his heuristic 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 reference to Runge in the above passage might suggest. On the contrary, as we have seen above (compare (48)), Hilbert also 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. The results of these first steps in turn shaped Hilbert’s further research. They affected a change of perspective from seeing electrodynamics and gravitation on an equal footing to his vision of deriving electromagnetism from gravitation. As a consequence, the structure of Hilbert’s original,

---

76 For a discussion of Runge’s work, see *Rowe (Forthcoming)*.

77 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*, p. 544.

non-covariant theory is strikingly close to that of Einstein's 1913/14 theory of gravitation in spite of the covariance of Hilbert's gravitational equations and the different physical interpretation that he gave to his equations.

Electromagnetism as an effect of gravitation: the core of Hilbert's theory

We come now 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 his elaboration of a special case of Noether's theorem. We shall show that, in the original version of Hilbert's theory, the corresponding results were actually part of a different physical framework that also effected their interpretation. In a later section, we shall see how these results were eventually transformed, especially 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 – which, however, is the case only under further assumptions that Hilbert did not explicitly state and that run counter to Mie's program. From the point of view of general relativity, he obtained the Maxwell equations as a consequence of the integrability conditions for the gravitational field equations with an electromagnetic source term. It thus appears as if he had merely treated a special case of Einstein's equations and expressed certain 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 equations. For him, his derivation was not just a specialization of a more general framework, but closely interwoven with other results of his theory that pointed to electromagnetism as an effect of gravitation. In particular, the equation, on the basis of which he argued that electrodynamics is a consequence of gravitation, was for him a result of four ingredients. Two of these incorporate other links between gravitation and electrodynamics, and all of them ultimately are based on his generally-covariant variational principle:

- a general theorem corresponding to the contracted Bianchi identies,
- 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 characteristic way in which the derivatives of the electrodynamic potentials enter Mie's Lagrangian.

In the proofs version of Paper 1, Hilbert's general theorem is introduced as[78]:

> 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 -

78 Proofs, p. 9, and *Hilbert 1915*, p. 399.

$$\frac{1}{\sqrt{g}}\sum_{\mu,\nu}[\sqrt{g}J]_{\mu\nu}h^{\mu\nu} \qquad (67)$$

eine Invariante dar; setzen wir in dieser Summe an Stelle von $h^{\mu\nu}$ den besonderen Tensor $p^{\mu\nu}$ ein und schreiben

$$\sum_{\mu,\nu}[\sqrt{g}J]_{\mu\nu}p^{\mu\nu} = \sum_{s,l}(i_s p^s + i_s^l p_l^s), \qquad (68)$$

wo alsdann die Ausdrücke

$$\begin{aligned} i_s &= \sum_{\mu,\nu}[\sqrt{g}J]_{\mu\nu}g_s^{\mu\nu}, \\ i_s^l &= -2\sum_{\mu}[\sqrt{g}J]_{\mu s}g^{\mu l}, \end{aligned} \qquad (69)$$

lediglich von den $g^{\mu\nu}$ und deren Ableitungen abhängen, so ist

$$i_s = \sum_l \frac{\partial i_s^l}{\partial w_l} \qquad (70)$$

in der Weise, daß diese Gleichung identisch für alle Argumente, nämlich die $g^{\mu\nu}$ und deren Ableitungen, erfüllt ist.

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

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

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

Hilbert next applies Theorem III to the gravitational field equations (51), and then uses his knowledge about the electrodynamic part of the Lagrangian (see the last two "ingredients" listed above) in order to extract the electrodynamic equations from an identity corresponding 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 of the Lagrangian, but proceeded to exploit it for insights into the electrodynamic part. For Hilbert, however, this was natural as he was convinced, on the basis of Theorem I, that generally-covariant equations of gravitation are impossible as a "stand-alone" theory. It simply made no sense, in his eyes, 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:[79]

79 See Proofs, p. 11, and *Hilbert 1915*, p. 405.

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

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 now 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 obtains an integrability condition for the field equations that can also be interpreted as representing energy-momentum conservation.

Hilbert, however, proceeded differently. He used Theorem III to further elaborate what he considered to be his crucial insight into the relation between Mie's energy-momentum tensor and the variational derivative of the electromagnetic part of the Lagrangian. Consequently he now focussed on (36), from which he attempted to extract the equations for the electromagnetic field. The left-hand side of this equation can in fact (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 then yields:

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

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 :

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

its vanishing constitutes the electromagnetic field equations. It is at this point that Hilbert makes use of the last ingredient listed above, the special way in which the derivatives of the potentials enter Mie's Lagrangian. In fact, 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:

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

While the right-hand side of this equation only involves the electrodynamic part of the Lagrangian, this is not the case for the left-hand side in view of (73). Therefore Hilbert once more uses the field equations, in the form of (72), in order 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 has thus essentially reached 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:

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

its first factor being antisymmetric, the second symmetric, in *s*, *m*.)

Finally, using (80), Hilbert obtains:

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

Summarizing what he had achieved, Hilbert claimed:[80]

> .... aus den Gravitationsgleichungen (4) [i.e. (51)] folgen in der Tat die vier von einander unabhängigen linearen Kombinationen (32) [i.e. (83)] der elektrodynamischen Grundgleichungen (5) [i.e. (52)] 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.*

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

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 the other three field equations hold everywheree. 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, 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).

Coming now to the second case, it is clear that the field equations can only hold algebraically by virtue of (83) if the second term vanishes, which implies that the theory is gauge invariant, i.e. 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 staking his claim. First of all, he did not introduce the additional assumptions that are required – and could not have introduced them because they violated his physical assumptions;[81] and second, he did not derive the identity for gauge-invariant electromagnetic Lagrangians that makes this argument work. As is illustrated by Klein's later work,

80 Proofs, p. 12; in *Hilbert 1915*, p. 406, "ganze" is corrected to "genaue" in the last sentence.

the derivation of these identities is closely related to a different perspective on the results assembled by Hilbert, a perspective in which electromagnetism is no longer, as in Hilbert's proofs version, perceived as an epiphenomenon of gravitation, but in which they are treated in parallel.[82]

In summary, Hilbert's claim that the electromagnetic equations are a consequence of the gravitational field equations turns out to be an interpretation forced upon his mathematical results by his overall program rather than being implied by them. This interpretation is, in any case, different from that given to the corresponding results in general relativity that are usually associated with Hilbert's work.

The deductive structure of the proofs version

Having attempted to reconstruct the line of reasoning that Hilbert followed when developing the original version of his theory, we now take a closer look at the way in which he presented his results in the Proofs. This serves as a review of the deductive structure of his theory, indicates which results were emphasized by Hilbert himself, and will allow a comparison between the Proofs and the published versions.

We begin by recalling the elements of the deductive structure of Hilbert's theory that he explicitly introduced:

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

Furthermore he made use of the following assumptions, introduced as part of his deductive structure without being designated explicitly:

- vanishing divergence of the energy expression (compare (63))
- the split of the Lagrangian into gravitational and electrodynamical part (compare (16))
- an assumption about the character of the electrodynamical part: it does not depend not on the derivatives of the metric tensor (compare (25)).

---

81 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).

82 Compare Klein's attempt to derive analogous equations for the gravitational and the electromagnetic potentials, from which then the Maxwell equations are derived, *Klein 1917*, pp. 472-473.

Hilbert’s main physical results, not labelled as theorems, are:

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

The exposition of Hilbert’s theory in the Proofs can be subdivided into four sections. In the following, we give these sections short titles, listing under each heading the basic elements of Hilbert’s theory introduced above:

1. basic framework[83]

   Axioms I and II, Theorem I, and the field equations for gravitation and electromagnetism
2. causality and the energy expression[84]

   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 parts, the structure of the electrodynamical part
3. basic theorems[85]

   Theorems II and III
4. implications for electromagnetism[86]

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

The sequence in which Hilbert presented the elements of his theory suggests that he considered its implications for electromagnetism as his central results. Indeed, the gravitational field equations are only briefly considered at the beginning, as part of the general framework and never explicitly presented, whereas Hilbert’s presentation concludes with the three results concerning Mie’s theory. The centrality of these electromagnetic implications is also clear from his introductory and concluding remarks. In fact, Hilbert’s initial discussion not only opens with Mie’s electrodynamics but closes with the promise of further elaboration of these consequences for electrodynamics:[87]

> Die tiefgreifenden Gedanken und originellen Begriffsbildungen vermöge derer Mie seine Elektrodynamik aufbaut, und die gewaltigen Problemstellungen von Einstein

83 Proofs, pp. 1-3.

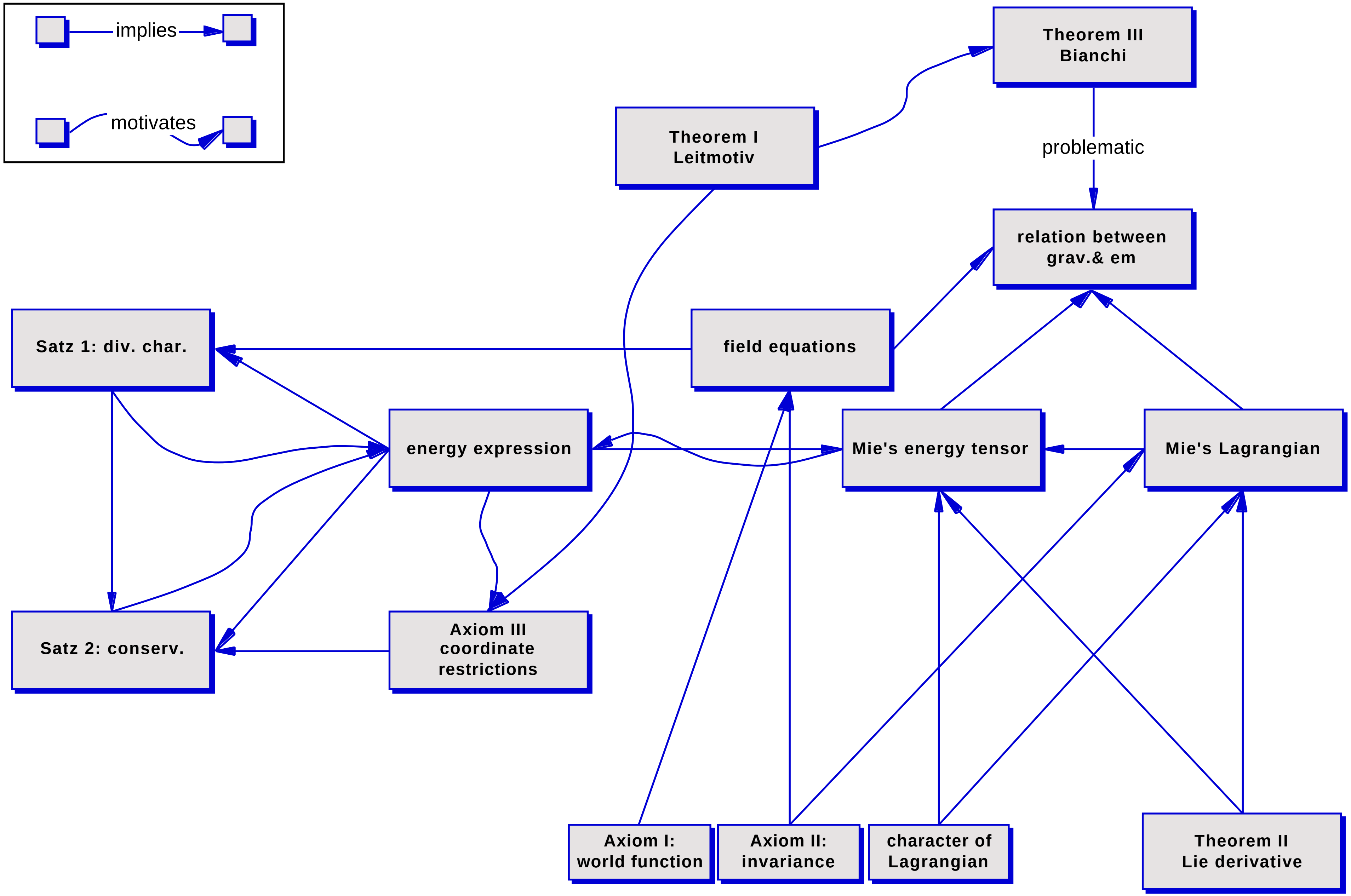
84 Proofs, pp. 3-8.

85 Proofs, pp. 8-9.

86 Proofs, pp. 9-13.

87 Proofs, p. 1.

Deductive structure of proofs version



> 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.

In his conclusion, Hilbert makes clear what his hopes were: to contribute to the solution of the riddles of atomic physics:[88]

> 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.

Hilbert's final remarks about the status of his theory vis à vis Einstein's work on gravitation strikingly parallels Minkowski's assessment of the relation of his four-dimensional formulation to Einstein's special relativity; that is, in both cases, 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.[89]

Diagram 1 gives a graphical overview of the deductive structure of Hilbert's theory, connecting the main elements listed above by arrows which either represent mathematical implications (straight arrows) or inferences based on heuristic reasoning (curved arrows). The figure shows that, 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 other results of Hilbert's theory, they depend on practically all the other parts of this theory. But 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.

88 Proofs, p. 13.

89 See *Walter 1998.*

The analysis of the deductive structure of Hilbert's theory thus confirms that this theorem is indeed the "Leitmotiv" of Hilbert's theory. The two principal clusters of results are obviously also related to what he considered to be 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 even mentioned by Hilbert. Einstein's imprint on Hilbert's theory had more a mathematical or structural than a physical character.

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

What Einstein could learn from Hilbert

The Hilbert-Einstein correspondence begins with a letter from Einstein dated 7 November 1915.[90] It was the beginning of a period, during which Einstein's theory of gravitation underwent several dramatic changes documented by four papers presented to the Prussian Academy and culminating in the definitive version of the field equations in a paper submitted on 25 November 1915.[91] 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;[92] he included the proofs of this note in his letter to Hilbert. In spite of the 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 particular, in both theories, the requirement of energy-momentum conservation is linked to a restriction to adapted coordinate systems. In Einstein's 4 November version, this restriction implies the following equation:[93]

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

As Einstein pointed out, one immediate consequence for the choice of an adapted coordinate system was obvious from this equation:[94]

> 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.

That the scalar [i.e. the trace] of the energy-momentum tensor cannot vanish is obvious if one considers Einstein's standard example of an incoherent swarm of particles ("dust") as

---

90 Einstein to David Hilbert, 7 November 1915, *Schulmann et al. 1998*, p. 191.

91 See *Einstein 1915e*.

92 *Einstein 1915a*.

93 *Einstein 1915a*, p. 785.

94 *Einstein 1915a*, p. 785.

the source of the gravitational field, for which the trace equals the mass density of the dust. Condition (87) itself, however, is rather curious since its physical meaning was entirely obscure. It was therefore incumbent upon Einstein to either modify his theory once more in order to get rid of this condition, or to find a physical interpretation for it. He soon succeded in doing both and formulated his new view in an addendum, published on 11 November.[95]

On 12 November 1915 he reported his success to Hilbert:[96]

> 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!

What had happened? Einstein had meanwhile noticed that the condition $\sum_{\sigma} T_{\sigma}^{\sigma} = 0$, implied by setting $\sqrt{-g} = 1$ in (87), can be related to an electromagnetic theory of matter since the vanishing of its trace is a characteristic property of the electromagnetic energy-momentum tensor in Maxwell's theory. Thus, if one assumes matter to be of electromagnetic origin, the vanishing of its trace becomes a fundamental property of the energy-momentum tensor, implying $\sqrt{-g} = 1$. This has two important consequences: Condition (87) is no longer an inexplicable restriction on the admissible coordinate systems, and the equations of 4 November can be seen as a peculiar 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 convient choice of a coordinate system.

The core of Einstein's new theory is strikingly simple. The left-hand side of the gravitational field equations is now given by the Ricci tensor and the right-hand side by an energy-momentum tensor whose trace has to vanish:[97]

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

What distinguishes these field equations from the final equations of general relativity 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:[98]

95 *Einstein 1915b.*

96 Einstein to David Hilbert, 12 November 1915, *Schulmann et al. 1998*, p. 194.

97 See *Einstein 1915b*, p. 801 and p. 800.

98 See *Einstein 1915e,* p. 845.

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

In the winter of 1912/13, Einstein had considered the linearized form of these field equations, but discarded it because he found they were not compatible with his expectation of how the Newtonian limit should result from the theory.[99] At that time, he had also considered field equations of the form (88), which he had also rejected, however, because they imply, when taken together with the requirement of energy-momentum conservation, the condition $\sum_\sigma T^\sigma_\sigma = 0$. At the time, this condition seemed to him unacceptable because (as pointed out above) 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 as 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 had earlier led him to reject this very theory? What impelled Einstein's change of perspective in November 1915?

The solution to this puzzle seems to lie in the scientific context within which Einstein formulated his new approach and, in particular, in his exchange with Hilbert. As will become evident, it would have been entirely uncharacteristic for him to adopt the new approach so readily had it not been for the current discussions of the electrodynamic world view and his feeling that he had entered into a competition with Hilbert.[100]

In his addendum, Einstein directly referred to the supporters of the electrodynamic world view:[101]

> 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.

Only the context of this discussion lends credibility to 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, in addition, makes it clear that this perspective was shaped by Einstein's rivalry with Hilbert. We have already cited Einstein's letter to Hilbert where he writes:[102]

---

99 See Doc. 10 of *Klein et al. 1995* and *Renn and Sauer 1998*.

100 For a discussion of Hilbert's reaction to what he must have seen as an intrusion by Einstein into his domain, see *Sauer 1999*, pp. 542-543.

101 *Einstein 1915b*, p. 799.

> 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!

And when, in a letter of 14 November to Einstein, Hilbert claimed to have achieved a unification of gravitation and electromagnetism, Einstein responded that this had been his goal, too:[103]

> Ihre Untersuchung interessiert mich gewaltig, zumal ich mir schon oft das Gehirn zermartert habe, um eine Brücke zwischen Gravitation und Elektromagnetik zu schlagen.

Somewhat later (after the calculation of the perihelion shift on the basis of the new theory), he expressed himself similarly to his friend Michele Besso:[104]

> 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.

When, however, one examines Einstein's earlier writings on gravitation, both published and unpublished, one finds no trace of attempts to unify gravitation and electromagnetism. He had never been an adherent of the electromagnetic world view. On the contrary, in 1913, he was apparently disinterested in if not hostile to Mie's attempt at a unification of gravitation and electrodynamics, not even finding it worth mentioning when reviewing contemporary gravitation theories.[105]

Soon afterwards, right after completion of the final version of general relativity, Einstein returned to his earlier view that general relativity could make no assertions about the structure of matter. He wrote to Arnold Sommerfeld:[106]

> 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's pursuit of a relation between gravitation and electromagnetism was, then, merely a short-lived episode in his search for a relativistic theory of gravitation that can be dated rather precisely to mid-November 1915. That the possibility of solving the problems of a theory of gravitation based on the Ricci tensor by a new model of matter was indeed a recent idea is confirmed by a footnote in the addendum:[107]

---

102 Einstein to David Hilbert, 12 November 1915, *Schulmann et al. 1998*, p. 194.

103 Einstein to David Hilbert, 15 November 1915, *Schulmann et al. 1998*, p. 199.

104 Einstein to Michele Besso, 17 November 1915 , *Schulmann et al. 1998*, p. 201.

105 See *Einstein 1913*.

106 Einstein to Arnold Sommerfeld, 9 December 1915, *Schulmann et al. 1998*, p. 216.

> Bei Niederschrift der früheren Mitteilung war mir die prinzipielle Zulässigkeit der Hypothese $\sum T_\mu^\mu = 0$ noch nicht zu Bewußtsein gekommen.

It thus seems quite likely 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.

With hindsight, Einstein's adoption of an electromagnetic theory of matter might appear as a bizarre and unnecessary detour. A closer analysis of the last steps of Einstein's path to general relativity shows, however, that these steps 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.[108] The condition $\sqrt{-g} = 1$, implied via (87) by the assumption of an electromagnetic origin of matter, turned out to be essential for this calculation. Einstein himself considered this success as a striking confirmation of his audacious hypothesis on the constitution of matter, which also definitely favored this theory over that of 4 November.[109] Furthermore, the 11 November theory turned out to be the basis for the development of a new understanding of the Newtonian limit, which finally made it possible for Einstein to accept the field equations of general relativity as the definitive solution to the problem of gravitation. Ironically, perhaps the most important contribution of Hilbert to general relativity consisted in enhancing the credibility of a speculative and ultimately untenable physical hypothesis, which nevertheless provided Einstein with a perspective that guided the final mathematical steps towards the formulation of his theory.

Einstein submitted his paper on the perihelion shift of Mercury on 18 November 1915. In a footnote appended after the paper's completion, Einstein observed that the hypothesis of an electromagnetic origin of matter is, in fact, unnecessary for the perihelion shift calculation, and announced a further modification of his field equations, this time yielding the definitive version.[110] On the same day, Einstein wrote a letter to Hilbert, acknowledging the receipt of Hilbert's work including a system of field equations:[111]

> 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.

---

107 *Einstein 1915b*, p. 800.

108 See *Einstein 1915c*.

109 See *Einstein 1915d*, the abstract of the paper on the perihelion shift, probably due to Einstein himself, in the summary of 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."

110 See *Einstein 1915c*, p. 831.

111 Einstein to David Hilbert, 18 November 1915, *Schulmann et al. 1998*, pp. 201-202.

Einstein emphasized, however, that the real difficulty lay not in postulating generally covariant gravitational field equations, but in showing their agreement with physical requirements such as the Newtonian limit; stressing his priority, he also mentioned that he had considered these equations three years ago:

> 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.

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

What Hilbert could learn from Einstein

When Hilbert received Einstein's letter of 12 November, announcing his insight into a fundamental role of gravitation for the constitution of matter, he must have seen the latter's intention to develop his theory of gravitation in the same direction that he, Hilbert, was pursuing, as a threat for his priority.[112] At any rate, Hilbert hastened to present his results publicly. In his response of 13 November, he gave a brief sketch of his theory and announced a seminar on it for the 16th of November:[113]

> 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 interessirt 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: $E = \sum(e_s t^s + e_{ih} t^{ih})$, [the t^s corresponds to p^s in Hilbert's papers, etc.] die ebenfalls eine allgemeine Invariante ist [compare (56)], 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

112 This aspect of the Hilbert-Einstein relationship was first discussed in *Sauer 1999*, where also the chronology of events is carefully reconstructed.

113 David Hilbert to Einstein, 13 November 1915, *Schulmann et al. 1998*, p. 195.

> gewöhnliche elektrische Energie herauskommt, wenn man eine gewisse absolute Invariante mit den Gravitationspotentialen differenziert und dann $g = 0, 1$ setzt.

Hilbert's letter presents the essential elements of his theory as we have reconstructed them from the Proofs. The way in which he refers to "the missing space-time equations" suggests that he assumed these equations and their relation to the energy concept to be a problem common to his theory and Einstein's.

Einstein responded in a letter of 15 November 1915, declining the invitation to come to Göttingen for reasons of health.[114] Instead, he asked Hilbert to send him proofs of his paper. As mentioned above, by 18 November Hilbert had fulfilled Einstein's request, although he could not have send his first Proofs, which are dated 6 December. So he must have sent a manuscript, probably corresponding to the talk he gave on 20 November. Since the proofs version of Hilbert's paper is also dated 20 November, this manuscript may well have presented practically the same version of his theory as the proofs version. On 19 November, a day after Einstein announced his successful calculation of the perihelion shift of Mercury to Hilbert, the latter congratulated Einstein, making it clear once more that the physical problems he intended to solve were of a rather different nature, concerned with the microcosm:[115]

> 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.

There can be little doubt that Einstein fulfilled Hilbert's request to keep him informed. His definitive paper on the field equations, submitted on 25 November and published on 2 December, must have been on Hilbert's desk within a day or two. In this paper Einstein showed that, in contrast to all earlier versions of his theory, energy-momentum conservation does not imply additional coordinate restrictions on the field equations (89). He also made clear that, contrary to his earlier belief, these field equations do fulfill the requirement of having a Newtonian limit. Finally, Einstein left no doubt that these equations 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, at least as Hilbert perceived this scope at the end of 1915. 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 con-

---

114 Einstein to David Hilbert , 15 November 1915, *Schulmann et al. 1998*, p. 199.

115 David Hilbert to Einstein, 19 November 1915, *Schulmann et al. 1998*, p. 202.

servation was structured by a logic paralleling that of Einstein's non-covariant theory of gravitation. Moreover, Hilbert's "Leitmotiv," Theorem I, was motivated by Einstein's hole argument that generally-covariant field equations cannot have unique solutions. In his definitive paper of 25 November, Einstein did not explicitly mention the hole argument, but simply took it for granted that his new generally-covariant field equations do not suffer from the difficulties he had earlier envisioned.[116] Hilbert could check that Einstein's definitive field equations were actually compatible with the equations that follow from Hilbert's variational principle, which he had not explicitly calculated – at least not in the Proofs. While the formal equivalence of Einstein's and Hilbert's field equations was certainly reassuring for Hilbert, the fact that the hole argument evidently no longer troubled Einstein must have let Hilbert to question his "Leitmotiv," with its double role of motivating coordinate restrictions and providing the link between gravitation and electromagnetism.

It thus seems clear that Einstein's paper of 25 November 1915 represented a major challenge to Hilbert's theory. As we shall discuss in more detail below, the results of Einstein's paper, in particular the new field equations, their compatibility with the requirements of the Newtonian limit and energy-momentum conservation, and the support of these field equations by astronomical evidence, posed quite different problems for Hilbert's theory. No doubt, the explicit form that Einstein gave for the field equations was by far the easiest result to assimilate to Hilbert's theory. As we shall see when discussing the published version of Hilbert's paper, it turns out that, while Einstein temporarily took over Hilbert's physical perspective, Hilbert appears to have taken this crucial mathematical result – the explicit form of the field equations – from Einstein's theory.

Cooperation in the form of competition

In a situation such as we have described, in which the interactions between several people working on the same problem change the way in which each of them proceeds, it is not easy for the individuals themselves 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 the new theory of gravitation, he also resented the way in which Hilbert took over some of his results without, as Einstein saw it, giving him due credit. Thus Einstein wrote to his friend Heinrich Zangger on 26 November 1915 with regard to the theory he had just completed:[117]

> Die Theorie ist von unvergleichlicher Schönheit. Aber nur *ein* Kollege hat sie wirklich verstanden und der eine sucht sie auf geschickte Weise zu "nostrifizieren" (Abra-

116 The fact that these equations are supported by Einstein's calculation of the perihelion shift made it impossible for Hilbert simply to discard them as unsound because of the hole argument.

117 Einstein to Heinrich Zangger, 26 November 1915, *Schulmann et al. 1998*, p. 205. See the discussion of the term "nostrification" above.

> ham'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'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 the service of Einstein, who supplied the physical heuristics.[118] What Hilbert offered was not cooperation but competition. Hilbert may well have been upset that Einstein anticipated in print, in his paper of 11 November, what Hilbert felt to be his idea of a close link between gravitation and the structure of matter. Ultimately even more disturbing to Hilbert may have been the fact that, contrary to what Hilbert had asserted in the Proofs, Einstein's final formulation of his theory required no restriction on general covariance, not even from conservation of energy. But it is not clear exactly when this insight led him to abandon all non-covariant elements of his program, in particular his entire approach to the energy problem and the consequent restriction to a preferred class of coordinate systems.[119]

Hilbert evidently came know of Einstein's resentments 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, in reaction to Einstein's complaints, he began to introduce changes into his Proofs of 6 December that are documented by handwritten marginalia in the Proofs. These changes testify not only to Hilbert's acknowledgement of Einstein's priority but also his attempt to placate Einstein. At the same time, Hilbert's notes provide an indication of what the content of Einstein's complaint may have been. He began by revising the programmatic statement in the introduction of his paper (his insertion is rendered here in italics):[120]

> 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.

The insertion "wesentlich" was probably motivated by Hilbert's recognition that his theory actually presupposed more assumptions of substantial content than those in his three axioms. As our analysis of the deductive structure of his theory indicates, the assumption of a split of the Lagrangian into gravitational and electromagnetic parts, as well as the assumption that the latter does not depend on derivatives of the metric are among these additional presuppositions. A further assumption, also not represented by an axiom, was

118 See the editorial note "Einstein on Gravitation and Relativity: The Collaboration with Marcel Grossmann" in *Klein et al. 1995*, pp. 294-301.

119 According to *Sauer 1999*, p. 562, Hilbert had found the new energy expression by 25 January 1916.

120 Proofs, p. 1.

the requirement that the gravitational part of the Lagrangian involve no derivatives of the metric higher than second order. Einstein had justified this requirement by the necessity for the new theory of gravitation to have a Newtonian limit, and it may well 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 indicated. Remarkably, in characterizing his system of equations, Hilbert deleted the word “neu,” a clear indication that he had meanwhile seen Einstein’s paper and recognized that the equations implied by his own variational principle are formally equivalent to those which Einstein had explicitly written down (because of where the trace term occurs), if Hilbert’s stress-energy tensor is substituted for the unspecified one on the right-hand side of Einstein’s field equations.

The next change Hilbert introduced was probably 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 had been confined to asking the right questions, while Hilbert’s merit was to provide the answers.

In the new version, Hilbert’s description of these gravitational potentials reads (his insertion is again rendered here in italics):[121]

> Die das Geschehen in w_s charakterisierenden Größen seien:
>
> 1) die zehn *von Einstein 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.

The next change that Hilbert introduced represents an even more far-going recogition that he could not simply claim the results he presented as parts of “his theory,” as if this theory had nothing substantial in common with that of Einstein. The change again occurs in connection with a programmatic statement of Hilbert’s:[122]

> Das Leitmotiv für den Aufbau ~~meiner~~ *der* Theorie liefert der folgende mathematische Satz, dessen Beweis ich an einer anderen Stelle darlegen werde.

The final marginal note made by Hilbert just consists in an exclamation mark next to a minor correction of the energy expression (39), and is perhaps evidence that he had identified this expression as the central problem in the Proofs of his paper. While Hilbert’s first corrections were probably still intended as revisions of a text that was going to remain basically unchanged, this exclamation mark signals the abandonment of such

121 Proofs, p. 1.

122 Proofs, p. 2.

attempted revisions. Perhaps at this point 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 a reaction to Einstein's results and not a consequence of an internal dynamics of the development of his theory.[123] Indeed, as our analysis of the deductive structure of Hilbert's theory shows, his assertions concerning energy are 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 the 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 must have left Hilbert at a loss, as we have argued above. But precisely the way in which energy-momentum conservation was connected to other results of his theory also offered a guide towards its modification in the direction indicated by Einstein. In particular, Hilbert had to find a new energy expression that does not imply a restriction on the choice of coordinates but is still connected with Mie's energy-momentum tensor. Precisely the fact that, as we have seen, his energy expression was decoupled from the physical consequences of Hilbert's theory made such a modification conceivable. For the time being, in any case, Hilbert renounced publication of his results and, at some point, began to rework his theory. Only by early 1916 had he arrived at results that made it possible to rewrite his paper and submit it for publication; by mid-February 1916, Hilbert's first communication, which we will discuss in the following section, was in press.[124]

Meanwhile, having emerged victorious from the exchange of November 1915, Einstein offered a reconciliation to Hilbert:[125]

> 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.

---

123 For a different view, see *Sauer 1999*, p. 570.

124 For a detailed discussion of the chronology, see the reconstruction in *Sauer 1999*, pp. 560-565.

125 Einstein to David Hilbert, 20 December 1915, *Schulmann et al. 1998*, p. 222. The "schäbige[.] Welt" probably refers to World War I – on the background of Einstein's and Hilbert's critical attitude to the war.

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

The new energy concept – an intermediary solution

As we have seen above, a modification of Hilbert's treatment of energy-momentum conservation was the most urgent step necessitated by the confrontation of his theory with Einstein's results of 25 November 1915. These results, together with Hilbert's own program, indicated the way in which such a modification had to be carried out. First of all, the energy-momentum conservation law should no longer give rise to coordinate restrictions but should itself be an invariant equation. Second, it should still be possible to recover Mie's energy-momentum tensor from the modified energy expression; otherwise the link between gravitation and electromagnetism, fundamental to Hilbert's program, would have been endangered. Third, in accordance with Hilbert's general understanding of energy-momentum conservation in physics, the new energy concept should still satisfy a divergence equation. In this section, we shall show that Hilbert's modification of his energy expression was guided by these criteria, while the relation of this energy expression to a physical understanding of energy-momentum conservation remained as tenuous as ever.[126] In the following section, we shall study the effect of the new energy concept on the deductive structure of Hilbert's theory.

The statement introducing his discussion of the energy concept in the published version of his paper emphasizes that now only axioms I and II are required for its construction[127]:

> 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.

This emphasis only makes sense when contrasted with the treatment in the proofs version in which the energy concept is closely related to axiom III (dropped in the published version). Hilbert then proceeds exactly as in the Proofs, introducing a polarization of the Lagrangian with respect to only the gravitational variables (compare 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 distinction from (37), however, Hilbert considered $\sqrt{g}H$ instead of H. Clearly, his aim was to form an equation analogous to (45), but now with only a divergence term on the right-hand side. Indeed, in view of the relation:

---

126 For a discussion of Hilbert's concept of energy, see also *Sauer 1999*, pp. 548-550, which stresses the mathematical roots of this concept but leaves open the question of its relation to physics.

127 *Hilbert 1915*, p. 400.

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

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

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

This equation is of just the required form since the right-hand side vanishes due to the field equations.

The way in which Hilbert obtains the last equation closely parallels that used in the proofs version, i.e. by splitting off divergence terms. He starts out by noting that:

$$a^l = \sum_{\mu,\nu,k}\frac{\partial H}{\partial g^{\mu\nu}_{kl}}A^{\mu\nu}_k \tag{93}$$

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

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:

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

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

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

obtaining the desired result, (92).

He next forms the expression analogous to (92) for the electromagnetic variables (compare 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 into account 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 last step consists in rewriting also the left-hand side of this equation as a divergence. For this purpose Hilbert made use of (91), which he expands as:

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

and of Theorem II (compare (22)).[128] In this way he 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 furthermore, 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 the final result of the energy calculation, since it gives an equation that does indeed satisfy two of the three criteria mentioned above, being an invariant divergence. But it fails the criterion that it yield a relation to Mie's energy-momentum tensor. In order to establish this relation, Hilbert adds yet another term d^l to the expression in the parenthesis in (102), where:

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

It is a contravariant vector (because:

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

is an antisymmetric tensor) and satisfies the identity:

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

Hilbert concluded:[129]

> Definieren wir nunmehr
>
> $$e^l = Hp^l - a^l - b^l - c^l - d^l \tag{106}$$

128 In the published version, it is only for this purpose that this form of Theorem II is explicitly introduced. It is, however, likely that (23) already had been derived from this form.

129 *Hilbert 1915*, p. 402.

> 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*
>
> $$\sum_l \frac{\partial \sqrt{g} e^l}{\partial w_l} = 0 \tag{107}$$
>
> erfüllt.

While Hilbert does not explicitly introduce the condition that his energy vector be related to Mie's energy-momentum tensor, that is clearly both the guiding principle and the outcome of his calculation. Apparently, he wanted to present this connection as the result of an independently-justified definition of energy.

In effect, starting from (106) and taking into account the definitions (98) and (103), Hilbert obtains for the electromagnetic contribution to the energy, i.e. that part originating from the electromagnetic part of the Lagrangian L:

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

He now uses the field equations and (27) to show that this expression can be rewritten as:

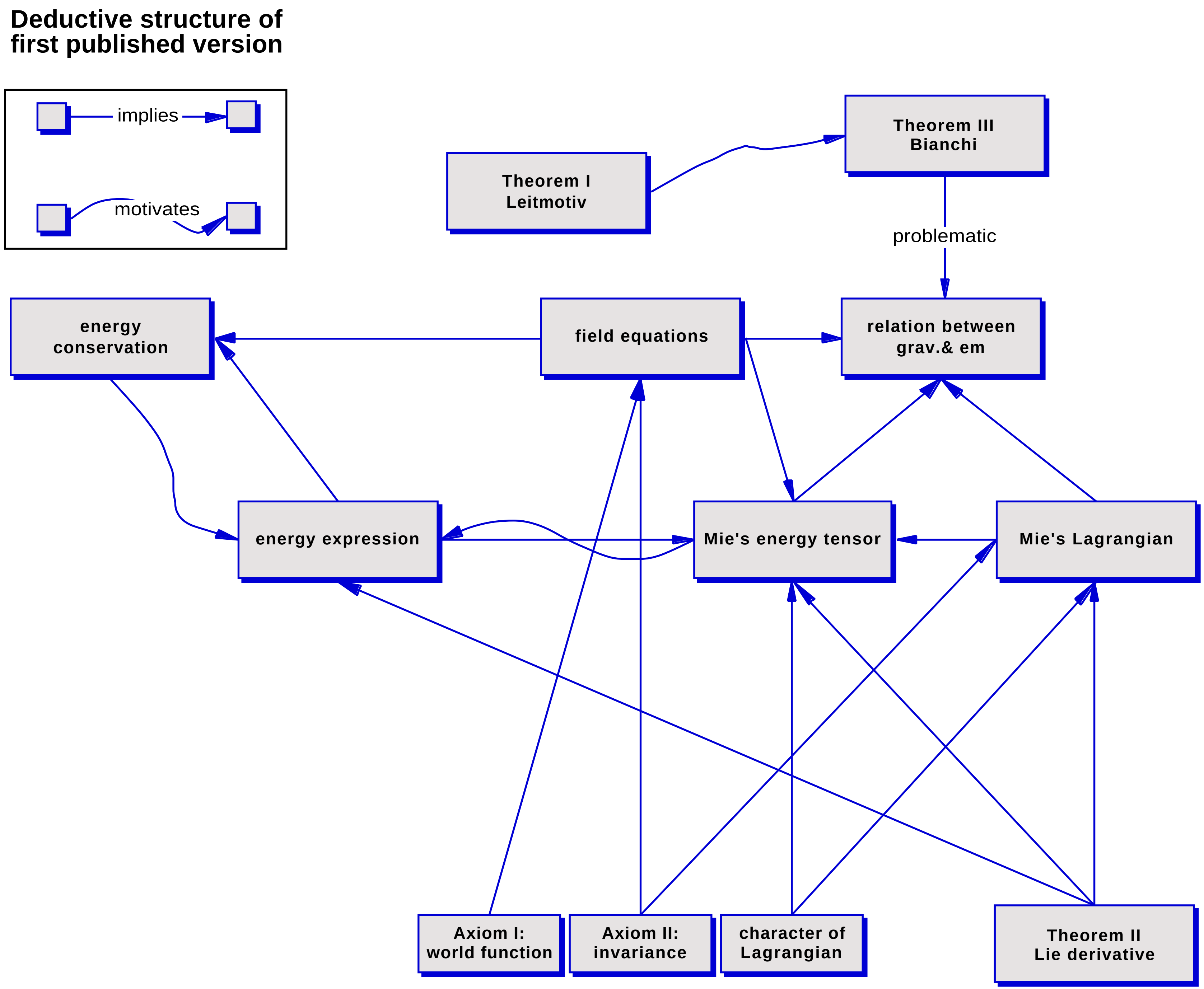
$$\sum_{s,k} \left(L \delta^l_s - \frac{\partial L}{\partial M_{lk}} M_{sk} - \frac{\partial L}{\partial q_l} q_s \right) p^s, \tag{109}$$

which, except for p super s, corresponds to the right-hand side of (36), that is to the generally covariant generalization of Mie's electromagnetic energy-momentum tensor.

In contradistinction to the Proofs, Theorem II and (36) no longer explicitly enter the demonstration of the relation between the energy concept and Mie's energy tensor (Theorem II only enters implicitly by determining the form in which the electromagnetic variables enter the Lagrangian, compare (27)). Hilbert still needed Theorem II in deriving his "first result," that is, in showing that this energy-momentum tensor is equal to:

$$-\frac{2}{\sqrt{g}} \sum_{\mu,s} \frac{\partial \sqrt{g} L}{\partial g^{\mu s}} g^{\mu l}, \tag{110}$$

i.e., 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 in the form (72), the electromagnetic energy and also the 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 but not on the electromagnetic potentials and their derivatives. This consequence now follows only with the help of Theorem II, whereas, in the Proofs, it had been an immediate consequence of the definition of the energy and of the field equations (compare (49)).


Deductive structure of first published version
implies
motivates
Theorem I
Leitmotiv
Theorem III
Bianchi
problematic
energy conservation
field equations
relation between grav.& em
energy expression
Mie's energy tensor
Mie's Lagrangian
Axiom I: world function
Axiom II: invariance
character of Lagrangian
Theorem II
Lie derivative

While Hilbert had succeeded in complying with his heuristic criteria and also with the new challenge of deriving an invariant energy equation, the status of this equation within the deductive structure of his theory had become more precarious. An analysis of the deductive structure of Hilbert's theory as presented in Paper 1 (see the diagram) shows that this theory still comprises two main clusters of results, those concerning the implications of gravitation for electromagnetism and those concerning energy conservation. But the latter cluster had now become even more isolated from the rest of his theory than before. 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 by the link with Mie's energy-momentum tensor. The new energy concept plays no role at all 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 establishment in the 19th century. Therefore, it should not come as a surprise to learn that Hilbert's new energy concept turned out to play only a transitional role in his theory, and was eventually replaced by the understanding of energy-momentum conservation developed in the work of Einstein, Klein, Noether, and others.[130]

In fact, neither the physical meaning nor the mathematical status of Hilbert's new energy concept was entirely clear. From a physical point of view, Hilbert 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 equations in a gravitational field that Einstein had established in 1913 (compare (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.[131] Concerning the mathematical status of (107), in 1917 Emmy Noether and Felix Klein found that Hilbert's equation actually is an identity, and not just a consequence of the field equations, as is the case for conservation equations in classical physics.[132] As a matter of fact, similar identities follow for the Lagrangian of any generally covariant variational problem. As a consequence, Hilbert's counting of equations no longer works, since 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. So Einstein's abandonment of these coordinate restrictions, and the further exploration of energy-momentum conservation by Noether, Klein, Einstein, and others, confronted Hilbert's approach with a severe challenge: They put into question

---

130 For discussion, see *Rowe (Forthcoming).*

131 See *Klein 1918a*, pp. 179-185.

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

the organization of his theory into two more-or-less independent domains, dealing, respectively, with energy-momentum conservation and the implications of gravitation for electromagnetism. Below we shall argue that Hilbert responded to this challenge by further adapting his theory to the framework of general relativity.

Hilbert's reorganization of his theory in Paper 1

The challenge presented by Einstein's adoption of generally covariant field equations and consequent abandonment of coordinate restrictions made it necessary for Hilbert to reorganize his theory. In particular, he had to demonstrate the compatibility between his variational principle and Einstein's field equations (which had received striking confirmation by Einstein's derivation of Mercury's perihelion shift from these equations), and he had to completely rework his treatment of energy conservation. Hilbert shifted both issues to the end of Paper 1. Energy conservation was no longer tied to Theorem I and its heuristic consequences, as had been the case in the Proofs, but was treated among the results of Hilbert's theory. The structure of Paper 1 is thus:[133]

1. basic setting[134]

   Axioms I and II, Theorem I, and the combined field equations of gravitation and electromagnetism for an arbitrary Lagrangian
2. basic theorems[135]

   Theorems II and III
3. the new energy expression and derivation of the new energy equation[136]
4. implications for the relation between electromagnetism and gravitation[137]

   the split of the Lagrangian, the character of the gravitational and the electrodynamical parts of the Lagrangian, the form of Mie's Lagrangian, the relation between Mie's energy tensor and Mie's Lagrangian, the explicit form of the gravitational field equations, and the relation between electromagnetic and gravitational equations.

Apart from the technical and structural revisions induced by the new energy expression, practically all other changes in Hilbert's paper concern the relation of his theory to that of Einstein. Throughout Paper 1, Hilbert followed a tendency already visible in the marginal additions to the Proofs, that is, greater emphasis on Einstein's contribution while maintaining his claim to have developed an independent approach. In the opening paragraph of his paper, Hilbert changed the order in which he mentioned Mie and Einstein. In the Proofs he wrote:[138]

---

133 For a sketch of Hilbert's revisions of his first paper, see also *Corry 1999a*, pp. 517-522.

134 *Hilbert 1915*, pp. 395-398.

135 *Hilbert 1915*, pp. 398-400.

136 *Hilbert 1915*, pp. 400-402.

137 *Hilbert 1915*, pp. 402-407.

138 Proofs, p. 1.

> 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.

In Paper 1 we read instead:[139]

> 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.

He appended a footnote listing all of Einstein's publications on general relativity starting with his major review paper of 1914, and including the definitive paper submitted on 25 November. Although Hilbert thus makes clear that he must have revised his own 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[140]), which remains "Vorgelegt in der Sitzung vom 20. November 1915." Inevitably this created the false impression that he had not introduced any substantial changes in the Proofs.

The next sentence makes it clear that Hilbert had not renounced his claim to having solved the problems posed by Mie and Einstein, combining this claim with a more explicit recognition of what he considered to be the achievements of his predecessors. In the corrected Proofs this sentence reads:[141]

> 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.

In Paper 1, Hilbert reiterated the claim that he had introduced a "new" system of equations but now mentioned Einstein and Mie once more:[142]

> 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.

Similarly, although he had changed the reference to "his theory" in a marginal note in the proofs version to "the theory," he now returned to the original claim:[143]

---

139 *Hilbert 1915*, p. 395.

140 See *Klein 1918a*; *Noether 1918*.

141 Proofs, p. 1.

142 *Hilbert 1915*, p. 395.

143 *Hilbert 1915*, p. 396.

> Das Leitmotiv für den Aufbau meiner Theorie liefert der folgende mathematische Satz, dessen Beweis ich an einer anderen Stelle darlegen werde.

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

The not-insignificant result that Hilbert's variational principle gives rise to gravitational field equations formally equivalent to those of Einstein's theory is rather hidden in Hilbert's presentation – it only appears as an intermediate result of his demonstration that the electromagnetic field equations are a consequence of the gravitational equations. The newly-introduced passage reads:[144]

> Unter Verwendung der vorhin eingeführten Bezeichungsweise für die Variationsableitungen bezüglich der $g^{\mu\nu}$ erhalten die Gravitationsgleichungen wegen (20) [i.e. (16)] die Gestalt
>
> $$[\sqrt{g}K]_{\mu\nu} + \frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}} = 0. \tag{111}$$
>
> Das erste Glied linker Hand wird
>
> $$[\sqrt{g}K]_{\mu\nu} = \sqrt{g}\left(K_{\mu\nu} - \frac{1}{2}Kg_{\mu\nu}\right), \tag{112}$$
>
> 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's attempt to avoid calculation of $[\sqrt{g}K]_{\mu\nu}$ is untenable, because there are many other tensors of second rank and many other invariants that can be constructed from the Riemann tensor. Even if one requires the tensors and invariants to be linear in the Riemann tensor, the crucial coefficient of the trace term remains undetermined by such an argument. (In fact, Hilbert later withdrew it – see below.) The explicit form of the field equations given by Hilbert in Paper 1 – not given in the proofs version – appears to be a response to Einstein's prior, decisive publication of 25 November. But in a footnote appended to the above passage, Hilbert only gives a generic reference to all four of Einstein's 1915 Academy publications. His cautious formulation concerning the apparent

144 *Hilbert 1915*, pp. 404-405.

agreement between his and Einstein's results was probably motivated by their belonging to different frameworks, but it added to the impression that Hilbert had actually arrived independently at the explicit form of the gravitational field equations.

In the concluding paragraph of Paper 1, Hilbert acknowledges his debt to Einstein in a more indirect way. At the beginning of this paragraph of the Proofs, he had given the impression that Einstein posed the problems while he, Hilbert, offered the solutions:[145]

> 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 ...

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:"[146]

> 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 ...

Einstein's energy in Hilbert's 1924 theory

In 1924 Hilbert published another revised version, comprising Papers 1 and 2.[147] Meanwhile important developments had taken place in physics, such as the rapid progress of quantum physics, which also changed the scientific context of Hilbert's results. But it was undoubtedly the further clarification of the meaning of energy-momentum conservation in general relativity, already mentioned in the preceding sections, that affected his theory most directly. In fact, there was a correspondence between Hilbert and Klein (published in part in 1918[148]), in which 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 the effect of this development on Hilbert's 1924 revision of his theory. In spite of the reassertion of his programmatic goal of providing foundations for all of physics, this theory now was, in effect, transformed into a variation on the themes of general relativity.

On a purely technical level, Hilbert's revisions of his first paper appear to be rather modest; the most important one concerns Theorem III (the contracted Bianchi identities). Fol-

145 Proofs, p. 13.

146 *Hilbert 1915*, p. 407.

147 *Hilbert 1924*. In the following, we will refer to the 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).

148 See *Klein 1917*.

lowing a suggestion by Klein,[149] Hilbert extended this theorem, now labelled Theorem 2, to include the electromagnetic variables:[150]

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 $\sqrt{g}J$ bezüglich $g^{\mu\nu}$ mit $[\sqrt{g}J]_{\mu\nu}$, bez. q_μ mit $[\sqrt{g}J]_\mu$ bezeichnet werden, und wenn ferner zur Abkürzung:

$$i_s = \sum_{\mu,\nu}[\sqrt{g}J]_{\mu\nu}g_s^{\mu\nu} + [\sqrt{g}J]_\mu q_\mu$$
$$i_s^l = -2\sum_\mu[\sqrt{g}J]_{\mu s}g^{\mu l} + [\sqrt{g}J]_l q_s, \tag{113}$$

gesetzt wird, so gelten die Identitäten

$$i_s = \sum_l \frac{\partial i_s^l}{\partial w_l} \qquad (s = 1, 2, 3, 4). \tag{114}$$

He revised the proof of this theorem accordingly.

A second, apparently trivial, but actually significant change concerns the gravitational field equations. Hilbert now treated them more carefully, tacitly withdrawing his claim that no derivation was needed by sketching a derivation of their explicit form and writing them, like Einstein, with the energy-momentum tensor as source. As in his earlier versions, he derived (72) but now in the form:[151]

$$[\sqrt{g}K]_{\mu\nu} = -\frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}}. \tag{115}$$

After writing down the electromagnetic field equations, Hilbert proceded to sketch the following derivation of (115):[152]

Um den Ausdruck $[\sqrt{g}K]_{\mu\nu}$ 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:

$$[\sqrt{g}K]_{\mu\nu} = \sqrt{g}\left(K_{\mu\nu} - \frac{1}{2}Kg_{\mu\nu}\right). \tag{116}$$

Führen wir noch für den Tensor

$$-\frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}} \tag{117}$$

die Bezeichnung $T_{\mu\nu}$ ein, so lauten die Gravitationsgleichungen

---

149 *Klein 1917*, pp. 471-472.

150 *Hilbert 1924*, p. 5.

151 See *Hilbert 1924*, p. 7.

152 See *Hilbert 1924*, pp. 7-8.

$$K_{\mu\nu} - \frac{1}{2}Kg_{\mu\nu} = T_{\mu\nu}. \tag{118}$$

Although the introduction of Einstein's notation for the energy-momentum tensor may only appear as an adaptation of Hilbert's notation to the by-then standard usage, it actually effected a major revision of the deductive structure of his theory. The energy-momentum tensor now became the central knot binding together the physical implications of Hilbert's theory.

First of all, it served, as Hilbert's earlier energy expressions had 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 Hilbert's energy-expression. The new energy expression that Hilbert now took over from Einstein was, on the other hand, supported by much more than such an isolated result. It had emerged from the development of special-relativistic continuum physics by Herglotz, Laue, Nordström, and others, and been validated by numerous applications to various areas of physics, including, in particular, general relativity.

By introducing the equation:

$$T_{\mu\nu} = -\frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}}, \tag{119}$$

Hilbert returned, in a sense, to the approach of the Proofs, establishing a relation between the energy concept and the derivative of the electromagnetic Lagrangian (compare (49), but again without making clear that this relation does not single out Mie's theory but actually holds more generally. Introducing the notations:

$$\frac{\partial L}{\partial q_{sk}} = \frac{\partial L}{\partial M_{ks}} = H^{ks}, \tag{120}$$

and:

$$\frac{\partial L}{\partial q_k} = r^k, \tag{121}$$

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

$$-\frac{2}{\sqrt{g}}\sum_{\mu}\frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}}g^{\mu m} = L\delta_{\nu}^{m} - \sum_{s} H^{ms}M_{\nu s} - r^{m}q_{\nu}, \tag{122}$$

(compare (36)). On the basis of this equation, Hilbert again 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:[153]

> Demnach ergibt sich für $T_{\mu\nu}$ die Darstellung:

153 *Hilbert 1924*, p. 9.

$$T_{\mu\nu} = \sum_{\mu} g_{\mu m} T^{m}{}_{\nu}$$

$$T^{m}{}_{\nu} = \frac{1}{2}\left\{ L\delta^{m}_{\nu} - \sum_{s} H^{ms} M_{\nu s} - r^{m} q_{\nu} \right\}. \qquad (123)$$

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.

While Hilbert's claim thus remains unchanged, in actuality what he had done was to specialize the source term in Einstein's field equations. The nature of this source term can be expressed on the level of the Lagrangian or on the level of the energy-momentum tensor, and these two ways of expression are obviously equivalent to each other – but this relation is in no way peculiar Mie's theory. The fact that the energy expression introduced in Paper 1 was specifically chosen so as to give rise to Mie's energy-momentum tensor had obscured this circumstance, which was now made rather obvious by the introduction of Einstein's energy-momentum tensor. But it was nevertheless difficult for Hilbert to draw this consequence because it contradicted his program, according to which electromagnetism should arise as an effect of gravitation.

The situation is 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 the recognition of the close relation between the contracted Bianchi identities and energy-momentum conservation in general relativity, it had become unavoidable for Hilbert to reconsider the link he believed he had established between the two groups of field equations. In this way, energy-momentum conservation also played an ever-more central role in Hilbert's approach, turning the link between gravitation and electromagnetism into a mere by-product, obtained not because of any deep intrinsic connection between these two areas of physics but merely due to the introduction of additional variables representing the electromagnetic potentials into the variational principle. With the same logic one could argue that any form of matter giving rise to a stress-energy tensor deriveable from a Lagrangian arises as an effect of gravitation.

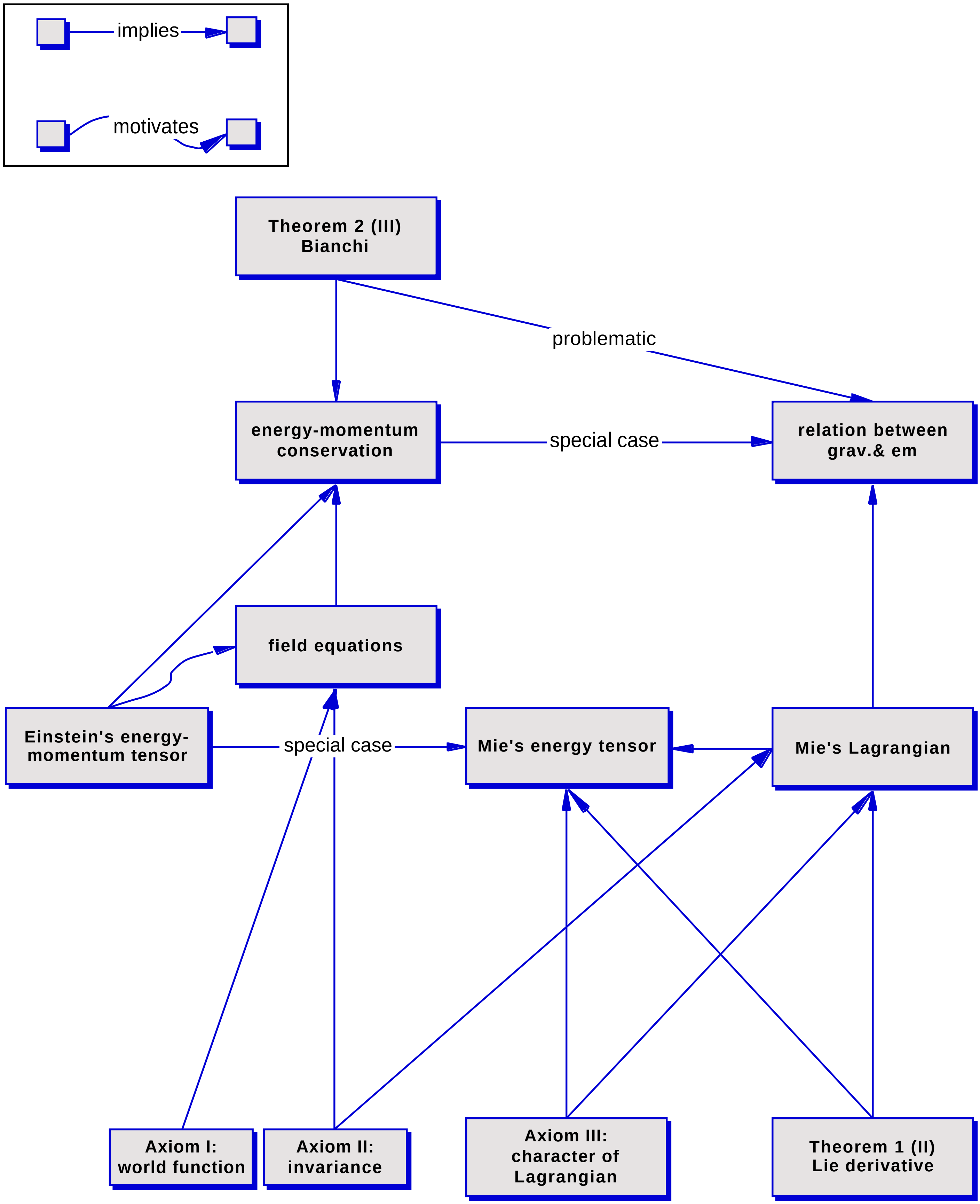
This is reflected in the new way that Hilbert used to obtained the desired link between gravitation and electromagnetism. Following Klein's suggestion, in Part 1 Hilbert treated the contracted Bianchi identities in parallel for both the gravitational and the electromagnetic variables:[154]

Die Anwendung des Theorems 2 auf die Invariante *K* liefert:

---

154 *Hilbert 1924*, pp. 10-11.

Deductive structure of second published version



$$\sum_{\mu\nu}[\sqrt{g}K]_{\mu\nu}g_s^{\mu\nu} + 2\sum_m \frac{\partial}{\partial x_m}\left(\sum_\mu [\sqrt{g}K]_{\mu s}g^{\mu m}\right) = 0. \tag{124}$$

Die Anwendung auf L ergibt

$$\begin{aligned}&\sum_{\mu\nu}(-\sqrt{g}T_{\mu\nu})g_s^{\mu\nu} + 2\sum_m \frac{\partial}{\partial x_m}(-\sqrt{g}T_s^{\ m}) \\ &+\sum_\mu [\sqrt{g}L]_\mu q_{\mu s} - \sum_\mu \frac{\partial}{\partial x_\mu}\left(\sum_\mu [\sqrt{g}L]_\mu q_s\right) = 0 \qquad (s = 1, 2, 3, 4).\end{aligned} \tag{125}$$

Orginally, he had only derived the first set of identities and made use of them in order to derive (83). Now Hilbert showed that both identities yield the equation for energy-momentum conservation that had been central to Einstein’s work since 1912. Hilbert also made it clear, following the work of Einstein and others, that this equation is related to the equation of motion for the source, and that it represents a generalization of energy-momentum conservation in special relativity:[155]

Als Folge der elektrodynamischen Grundgleichungen erhalten wir hieraus:

$$\sum_{\mu\nu}\sqrt{g}T_{\mu\nu}g_s^{\mu\nu} + 2\sum_m \frac{\partial}{\partial x_m}\sqrt{g}T_s^{\ m} = 0 \tag{126}$$

Diese Gleichungen ergeben sich auch als Folge der Gravitationsgleichungen, auf Grund von (15a) [i.e. (124)]. 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

$$\sum \frac{\partial T_s^{\ m}}{\partial x_m} = 0 \tag{127}$$

welche die Erhaltung von Energie und Impuls ausdrücken.

Hilbert had thus anchored his theory in the same physical foundation that had provided Einstein’s work general relativity with a stable point of reference. Only after having secured this foundation, did Hilbert turn to his original goal, the link between gravitation and electromagnetism, whose problematic character we have discussed above (compare the comments on (83)):[156]

Aus den Gleichungen (16) [i.e. (126)] folgt auf Grund der Identitäten (15b) [i.e. (125)]:

$$\sum_\mu [\sqrt{g}L]_\mu q_{\mu s} - \sum_\mu \frac{\partial}{\partial x_\mu}([\sqrt{g}L]_\mu q_s) = 0 \tag{128}$$

oder

155 *Hilbert 1924*, p. 11.

156 *Hilbert 1924*, p. 11.

$$\sum_{\mu} \left\{ M_{\mu s}[\sqrt{g}L]_{\mu} + q_s \frac{\partial}{\partial x_{\mu}}[\sqrt{g}L]_{\mu} \right\} = 0, \qquad (129)$$

> 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.

A glance at the deductive structure of Part 1 (see diagram) shows both the fundamental changes with respect to Paper 1 and the central role of Einstein's energy-momentum tensor in this reorganization. In fact, this energy-momentum tensor suggests the particular form in which Hilbert rewrote the gravitational field equations, it establishes the link between gravitation and electromagnetism (in terms of the choice of a specific source), and, of course, it is fundamental to Hilbert's new formulation of energy-momentum conservation.

The deductive structure of his revised theory now has a kernel, consisting of his variational principle, field equations, and energy-momentum conservation, that is – not only from a purely formal perspective, but also from a physical one – fully equivalent to Einstein's formulation of general relativity. Clearly, Hilbert's deductive presentation gives a greater emphasis to the variational principle than Einstein's formulation, and the mathematically more elegant form in which his variational principle is formulated, using the Ricci scalar, contributes to this emphasis. Therefore, this formulation of general relativity is, in the end, rightly associated with Hilbert's name. But, on the other hand, Hilbert's original programmatic 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). This claim is supported by our analysis of the deductive structure of Hilbert's revised theory. The intrinsic links between the main components substantiating Hilbert's claim of a special relation between Mie's and Einstein's theory have been weakened; they are only held together by the choice of a specific source. The link that Hilbert wanted to establish between electromagnetism and gravitation is thus no longer central to an approach alternative to that of Einstein, but has become little more than an attempt to fill the general framework established by Einstein with a specific physical content, represented by Mie's electrodynamics, an attempt that is now fully based on the firm foundations of general relativity.

A scientist's history

It is not usual for scientists to trace with great care the often only small and gradual conceptual transformations that scientific insights undergo in the course of historical development. Instead of undertaking such a demanding enterprise with little promise of new scientific results, they rather tend to keep past insights "alive," which usually means reinterpreting them in the light of their present and prospective future uses, rather than in the light of past achievements, let alone past failures. This natural tendency is, it seems, particularly strong where their own past insights are concerned. And, as we shall see now, it was inescapable for somebody like 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 version of his theory not as a revision of his original version, comprising conceptual adjustements and a reorganization of the deductive structure, but basically as a reprint of his earlier work:[157]

> 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.

In fact, the organization of Part 1, as compared to Paper 1, has not undergone major changes and seems to represent simply a tightening up of his earlier version; it can be subdivided into the following sections:

1. general introduction[158]
2. basic setting[159]

   Axioms I and II, field equations of electromagnetism and gravitation
3. basic theorems[160]

   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[161]

   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 equations.

The most noteworthy changes in the order of presentation are: the new introductory section and the integration of the treatment of energy-momentum conservation with other results of Hilbert's theory towards the end of the paper. Another conspicuous change is the fact that Hilbert's *Leitmotiv*, Theorem I of Paper 1, has now lost its central place, in spite of the fact that it had meanwhile been proven by Emmy Noether. But, as we have seen, it no longer played the key heuristic role for Hilbert that it had originally played in the Proofs, combining essential elements of Einstein's and Mie's theories. In any case, the

---

157 *Hilbert 1924*, p. 1.

158 *Hilbert 1924*, pp. 1-2.

159 *Hilbert 1924*, pp. 2-4.

160 *Hilbert 1924*, pp. 4-7.

161 *Hilbert 1924*, pp. 7-11.

preceding discussion should have made it clear that the rather unchanged form of presentation of Hilbert's work in fact hides major changes in the substance of his theory.

These changes are also reflected in the introductory section of Part 1, even if again combined with an attempt to play them down.

While in his earlier versions 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 announced his results as essentially providing a most simple and natural representation of Einstein's general relativity, completed only in a formal respect:[162]

> 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.

In view of the overwhelming impact of Einstein's general relativity, there was hardly any role left for Mie in Hilbert's new version. In fact, in the passage just cited, Mie is no longer presented as having posed problems of a similar profundity to those of Einstein, but rather as someone who had merely served to inspire Hilbert's esthetically more pleasing representation of general relativity, as well as a certain, unspecified completion of the theory "in formal respect."

In Hilbert's introduction, Mie has essentially become a man of the past. His name is no longer connected with the hope of gaining concrete insights into microphysics going beyond those of classical electrodynamics. Instead of attributing a specific role in the contemporary scientific discussion to Mie, Hilbert removes him from these down-to-earth struggles in order to elevate him into the role of one of the lofty founding fathers of field theory:[163]

> 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.

The immediate sequel to this passage, in which neither Einstein nor Minkowski are mentioned, even makes it appear as if it were Mie who had also introduced the idea of treating the space-time continuum as a field:[164]

---

162 *Hilbert 1924*, pp. 1-2. Evidently, the gradual change in Hilbert's theory was accompanied by a gradual change of his attitude with regard to Einstein's achievement by which he was increasingly impressed, see *Corry 1999a*, pp. 522-525.

163 *Hilbert 1924*, p. 1.

164 *Hilbert 1924*, p. 1.

> 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.

While Mie was thus exalted to the heaven of the founding fathers, which otherwise remained rather empty, this left room for relating Hilbert's efforts to other contemporary attempts at a unified theory of gravitation and electromagnetism, which he generously adopted as off-springs of his own contribution, although this view was hardly shared by his contemporaries, as we shall discuss below:[165]

> 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.

In this passage Hilbert leaves it open exactly to which equations of his theory he is referring. Given his references to Weyl and Einstein, he can only mean the two sets of field equations (51) and (52), which rather obviously are ingredients of any attempted unification of gravitation and electromagnetism; and, in any case, do not represent the unique feature of his own approach, which is the specific connection he introduced between these two sets of equations (compare (83)) which is for Hilbert the complete mathematical expression of the supposed character of electrodynamics as a phenomenon that follows from gravitation). But, as we have seen, in view of the results of general relativity this specific connection had become highly problematic and indeed was not taken up by either Weyl or Einstein.

In any case, it was highly uncertain whether Weyl's and Einstein's attempts at a unification were any more fortunate than Hilbert's. In his concluding paragraph Hilbert himself expressed this hesitation which was justified by the rapid progress of quantum physics, on the one hand, and the lack of concrete physical results of such unified theories, on the other:[166]

> 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.

165 *Hilbert 1924*, p. 2.

166 *Hilbert 1924*, p. 2.

But in spite of this hesitation, in the preceding paragraph Hilbert nevertheless showed himself convinced that “his theory” would last and indicated several possible ways that this might occur. He expressed the belief that this theory would be of programmatic significance for future developments in physics and, even if that should turn out not to be the case, there was, at least, philosophical benefit to be drawn from it:[167]

> 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.

The fact that his theory is actually not exclusively constructed out of these axioms, but also depends rather crucially on other physical concepts involved, such as the concept of energy, and that his theory may change in content as well structure if these concepts change their meaning, – all of that evidently remained outside the horizon of Hilbert’s epistemological understanding.

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

From Paper 1 to Paper 2

When Hilbert published his first paper in early 1916 he still had hopes that his unification of electromagnetism and gravitation would provide a basis for solving some of the riddles of microphysics as they were then discussed. In the introduction of this paper he announced:[168]

> Die genauere Ausführung sowie vor Allem die spezielle Anwendung meiner Grundgleichungen auf die fundamentalen Fragen der Elektrizitätslehre behalte ich späteren Mitteilungen vor.

And in the conclusion he added:[169]

> ... 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 ...

Clearly, he wanted to dedicate a second communication to the physical consequences of his theory. By the beginning of March 1916 he had submitted such a second installment of his theory which was then, however, withdrawn – no trace of it remains.[170] All that

---

167 *Hilbert 1924*, p. 2.

168 *Hilbert 1915*, p. 395.

169 *Hilbert 1915*, p. 407.

remains are the notes of Hilbert's SS 1916 and WS 1916/17 Lectures, as well as those of his related Causality Lecture. In the WS 1916/17 Lectures Hilbert gave his students hints at how he thought that his theory would lead to a modification of Maxwell's equations near the sources. While this was clearly still related to Hilbert's original project, the bulk of these notes actually testify to his careful study of the current work by Einstein and others on general relativity, and also contain his own original contributions to that project. When he eventually submitted his second communication to the Göttingen Academy at the end of December 1916, this work occupied the entire paper, which thus was completely different from what he had earlier announced.[171] As Hilbert's lecture notes are important for understanding the transition from his original project to what he actually published as Paper 2, as well as the contents of this paper, we shall discuss them here in some detail.[172] One of the most remarkable features of these notes is the opennes and informality with which Hilbert shares the discussion of unsolved problems with his students. He later even explicitly stated that this was a central goal of his lectures:[173]

> In my 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 student keeping clean and 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.

The causality quandary

The lecture notes make it clear that Hilbert was still in a quandary about how to treat the causality issue. As Einstein had shown, generally covariant field equations do not need to be supplemented by additional, non covariant, equations in order to arrive at a satisfactory theory of gravitation; and the same should be the case when electromagnetism is included. Yet, Hilbert's Theorem I, which implies the existence of four differential relations between the fourteen generally-covariant field equations for the gravitational and electromagnetic fields, shows that these equations cannot uniquely determine the evolution of all 14 field variables (the ten gravitational $g_{\mu\nu}$ and the four electromagnetic q_{μ}) off an initial hypersurface $x_4 = 0$ – which should be obvious anyway, since it is always possible by means of a coordinate transformation to introduce four arbitrary functions into a given solution to generally covariant field equations. That is to say, Hilbert's Proofs argument, based on causality, against general covariance seemed to remain valid.

An index of his continued perplexity is found in the typescript notes of his SS 1916 Lectures. The bulk of these lecture notes deal with special relativity (which he calls "die

170 See the discussion in *Sauer 1999*, p. 560, note 129.

171 *Hilbert 1917*.

172 The importance of Hilbert's lectures has been emphasized by Leo Corry in his publications.

173 *Reidmeister 1971*, pp. 79-82; transl. by Leo Corry.

kleine Relativität”): kinematics, and vector and 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 field equations, he explicitly notes that the causality problem is as yet unsolved:[174]

> Dies sind 14 Gleichungen für die 14 unbekannten Funktionen $g^{\mu\nu}$ und q_h $(\mu, \nu, h = 1 \quad 4)$. Das Kausalitätsprinzip kann erfüllt sein, oder nicht (Die Theorie hat diesen Punkt noch nicht aufgeklärt). Jedenfalls lässt sich auf die Gültigkeit dieses Prinzips nicht wie im Falle der Mie'schen Theorie durch einfach 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 der Zahl der aus dem Hamiltonschen Prinzip fliessendem 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.

It appears that he still had not resolved the causality problem when he continued the lectures in the winter semester of 1916-1917. The WS 1916/17 Lecture notes contain among other things much raw material for Paper 2. For example, they discuss causal relations between events within a given space-time in ways that very much resemble the treatment in that paper.[175] Yet they contain no discussion at all of the causality question for the fields themselves.

The answer to this causality problem given in Paper 2 does appear in the typescript (unfortunately undated) of his Causality Lecture. From the contents of these notes, it is reasonable to conjecture that they represent Hilbert's first exposition of his newly-found solution to this problem. After discussing the problem for his generally covariant system of equations and constructing an example to illustrate its nature (pp. 1-5), he comments:[176]

> 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.

---

174 SS 1916 Lectures, p. 110.

175 Compare Chapter XIII of the notes, *Einiges über das Kausalitätsprinzip in der Physik*, pp. 97-103, with pp. 57-59 of Paper 2, both discussed below).

176 Causality Lecture, p. 5. As we have discussed above, in his “Entwurf” theory Einstein did not actually first set up a system of generally covariant equations which he then restricted by non-invariant conditions, but rather started from non-generally covariant field equations right away. But as we have also discussed, Einstein considered the possibility that these equations do have a generally covariant counterpart from which they can be obtained by imposing non-invariant conditions as described here by Hilbert.

A similar comment appears in Paper 2:[177]

> 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.

Note that, neither in Paper 2 nor in any later publication, does Hilbert make the claim that this procedure (which it will be remembered he himself had followed in the Proofs) is "mathematisch falsch," as he does in the lecture notes, which strongly suggests that the notes precede Paper 2.

This suggested temporal sequence between the Causality Lecture and Paper 2 is confirmed by another significant passage: In his lecture, Hilbert makes a comparison between the problem created by general covariance in physical theories and that created by parameter invariance in the calculus of variations:[178]

> Auf die Schwierigkeit, zwischen einer sinnvollen und einer sinnlosen Behauptung unterscheiden zu müssen, stösst man übrigens auch in die 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 ein willkürliche Funktion von *p* ersetzt.

This particular comparison may well have played a role in his developing understanding of the causality problem. Yet the corresponding passage in Paper 2 makes a much more general comparison:[179]

> 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.

This argument is so much more general that it is hard to believe that, once he had given it, Hilbert would have later returned to its restricted application to extremization of curves. We take this as a further strong argument for the priority of the Causality Lecture notes.

In his lecture notes, Hilbert claims that the causality quandary can be resolved by an appropriate understanding of physically meaningful statements:[180]

> 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

177 *Hilbert 1917*, p. 61.

178 Causality Lecture, p. 8.

179 *Hilbert 1917*, p. 61.

180 Causality Lecture, pp. 5-6.

> 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.

In the last clause[181]

> es hat .. 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

one hears distant echoes of Einstein's assertion:[182]

> 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 x_4 entspricht. Zwei koinzidierenden Punktereignissen entspricht dasselbe Wertsystem der Variablen x_1 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.

One may indeed suspect that perusal of Einstein's expository paper *Die Grundlage der allgemeinen Relativitätstheorie*, published on 11 May 1916 and referred to in Hilbert's WS 1916/17 Lectures,[183] contributed to his new understanding of the causality problem.

The sequel of Hilbert’s argument shows, however, that his understanding of a physically meaningful statement actually differs from that of Einstein. Whereas Einstein had turned the uniqueness problem for solutions of a generally covariant theory 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 being those for which no such ambiguities arise, whether these statements employ coordinate systems or not. In his Causality Lecture, Hilbert claims to demonstrate a “causality principle,” formulated in terms of physically meaningful statements:[184]

> 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.

In order to establish this principle, he considers an arbitrary set of generally-covariant field equations (which he calls “ein System invarianter Gleichungen”) involving the metric tensor and the electromagnetic potentials and their derivatives.[185] He specifies the val-

---

181 Causality Lecture, p. 5.

182 *Einstein 1916a*, pp. 776-777.

183 See WS 1916/17 Lectures, p. 112.

184 Causality Lecture, pp. 5-6.

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

ues of these fields and their derivatives on the hypersurface $t = 0$, 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 (as he puts it: "die Transformation soll die Gegenwart ungeändert lassen"). He then defines a physically meaningful statement in terms of its unique determination by a Cauchy problem, 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:[186]

> 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 hingeben 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.

This passage makes it clear, however, that Hilbert's attempt at a definition of physically meaningful statements and clarification of the problem of causality was flawed by the still unrecognized intricacies of the Cauchy problem in general relativity. In fact, he evidently failed to realize that the classical notion of freely choosable initial values no longer works since some of the generally covariant field equations function act as constraints on the data that can be given on the initial hypersurface, rather than as evolution equations for that data off that surface. In the next section, we will discuss Hilbert's treatment of the problem of causality in Paper 2 and encounter further evidence for his neglect of Einstein's insight that, in general relativity, coordinate systems serve as mathematical devices for the description of space-time coincidences and have no physical significance of their own.

Hilbert at work on general relativity

Hilbert's choice of topics in Paper 2 shows that his original goal of developing a unified gravito-electromagnetic theory, with the immediate 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. Of course, Hilbert's own 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, these equations are just that subclass of solutions to his fourteen "unified" field equations, in which the electromagnetic potentials are set equal to zero. This makes them formally equivalent to the class of solutions to Einstein's field equations with vanishing stress-energy tensor – either 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 conflated Einstein's and Hilbert's programs. They

186 Causality Lecture, pp. 6-7.

could and did treat Hilbert's work in Paper 2 as a contribution to the development of the general theory of relativity, and this is how the work came to be assimilated into the relativistic tradition, as we shall discuss in more detail below.

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

1. the metric tensor and measurement of its components;[187]
2. the characteristics and bicharacteristics of the Hamilton-Jacobi equation corresponding to the metric tensor;[188]
3. the causality problem for events in a space-time with given metric;[189]
4. the causality problem for the field equations determining the metric tensor;[190]
5. Euclidean geometry as a solution to the field equations - in particular, the investigation of conditions that characterize it as a unique solution;[191] and
6. the Schwarzschild solution, its derivation,[192] and determination of the paths of (massive) particles and light rays in it.[193]

We shall now briefly indicate the nature of Hilbert's work on each of these six topics.

1) The metric tensor and its measurement: First of all, Hilbert switches from the use of the symbol *w* for the four coordinates (one of which is imaginary) to *x*, all of which are real, perhaps under the influence of Einstein's consistent use of such real coordinates; and in any case emphasizing that the $g_{\mu\nu}$'s, now all real, provide the “Massbestimmung einer Pseudogeometrie.”[194] He divides the elements (“Stücke”) of all curves into three classes: timelike, in which case an element measures the proper time; spacelike, measuring an element of length; and null, an element of a light path. He introduces two ideal measuring instruments, a measuring tape (“Maßfaden”) for lengths and a light clock (“Lichtuhr”) for proper times. Then Hilbert makes a comment that suggests, in spite of his remarks in Paper 1 and the lecture notes on causality (see above), that he still has a lingering belief that there is some objective significance to the choice of a coordinate system, even independently of the metric tensor:[195]

> 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.

---

187 *Hilbert 1917*, pp. 53-55.

188 *Hilbert 1917*, pp. 56-57.

189 *Hilbert 1917*, pp. 57-59.

190 *Hilbert 1917*, pp. 59-63.

191 *Hilbert 1917*, pp. 63-66 and p. 70.

192 *Hilbert 1917*, pp. 67-70.

193 *Hilbert 1917*, pp. 70-76.

194 *Hilbert 1917*, p. 54.

195 *Hilbert 1917*, p. 55.

This first theme ends with some comments on a possible axiomatic construction ("Aufbau") of the pseudogeometry. He suggests the need for two axioms:[196]

> 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 [p, where $x_s = x_s(p)$ is the parametric representation of a curve] ist;...
>
> Zweitens ist ein Axiom erforderlich, wonach die Sätze der pseudo-Euklidischen Geometrie d.h. das alte Relativitätsprinzip im Unendlichkleinen gelten soll;

2) Characteristics and bicharacteristics: Hilbert defines the null cone at each point, and points out that the Monge differential equation:[197]

$$g_{\mu\nu}\frac{dx_\mu}{dp}\frac{dx_\nu}{dp} = 0, \tag{130}$$

and the corresponding Hamilton-Jacobi partial differential equation:

$$g^{\mu\nu}\frac{\partial f}{\partial x_\mu}\frac{\partial f}{\partial x_\nu} = 0, \tag{131}$$

belong to 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 contemporary texts extend the term "null cone" from flat to non flat space-times, but we shall use the term "conoid") emanating from that point. He points out that the equation for this conoid is a solution to the Hamilton-Jacobi equation; and that all timelike world lines emanating from a world point lie inside its conoid, which forms their boundary.

While the discussion of the first two topics in Paper 2 is rather brief, they are treated much more extensively in Hilbert's WS 1916/17 Lectures. Chapter IX of these notes (pp. 69-80) is entitled "Die Monge'sche Differentialgleichung," but it also treats the Hamilton-Jacobi equation and the theory of characteristics, empasizing their relation to the Cauchy problem, and the reciprocal relation between integral surfaces of the Hamilton-Jacobi equation and null curves. Here the null conoids are called "transzendentale Kegelfläche." Chapters X (pp. 80-82) on "Die vierdimensionale eigentliche u. Pseudogeometrie" and XI (pp. 82-97) on the "Zusammenhang mit der Wirklichkeit" cover the material in the first section of Paper 2 in much greater detail. In particular, the measuring tape ("Massfaden") is discussed in section 38 (pp. 85-86 and pp. 91-92), and the light clock (which had already been introduced in the context of special relativity, see the SS 1916 Lectures, pp. 6-10), is reintroduced in section 44 (pp. 93-94), "Axiomatische Definition der Lichtuhr." Both are then used to determine the components of the metric tensor as functions of

196 *Hilbert 1917*, p. 56.

197 *Hilbert 1917*, p. 56.

the coordinates, “sobald nur ein bestimmtes Raum-Zeit Koordinatensystem x_i eingeführt worden ist” (p. 95). In many ways Hilbert’s discussion of these topics in Paper 2 reads like a brief precis of the lecture notes; it becomes much more intelligible if these are consulted.

3) Causality relation between events: [198] In accordance with an implicit requirement that three of the coordinates be spacelike and one timelike, Hilbert imposes corresponding requirements on the components of the metric tensor. But he has a curious way of motivating this demand:[199]

> 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) [equation number in the original; the conditions Hilbert imposes on the metric tensor] zu unseren Grundgleichungen hinzunehmen d.h. uns auf den Gebrauch *eigentlicher* Raum-zeitkoordinaten beschränken.

Again, we see that he seems to believe that the choice of a coordinate system has some residual physical significance; here, that it must reflect the relations of cause and effect between events on the same timelike worldline. He defines a proper (“eigentliches”) coordinate system as one, in which (in effect) the first three coordinates are spacelike and the fourth timelike in nature; and transformations between such proper coordinate systems are also called proper. Actually, given Hilbert’s stated goal of restricting the choice of coordinates to those that reflect the causal order on all timelike worldlines, his conditions are sufficient to realize this goal but not necessary since they exclude retarded null coordinates, which also preserve this causal order.

4) Causality problem for the field equations: As noted earlier, Hilbert here follows his Causality Lecture. In Paper 2 he formulates the argument as follows:[200]

> 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 gegen-

---

198 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 this time in no more detail than, the text of Paper 2.

199 *Hilbert 1917*, p. 57 and p. 58.

200 *Hilbert 1917*, p. 61.

> über 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.

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 we see that this is not what he actually says. First of all, if his words were so interpreted, they would stand in flagrant contradiction to his earlier statements (cited above in connection with the measurement of the metric and the causal relation between events), which presuppose attaching some residual physical meaning to the choice of coordinates. Secondly, it must be emphasized that his very definition of physical meaning ("physikalischer Sinn") involves a class of coordinate systems that leave the coordinates on the initial hypersurface ("die Gegenwart") unchanged. In the third place, Hilbert uses a Gaussian coordinate system, which he had introduced earlier,[201] in order to prove his assertion of the causality principle.[202]

His proof is essentially a brief discussion of the Cauchy problem for the 14 field equations in a Gaussian coordinate system. One of us has discussed this aspect of his work elsewhere,[203] 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 these ten. Gaussian coordinates then 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 he has a system in Cauchy normal form. This treatment is erroneous on several counts, but we shall postpone discussion of this question until the next section. More relevant to present topic is Hilbert's statement:[204]

> Da das Gaußische Koordinatsystem selbst eindeutig festgelegt ist, so sind auch alle auf dieses Koordinatensystem bezogenen Aussagen über jene Potentiale (34) [equation number in the original; the ten potentials mentioned above] von invariantem Charakter.

He never considers the question of the lack of invariance of the initial data under coordinate transformations on the initial hypersurface (three-dimensional hypersurface diffeomorphisms in modern terminology), again demonstrating that he is not thinking in terms of invariants, but is still tied to the use of particular coordinate systems.

Finally, his discussion of how to implement the demand for physically meaningful assertions depends heavily on the choice of a coordinate system. He remarks that:[205]

---

201 See *Hilbert 1917*, pp. 58-59.

202 See *Hilbert 1917*, pp. 61-62.

203 *Stachel 1992*.

204 *Hilbert 1917*, p. 62.

> Die Formen in denen physikalisch sinnvolle d.h. invariante Aussagen mathematisch zum Ausdruck gebracht werden können, sind sehr mannigfaltig,

and proceeds to discuss three of them:[206]

> Erstens. Dies kann mittelst eines invarianten Koordinatensystem geschehen. ...
>
> Zweitens. Die Aussage, wonach sich ein Koordinatensystem finden läßt, in welchem die 14 Potentiale gμn , qs 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.

The first two forms explicitly depend on the choice of a coordinate system (not necessarily unique). As examples of the first, he cites Gaussian and Riemannian coordinates. It is true that, discussing the second, he notes:[207]

> Der mathematische invariante Ausdruck für eine solche Aussage wird durch Elimination der Koordinaten aus jenen Beziehungen erhalten.

But he does not give an example of this procedure, nor does he suggest the most obvious way of realizing his goal, if indeed he had in mind a coordinate-independent solution to the problem: the use of invariants as coordinates. In particular, as Kretschmann noted a few years later, 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.[208] The use of such coordinates was taken up again by Arthur Komar in the 1960's, and hence they are often called Kretschmann-Komar coordinates.[209]

One might think that Hilbert meant something like this by his third suggested form. However, the example he cites makes it clear that he had something else in mind:[210]

> Ein Beispiel dafür sind die Einsteinschen Impuls-Energiegleichungen vom Divergenz Character. Obwohl nämlich die Einsteinsche Energie [that is, the gravitational energy-momentum pseudotensor] 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.

---

205 *Hilbert 1917*, p. 62.

206 *Hilbert 1917*, pp. 62-63.

207 *Hilbert 1917*, pp. 62-63.

208 See *Kretschmann 1917.*

209 See *Komar 1958*.

210 *Hilbert 1917*, p. 63.

Rather than anything invariant, apparently he has in mind here such 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 the flaws of his original approach, in particular the physical significance he ascribed to coordinate systems and his claim that the identities following from Theorem I represent a coupling between 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, as we shall discuss in the following.

5) Euclidean geometry: This section of Paper 2 opens with some extremely interesting general comments contrasting the role of geometry in what Hilbert calls the old and the new physics:[211]

> 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 Gestze nämlich die Grundgleichungen (4) und (5) [equation numbers in the original; the ten gravitational and four electromagnetic field equations] 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 declares:[212]

> 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.

He later formulates this question more precisely:[213]

> Die oben genannte geometrische Frage läuft darauf hinaus, zu untersuchen, ob und unter welchen Voraussetzungen die vierdimensionale Euklidische Pseudogeometrie [i.e., the Minkowski metric]... eine Lösung der physikalischen Grundgleichungen bez. die einzige reguläre Lösung derselben ist.

Hilbert thus takes up a challenge that emerged with the development of non-Euclidean geometry in the 19th century and was taken seriously by such eminent mathematicians as Gauss and Riemann: the question of the relation between geometry and physical reality.

211 *Hilbert 1917*, pp. 63-64.

212 *Hilbert 1917*, p. 63.

213 *Hilbert 1917*, p. 64.

Remarkably, this question was, for a number of reasons, not central to Einstein's heuristics. Consequently, he had, by this time, never really addressed the question posed by Hilbert, concerning the conditions under which Minkowski space 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 empty-space solutions to his equations. In effect, this question was a rather embarassing one for Einstein since such solutions display inertial properties of test particles even in the absense of matter, a feature that, for a long time, he had difficulties in accepting because of his Machian conviction that inertial effects must be due to the interaction of masses.[214] By establishing a connection between the mathematical tradition of questioning the geometry of physical space and general relativity, Hilbert thus made a significant contribution to the foundations of general relativity.

In the attempt to answer the question of the relation between Minkowski space and his equations, Hilbert first of all notes that, if the electrodynamic potentials vanish, then the Minkowski metric is a solution to the resulting equations, i.e., to the vanishing of what we now call the Einstein tensor.[215] He then poses the question: under what circumstances is the converse true, i.e., under what conditions is the Minkowski metric the only regular solution to these equations? Hilbert proves several results here. 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, he here goes back to use of the imaginary time coordinate w_4) and behave regularly at infinity, then they must vanish. 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:

- pp. 104-106, in the table of contents (p. 197) is given the title: “Der Sinn der Frage: Gilt die Euklidische Geometrie?”
- pp. 109-111, 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 the title “Aufstellung der Grundgleichungen beim Fehlen der Materie” in the table of contents (p. 197); and
- pp. 111-112, with the handwritten title “Zwei Sätze über die Gültigkit 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 (p. 197).

---

214 For historical discussion, see *Renn 1994*.

215 “wenn alle Elektrizität entfernt ist, so ist die pseudo-Euklidische Geometrie möglich” See *Hilbert 1917*, p. 64.

The lecture notes make Hilbert's motivation for a discussion of the empty-space field equations in general, and of the Schwarzschild metric in particular, much clearer than does Paper 2 itself. In the notes, Hilbert actually introduces the field equations in Section 51,[216] sandwiched between 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:[217]

> Wir wollen das Resultat unserer Rechnung vorwegnehmen: unsere physikalischen Grundgleichungen haben im allgemeinen keineswegs [the Minkowski metric] zu 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.

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

> 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.

We see that, at this point, the problem of the status of geometry, broadened from three-dimensional geometry to four dimensional pseudo-geometry – and in particular of the status of Euclidean geometry, similarly broadened to four-dimensional Minkowski pseudo-geometry – plays a central role in Hilbert's thinking about his program. This problem, rooted as it was in the mathematical tradition since Gauss, thus led him naturally to consider what we would call the empty-space Einstein field equations. He hoped that the absence of matter and radiation 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):[218]

> 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) [equation number in the original; the vanishing of the Einstein tensor] haben $g_{\mu\nu} = \delta_{\mu\nu}$ als einzige Lösung.

Of course, Hilbert was *not* able to establish this theorem, since it is not true. He elaborates on what he means by "regular" in his discussion of the Schwarzschild metric, which we consider below.

---

216 WS 1916/17 Lectures, pp. 106-109 (p. 107 is missing from the typescript).

217 WS 1916/17 Lectures, p. 106.

218 WS 1916/17 Lectures, pp. 111-112.

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:[219]

> 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 einzige Lösungen der Gravitationsgleichungen (3) [equation number in the original].
>
> Von diesem Satz kann ich schon jetzt so viel beweisen, dass in der Nachbarschaft der Euklidischen Geometrie sicher keine Lösung dieser Gleichungen vorhanden sind.

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.[220]

6) The Schwarzschild solution: Although the Schwarzschild solution had already been published in 1916,[221] Hilbert nevertheless dedicates considerable space to it, both in his lecture notes and in Paper 2. He does so because he uses the Schwarzschild solution in the context of his effort to exploit the new tools of general relativity for addressing the foundational questions of geometry raised in the mathematical tradition. In particular, he introduces, in his lecture notes, a number of assumptions on the metric tensor in order to prove a theorem on the uniqueness of Euclidean geometry:[222]

> 1) Es sei wieder $g_{\mu\nu}$ unabhängig von *t*.
>
> 2) Es sei $g_{\nu 4} = 0 \qquad \nu = 1, 2, 3$ [interpolated by hand: “d.h. Gauss’sches Coordinatensystem 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 die Bedingungen 1-3, so ist die Euklidische Geometrie die einzige Lösung der physikalischen Grundgleichungen.

It is the proof of this theorem that leads him to consider the problem of spherically symmetric solutions to the empty-space Einstein field equations, a problem that Hilbert notes

---

219 WS 1916/17 Lectures, p. 112.

220 See *Lichnerowicz 1946.*

221 *Schwarzschild 1916.*

222 WS 1916/17 Lectures, p. 113.

had previously been treated by Einstein (in the linear approximation) and by Schwarzschild (exactly). He claims for his own calculations only that they are "auf ein Minimum reduziert" compared to those of others.[223] Indeed, he works from his variational principle for the field equations (see above). Hermann Weyl gave a similar variational derivation of the Schwarzschild solution in 1917;[224] the section of his book *Raum-Zeit-Materie* on the Schwarzschild metric includes a general reference to Hilbert's Paper 2 (which reproduces Hilbert's variational derivation).[225] But Pauli, in his magisterial survey of the theory of relativity mentions only Weyl's paper, and this probably contributed to the neglect of Hilbert's contribution in most later discussions which, moreover, did not share Hilbert's epistemological concerns.[226]

At any rate, in Paper 2, Hilbert derives the Schwarzschild metric from the same three assumptions, emphasizing that:[227]

> 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.

This point was lost in many later derivations of the Schwarzschild metric, which continue to impose unnecessary boundary conditions on the metric. However, Hilbert did not seem to realize that the assumption of time-independence is also unnecessary, as shown by Birkhoff in 1923 (the assertion that the Schwarzschild solution is the only spherically symmetric solution to the empty-space Einstein equations is often referred to as Birkhoff's theorem).[228]

Hilbert's discussion of the Schwarzschild solution also confronted him with the problem of its singularities, immediately raising the question of the relation between this peculiar feature, evidently due to the presence of masses, and Hilbert's theory of matter. In his lecture notes, after establishing the Schwarzschild metric, he writes:[229]

> Nach unserer Auffassung vom Wesen der Materie könne 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 veschiedenen Determinante haben. Wir bezeichnen ferner eine einzelne Funktion als

223 WS 1916/17 Lectures, p. 113.

224 *Weyl 1917*.

225 *Weyl 1918a and 1918b*, p. 230 (note 9), *Weyl 1923*, p. 250 (note 19).

226 *Pauli 1921*.

227 *Hilbert 1917*, p. 67. For the derivation, see pp. 67-70.

228 See *Birkhoff 1923*, pp. 253-256.

229 WS 1916/17 Lectures, p. 118.

> 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..

It is curious that Hilbert identifies regularity with infinite differentiability and continuity of all derivatives, or even with analyticity. Either of these requirements is much too strong: each precludes gravitational radiation that carries new information, for example gravitational shock waves. But at least Hilbert attempted to define a singularity of the gravitational field. According to his definition, the Schwarzschild solution has singularities at $r = 0$ and at the Schwarzschild radius. He remarks:[230]

> 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.

But Hilbert now returns to his original motif:[231]

> Wir müssen also, um singularitätenfreie Lösungen zu erhalten, *a* [i.e., the mass parameter] = 0 annehmen. [This leads to the Minkowski metric]... 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 Relativitätsprinzip in der Physik tatsächlich, und für *t* = const ist in der Welt die Euklidische Geometrie wirklich realisiert.

He then turns to the justification for his consideration of the case $a \neq 0$:[232]

> 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.

In the sequel, Hilbert no longer just uses the Schwarzschild solution as a means for discussing foundational problems of geometry but also explores its physical significance for describing the behavior of matter in space and time. His own conception of matter, based on Mie's theory, plays, however, no significant role in this discussion. Its role is taken instead by assumptions that Hilbert assimilated from Einstein's work on general relativity, such as the geodesic postulate for the motion of particles.

In his lecture notes, he emphasizes the extraordinary difficulty of integrating the 14 general physical equations, even for "the simple special case when they go over to $K_{\mu\nu} = 0$:" [233]

---

230 WS 1916/17 Lectures, pp. 118-119.

231 WS 1916/17 Lectures, p. 119.

232 WS 1916/17 Lectures, p. 120.

233 WS 1916/17 Lectures, pp. 120-121.

> 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 der in der Umgebung dieser Stelle kennen, so würden die $g_{\mu\nu}$ wenn wir die neutrale Masse immer mehr gegen einen Massenpunkt hin degenerieren lassen, in diesem Punkte eine Singularität aufweisen. Eine solche müssten wir als erlaubt ansehen in dem Sinne, dass die $g_{\mu\nu}$ ausserhalb der nächsten Umgebung der Singularität den in der Natur wirklich realisierten Verlauf richtig wiedergeben. Eine solche Singularität müssen wir nun in (8) [equation number in the original; 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, angenäherte 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 Zeitlinie ist.
>
> Axiom II: Die Lichtbewegung im Gravitationsfeld wird durch eine geodätische Nullinie dargestellt.
>
> Axiom III.: Eine singuläre Stelle der Massbestimmung ist äquivalent einem Gravitationszentrum.

Hilbert calls the first two axioms, essentially adopted from Einstein's work, a "rational generalization" of the behavior of massive particles and light in the "old physics," in which the metric tensor takes the limiting Minkowski values (note that he uses an imaginary coordinate, so the Minkowski metric again takes the form $\delta_{\mu\nu}$). He states that the Newtonian law of gravitational attraction and the resulting Keplerian laws of planetary motion follow from these axioms "in the first approximation." In this way, Hilbert had integrated into his own theory the essential building blocks of physical knowledge on which Einstein's pathway to general relativity had been 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 the knowledge of classical physics such as the Newtonian law, but also of Euclidean geometry as a physical interpretation of space:[234]

> 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

234 WS 1916/17 Lectures, p. 122.

> 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.

We thus 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 timelike and null geodesics of the Schwarzschild metric, leading to discussions of two general-relativistic effects that Einstein had already considered: the perihelion precession of planets and the deflection of light by the sun's gravitational field. This discussion occupies almost all of the rest of this chapter of his lecture notes.[235] After a short discussion of the dimensions of various physical quantities,[236] he discusses the behavior of measuring threads and clocks in the Schwarzschild gravitational field,[237] and concludes the chapter with a discussion of the third general-relativistic effect treated by Einstein, the red shift of spectral lines.[238]

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 timelike 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 curious 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.

Revisions of Paper 2

As was the case for Paper 1, Paper 2 was republished twice: Indeed, the two papers were combined in the 1924 version in *Mathematische Annalen*, in which Paper 2 becomes Part 2 of *Die Grundlagen der Physik* (we shall simply refer to this version as "Part 2");[239] the reprint of this version in the *Gesammelte Abhandlungen* was edited by others, presumably under Hilbert's supervision (we shall refer to this version as "Part 2-GA").[240] In contrast to Paper 1, Hilbert's additions and corrections to Paper 2 are less substantial, as is to be expected, for a paper that was already written largely within the tradition of general rela-

---

235 WS 1916/17 Lectures, pp. 122-156.

236 WS 1916/17 Lectures, pp. 156-158.

237 WS 1916/17 Lectures, pp. 159-163.

238 WS 1916/17 Lectures, pp. 163-166.

239 *Hilbert 1924*, pp. 11-32.

240 *Hilbert 1935*, pp. 268-289.

tivity. Most changes concern minor improvements, e.g. in connection with the recent literature on the theory. Three changes are, however, significant, one introduced by Hilbert himself at the beginning of Paper 2, the other two by the editors of the *Gesammelte Abhandlungen*. The first concerns Hilbert's view of the relation between Paper 1 and Paper 2, the second the Cauchy problem at the heart of Hilbert's understanding of causality, and the third his understanding of invariant statements. In the following, we will provide an overview of these revisions, both major and minor.

The first significant change occurs right at the beginning, in the stated goal: Paper 2 states that "erscheint es nötig, einige allgemeinere Fragen sowohl logischer wie physikalischer Natur zu erörtern."[241] Part 2 states: "Es soll nun der Zusammenhang der Theorie mit der Erfahrung näher erörtert werden."[242] This revision confirms our interpretation of Paper 2 as resulting, in its original version, from the tension between Hilbert's agony over 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 addressed by contemporary general relativity.

We shall now discuss the changes in Part 2 for each of the six topics discussed in Paper 2 (see above):

1. The metric tensor and measurement of its components: in the discussion of the measurement of the metric tensor, Part 2 drops all reference to "Messfaden." The treatment is entirely in terms of the "Lichtuhr," but otherwise runs quite parallel to that in Paper 2.[243]
2. The characteristics and bicharacteristics of the Hamilton-Jacobi equation corresponding to the metric tensor: the discussion remains unchanged.[244]
3. The causality problem for events in a space-time with given metric: the discussion remains unchanged.[245]
4. There are several changes in the discussion of the causality problem for the field equations,[246] first of all in the wording with which Hilbert introduces the problem:[247]

   > Unsere Grundgleichungen (4) und (5) [equation numbers in the original; the gravitational and the electromagnetic field equations] der Physik sind nun keineswegs von der oben charakterisierten Art [Cauchy normal form]; vielmehr sind, wie ich gezeigt habe, vier von ihnen eine Folge der übrigen ...

241 *Hilbert 1917*, p. 53.

242 *Hilbert 1924*, p. 11.

243 *Hilbert 1924*, pp. 11-13.

244 *Hilbert 1924*, pp. 13-14.

245 *Hilbert 1924*, pp. 14-16.

246 *Hilbert 1924*, pp. 16-19.

247 *Hilbert 1924*, p. 16.

Hier “wie ich gezeigt habe” replaces “nach Theorem I” on p. 59 of Paper 2. On p. 16 of Part 2 Hilbert says that if there were 4 additional invariant equations, then the systen of the equations in Gaussian normal coordinates “ein überbestimmtes System bilden würde,” replacing “untereinander in Widerspruch ständen” on p. 60 of Paper 2.

In the discussion of the first way in which “physically meaningful, i.e., invariant assertions can be expressed mathematically,”[248] he corrects a number of the equations in his example. His discussion of the third way is considerably curtailed. It now reads:[249]

> 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.

In Paper 2, the sentence had ended with “...gültig sein soll,” and the paragraph had continued with the example of Einstein’s gravitational energy-momentum complex.

5. His discussion of Euclidean geometry is essentially the same, except that he no longer returns to an imaginary time coordinate with Euclidean metric in his discussion of gravitational perturbations.[250]
6. His discussion of the Schwarzschild solution is also essentially the same.[251] He adds a footnote to his axiom that light rays follow null geodesics:[252]

> Laue hat für den Spezialfall $L = \alpha Q$ [i.e., for the usual Maxwell equations] gezeigt, wie man diesen Satz aus den elektrodynamischen Gleichungen durch Grenzübergang zur Wellenlänge Null ableiten kann

followed by a reference to Laue’s 1920 paper that indicates that Hilbert was keeping up with the relativity literature to some extent.[253] He also dropped a rather trivial footnote to the axiom that massive particles follow timelike worldlines:[254]

> Dieser letzte einschränkende Zusatz [i.e., “Zeitlinie”] findet sich weder bei Einstein noch bei Schwarzschild.

He adds a more careful discussion of the circular geodesics for the case when their radius equals the Schwarzschild radius,[255] but otherwise the discussion of geodesics remains the same.

When the second, combined version of his two papers on *Die Grundlagen der Physik* was republished in his *Gesammelte Abhandlungen*, the editors introduced two extremely significant changes that effectively retract the last remnants of Hilbert’s attempt to provide a

248 *Hilbert 1917*, p. 62; *Hilbert 1924*, p. 18.

249 *Hilbert 1924*, p. 19.

250 *Hilbert 1924*, pp. 19-23, and p. 26.

251 *Hilbert 1924*, pp. 23-32.

252 *Hilbert 1924*, p. 27.

253 *Laue 1920*.

254 *Hilbert 1917*, p. 71.

255 *Hilbert 1924*, p. 30, compared to *Hilbert 1917*, p. 75.

solution to the causality problem for his original theory. These changes appear in Part 2-GA in the form of footnotes marked “Anm[erkung] d[er] H[erausgeber],” (as well as some more trivial ones, which we shall not discuss). The first significant change occurs in the discussion of the causality principle for generally covariant field equations.[256] The sentence:[257]

> Da das Gaußische Koordinatensystem selbst eindeutig festgelegt ist, so sind auch alle auf dieses Koordinatensystem bezogenen Aussagen über jene Potentiale (24) [equation number in the original] von invariantem Charakter.

is dropped; and a lengthy footnote is added.[258] This footnote makes clear that the editors correctly understood the nature of the ten gravitational and four electromagnetic field equations. Only six of the gravitational and three of the electromagnetic equations involve second time derivatives of the corresponding spatial components of the metric tensor and electromagnetic potentials, and thus their values, together with those of their first time derivatives, on the initial hypersurface determine their evolution off that hypersurface. These initial values are subject to constraints set by the remaining four gravitional equations and the one electromagnetic equation; due to the differential identities between all the field equations, if these constraint equations are initially satisfied, they will hold off the initial hypersurface by virtue of the remaining field equations. In fact, the footnote of the editors culminates in the statement:[259]

> 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.*

They also understand that in the gauge-invariant electromagnetic case, it is only the fields and not the potentials that are determined by the field equations. While the editor's addition thus represents a lucid account of the Cauchy problem in general relativity, it demonstrates at the same time that Hilbert's attempt to formulate a principle of causality for this theory in terms of the classical notion of initial data (i.e. values that can be freely chosen at any given moment in time and then determine the future) had failed because he had not taken into account the constraints on the initial data that he had considered.

The second major addition to Part 2 occurs in the discussion of the first way of satisfying the requirement that physically meaningful assertions be invariant, namely by use of an invariant coordinate system.[260] A footnote is added that actually undermines the claims made in Hilbert's paper; it reads:[261]

---

256 *Hilbert 1924*, p. 18-19, *Hilbert 1935*, pp. 275-277.

257 *Hilbert 1924*, p. 18.

258 *Hilbert 1935*, pp. 275-277.

259 *Hilbert 1935*, p. 277.

260 *Hilbert 1924*, pp. 18-19.

261 *Hilbert 1935*, p. 277.

> 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) [equation number in the original] 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) [equation number in the original; the conditions for a "proper" coordinate system] eingeführt wurde.

Thus, the editors of Hilbert's papers corrected two major mathematical errors that survived Hilbert's own first revision of Paper 2. Since he was still active when this edition of his papers was published, it may be assumed that these changes were made with his consent, if not his participation.

7. THE FADING AWAY OF HILBERT'S POINT OF VIEW IN THE PHYSICS AND MATHEMATICS COMMUNITIES

Einstein and Weyl set the tone very early for the particular 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 works by Einstein and Weyl in question 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*, the standard source for original work on the development of relativity.[262] 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 have had their impression of the scope and history of relativity shaped by this book.

We shall first discuss Einstein's two mentions of Hilbert in 1916 papers. (His third occurs in a paper published in 1919, and is related to Weyl's 1918 paper, so we shall discuss it

262 See *Blumenthal 1913, 1919, 1923, and 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 five additional papers by Einstein on general relativity, the fifth edition added Weyl's first paper on his unified theory of gravitation and electromagnetism. This edition is the basis of the editions currently in print, and of the translations into other languages. It would be interesting to know how Blumenthal made his choice of papers to include in what became the canonical source book on relativity.

afterwards.) The first comment by Einstein refers to energy-momentum conservation and the second to the derivation of the field equations from a variational principle. 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 these results. This assimilation, in turn, assigned to 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:[263]

> 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

A footnote adds a reference to Paper 1.[264] Thus, Einstein integrated what Hilbert regarded as an outstanding achievement of his theory into the general theory of relativity as a particular case of an important but subordinate general result. Hilbert's interpretation of this result as embodying a coupling between gravitation and electromagnetism is not even mentioned.

In the same year Einstein also published his own derivation of the generally covariant gravitational field equations from a variational principle. While he had given a non-invariant "Hamiltonian" (= Lagrangian) for the field equations modulo the coordinate condition $\sqrt{-g} = 1$ in the 1916 review paper, he now proceeded in a manner rather reminiscent of Hilbert's in Paper 1. He introduces the same gravitational variables (the $g_{\mu\nu}$ and their first and second derivatives), but Einstein's $q_{(\rho)}$ "beschreiben die Materie (inklusive elektromagnetisches Feld)" and are hence arbitrary in number and have unspecified transformation properties. This straightforward generalization allowed Einstein to transform Hilbert's variational derivation into a contribution to general relativity, without having to adopt the latter's perspective on this derivation as providing a synthesis between a theory of gravitation and a specific theory of matter. On the contrary, Einstein's generalization of Hilbert's derivation made it possible to regard the latter as merely representing a problematic special case.

Einstein prefaced his calculations with some observations intended to place his work in context:[265]

263 *Einstein 1916a*, p. 810.

264 The reference is to "p. 3," probably in a separately paginated off-print that Hilbert circulated; see the discussion in *Sauer 1999*.

265 *Einstein 1916b*, p. 1111.

> 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.

Thus Einstein both gave Hilbert credit for his accomplishment, and clearly circumscribed its nature: Like Lorentz, Hilbert was supposedly looking for a variational derivation of the general-relativistic field equations, but made assumptions about the constitution of matter that were too special. In an earlier, unpuplished draft related to this paper, Einstein's tone was sharper:[266]

> 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.

In private correspondence, he was not only much harsher but also gave his reasons for disregarding Hilbert's point of view:[267]

> Der Hilbertsche Ansatz für die Materie erscheint mir kindlich, im Sinne des Kindes, das keine Tücken 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.

Evidently, Einstein had a clear perception of the diverse status of the knowledge underlying general relativity, on the one hand, and Hilbert's theory, on the other hand. From the point of view publicly adopted by Einstein, Hilbert's other 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 formulated them.

One would expect that Hilbert's work played a prominent role also in the developing field of unified field theory, in particular in view of his own claims in this regard. But his fate was that of a transitional figure whose role was eclipsed both by his predecessors and his followers. His achievements were rather perceived as odd contributions to general relativity 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 considered one of the founding fathers of general relativity.

---

266 See note 3 to Doc. 31 in *Kox et al. 1996*, p. 346.

267 Einstein to Hermann Weyl, 23 November 1916, *Schulmann et al. 1998*, pp. 365-366.

In his first contribution to the study of unified field theories, Hermann Weyl assigned a quite definite place to Hilbert, if largely by omission. After presenting his generalization of Riemannian geometry to include what he called “gauge invariance” (Eichinvarianz), Weyl turned to unified field theory:[268]

> Von der Geometrie zur Physik übergehend, haben wir nach dem Vorbild der Mieschen Theorie [references to Mies’s papers of 1912/13 and Weyl’s recently-published Raum-Zeit-Materie] anzunehmen, daß die gesamte Gesetzmäßigkeit der Natur auf einer bestimmten Integralinvariante, der Wirkungsgröße
>
> $$\int W d\omega = \int \mathrm{W}\, dx \qquad (\mathrm{W} = W\sqrt{g})$$
>
> 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.

In spite of its obvious relevance, there is no mention of Hilbert in this context. The only mention comes in what we shall refer to as “the litany” since it or a similar list occurs so frequently in the subsequent literature:[269]

> Wir werden nämlich zeigen: in der gleiche 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.

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 this paper in *Das Relativitätsprinzip*, Weyl notes that:[270]

> 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 gelöst.[271]

without any mention that this is the solution to precisely the problem raised by Hilbert’s Ansatz for the invariant Lagrangian first introduced in Paper 1. Little wonder if those who

---

268 *Weyl 1918c*, p. 475.

269 *Weyl 1918c*, p. 475.

270 *Blumenthal 1974* (this seventh edition is an unchanged reprint of the fifth edition of 1923), note 2 , p. 159.

271 *Weitzenböck 1920* 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.

learned their history of relativity via *Das Relativitätsprinzip* had no idea of Hilbert's actual aims and little more of his achievements.

Hilbert fared a little better in Weyl's *Raum-Zeit-Materie*, the first treatise to appear on general relativity.[272] The discussion of the energy-momentum tensor in the first edition (Section 27) credits Hilbert with having shown that:[273]

> [...] die Miesche Elektrodynamik von den Voraussetzungen der speziellen auf die der allgemeinen Relativitätstheorie übertragen werden [kann]. Dies is von Hilbert durchgeführt worden.

Footnote 5 cites Paper 1 and adds:[274]

> 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.

Curiously, both this textual reference to Hilbert and the footnote disappear from all later editions (but see discussion below of the fifth edition). Presumably because Weyl had already mentioned Hilbert in this connection, 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,[275] and remains there. In his discussion of causality for generally covariant field equations in the first edition, Weyl credits Papers I and II;[276] but, again, this note disappears from all later editions. Paper 2 is also cited in the first edition in connection with the Schwarzschild solution,[277] and the introduction of geodesic normal coordinates.[278]

The third edition carries over these references to Paper 2 and adds one in connection with linearized gravitational waves;[279] and the fourth edition (the one that was translated into French and English) includes all these footnotes. Perhaps some questions were raised concerning Weyl's treatment of Hilbert; at any rate, in the fifth edition, the footnote to the litany citing Hilbert adds what again amounts to crediting Hilbert with a contribution to general relativity, rather than to the study of unified field theories:[280]

---

272 *Weyl 1918a, 1918b, 1919, 1921, and 1923*. The second edition of 1918 was unchanged, the fourth of 1921 is the one translated into English and French; the fifth of 1923 has been thereafter reprinted without change.

273 *Weyl 1918a and 1918b*, p. 184.

274 *Weyl 1918a and 1918b*, p. 230.

275 *Weyl 1919*, note 8, p. 266.

276 *Weyl 1918a and 1918b*, p. 190 (and note 9 on p. 230).

277 *Weyl 1918a and 1918b*, note 15, p. 230.

278 *Weyl 1918a and 1918b*, note 21, p. 230.

279 *Weyl 1919*, note 14, p. 266.

280 *Weyl 1923*, note 10, p. 329.

> In der 1. Mitteilung stellte Hilbert gleichzeitig und unabhängig von Einstein die invarianten Feldgleichungen auf, aber im Rahmen der hypothetischen Mieschen Theorie der Materie.

In fact, 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 early source, still used (mainly in the English translation of 1956 with additional notes) by physicists and mathematicians for historical as well as technical information about relativity and unified field theories.[281] 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 breakthough, Pauli adds in a footnote:[282]

> 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, buit are specifically based on Mie’s theory of matter ... .

In his discussion of invariant variational principles in section 23, Pauli recites the litany, citing “investigations by Lorentz, Hilbert, Einstein, Weyl and Klein on the role of Hamilton’s Principle in the general theory of relativity.”[283]

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, concluding:[284]

> 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 perhaps the most striking example of falsely crediting Hilbert with a contribution to general relativity while neglecting Hilbert’s actual achievements. To make matters worse, Pauli then adds a footnote curiously citing Paper 1, rather than Paper 2; after also crediting Mach with a version of this insight, he adds:[285]

> 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 *Die formalen Grundlagen*].

---

281 *Pauli 1921* and *Pauli 1958*.

282 *Pauli 1921* (section 50, cited from translation in *Pauli 1958*, p. 145, note 277).

283 *Pauli 1921* (cited from translation in *Pauli 1958*, p. 68).

284 *Pauli 1921* (cited from translation in *Pauli 1958*, p. 160).

285 *Pauli 1921* (cited from translation in *Pauli 1958*, p. 160, note 315).

Pauli also acknowledges various real contributions by Hilbert to general relativity in Paper 2.[286] But his discussion of unified field theories (Part V), like Weyl's, jumps from Mie (Section 64) to Weyl (Section 65) without even a mention of Hilbert.

We can get some idea of how the Einstein-Weyl line as canonized by Pauli was propagated by examining a couple of early treatises on relativity by non-German authors. Jean Bequerel'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," Bequerel asserts:[287]

> Lorentz et Hilbert [references to Papers I and II], puis Einstein, ont reussi à presenter les équations générales de la theorie de la gravitation comme des conséquences d'un unique principe d'action stationnaire.

Then follows Section 103 on "Méthode de Lorentz et d'Hilbert."[288] Paper 2 is cited in connection with linearized gravitational waves,[289] 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."[290]

Eddington's treatise, *The Mathematical Theory of Relativity*, was widely read, cited and studied until recently by students, and translated into French and German.[291] The two English editions contain Papers I and II in the bibliography, with a reference to Section 61 on "A Property of Invariants,"[292] which demonstrates the theorem:[293]

> 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.[294] On p. 114, footnote 1 he 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. On p. 183, he credits the same paper for the first strict proof

---

286 See *Pauli 1921*, section 13, for Axiom II, section 22, for discussion of restrictions on coordinate systems if three coordinates are to be spacelike and one timelike, and section 60 for the proof that linearized harmonic coordinate conditions may always be imposed.

287 *Bequerel 1922*, p. 256.

288 *Bequerel 1922*, pp. 257-262.

289 *Bequerel 1922*, p. 216.

290 *Bequerel 1922*, pp. 309-335.

291 *Eddington 1923 and 1924*.

292 See *Eddington 1924*, p. 264, where he writes "wherever possible the subject matter is indicated by references to the sections in this book chiefly concerned."

293 See *Eddington 1924*, pp. 140-141.

294 *Eddington 1925*.

that one can always satisfy the linearized harmonic coordinate conditions by an infinitesimal coordinate transformation. And that is it.

We see that, with minor variations within the acceptable limits, the Einstein-Weyl line on Hilbert's role is 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 formulation of the field equations of general relativity. In fact, he did not formulate these equations at all but developed, at the end of 1915, a theory of gravitation and electromagnetism that is incompatible with Einstein's general relativity. This theory can, nevertheless, hardly be considered an achievement in its own right, parallel to that of Einstein's creation of general relativity and to be judged by independent criteria. Not only is the dependence on Hilbert's theory from and similarity to Einstein's earlier, non-covariant "Entwurf" theory of gravitation too striking; but both its contemporary reception as a contribution to general relativity and Hilbert's own gradual transformation of his theory into such a contribution are evidence to 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 looser – 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 to nor did achieve and ignores, on the other hand, 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.

General relativity, as it was formulated by Einstein in 1915, incorporates knowledge of classical mechanics, electrodynamics, the special theory of relativity, and planetary astronomy, as well as knowledge from mathematical traditions such as non-Euclidean geometry and absolute differential calculus, and it integrates this knowledge into a single, coherent conceptual framework centered around new concepts of space, time, and gravitation. Without this enormous body of knowledge as the underpinning of general relativity, it would be hard to explain the theory's impressive stability and powerful role even in today's physics. The integration of knowledge at the roots of general relativity is the result of a long-winded and conflict-laiden process to which not only Einstein but also many other scientists contributed. From the point of view of historical epistemology, the resulting transformation of the knowledge structures of classical physics is, however, a collective process in an even deeper sense.[295] It involves a substantial, shared knowledge-base, structured by fundamental concepts, models, heuristics etc. which are transmitted by social structures, such as institutions, and by material representations, such as textbooks, and which are individually appropriated by learning processes. While individual thinking is usually governed to a large degree by these shared resources, it also affects these resources, amplifying them and occasionally even changing their epistemic structures. On the basis of an historical epistemology that takes into account this interplay between shared knowledge resources and individual thinking, the emergence and fading away of a theory such as Hilbert's becomes understandable as an aspect of the integration of knowledge that gave rise to general relativity.

In the following, we will first look at some of the resources available in the shared knowledge of the time for formulating gravitational field equations such as those of Einstein and Hilbert, thus answering, in our terms, the question of where alternative solutions (or attempted solutions) to the same problem come from. We then describe the interplay between individual thinking and knowledge resources that led to the establishment of general relativity and to the transformation of Hilbert's theory into a contribution to general relativity, with the intention of explaining the fading-away of Hilbert's theory. It will become clear that, in both cases, essentially the same equilibration mechanism is at work. In the first case, it integrates the various components of the shared knowledge and results in the creation of a stable epistemic structure, general relativity, which represents the integrated knowledge. In the second case, the same process disintegrates the various components of the shared knowledge brought together in a temporary construct, Hilbert's theory, and rearranges them as parts of the more stable structure.

In order to address the problem that, in late 1915, occupied both Einstein and Hilbert, the formulation of differential equations governing the gravitational potential represented by the metric tensor, the knowledge available at the time offered a limited number of pathways. Principal alternatives for possible solutions were embodied in the fundamentally different models underlying contemporary field theories, mostly of electrodynamics. Among these models was that of conceiving all physical phenomena, including matter, in terms of fields, the "monistic model," and that of a dualism of fields and matter, the "fields-with-matter-as-source model." The first model was the basis for attempts to formulate "an electromagnetic world picture," which, however, remained fragmentary and never succeeded in accounting, in its terms, for contemporary physical knowledge. The second model was the basis for Lorentz's formulation of electron theory, the epitome of classical electrodynamics. It is based on the idea that matter acts as source for electrody-

295 Compare, also for the following, the discussion of the notion of creativity in *Czikszentmihalyi 1988*, discussed in *Stachel 1994*: "All of the definitions ... 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."

namic fields which, in turn, affect the motion of material bodies. Rather than having to reconstruct the knowledge embodied in classical mechanics in terms of electrodynamic field concepts, a challenging task associated with the program of an electrodynamic world picture, Lorentz's electron theory with its dualistic model successfully integrates the classical knowledge about electromagnetic and mechanical phenomena. While the first model became the core of Hilbert's approach, who attempted to create a unified field theory, Einstein's search for the gravitational field equations was guided by the second model. The qualitative features of these two models account to a large extent for the differences between Hilbert's and Einstein's approaches, including their unequal capacity to integrate the available physical knowledge in their theories. Whereas the knowledge about matter available to Hilbert's theory was modest, being essentially represented by Mie's speculative theory, Einstein's source-term for the gravitational field equations embodied the vast knowledge about matter represented by special-relativistic continuum theory, including energy-momentum conservation.

The space of alternative pathways to solving the problem of the gravitational field equations is, of course, not exhausted by the different qualitative models of the interaction between fields and matter available for these pathways. Contemporary mathematics also offered a reservoir of different tools that could be used for addressing the problem of the field equations. The series of attempts to formulate a theory of gravitation between 1912 and 1915, including contributions by Abraham, Einstein, Nordström, and Hilbert, illustrates the dependence of a particular version on the mathematical formalism employed, ranging from linear, partial differential equations for a single variable to the absolute differential calculus applied to the metric tensor. As did the models discussed above, so also different formalisms showed different capacities for integrating the available physical knowledge about matter and gravitation, e.g. that embodied in Newtonian gravitation theory or in the observational results on the Mercury perihelion shift. In order to explore a formalism's capacity of integrating knowledge, it needs to be elaborated and its consequences interpreted, if possible, as representations of physical knowledge. Different degrees of elaboration and interpretation, the "exploration depth" of a given formalism, may determine its acceptability as an adequate representation of the physical problem at hand. When, for instance, Einstein found in early 1913 that he was unable to recover the Newtonian limit from generally covariant field equations – a problem that involves the exploration of both technical and conceptual aspects, he decided in favor of the non-covariant "Entwurf" theory. At the end of 1915, on the background of a much increased "exploration depth" of the same formalism, he decided instead in favor of generally covariant equations.

Which models or mathematical formalisms are employed or favored in a given historical situation depends on many factors, among them their social accessibility and group-specific epistemological preferences ("images of knowledge," Elkana) that, to certain people, make some of these knowledge resources appear more attractive than others. Concerning the accessibility of resources to specific groups, it was, for instance, natural for a mathematician of Hilbert's caliber to start from a generally covariant variational principle, while Einstein, for a time, seriously considered developing his own, “pedestrian” differential calculus for dealing with the metric tensor, being ignorant of the appropriate mathematical resources.[296] Concerning the role of images of knowledge, it is clear that the monistic model of a pure field theory must have looked more appealing to Hilbert, a mathematician in search for an axiomatic foundation of all physics, than the conceptually more clumsy dualistic model. The dualistic model of fields-with-matter-as-source, on the other hand, was a more natural starting point for physicists such as Abraham, Einstein, and Nordström who were familiar with the extraordinary success of this model in the domain of electromagnetism. Similarly, images of knowledge also determine decisions on the exploration depth and exploration direction of a given formalism. While the question of the Newtonian limit, for instance, was crucial to the physicist Einstein, Hilbert did not bother at all, in his original approach, with this problem.

Individual constructs created by individual scientists, e.g. Hilbert’s paper proposing an axiomatic foundation of physics, are largely a matter of contingency, but their building blocks (concepts, models, techniques) are, in any case, taken from the reservoir of the socially available knowledge that is characteristic for a given historical situation. The neglect of this shared background in a historical study typically either induces problematic divisions (“incommensurability,” Kuhn) or necessitates the building of artificial bridges (“trading zones,” Galison) between seemingly disparate entities. The reservoir of shared background knowledge accounts, however, for more than just the communicability of individual results such as those of Hilbert and Einstein. Given that individual contributions are assimilated to shared knowledge by various processes of communication and intellectual digestion, an equilibration process takes place between the individual constructs and the shared knowledge-reservoir. It is this equilibration process that decides on the stability of an individual contribution, its longivity (the case of general relativity) 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, which, 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; otherwise they will be transmitted as part of the shared knowledge. Usually, individual contributions are not assimilated wholesale to the shared knowledge but only in a piecemeal fashion. Thus, 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

---

296 See his calculations (e.g. on p. 15) in the Zurich Notebook, partially published as Doc. 10 of *Klein et al. 1995*.

shared knowledge. The reception of Hilbert's contribution is thus not different from that of most scientific contributions that become assimilated to the great banquet of the 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 just overpower any impact from individual contributions by the sheer mass of the 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 affecting 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. He had, over a period of several years, attempted not only to reconcile the knowledge of classical physics concerning gravitation with the special-relativistic requirement of the finite propagation speed of physical interactions, but also with insights into the affinity between 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, 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 he kept fixed throughout his search for the gravitational field equations. The affinity between gravitation and inertia as expressed by the equivalence principle, on the other hand, could be given various mathematical representations; for Einstein the most natural one was the demand for general covariance of the field equations. 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.[297] On the one hand, he started from those physical principles which embody the vast knowledge of classical and special-relativistic physics and explored the consequences of their mathematical representations in 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 would integrate also 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 modifica-

297 See *Renn and Sauer 1998*.

tions eventually became more than just personal idiosyncracies and 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 contributions of being dissolved and assimilated to the structures of the shared knowledge. Even if he had, in 1915, 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, that is, 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. This difference in reacting to a similar candidate for solving the problem of the gravitational field equations obviously does not reveal any different in the epistemic status of Hilbert's theory compared to Einstein's intermediate version but only a different attitude with regard to a given exploration depth, motivated by the different image of knowledge Hilbert associated with his endeavor. For the fate of a theory in the life of the scientific community such motivations make little difference. In fact, the subsequent elaborations, revisions, and transformations of Hilbert's result testify to an equilibration process similar to that undergone also by Einstein's intermediate versions in which ever new elements of the 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 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." For the history of knowledge, 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.

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 the colleagues with whom we had the opportunity of working in this context, in particular Leo Corry, Tilman Sauer, and Matthias Schemmel, for the discussions, the support, and the criticism, all of it intensive, which have made this collaboration a memorable and fruitful experience for all 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 Schöpflin, for their creative and indefatigable support in all library-related activities of our work. We would like to 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. We finally would like to thank Ze’ev Rosenkranz and the Einstein Archives, The Hebrew University of Jerusalem, for permission to quote from Einstein’s letters.

10. BIBLIOGRAPHY

Belinfante 1939 Belinfante, F. J. "Spin of Mesons." *Physica* 6 :887-898.

Bequerel 1922 Bequerel, Henri. *Le Principe de la Relativité et la Théorie de la Gravitation*. Paris: Gauthier-Villars.

Birkhoff and Langer 1923 Birkhoff, George D., and Rudolph E. Langer. *Relativity and Modern Physics*. Cambridge, Ma.: Harvard University Press.

Blumenthal 1913 Blumenthal, Otto, ed. *Das Relativitätsprinzip*. 1st ed. Leipzig, Berlin: Teubner.

Blumenthal 1919 Blumenthal, Otto, ed. *Das Relativitätsprinzip*. 3rd ed. Leipzig, Berlin: Teubner.

Blumenthal 1923 Blumenthal, Otto, ed. *Das Relativitätsprinzip*. 5th ed. Leipzig, Berlin: Teubner.

Blumenthal 1974 Blumenthal, Otto, ed. *Das Relativitätsprinzip*. 7th ed. Darmstadt: Wissenschaftliche Buchgesellschaft.

Born 1914 Born, Max. "Der Impuls-Energie-Satz in der Elektrodynamik von Gustav Mie." *Königliche Gesellschaft der Wissenschaften zu Göttingen. Nachrichten* (1914):23-36.

Born and Infeld 1934 Born, Max, and Leopold Infeld. "Foundations of the New Field Theory." *Royal Society of London. Proceedings* A 144:425-451.

Caratheodory 1935 Caratheodory, Constantin. *Variationsrechnung und partielle Differentialgleichungen erster Ordnung.* Leipzig, Berlin: B. G. Teubner.

Corry 1997 Corry, Leo. "David Hilbert and the Axiomatization of Physics (1894-1905)." *Archive for History of Exact Sciences* 51:83-198.

Corry 1999a Corry, Leo. "David Hilbert between Mechanical and Electromagnetic Reductionism (1910-1915)." *Archive for History of Exact Sciences* 53:489-527.

Corry 1999b Corry, Leo. "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.

Corry (Forthcoming) Corry, Leo. "Hilbert and Physics (1900-1915)." In *The Visual World: Physics and Mathematics (1900-1930)*, edited by J. J. Gray. Oxford: Oxford University Press, 145-188.

Corry, Renn, and Stachel 1997 Corry, Leo, Jürgen Renn, and John Stachel, eds. *Belated Decision in the Hilbert-Einstein Priority Dispute*. Vol. 278, Science.

Cziksentmihalyi 1988 Cziksentmihalyi, Mihalyi. "Society, Culture and Person: a Systems View of Creativity." In *The Nature of Creativity*, edited by Robert J. Sternberg. Cambridge: Cambridge University Press.

Earman and Glymour 1978 Earman, John, and Clark Glymour. "Einstein and Hilbert: Two Months in the History of General Relativity." *Archive for History of Exact Sciences* 19:291-308.

Eddington 1923 Eddington, Arthur Stanley. *The Mathematical Theory of Relativity.* Cambridge: The University Press.

Eddington 1924 Eddington, Arthur Stanley. *The Mathematical Theory of Relativity.* 2nd ed. Cambridge: The University Press.

Eddington 1925 Eddington, Arthur Stanley. *Relativitätstheorie in Mathematischer Behandlung.* Translated by Alexander Ostrowski Harry Schmidt. Berlin: Springer.

Einstein 1913 Einstein, Albert. "Zum gegenwärtigen Stande des Gravitationsproblems." *Physikalische Zeitschrift* 14 (25):1249-1262.

Einstein 1914a Einstein, Albert. "Prinzipielles zur verallgemeinerten Relativitätstheorie und Gravitationstheorie." *Physikalische Zeitschrift* 15:176-180.

Einstein 1914b Einstein, Albert. "Die formale Grundlage der allgemeinen Relativitätstheorie." *Königlich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (1914) (XLI):1030-1085.

Einstein 1915a Einstein, Albert. "Zur allgemeinen Relativitätstheorie." *Königlich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (1915) (XLIV):778-786.

Einstein 1915b Einstein, Albert. "Zur allgemeinen Relativitätstheorie (Nachtrag)." *Königlich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (1915) (XLVI):799-801.

Einstein 1915c Einstein, Albert. "Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie." *Königlich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (1915) (XLVII):831-839.

Einstein 1915d [Einstein, Albert]. [Zusammenfassung der Mitteilung "Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie."] *Königlich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (1915) (XLVII):803.

Einstein 1915e Einstein, Albert. "Die Feldgleichungen der Gravitation." *Königlich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (1915) (XLVIII-XLIX):844-847.

Einstein 1916a Einstein, Albert. "Die Grundlage der allgemeinen Relativitätstheorie." *Annalen der Physik* 49 (7):769-822.

Einstein 1916b Einstein, Albert. "Hamiltonsches Prinzip und allgemeine Relativitätstheorie." *Königlich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (1916) (XLII):1111-1116.

Einstein and Grossmann 1914 Einstein, Albert, and Marcel Grossmann. "Kovarianzeigenschaften der Feldgleichungen der auf die verallgemeinerte Relativitätstheorie gegründeten Gravitationstheorie." *Zeitschrift für Mathematik und Physik* 63 (1 / 2):215-225.

Fölsing 1997 Fölsing, Albrecht. *Albert Einstein: a biography.* New York: Viking.

Frei 1985 Frei, Günther, ed. *Der Briefwechsel David Hilbert-Felix Klein (1886-1918).* Vol. 19, *Arbeiten aus der Niedersächsischen Staats- und Universitätsbibliothek Göttingen*. Göttingen: Vandenhoeck & Ruprecht.

Goenner et al. 1999 Goenner, Hubert, Jürgen Renn, Jim Ritter, and Tilman Sauer, eds. *The Expanding Worlds of General Relativity.* Vol. 7, *Einstein Studies*. Boston, Basel, Berlin: Birkhäuser.

Guth 1970 Guth, E. "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.

Hilbert 1905 Hilbert, David. "Logische Prinzipien des mathematischen Denkens." Ms. Vorlesung SS 1905, annotated by E. Hellinger, Bibliothek des Mathematischen Seminars, Göttingen.

Hilbert 1912-13 Hilbert, David. "Molekulartheorie der Materie." Ms. Vorlesung WS 1912-13, annotated by M. Born, Nachlass Max Born #1817, Stadtbibliothek Berlin.

Hilbert 1913 Hilbert, David. "Elektronentheorie." Ms. Vorlesung SS 1913, Bibliothek des Mathematischen Seminars, Göttingen.

Hilbert 1915 Hilbert, David. "Die Grundlagen der Physik. (Erste Mitteilung)." *Königliche Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Nachrichten* (1915):395-407.

Hilbert 1917 Hilbert, David. "Die Grundlagen der Physik (Zweite Mitteilung)." *Königliche Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Nachrichten* (1917):53-76.

Hilbert 1924 Hilbert, David. "Die Grundlagen der Physik." *Mathematische Annalen* 92:1-32.

Hilbert 1932-35 Hilbert, David, ed. *Gesammelte Abhandlungen*. 3 vols. Berlin: Springer.

Hilbert 1971 Hilbert, David. "Über meine Tätigkeit in Göttingen." In *Hilbert-Gedenkenband*, edited by K. Reidemeister. Berlin, Heidelberg, New York: Springer, 79-82.

Howard and Norton 1993 Howard, Don, and John D. Norton. "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.

Kerschensteiner 1887 Kerschensteiner, ed. *Paul Gordan's Vorlesungen über Invariantentheorie. Zweiter Band: Binäre Formen.* Leipzig: Teubner.

Klein 1917 Klein, Felix. "Zu Hilberts erster Note über die Grundlagen der Physik." *Königliche Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Nachrichten* (1917):469-482.

Klein 1918a Klein, Felix. "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.

Klein 1918b Klein, Felix. "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.

Klein 1921 Klein, Felix. "Zu Hilberts erster Note über die Grundlagen der Physik." In *Gesammelte Mathematische Abhandlungen*, edited by R. Fricke A. Ostrowski. Berlin: Julius Springer, 553-567.

Klein et al. 1995 Klein, Martin J., Anne J. Kox, Jürgen Renn, and Robert Schulmann, eds. *The Collected Papers of Albert Einstein. Volume 4. The Swiss Years: Writings, 1912-1914.* Princeton: Princeton University Press.

Komar 1958 Komar, Arthur. "Construction of a Complete Set of Independent Observables in the General Theory of Relativity." *Physical Review 111*:1182-1187.

Kox, Klein, and Schulmann 1996 Kox, Anne J., Martin J. Klein, and Robert Schulmann, eds. *The Collected Papers of Albert Einstein. Volume 6. The Berlin Years: Writings, 1914-1917.* Princeton, N.J.: Princeton University Press.

Kretschmann 1917 Kretschmann, Erich. "Über den physikalischen Sinn der Relativitätspostulate, A. Einsteins neue und seine ursprüngliche Relativitätstheorie." *Annalen der Physik* 53 (16):575-614.

Laue 1920 Laue, Max von. "Theoretisches über neuere optische Beobachtungen zur Relativitätstheorie." *Physikalische Zeitschrift* 21:659-662.

Lichnerowicz 1946 Lichnerowicz, André. "Sur le caractère euclidien d'espaces-temps extérieurs statiques partout réguliers." *Academie des Sciences (Paris). Comptes Rendus* 222:432-436.

MacCullagh, Jellet, and Haughton 1880 MacCullagh, James, John Hewitt Jellet, and Samuel Haughton, eds. *The Collected Works of James MacCullagh.* Dublin: Figgis Hodges.

Mehra 1974 Mehra, Jagdish. *Einstein, Hilbert, and the Theory of Gravitation. Historical Origins of General Relativity Theory.* Dordrecht, Boston: D. Reidel Publishing Company.

Mie 1912a Mie, Gustav. "Grundlagen einer Theorie der Materie. Erste Mitteilung." *Annalen der Physik* 37:511-534.

Mie 1912b Mie, Gustav. "Grundlagen einer Theorie der Materie. Zweite Mitteilung." *Annalen der Physik* 39:1-40.

Mie 1913 Mie, Gustav. "Grundlagen einer Theorie der Materie. Dritte Mitteilung." *Annalen der Physik* 40:1-66.

Noether 1918 Noether, Emmy. "Invariante Variationsprobleme." *Königliche Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Nachrichten* (1918):235-257.

Norton 1984 Norton, John D. "How Einstein Found His Field Equations, 1912-1915." *Historical Studies in the Physical Sciences* 14:253-316.

Pais 1982 Pais, Abraham. *'Subtle is the Lord ...' The Science and the Life of Albert Einstein.* Oxford, New York, Toronto, Melbourne: Oxford University Press.

Pauli 1921 Pauli, Wolfgang. "Relativitätstheorie." In *Encyklopädie der mathematischen Wissenschaften, mit Einschluss ihrer Anwendungen,* edited by Arnold Sommerfeld. Leipzig: B. G. Teubner, 539-775.

Pauli 1958 Pauli, Wolfgang. *Theory of Relativity.* Translated by G. Field. London: Pergamon.

Renn 1994 Renn, Jürgen. "The Third Way to General Relativity." Preprint, *Max Planck Institute for the History of Science,* Berlin.

Renn and Sauer 1996 Renn, Jürgen, and Tilman Sauer. "Einsteins Züricher Notizbuch." *Physikalische Blätter* 52:865-872.

Renn and Sauer 1999 Renn, Jürgen, and Tilman Sauer. "Heuristics and Mathematical Representation in Einstein's Search for a Gravitational Field Equation." In *Goenner et al. 1999*, 87-125.

Rosenfeld 1940 Rosenfeld, Leon . "Sur le tenseur d'impulsion-énergie." *Mémoires de l'Academie royale de Belgique* 18 (16):1-30.

Rowe (Forthcoming) Rowe, David. "The Göttingen Response to General Relativity and Emmy Noether's Theorems." In *The Visual World: Geometry and Physics (1900-1930)*, edited by J. J. Gray. Oxford: Oxford University Press, .

Rowe 1989 Rowe, David. "Klein, Hilbert, and the Göttingen Mathematical Tradition." *Osiris* 5:186-213.

Sauer 1999 Sauer, Tilman. "The Relativity of Discovery: Hilbert's First Note on the Foundations of Physics." *Archive for History of Exact Sciences* 53:529-575.

Schouten 1935 Schouten, J. A. *Algebra und Übertragungslehre*. Vol. 1, *Einführung in die neueren Methoden der Differentialgeometrie*. Groningen, Batavia: P. Noordhoff.

Schulmann et al. 1998 Schulmann, Robert, Anne J. Kox, Michel Janssen, and József Illy, eds. T*he Collected Papers of Albert Einstein. Volume 8. The Berlin Years: Correspondence, 1914-1918*. Princeton, N.J.: Princeton University Press.

Schwarzschild 1916 Schwarzschild, Karl. "Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie." *Königlich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (1916) (VII):189-196.

Siegmund-Schultze 1998 Siegmund-Schultze, Reinhard. *Mathematiker auf der Flucht vor Hitler: Quellen und Studien zur Emigration einer Wissenschaft*. Vol. 10, *Dokumente zur Geschichte der Mathematik*. Braunschweig Wiesbaden: Vieweg.

Slebodzinski 1931 Slebodzinski, W. "Sur les equations de Hamilton." *Bulletin de l'Academie royale de Belgique* (5) (17):864-870.

Stachel 1989 Stachel, John. "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.

Stachel 1992 Stachel, John. "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.

Stachel 1994 Stachel, John. "Scientific Discoveries as Historical Artifacts." In *Current Trends in the Historiography of Science*, edited by Kostas Gavroglu. Dordrecht, Boston: Reidel, 139-148.

Thorne 1994 Thorne, Kip S. *Black Holes and Time Warps: Einstein's Outrageous Legacy.* New York, London: Norton.

Vizgin 1989 Vizgin, Vladimir P. "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.

Vizgin 1994 Vizgin, Vladimir P. *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 1999 Walter, Scott. "Minkowski, Mathematicians, and the Mathematical Theory of Relativity." In *Goenner et al. 1999*, 45-86.

Weitzenböck 1920 Weitzenböck, R. "Über die Wirkungsfunktion in der Weyl'schen Physik." *Akademie der Wissenschaften (Vienna). Mathematisch-naturwissenschaftliche Klasse. Sitzungsberichte* 129:683-696.

Weyl 1917 Weyl, Hermann. "Zur Gravitationstheorie." *Annalen der Physik* 54:117-145.

Weyl 1918a Weyl, Hermann. *Raum-Zeit-Materie. Vorlesungen über allgemeine Relativitätstheorie*. 1st ed. Berlin: Julius Springer.

Weyl 1918b Weyl, Hermann. *Raum-Zeit-Materie. Vorlesungen über allgemeine Relativitätstheorie*. 2nd ed. Berlin: Julius Springer.

Weyl 1918c Weyl, Hermann. "Gravitation und Elektrizität." *Königlich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (1918):465-480.

Weyl 1919 Weyl, Hermann. *Raum-Zeit-Materie. Vorlesungen über allgemeine Relativitätstheorie*. 3rd, revised ed. Berlin: Julius Springer.

Weyl 1921 Weyl, Hermann. *Raum-Zeit-Materie. Vorlesungen über allgemeine Relativitätstheorie.* 4th ed. Berlin: Julius Springer.

Weyl 1923 Weyl, Hermann. *Raum-Zeit-Materie. Vorlesungen über allgemeine Relativitätstheorie.* 5th, revised ed. Berlin: Julius Springer.