JÜRGEN RENN AND TILMAN SAUER

PATHWAYS OUT OF CLASSICAL PHYSICS

Einstein's Double Strategy in his Search for the Gravitational Field Equation

CONTENTS

Jürgen Renn (ed.). *The Genesis of General Relativity,* Vol. 1
Einstein's Zurich Notebook: *Introduction and Source*.

1. INTRODUCTION

1.1 The Incomplete Revolution

The relativity revolution was far from complete when Einstein published his path-breaking paper on the electrodynamics of moving bodies in 1905. It started with his reinterpretation of Lorentz's theory of electromagnetism in what may be called a "Copernicus process" in analogy to the transition from the Ptolemaic to the Copernican world system or to the transition from preclassical to classical mechanics.[1] In such a transition the formalism of an old theory is largely preserved while its semantics change.[2] Einstein's special theory of relativity of 1905 had altered the semantics of such fundamental concepts like space and time, velocity, force, energy, and momentum, but it had not touched Newton's law of gravitation. Since, however, according to special relativity, physical interactions cannot propagate faster than light, Newton's well-established theory of gravitation, based on instantaneous action-at-a-distance, was no longer acceptable after 1905. The relativity revolution was completed only when this conflict was resolved ten years later in November 1915 with Einstein's formulation of the general theory of relativity.

Neither the emergence of the special theory of relativity nor that of the general theory of relativity were isolated achievements. The virtual simultaneity of the beginning of the relativity revolution with Einstein's other breakthrough discoveries of 1905 indicate that his non-specialist outlook and, in particular, his youthful pursuit of atomistic ideas enabled him to activate the hidden potentials of highly specialized nineteenth-century physics that others, such as Henri Poincaré, had also exposed.[3] In 1907, Einstein first attempted to address the issue as to how to modify Newton's law of gravitation according to the new kinematic framework of special relativity, as did others, like Hermann Minkowski and Henri Poincaré.[4] But Einstein began to transcend the very special-relativistic framework in light of Galileo's insight that in a vacuum all bodies fall with the same acceleration. In 1912, to the amazement of colleagues like Max Abraham,[5] he abandoned the scalar gravitational potential of Newtonian physics in favor of a ten-component object—the metric tensor—the mathematics of which he subsequently began to explore with the help of his mathematician friend Marcel Grossmann. And he was able to formulate clear-cut criteria which a field equation for the metric tensor acting as a gravitational potential would have to satisfy. However, in the winter of 1912–1913, Einstein and Grossmann dis-

---

1 Cf. (Damerow et al. 2004, Renn 2004).

2 Such a change of semantics may be illustrated with the example of Lorentz's concept of local time. Originally merely a peripheral aspect of his theory, this auxiliary variable was reinterpreted by Einstein as the time actually measured by clocks in a moving reference system, thus assuming a central role in the new kinematics of special relativity. For an extensive treatment of the first phase of the relativity revolution compatible with this view, see, e.g., (Janssen 1995).

3 See (Renn 1993, 1997). For the parallelism between Einstein and Poincaré, see also (Galison 2003).

4 See Scott Walter's "Breaking in the 4-vectors ..." (in vol. 3 of this series) and (Katzir 2005).

5 See (Cattani and De Maria 1989a) and "The Summit Almost Scaled ..." (in vol. 3 of this series).

carded generally-covariant field equations based on the Riemann tensor, an expression that included second-order derivatives of the metric tensor. Einstein even believed to have a proof that such field equations had to be ruled out, although in hindsight these were the only acceptable mathematical solution. In spite of the skepticism of many of his physics colleagues but supported by the critical sympathy of mathematicians like Tullio Levi-Civita and David Hilbert,[6] Einstein stood by his original agenda and in late 1915 returned to field equations based on the Riemann tensor, finally formulating the general theory of relativity, a theory which became the basis of all subsequent developments in physics and astronomy.

Einstein's Zurich Notebook represents a uniquely valuable and, as it turns out, surprisingly coherent,[7] record of his thinking in an intermediate phase of the emergence of general relativity. The entries begin in mid-1912 and end in early 1913. His aim during this period was to create a relativistic theory of gravitation that makes sense from a physical point of view and that, at the same time, corresponds to a consistent mathematical framework based on the metric tensor. Central to his thinking was the problem of interpreting the physical knowledge on gravitation in terms of a generalization of the mathematical representation associated with Minkowski's four-dimensional spacetime. The main challenge he faced was to construct a field equation, on the one hand, that can be reduced by an appropriate specialization to the familiar Newtonian law of gravitation, and, on the other hand, that satisfies the requirements resulting from his ambitious program to formulate a relativistic theory of gravitation.

There is perhaps no single episode that better illustrates the conceptual turn associated with the genesis of general relativity than the fact that, in the Zurich Notebook, Einstein first wrote down a mathematical expression close to the correct field equation and then discarded it, only to return to it more than three years later. Why did he discard in the winter of 1912–1913 what appears in hindsight to be essentially the correct gravitational field equation, and what made this field equation acceptable in late 1915?[8] Our analysis of the Zurich Notebook has made it possible not only to answer these questions but, more generally, to resolve what might be called the three epistemic paradoxes raised by the genesis of general relativity:

The paradox of missing knowledge. How was it possible to create a theory such as general relativity that was capable of accounting for a wide range of phenomena which were only later discovered in the context of several revolutions of observational astronomy? If neither the expansion of the universe, black holes, gravitational lenses, nor gravitational radiation were known when Einstein set up the gravitational field equation, how could he nevertheless establish such a firm foun-

6 See Einstein's correspondence with Levi-Civita and Hilbert in (CPAE 8). For further discussion, see also (Cattani and De Maria 1989b) and (Corry 2004).

7 See the "Commentary ..." (in vol. 2 of this series) and especially our reconstruction in section 6 of the present chapter.

8 For further discussion of these two questions, see also "Commentary ..." sec. 5, and "Untying the Knot ..." (both in vol. 2 of this series).

dation for modern cosmology? Which knowledge granted such stability to a theory that did not initially seem superior to its competitors, since no phenomena were known at the time which could not also be explained with traditional physics?

The paradox of deceitful heuristics. After a tortuous search in the course of which he even temporarily abandoned hope of ever solving his problem, how was Einstein able to formulate the criteria for a gravitational field equation years before he established the solution? How could he establish a heuristic framework that would quickly lead him to a correct mathematical expression, and then to the conclusion that it was unacceptable, only to bring him back to essentially the same expression three years later?

The paradox of discontinuous progress. How could general relativity with its nonclassical consequences—such as the dependence of space and time on physical interactions—be the outcome of classical and special-relativistic physics although such features are incompatible with their conceptual frameworks?

Addressing the challenges which these paradoxes formulate requires taking into account all of the following dimensions that are crucial to a historical epistemology of scientific knowledge: the long-term character of knowledge development, the complex architecture of knowledge, and the intricate mechanisms of knowledge dynamics. In order to resolve these paradoxes and to adequately describe the reorganization of knowledge occurring between 1912 and 1915, we shall, in particular, make use of concepts from cognitive science, adapted to the description of the structures of shared knowledge resources such as those Einstein adopted from classical and special-relativistic physics. These concepts will be used to analyze the architecture of the knowledge relevant to Einstein's search for a gravitational field equation and to explain its restructuring as a result of the interaction with the mathematical representation of this knowledge.

We intend to show in the following that the history of Einstein's search for a gravitational field equation can, against the background of the Zurich Notebook, be written as that of a mutual adaptation of mathematical representation and physical meaning. The eventual success of this adaptation becomes intelligible only if it is conceived of as part of a long-term process of integrating intellectual resources relevant to Einstein's problem that were rooted in the shared knowledge of classical and special-relativistic physics.

Only by analyzing the complex architecture of these shared knowledge resources is it possible to understand in which sense classical and special-relativistic knowledge about gravitation and inertia, energy and momentum conservation, and the relation between different reference frames, was turned into a heuristic framework for Einstein's search. In the course of his work, elements of this heuristic framework crystallized into a double strategy that shaped his search in an essential way until he succeeded in formulating the definitive field equation of general relativity in November 1915.

The identification of the two components of this double strategy has not only allowed us to reconstruct Einstein's notes and calculations in the Zurich Notebook as traces of a surprisingly coherent research process, but also to analyze the dynamics of this process. It has become clear, in particular, how a combination of knowledge resources rooted in classical and special-relativistic physics could give rise to the theory of general relativity whose conceptual foundation is no longer compatible with the knowledge that formed the starting point of Einstein's search. In this way, the genesis of general relativity can be understood as resulting from a transformation of shared resources of knowledge, while Einstein's search for the gravitational field equation appears as an investigation of pathways out of classical physics.

In this introduction, we shall briefly recapitulate the essential elements of our story.[9] We begin with a review of the principal steps taken by Einstein towards a relativistic theory of gravitation between the years 1907 and 1912 before his research is documented in the Zurich Notebook.[10] Here our aim is to show that each of these steps highlighted knowledge resources that were relevant for addressing the challenge of constructing a relativistic theory of gravitation. The heuristics at work in the Zurich Notebook were the result of this prior research experience. We shall then offer a first description of the crucial role played by Einstein's double strategy for his heuristics and finally introduce the epistemological framework for our analysis of how exactly this strategy worked.

In the second section, we shall discuss what we will call Lorentz model, as the conceptual framework for Einstein's construction of a relativistic field theory of gravitation. In the third section, we shall examine the essential elements of his heuristics, showing in which sense these elements turned knowledge resources of classical and special-relativistic physics into key components of Einstein's search. In the fourth section, we shall analyze how this search process was structured by the way in which the Lorentz model functioned as a mental model in the sense of cognitive science. In the fifth section, we shall examine how the candidates for a gravitational field equation that Einstein considered in the course of his search fared in the light of the heuristic criteria he had established on the basis of his prior research experience. This discussion will help to understand why one and the same candidate fared differently depending on the depth to which Einstein had explored the implications of the mathematical representation. In the sixth section, we shall reconstruct Einstein's pathway, as documented in the Zurich Notebook, as a learning experience in which he passed from one candidate field equation to the other, building up strategic devices that would guide him until he reached his final result in 1915. In the seventh section, we

---

9 For Einstein's own account, see (Einstein 1933). The history of general relativity has been an intensive subject of research in the last decades, see, in particular, the contributions in (Stachel and Howard 1989–2006). Especially with respect to Einstein's own path, early contributions were (Hoffmann 1972, Lanczos 1972, Mehra 1974, Earman and Glymour 1978, Vizgin and Smorodinski 1979, Pais 1982, sec. IV., Stachel 1980, 1982), a groundbreaking paper was (Norton 1984). See also (Capria 2005, Howard and Norton 1992, Janssen 1999, 2005, Maltese 1991, Maltese and Orlando 1995, Miller 1992, Norton 1992a, 1992b, 1999, 2000, Renn 2005b, 2005c, Renn and Sauer 1996, 1999, 2003a, Sauer 2005b, Stachel 1987, 1989b, 1995, 2002, Vizgin 2001).

10 For detailed analyses of this part of the story, see "The First Two Acts" and "Classical Physics in Disarray …" (both in this volume).

shall turn to Einstein's elaboration of the so-called *Entwurf* theory, published in 1913 as the result of the research documented in the Zurich Notebook. It will be shown, in particular, how the work on this problematic theory created the preconditions for the conceptual changes of the final theory of general relativity. In the concluding eighth section, we shall review our reconstruction with a view to pinpointing the essential structures of this scientific revolution.

1.2 The Emergence of a Heuristic Framework

The incompatibility between Newton's theory of gravitation and the special theory of relativity of 1905 presented Einstein and his contemporaries with the task of constructing a relativistic theory of gravitation. Special relativity, for the purpose of our account, arose from the confrontation of classical mechanics and classical electrodynamics as two major knowledge blocks, i.e. from the confrontation of two highly elaborated, individually consistent, and empirically well-confirmed systems of knowledge whose simultaneous validity had nevertheless produced inconsistencies and contradictions. The newly established mathematical and conceptual framework of special relativity added to the physical knowledge available for dealing with the problem of a relativistic theory of gravitation. The knowledge blocks of classical mechanics and electrodynamics and of special relativity offered various points of departure for the continuation of the relativity revolution in coming to terms with the problem of gravitation.

The new spatio-temporal framework of special relativity suggested a plausible mathematical procedure for adapting the classical theory of gravitation to the requirements of a relativistic field theory. In classical physics, the Poisson equation determines the Newtonian gravitational potential by a given distribution of the masses that act as the sources of the gravitational field (which in turn can be derived from the gravitational potential).[11] This equation is not invariant with respect to the Lorentz transformations of special relativity. But the Poisson equation can easily be extended in a formal way to a relativistic field equation by adding a differential operator involving the time coordinate. The problem with this obvious generalization was that the resulting theory of gravitation no longer incorporates Galileo's principle according to which all bodies fall with the same acceleration. The most obvious way of bringing gravitation within the purview of the relativity revolution therefore came at the price of having to give up one of the fundamental insights of classical mechanics.

At this point, classical mechanics provided knowledge resources that were turned into an alternative heuristic starting point for the continuation of the relativity revolution. In 1907 Einstein formulated his principle of equivalence as a heuristic device

---

11 The Poisson equation, being an equation for the gravitational potential, should properly be called a potential equation. However, since the Einstein equations are commonly referred to as "field equations" rather than "potential equations," we will in the following loosely also refer to the Poisson equation as a "field equation."

that allowed him to incorporate Galileo's principle into a relativistic theory of gravitation.[12] The equivalence principle asserts that it is not possible to distinguish between a uniformly and rectilinearly accelerated reference frame without gravitational fields and an inertial system with a static and homogeneous gravitational field. Accordingly, the problem of a revision of the classical theory of gravitation became associated with that of a generalization of the relativity principle to accelerated motion, which henceforth constituted another heuristic guideline for Einstein's further research.[13]

Between 1907 and 1911 Einstein used the equivalence principle to derive several consequences of his yet to be formulated new gravitation theory.[14] By the spring of 1912, he made a first attempt at formulating a theory for a static but otherwise arbitrary gravitational field.[15] The gravitational field equation of this theory was a straightforward modification of the Poisson equation of classical physics. Since the Poisson equation embodies the classical knowledge of gravitation from Newtonian theory, it formed a crucial asset for Einstein's heuristics. The further elaboration of Einstein's theory of the static field met with great difficulties. He found that this theory was incompatible with the conservation of energy and momentum, another pillar of classical physics. This led to another key element of his heuristic framework, the requirement that the conservation laws must be fulfilled.

The various heuristic requirements serving as different starting points for the search for a relativistic theory of gravitation could lead into different directions, confronting it with different obstacles and different intermediate results, as well as leading perhaps to different solutions to the original problem. After finding a more or less satisfactory theory of the static field, Einstein further pursued the heuristics embodied in the equivalence principle and in the knowledge about field theory available in classical physics. This approach led him to consider uniformly rotating reference frames.[16] As with linearly accelerated motion, he sought to interpret the inertial forces occurring in such reference frames as generalized gravitational forces. This interpretation was made plausible by Mach's critical analysis of classical mechanics. But the conceptual and technical difficulties implied by the inclusion of rotating reference frames prevented, for the time being, the formulation of a gravitation theory that covered this more general case as well. In hindsight, it is clear that a response to the difficulties which Einstein encountered required the introduction of more sophisticated mathematical tools. The heuristics based on the equivalence principle led to

12 See (Einstein 1907). For historical discussion, see (Miller 1992).
13 For a discussion of the problematic relation between the equivalence principle and the generalization of the relativity principle, see, for instance, secs. 1.1.1–1.1.2 of "Commentary" (in vol. 2 of this series) and (Janssen 2005, 61–74).
14 See (Einstein 1911).
15 See (Einstein 1912b) and, for historical discussion, (CPAE 4, 122).
16 The crucial role of rotating reference frames in recognizing the role of non-Euclidean geometry was first discussed in (Stachel 1980), see also (Maltese and Orlando 1995).

substantial but isolated physical insights, and not to the kind of coherent mathematical framework necessary for formulating a relativistic field theory of gravitation.

A different path had meanwhile been followed by Max Abraham, who exploited heuristic clues of the four-dimensional mathematical framework established by Minkowski for special relativity.[17] Abraham succeeded in developing a comprehensive theory of gravitation through an *ad-hoc* modification of this framework. Einstein soon discovered weaknesses in Abraham's theory. After a controversy with Abraham, he realized that a successful application of Minkowski's formalism to the problem of gravitation called for a mathematical generalization of this formalism. In late spring 1912 Einstein found the appropriate starting point for such a generalization of Minkowski's formalism. In the appendix to the last paper he published before the considerations documented in the Zurich Notebook, he formulated the equation of motion in a static gravitational field in a form that suggested that a generalization of his theory of gravitation would involve non-Euclidean geometry as had been formulated by Gauss for curved surfaces. As early as summer 1912 Einstein succeeded in formulating a generally-covariant equation of motion for a test particle in an arbitrary gravitational field. In this equation, the gravitational potential is represented by a four-dimensional metric tensor, which became the key object for Einstein's further research in the following years.

The search for a relativistic gravitational field equation, which occupied Einstein for the following three years, also involved a new role of the heuristic clues that had so far guided the research of Einstein and his contemporaries. Initially, these heuristic clues were more or less isolated hints. They gradually turned into elements of a more systematic research program, characterized by what we have called Einstein's "double strategy." This double strategy allowed him to attack the problem of finding a gravitational field equation by bringing to bear on this problem the entire range of knowledge resources embodied in the various heuristic elements sketched above.

1.3 The Double Strategy

The mathematical difficulty of finding a field equation for the ten-component metric tensor representing the gravitational potential showed Einstein that he needed much more sophisticated mathematical methods than those available to him at that point. A mathematical formalism providing what Einstein's generalized theory of relativity required had been developed in the second half of the 19th century by Gauss, Riemann, and Christoffel. In a paper published in 1901 by Ricci and Levi-Civita on the so-called absolute differential calculus, the work of these mathematicians had been extended to an elaborate mathematical apparatus.[18] However, among physicists, the absolute differential calculus remained largely unknown for a considerable time. Einstein was certainly not familiar with it until mid-1912.[19] Only after his move from

---

17 For a more detailed treatment, see "The Summit Almost Scaled ..." (in vol. 3 of this series).
18 See (Ricci and Levi Civita 1901).

Prague to Zurich did he gain access to these mathematical methods through his contact with Marcel Grossmann. In October 1912, he wrote to Arnold Sommerfeld:

> I am now working exclusively on the gravitation problem and believe that I can overcome all difficulties with the help of a mathematician friend of mine here. But one thing is certain: never before in my life have I troubled myself over anything so much, and I have gained enormous respect for mathematics, whose more subtle parts I considered until now, in my ignorance, as pure luxury! Compared with this problem, the original theory of relativity is child's play.[20]

The mathematical difficulty of finding a satisfactory relativistic field equation also gave a new role to the physical requirements that such an equation had to satisfy. These physical requirements had to be translated into mathematical conditions to be satisfied by candidate field equations. They were thus also brought into systematic relations with each other. This translation was by no means unambiguous, since Einstein was exploring an as yet largely unknown territory of knowledge. At the same time, the theory had to preserve the physical knowledge on gravitation already available, and its relation to other parts of physics, and it had to be formulated as a mathematically consistent theory built, according to Einstein's insight of 1912, around the four-dimensional metric tensor. This combination of relatively clear-cut conditions and the incompleteness of the information needed to turn the situation into a fully determined mathematical problem was characteristic of Einstein's situation when he began the search for a field equation as documented in the Zurich Notebook. The search strategy that gradually emerged enabled a mutual adaptation of mathematical representation and physical concepts, and provided a heuristic device that eventually turned out to be the adequate response to this situation.

What conditions to be imposed on a relativistic gravitational field equation for the metric tensor had emerged from Einstein's prior research experience? From the mathematical point of view, the task was to find a differential operator of second order for the metric tensor covariant with respect to the largest possible class of coordinate transformations. The requirement that the candidate differential operator has to be of second order follows from the analogy with the classical theory of gravitation: the Poisson equation for the Newtonian gravitational potential is a differential equation of second order. The requirement of the covariance of this differential operator under a broad class of coordinate transformations represented for Einstein the goal of a gen-

---

19 The works by Bianchi (1910) and Wright (1908) probably served as Einstein's mathematical reference books. For historical discussion, see (Reich 1994).

20 "Ich beschäftige mich jetzt ausschliesslich mit dem Gravitationsproblem und glaube nun mit Hilfe eines hiesigen befreundeten Mathematikers aller Schwierigkeiten Herr zu werden. Aber das eine ist sicher, dass ich mich im Leben noch nicht annähernd so geplag[t] habe, und dass ich grosse Hochachtung für die Mathematik eingeflösst bekommen habe, die ich bis jetzt in ihren subtileren Teilen in meiner Einfalt für puren Luxus ansah! Gegen dies Problem ist die ursprüngliche Relativitätstheorie eine Kinderei." Einstein to Arnold Sommerfeld, 29 October 1912, (CPAE 5, Doc. 421). Unless otherwise noted, all translations are based on the English companion volumes to the *Collected Papers of Albert Einstein*.

eralized relativistic theory in which, if possible, all reference frames would be equivalent.[21] A further requirement was that the classical field equation emerge as a special case of the relativistic field equations under appropriate restrictive conditions, such as for weak and static fields. The heuristic framework furthermore included general physical principles such as Galileo's principle and the laws of energy and momentum conservation applying to the energy and momentum of the gravitational field as well.

These requirements formed the relatively stable framing conditions shaping Einstein's search for the gravitational field equation from its beginning in summer 1912 to the formulation of the eventual solution in late 1915. His main problem was to ensure the compatibility of these different heuristic components by integrating them into a coherent gravitation theory represented by a consistent mathematical framework. It turned out that again and again, in the course of his investigations, only some of Einstein's heuristic goals could be fully realized while others had to be given up or at least modified. If not all of his goals could be satisfied, the appropriate balance between the different heuristic requirements for a gravitational field theory could not be decided *a priori*. Their relative weight could only be judged by their concrete embodiment in candidate gravitational field theories.

Physical properties or mathematical statements could each be looked upon either as principles of construction for the building blocks of the theory or as criteria by which the acceptability of such building blocks could be checked. It is this double perspective that provided the basis for the double strategy that emerged in the course of Einstein's search for the gravitational field equation, as documented in the Zurich Notebook. Earlier the choice between physically or mathematically motivated expressions had been a choice between entirely different approaches to the problem of gravitation. Einstein's 1912 theory of static gravitational field was, for instance, motivated by physical considerations based on the equivalence principle, while Abraham's theory started from mathematical considerations related to Minkowski's formalism. In the course of Einstein's work documented in the Zurich Notebook, the two approaches gradually grew closer and turned into complementary strategies of a more or less systematic research program. In this research program the two approaches were distinguished mainly by the sequence in which the building blocks of the theory come into play. Einstein's "physical strategy" took the Newtonian limiting case as its starting point, then turned to the problem of the conservation of energy and momentum and only then examined the degree to which the principle of relativity is satisfied. His "mathematical strategy," took the principle of relativity as its starting point and only then turned to the Newtonian limiting case and the conservation of energy and momentum. The reconstruction of Einstein's notes in the Zurich Notebook has made it evident that his search for the gravitational field equation is to a large extent determined by the exploration of the possibilities offered by these alternatives.

---

21 See (Norton 1999) for a discussion about Einstein's ambiguity regarding the difference between invariance and covariance during this period.

Einstein's oscillation between these two strategies is characteristic not only of his approach in the notebook but of his entire struggle with the problem of gravitation between 1912 and 1915, a struggle that brought him from his 1912 static theory, via the *Entwurf* theory of 1913, to the final theory of general relativity.[22]

This oscillation between the physical and the mathematical strategy suggests that his search for the gravitational field equation was not just a matter of resolving a well-defined mathematical problem, but involved an interaction between mathematical representation and physical concepts that affects the structures of the mathematical and physical knowledge. Why else did Einstein's first attempts along the mathematical strategy in the winter of 1912–1913 fail, while his pursuit of the physical strategy seemed to be essentially successful, at least until the demise of the *Entwurf* theory in late 1915?[23] As we will show in detail, the completion of the general theory of relativity required, in addition to the appropriation of the available mathematical knowledge, a revision of foundational concepts of physics, the extent of which Einstein could hardly have foreseen at the beginning of his search. He initially believed that classical physics would provide the appropriate context for the theory to be found and attempted to formulate a gravitational field equation by immediate generalization of familiar Newtonian concepts. It eventually turned out to be more successful to construct a field equation corresponding to Einstein's program of integrating gravitation and relativity than to relate them to the conceptual foundations of classical theory.

1.4 The Epistemological Framework of the Analysis

How was it possible for Einstein to formulate a theory involving conceptual novelties on the basis of knowledge that was still anchored in the older conceptual foundation of classical physics? Such a development can hardly be described in terms of formal logic. As Einstein's investigative pathway illustrates, scientific conclusions can result in a reconceptualization of the premises on which these conclusions were based. Even in cases involving major restructuring of knowledge, science never starts from scratch. In fact, not only scientific knowledge but also the knowledge of large domains of human experience transmitted over generations is not simply lost when new scientific theories replace the old ones. In the case at hand, the knowledge of classical physics had to be preserved and exploited in a conceptual revolution, the outcome of which was a relativistic theory of gravitation whose far-reaching physical implications were largely unknown when it was created. But they eventually changed our understanding of the universe. An adequate description of the cognitive dynamics of the genesis of general relativity therefore requires an account of the knowledge that

---

22 Very similar characteristics of a physical and mathematical double strategy have also been identified in Einstein's later work on unified field theory, see (van Dongen 2002, 2004) and (Sauer 2006) for further discussion. See also the discussion in (Norton 2000).

23 And, as is argued in "Untying the Knot …" (in vol. 2 of this series), beyond the demise of this theory as well.

makes it understandable. We have to understand, first, how past experiences can enter inferences about matters for which only insufficient information is available, and, second, how conclusions can be corrected without eventually having to start from scratch each time a premise is found to be wanting, with the possibility that the whole deductive structure changes in the process. Such an approach is offered by an historical epistemology that integrates the methodology of historical analysis with a theoretical framework informed by philosophical epistemology and cognitive science.

In order to adequately account for the features of Einstein's search for the gravitational field equation described above, we will in the following make use in particular of the concept of a "mental model" and the concept of a "frame."[24] A mental model for us is an internal knowledge representation structure serving to simulate or anticipate the behavior of objects or processes. It possesses "terminals" or "slots" that can be filled with empirically gained information, but also with default assumptions resulting from prior experience. The default assumptions can be replaced in light of new information, so that inferences based on the model can be corrected without abandoning the model as a whole. Information is assimilated to the slots of a mental model in the form of "frames." These are chunks of knowledge which themselves are equipped with terminals and which have a well-defined meaning anchored in a given body of shared knowledge.

Mental models can, as a rule, be externally represented by material models which also serve as the element of continuity in their transmission from one generation to the next. The basic features of the field-theoretical model of distant causation, which will play a central role in our analysis, may, for instance, be represented by the material model of a magnet setting a piece of iron into motion by affecting the state of its environment. In addition, it may be represented by symbolic representations making use of natural and formal language. The internal architecture of a system of knowledge is constituted by a network of mental models and frames that can be linked by operations in the sense of mental acts typically corresponding to handling external representations, be they material arrangements or symbolic expressions. A sequence of such operations constitutes a procedure which typically has a goal, for instance of creating an *ad hoc* knowledge representation structure, which is called a "real-time construction" in cognitive science. A real-time construction may be exemplified by the geometrical construction typically accompanying the Euclidean proof of a geometrical theorem or by a set of mathematical expressions corresponding to checking a candidate field equation according to one of Einstein's heuristic principles.

Two fundamentally important types of mental acts are "chunking" and "reflection." By chunking different knowledge representation structures are combined into a unity. This often leads to a linguistic representation of the resulting chunk by a technical term designating, for instance, a particular procedure. By reflection, the usage

24 For the concepts of frame and mental model, see (Minsky 1975, 1987; Damerow 1996; Gentner and Stevens 1983; and Davis 1984). For a view on the potential of cognitive science and cognitive psychology for the history of science to which the present work is much indebted, see (Damerow 1996).

of knowledge representation structures becomes the object of reasoning; it typically presupposes an external representation of these structures, for instance by a technical term. Reflection obviously plays a crucial role in accommodating a system of knowledge to new experiences by changing its architecture. An example is what we call a "Copernicus process" in which the internal network of a system of knowledge is essentially preserved while originally peripheral elements take on a central role in the deductive structure. The status of such elements as being either peripheral or central is prescribed by a "control structure." A control structure is constituted by any knowledge representation structure serving to control the operation and to order other such structures. In this sense, Einstein's heuristic principles as well as his double strategy may be considered examples of such control structures. These elements of the architecture of knowledge can partially be captured by traditional epistemological terminology. A concept, for instance, may be understood as the linguistic representation of a mental model, a frame, or a particular terminal of a frame, while a theory is just one example of many conceivable control structures. We shall also use of these traditional terms, specifying their meaning in the context of the epistemological framework we have introduced whenever appropriate.

We claim that the shared knowledge of classical and special-relativistic physics can be conceived of in terms of this richer epistemological framework, and that it then becomes understandable how this knowledge could serve as a resource for Einstein's search for the gravitational field equation. We will argue that essential relations between fundamental concepts such as that between field and source remain the same to a great extent even though the concrete applications of these concepts differ considerably from their applications to a classical or a relativistic field equation. This structural stability turned the concepts and principles of classical and special-relativistic physics into guiding principles when Einstein entered unknown terrain, for instance, when he encountered a new expression generated by the elaboration of a mathematical formalism. None of these expressions by themselves constituted a new theory of gravitation. Only by complementing them with additional information based on the experience accumulated in classical and special-relativistic physics, as well as in the relevant branches of mathematics did such expressions become candidates for a gravitational field equation embedded in a full-fledged theory of gravitation. In the language of mental models, such past experience provided the default assumptions necessary to fill the gaps in the emerging framework of a relativistic theory of gravitation. Because of their nature as default assumptions, they could be given up again in the light of novel information without making it necessary to abandon the underlying mental models, which thus continue to play their heuristic role.

In this way we hope to render understandable how a gradual process of knowledge accumulation could overcome the very conceptual foundations that had formed its starting point. The concepts of classical physics shaped Einstein's search at its beginning, and made the physical strategy the most natural approach to exploit his heuristic principles for finding the gravitational field equation. In the context of the physical strategy, the default assumptions of the relevant mental models were supplied by the knowledge of classical and special-relativistic physics. The mathematical strategy, on the other hand, drew on default assumptions based on prior mathematical knowledge and led to candidate field equations whose compatibility with established physical knowledge was problematic. The gradual accumulation of knowledge fostered by both of these approaches enriched the network constituted by the mental models and frames relevant to a relativistic theory of gravitation. Eventually, the reorganization of this network by a Copernicus process became feasible.

In the following, we shall discuss in detail the essential aspects and phases of this process, Einstein's heuristic framework, the gradual accumulation of knowledge in the course of his research, the successive replacement of one candidate gravitational field equation by another, the switches back and forth between the physical and the mathematical strategy, and finally the reinterpretation of the results acquired in this way as aspects of one and the same transformation leading from the system of knowledge of classical to that of general-relativistic physics.

2. THE MENTAL MODEL OF FIELD THEORY

2.1 The Poisson Equation of Classical Mechanics and the Field Equation of General Relativity

The revision of Newton's theory of gravitation confronted Einstein with two fundamental problems. He needed to find an equation of motion for bodies in a gravitational field (the analogue of Lorentz's equation of motion of a charged body in an electromagnetic field) and to find a field equation determining the gravitational field itself (the generalization of the Poisson equation and the analogue of Maxwell's equations relating the electromagnetic field to its sources). These two problems presented themselves in terms of basic concepts and structures of classical physics and the special theory of relativity. These concepts and structures also provided an essential part of the intellectual resources for solving these problems.

The most fundamental structures of knowledge relevant to Einstein's search for a new theory of gravitation were incorporated in the understanding of an equation of motion and of a field equation in classical physics and in the special theory of relativity. This understanding involves concepts such as force, energy, momentum, potential, field, source, and mass.[25] Above all, however, it involves the mental model of field theory which had emerged, in its most mature, successful, and widely accepted form, in Lorentz's electron theory of electrodynamics and hence may also be referred to as the Lorentz model.[26] This model actually comprises two mental models with more ancient roots in the history of physics; one for a field equation and one for the equation of motion. The Lorentz model of a field equation, which will be at the center

25 For Einstein's account of the emergence of fundamental concepts of physics, including that of field, see (Einstein and Infeld 1938).

26 See (Lorentz 1895), and for historical discussion (Whittaker 1951, ch. XIII, 1953, ch. II, Buchwald 1985, Janssen 1995, Darrigol 2000, Janssen and Stachel 2004).

of our analysis of Einstein's search for the gravitational field equation, has slots for the source, the potential, and a differential operator acting on the potential. Default settings for these slots are provided by the classical theory of gravitation which describes the relation between gravitational source and gravitational potential in terms of the Poisson equation. In the classical case, the source-slot and the potential-slot of the frame are filled by scalar functions that can be subsumed under what we might call the potential-frame and the mass-density-frame, respectively. The default setting for the differential-operator-slot is the Laplace operator.

Before we come back to a more detailed examination of the structure of the Lorentz model, we want to justify the introduction of this model by examining some of the basic concepts and knowledge structures relating the Poisson equation in classical mechanics to the Einstein field equation of general relativity. We claim that these common features played an important role in the historical development linking the two equations so that their description by an overarching structure makes historical sense.

The Poisson equation of classical gravitation theory describes how gravitating matter generates a gravitational potential. This potential can then be related to the gravitational field and to the force acting on material particles exposed to it. The Poisson equation is

$$\Delta\varphi = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\varphi = 4\pi\kappa\rho \tag{1}$$

where the gravitational potential is denoted by $\varphi = \varphi(x, y, z)$, which is a function of spatial coordinates x, y, z, where $\rho = \rho(x, y, z)$ denotes the density of gravitating matter, and where κ is a constant. Δ is a linear second-order differential operator, known as the Laplace operator.

The gravitational interaction between material bodies in classical physics can, of course, also be treated directly on the basis of Newton's law of gravitation. This law states that an attractive force between two point particles acts instantaneously along the direction defined by the two bodies and its strength varies inversely proportional to the squared distance between the particles. This action-at-a-distance force can also be calculated from a local potential function φ which is then determined by the Poisson equation introduced above.[27] While the Poisson equation thus appears only as an alternate description of the same physical content as Newton's law, this equation suggests, at the same time, a different physical interpretation of gravitation. According to this interpretation, gravitation—represented by the potential φ and produced by some matter distribution ρ which acts as its source—fills the entire space and exerts its influence on matter locally as a force. By virtue of this interpretation, the Poisson equation can be considered as a first hint at a gravitational field theory, in particular at a time when the field theoretic framework established by Maxwell's electrodynamics

27 Recall (see note 11) that we are loosely referring to the Poisson equation as a "field equation" even though it should properly be called a "potential equation."

suggested a field-theoretic revision of Newtonian gravitation. Nevertheless, the Newtonian gravitational potential lacks two essential features required by a genuine physical field theory. First, the gravitational field does not propagate with a limited speed, a field-theoretical feature that became mandatory after the advent of the theory of special relativity. It also does not describe some expected dynamical effects of gravitation such as dragging effects due to moving masses ("gravitational induction") or gravitational waves.

The Einstein equation stands at the end of a historical process in which the wish to conceive of the gravitational interaction in a truly field-theoretic manner played a significant heuristic role. The Poisson equation and the Einstein equation share a number of common features, in spite of the long and sometimes circuitous discovery process separating the two. In general relativity, the gravitational interaction is also determined by a second-order partial differential equation —the Einstein equation— which relates the gravitational potential to its source.

The Einstein field equation

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R = \bar{\kappa}T^{\mu\nu} \tag{2}$$

written in terms of the Ricci tensor $R^{\mu\nu}$ and the Riemann scalar R can also be written explicitly as

$$\begin{aligned} G^{\mu\nu} = {} & \sum_{i\kappa}\left(g^{\mu i}g^{\nu\kappa} - \frac{1}{2}g^{\mu\nu}g^{i\kappa}\right) \\ & \left\{\frac{1}{2}g^{lm}\left(\frac{\partial^2 g_{im}}{\partial x^\kappa \partial x^l} + \frac{\partial^2 g_{\kappa m}}{\partial x^i \partial x^l} - \frac{\partial^2 g_{i\kappa}}{\partial x^m \partial x^l} - \frac{\partial^2 g_{lm}}{\partial x^i \partial x^\kappa}\right)\right. \\ & + \frac{1}{2}\frac{\partial g^{lm}}{\partial x^l}\left(\frac{\partial g_{im}}{\partial x^\kappa} + \frac{\partial g_{\kappa m}}{\partial x^i} - \frac{\partial g_{i\kappa}}{\partial x^m}\right) - \frac{1}{2}\frac{\partial g^{lm}}{\partial x^\kappa}\frac{\partial g_{lm}}{\partial x^i} \\ & - \frac{1}{4}g^{jm}g^{ln}\left[\left(\frac{\partial g_{ml}}{\partial x^i} + \frac{\partial g_{im}}{\partial x^l} - \frac{\partial g_{li}}{\partial x^m}\right)\left(\frac{\partial g_{jn}}{\partial x^\kappa} + \frac{\partial g_{\kappa n}}{\partial x^j} - \frac{\partial g_{\kappa j}}{\partial x^n}\right)\right. \\ & \left.\left. - \left(\frac{\partial g_{im}}{\partial x^\kappa} + \frac{\partial g_{\kappa m}}{\partial x^j} - \frac{\partial g_{i\kappa}}{\partial x^m}\right)\frac{\partial g_{ln}}{\partial x^j}\right]\right\} = \bar{\kappa}T^{\mu\nu} \end{aligned} \tag{3}$$

where the gravitational potentials are denoted by $g_{\mu\nu}$ and $g^{\mu\nu}$, which are functions of the spacetime coordinates x_i, $(i) = 1, 2, 3, 4$, and where $\bar{\kappa} = -\frac{8\pi\kappa}{c^4}$ is a constant.

Like the left-hand side of the Poisson equation, the Einstein tensor $G^{\mu\nu}$ is a second-order differential operator applied to the gravitational potential, even though the operator in this case is much more complicated than the Laplace operator. $T^{\mu\nu}$ denotes the so-called stress-energy or energy-momentum tensor and corresponds to another element familiar from the Poisson equation, the role of matter as gravitating source. The mass density ρ which functions as the gravitating source in the Poisson equation reappears, for instance, in the following example of an energy-momentum tensor:

$$T^{\mu\nu} = \rho \frac{dx^{\mu}}{ds}\frac{dx^{\nu}}{ds}, \qquad (4)$$

which describes a dust-like cloud of material particles acting as the source of the gravitational field where ρ stands for the mass-density of the swarm and dx^{μ}/ds denotes the special-relativistic four-velocity of the dust particles.

2.2 The Lorentz Model of a Field Equation

In our introduction of the Poisson equation as the point of departure in classical physics for a development eventually leading to the Einstein equations, we have emphasized their common features. One such basic feature is that both equations establish a relation between matter and gravitational potential; a second feature is that both equations relate the action of gravitation to its source by second-order partial differential equations. Such common features are more than distant mathematical similarities or analogies perceived only in hindsight. We claim that such similarities guided the historical development linking the two equations. These similarities, we believe, correspond to structural properties following from the basic mental model shaping the thinking process connected with this development. This interpretation is corroborated by the historical observation that the development from the Poisson to the Einstein equation went through a number of intermediate field equations of the same fundamental structure. We will show that they can all be interpreted as instantiations of the mental model of a field equation, which was modified, again and again, in response to inconsistencies by replacing a minimal number of specific features, while all other components retained their “default” settings. In spite of the inherently conservative structure of this development its outcome entailed fundamental changes in the conceptual structure of classical physics including the original mental model of a gravitational field equation itself.

We will write the basic structure of the mental model of a field equation implemented in the context of gravitational theory symbolically as:

OP(POT) = SOURCE. (I)

This equation is meant to symbolize a structure of shared physical knowledge according to which a source **SOURCE** generates a potential **POT**, related to each other by a differential equation with a second-order differential operator **OP** acting on the potential. We justify the introduction of our symbolic notation by the observation that the same knowledge structure can be found in such different cases as the Poisson equation, Einstein’s intermediate equations for the gravitational potential, the Laplace equation for the electrostatic potential, and the four-dimensional potential formulation of Maxwell's equations. Correspondingly, **OP**, **POT**, and **SOURCE** can be instantiated in many different ways, such as the Laplace or the d'Alembert operators for **OP**, mass density or the energy-momentum tensor for **SOURCE**, Newton's gravitational potential or the metric tensor for **POT**. Notwithstanding the different contexts for each of those instantiations, we find an overarching conceptual structure relevant for each of them. It is the role of these overarching structures in guiding physical reasoning in a qualitative way that we wish to describe in terms of mental models and frames and that we wish to capture in our symbolic notation. We discuss the relevant instantiations for the frame of the field equation in somewhat greater detail.

Before the crucial phase of Einstein's search for a gravitational field equation in the years 1912 – 1915, the mental model of a field equation essentially covered two physical structures, that shaped Einstein's conceptual background in his search: the Poisson equation of classical mechanics and the potential equations of electrodynamics. In the latter, the structure even appears twice, once in electrostatics, in a simple form analogous to that in classical mechanics, and once in a more complex version extended to cover the dynamical aspects of the electromagnetic field as well. In electrostatics, the electrostatic potential φ_e is generated by an electric charge density ρ_e according to[28]

$$\Delta\varphi_e = -4\pi\rho_e. \tag{5}$$

The more extended version, which covers this equation as a special case under certain conditions, is the four-dimensional potential formulation of Maxwell's equations at the core of classical electrodynamics. The four-dimensional, special-relativistic formulation of electrodynamics was developed beginning in 1908 by Minkowski, Laue, and Sommerfeld[29] and was quickly established as a standard.[30] In this framework, the inhomogeneous Maxwell equations can also be written in a potential formulation:[31]

$$\Box\phi^\mu = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\phi^\mu = -\frac{4\pi}{c} j^\mu \tag{6}$$

---

28 The minus sign which does not appear in the Poisson equation of classical mechanics given in eq. (1) reflects the fact that the gravitational interaction is attractive whereas the electrostatic interaction of two charges of equal sign is repulsive.

29 See (Minkowski 1908, Laue 1911, and Sommerfeld 1910a; 1910b).

30 For historical studies, see (Reich 1994, Walter 1999).

31 See, e.g., (Laue 1911; 1913 § 19). The potential formulation of Maxwell's equation given in eq. (6) presupposes a gauge fixing of the form

$$\partial_\mu\varphi^\mu = \mathrm{div}\boldsymbol{A} + \frac{1}{c}\frac{\partial\varphi_e}{\partial t} = 0$$

(Lorentz gauge). Together with this gauge condition eq. (6) represents a fully equivalent representation of Maxwell's equations.

where $\Box$ is the d'Alembert operator, $\phi^{\mu} = (\varphi_e, \boldsymbol{A})$ the electromagnetic four-potential composed of a scalar electric potential φ_e, and a vector magnetic potential $\boldsymbol{A} = (A^x, A^y, A^z)$, and $j^{\mu} = (\rho_e c, \rho_e \boldsymbol{v})$ is the four-current, composed of the electric charge density ρ_e and the velocity vector $\boldsymbol{v} = (v^x, v^y, v^z)$ acting as a source of the potential.

As we shall discuss in more detail below,[32] the relation between electrostatics and electrodynamics provided Einstein and his contemporaries with a basis for an understanding of how Newton's theory of gravitation might be elaborated into a field theory satisfying the requirements of the relativity theory of 1905. Einstein explicitly compared the task of building a relativistic theory of gravitation to the task of developing the entire theory of electromagnetism knowing only Coulomb's law, and found it just as formidable.[33]

Concrete instantiations of the general structure (I) make it clear that there are profound differences between them that are not represented by the simple symbolic equation. A major difference between the gravitational or electrostatic Poisson equation (1) resp. (5) and the full electrodynamic wave equation (6) concerns, for instance, the behavior of the equations under coordinate transformations. The simple mathematical form of those two equations is valid only if specific systems of coordinates are used. The same equations rewritten for a different coordinate system would, in general, change their appearance, unless the new system of coordinates is related to the old one by a coordinate transformation of the appropriate covariance group. This group of admissible coordinate transformations is a mathematical feature of the equation, that is, of the differential operator as well as of the source-term appearing in the equation. It also expresses the validity of a relativity principle for the relevant physical theories, be they those of classical or special-relativistic physics. Coordinate systems can be associated with observers in different locations and in different states of motion, and covariance with regard to coordinate transformations can be associated with the independence of physical phenomena of the perspectives of these different observers.[34]

The Laplace operator, appearing in the electrostatic as well as in the gravitational Poisson equation, retains its form if Galilean coordinate transformations are used

32 See section on "correspondence principle," p. 148.

33 See (Einstein 1913). For more evidence that Einstein conceived the problem of gravitation in analogy with electrodynamics, consider e.g. the title of (Einstein 1912a): "Is there a Gravitational Effect Which Is Analogous to Electrodynamic Induction?" ("Gibt es eine Gravitationswirkung, die der elektrodynamischen Induktionswirkung analog ist?") or a number of references to the analogy with electrodynamics in Einstein's contemporary correspondence; see, e.g., Einstein to Paul Ehrenfest, before 20 June 1912: "A rotating ring does not generate a static field in this sense, even though it is a temporally invariant field. [...] My case corresponds to the electrostatic field in the theory of electricity, whereas the more general static case would also include the analogue of the static magnetic field. I haven't got as far as that yet." ("Ein sich drehender Ring erzeugt nicht ein statisches Feld in diesem Sinne, obwohl es ein zeitlich unveränderliches Feld ist. [...] Mein Fall entspricht in der Elektrizitätstheorie dem elektrostatischen Felde, wogegen der allgemeinere statische Fall noch das Analogon des statischen Magnetfeldes mit einschliessen würde. So weit bin ich noch nicht.") (CPAE 5, Doc. 409).

which relate the inertial reference frames of classical mechanics to each other; these inertial transformations express a symmetry of Newtonian spacetime. The d'Alembertian operator appearing in special relativistic electrodynamics, on the other hand, is invariant under the Lorentz transformations which relate inertial reference frames of special relativity to each other and express a symmetry of four-dimensional Minkowski spacetime.

This illustrates that there can still be profound differences between various instantiations of the structure (I). But whatever these differences may be, insofar as the relation between **OP**, **POT**, and **SOURCE** is cancelled, the corresponding frames enter the same network of relatively stable relations to other physical concepts such as field and force. The concept of a field, in particular, is related to the concept of potential appearing in this mental model by a structure according to which a field **FIELD** is derived from a potential **POT** by some differential operation **GRAD**. This relation can be written in symbolic notation as

$$\textbf{FIELD} = -\,\textbf{GRAD(POT)}. \qquad \textbf{(II)}$$

In classical mechanics the equation relating the gravitational potential φ to the gravitational field $\boldsymbol{g}$ is given by

$$\boldsymbol{g} = -\mathrm{grad}\varphi. \qquad (7)$$

In electrostatics a similar equation holds for the electric field $\boldsymbol{E}$ derived from the electrostatic potential φ_e by

$$\boldsymbol{E} = -\mathrm{grad}\varphi_e. \qquad (8)$$

In the case of electrodynamics the same structural relation reappears, albeit in a somewhat more complex form. The components of φ_e and $\boldsymbol{A}$ of the four-potential φ^{μ} are related to the electric field $\boldsymbol{E}$ and the magnetic field $\boldsymbol{B}$ by

$$\boldsymbol{E} = -\,\mathrm{grad}\varphi - \frac{d}{dt}\boldsymbol{A}, \qquad (9)$$

and

$$\boldsymbol{B} = \mathrm{curl}\boldsymbol{A}. \qquad (10)$$

These equations can be combined in a tensor equation for the electromagnetic field tensor $F_{\mu\nu}$

---

34 We add a note of caution here. From the point of view of modern coordinate-free descriptions of physical theories, the covariance of a particular equation under a specific group of coordinate transformations can be understood as expressing a symmetry property of the underlying spacetime manifold, if the coordinate systems related by the transformations are those associated with so-called geodesic observers. In this case, the geodesic observer field is also a Killing vector field of the manifold, see (O'Neill 1983, 358–362). See also the discussion in (Salmon et al. 1999, chap. 5) and (Norton 1992b). At the time, however, the validity of a physical principle of relativity was directly associated with the covariance of the corresponding equations under coordinate transformations, without considering the symmetry properties of a manifold independently from its coordinate representation.

$$F_{\mu\nu} = \frac{\partial\phi_\mu}{\partial x^\nu} - \frac{\partial\phi_\nu}{\partial x^\mu}, \tag{11}$$

which may be considered as another instantiation of (II).

The discussion of the foregoing examples should make it clear that any concrete meaning of our symbolic equations is context-dependent as is only fitting for relations between frames in the sense introduced above. Entities such as **OP**, **POT**, **SOURCE**, **FIELD**, **GRAD** but also operations such as multiplication can take on entirely different mathematical meanings in different contexts. These symbolic operators may inherit different default-settings from different frameworks of reasoning. There is no *a priori* guarantee that the resulting concrete expressions can still be subsumed under one overarching theory. If one looks, however, at the function of these frames as heuristic devices that guided Einstein's pathway out of classical physics, this obviously was precisely their strength.

In summary, the stability of the mental model of a field equation is a consequence of its embedding in a network of physical concepts covering a broad spectrum of physical knowledge. Specifically, the concepts of potential and mass have stable relations to such concepts as field, force, energy, momentum, and motion. Furthermore, all instantiations of the Lorentz model we have encountered in classical and special-relativistic physics include a second-order differential operator **OP** and can be characterized by symbolic relations between the associated physical concepts such as (II). For each instantiation, the mental model of a field equation acquires local stability also through the representation in terms of mathematical concepts that in themselves are interconnected in an elaborated network allowing for formal manipulations of the mathematical expressions using well-known formal rules.

The various slots in our symbolic equations can be filled with objects of very different mathematical character, **POT** and **SOURCE** may be instantiated by scalar or vectorial objects, which behave differently under coordinate transformations; the corresponding differential operator **OP** may be the Laplacian or the d'Alembertian operator. Physically, potential and mass enter the stable conceptual relation described above, but are at the same time connected with quite different physical concepts and hence quite different physical phenomena. Thus, the potential **POT** could be instantiated to the potential of gravitational, electrostatic, or electrodynamic interaction, and the source-term **SOURCE** could be gravitating mass-density, electric charge-density, or electric current.

The Einstein equation introduced in the beginning of this section emerged, as we shall see, in a process that started from the Poisson equation of classical mechanics and proceeded via intermediate field equations that are *all* structured by what we have called the Lorentz model of a field equation. It is therefore no accident that the Einstein equation also displays features of this model. We shall show that the Einstein equation came about only as the result of a complicated process of adaptations of the original mental model demanding a number of variations ("changes of default settings") that at each step had to fulfill different, and often conflicting requirements. A consistent solution for meeting those requirements was reached only with the final theory of general relativity. In this theory, however, the field equation has implications that, as we shall see, challenge the original mental model.

2.3 The Lorentz Model of an Equation of Motion

The field-theoretic model comprises not only a structure shaping the understanding of a field equation but also a scheme determining the meaning of an equation of motion. In classical physics a field equation must be complemented by an equation of motion. Their complementarity derives from the way in which interactions are split into cause and effect in the Lorentz model. In classical mechanics, the concept of force allows one to separate the generic features of the action of some agent, to be described in terms of a general force law, from its specific effect on a given physical object, to be described in terms of a change of its state of motion. A similar structure is characteristic of Maxwellian electrodynamics, especially in Lorentz's electron theory. The field equation describes how sources, represented in our symbolic equation by the **SOURCE**-frame, affect the state of the surrounding space, represented by the **POT**-frame or the **FIELD**-frame. The equation of motion describes the effect of the thus affected space on physical objects in it. From the perspective of classical mechanics, a field equation is therefore nothing but a specific way of prescribing a general force law. What is required is a bridge between the concept of field and that of force.

According to classical mechanics, the effect of a force is a deviation (to be observed within an inertial frame of reference) from a state of rest or a state of uniform rectilinear motion described in terms of an **ACCELERATION**-frame. The magnitude of the acceleration depends not only on the force (characterized in the following by a **FORCE**-frame) but also on the reactive properties of the physical object exposed to it; these properties will be summarily described by the inertial mass frame, $\mathbf{MASS_{IN}}$. In short, an equation of motion according to classical and special-relativistic physics, complies with a mental model of causation that may be called the "acceleration-implies-force model" and takes the form:

$$\mathbf{FORCE = MASS_{IN}\ x\ ACCELERATION.} \qquad \mathbf{(III)}$$

In classical mechanics this relation corresponds to Newton's

$$\boldsymbol{F} = m \cdot \boldsymbol{a}, \qquad (12)$$

where m is the inertial mass of a material particle, a its acceleration in three-space and F a classical force. The special relativistic generalization of this relation is

$$F^{\mu} = m\frac{du^{\mu}}{ds}, \qquad (13)$$

where m is the rest mass, u^{μ} the four-velocity and s the proper time. Here F^{μ} denotes the force as a four-vector.

The structure of the symbolic equation (III) also complies with that of a much more general and much older mental model of causation rooted in intuitive physics, the "force-implies-motion model," which thus serves as a "higher-order model" for the Newtonian relation (III):[35] According to this higher-order model of causation, the effect of an action (here **ACCELERATION**), depends on the strength of the action (here **FORCE**) as well as on the resistance to the action (here **MASS$_{\mathbf{IN}}$**).

How can the acceleration-implies-force model belonging to the core of Newtonian mechanics be integrated with the concept of field at the center of the field-theoretical model? In order to bridge the two models one needs a specification of the relation between **POT** or **FIELD**, describing the local state of the surrounding space, and **FORCE**, describing the role of this space as an agent determining the motion of matter. In classical field theory, this bridge relation is given by the notion that the field is tantamount to a local force. The force experienced by a particle in a field is proportional to the strength of the field at the point of the particle in space and time. It is also proportional to that quality of the particle that responds to the particular field, be it its gravitational mass, its electric charge, or its magnetic moment. We capture the relation between **FORCE** and **FIELD** by the symbolic relation

$$\textbf{FORCE = CHARGE x FIELD} \qquad \textbf{(IV)}$$

At this point, our symbolic relations allow us to describe a possible inference on the level of qualitative physical reasoning. We may use relation (II) between **FORCE** and **POT**, to derive a relation between **FIELD** and **POT**

$$\textbf{FORCE = – CHARGE x GRAD(POT).} \qquad \textbf{(V)}$$

From a different perspective, one may also look at the set of relations (II), (IV), and (V) as the expression for a conceptual network on the level of qualitative physical reasoning in which the frames **FORCE**, **FIELD**, **CHARGE** and **POT** are related to each other.

In electrostatics the **CHARGE**-frame is instantiated by the charge density ρ_e,

$$\textbf{CHARGE} \xrightarrow{\text{electrostatics}} \rho_e, \qquad \textbf{(VI)}$$

and the force density acting on ρ_e and determined by an electric field $\boldsymbol{E}$ derived from the electrostatic potential φ_e is given by:

$$\boldsymbol{F}_e = \rho_e \boldsymbol{E} = -\rho_e \mathrm{grad}\varphi_e. \qquad (14)$$

In Newtonian gravitational theory the default setting of the **CHARGE** frame is the so-called "passive gravitational mass":

35 For the force-implies-motion model, see (Gentner and Stevens 1983; Renn 2000), and also "Classical Physics in Disarray …" (in this volume).

$$\textbf{CHARGE} \xrightarrow[\text{gravitation}]{\text{Newtonian}} m_p \text{ resp. } \rho_p . \qquad \textbf{(VII)}$$

Accordingly, the force density acting on a mass density ρ due to a gravitational field $\boldsymbol{g}$ that can be derived from a gravitational potential φ is given by:

$$\boldsymbol{F} = \rho \boldsymbol{g} = -\rho \operatorname{grad}\varphi. \tag{15}$$

In the case of electrodynamics the same structural relation holds in terms of the electromagnetic field tensor $F^{\mu\nu}$ expressed in terms of a generalized electrodynamic potential in (11). The four-force density K^{μ} is given by

$$K^{\mu} = j_{\nu} F^{\mu\nu} . \tag{16}$$

This equation again exhibits the structure **FORCE = CHARGE x FIELD**, even though the multiplication of our symbolic equation is realized in this case by a four-dimensional contraction.

The discussion of the bridge relation required to integrate the acceleration-implies-force model with the field concept makes the intrinsic complexity of the Lorentz model particularly evident. This complexity stands in striking contrast to certain elementary features of gravitational interactions. It is mainly due to the fact that the Lorentz model results from the integration of mental models referring to two kinds of physical substances, the model of an extended, space-filling physical medium traditionally labelled as "aether" and the model of matter constituted by particles. The relation between field and force given by (IV) mediates between these models and at the same time points to the conceptual intricacies resulting from their integration. For instance, what at first sight merely seems to be a problem of two bodies moving about their common center of gravity, say of the sun and a planet, appears, from the perspective of the Lorentz model as the consequence of a field generated by one body which is then felt by the other body as a force that in turn is the cause of its motion.

As a consequence of this construction, both the concept of force and the concept of mass take on connotations, which they did not possess independently in the more elementary models. In classical mechanics, for instance, the concept of force comprises actions at a distance, typically between particles. In the context of the field-theoretical model, it applies exclusively to local interactions, a rather artificial limitation from the point of view of Newtonian physics. Similarly, while the Newtonian concept of force entails a reciprocity of the interaction it describes, expressed in Newton's *actio = reactio*, such a reciprocity is less evident for an interaction that is conceived to relate a state of space, characterized by the **FIELD**-frame or the **POT**-frame, to changes of the state of motion of a physical object, characterized by the **ACCELERATION**-frame. As a matter of fact, theories such as Lorentz's electron theory violate this reciprocity and *actio = reactio* no longer holds for the interaction between ether and charged matter.

The conceptual intricacies implied by the Lorentz model for the concept of mass are even more serious. Mass may be conceived as a "source" causing changes of the state of space or of the "aether" according to (**I**). We thus have "active gravitational mass" in the case of gravitational interaction:

$$\textbf{SOURCE} \xrightarrow{\text{gravitation}} m_a \text{ resp. } \rho_a . \qquad \textbf{(VIII)}$$

Mass may also, according to (V) and (VII), be conceived as a passive property of a physical object exposed to the resulting field, determining the degree to which the field locally acts as a force (**CHARGE** or "passive gravitational mass" in the case of gravitation); it may finally be conceived, according to (III), as "inertial mass" $\textbf{MASS}_{\textbf{IN}}$, i.e., as resistance to **ACCELERATION**. In classical electrostatics, these magnitudes are represented by electrical charge and inertial mass, respectively, and can vary independently from each other. In classical gravitation theory, gravitational and inertial mass happen to coincide empirically. In this case we are thus entitled to introduce a generic **MASS**-frame for which we have:

$$\textbf{MASS} = \textbf{MASS}_{\textbf{IN}}, \qquad \textbf{(IX)}$$

which may hence be instantiated by inertial, or active gravitational, or passive gravitational mass.

The integration of different mental models within the field-theoretical model produces conceptual distinctions that may actually not be warranted by the available knowledge of the interactions it describes. The emergence of conceptual distinctions as an artefact of a theoretical framework was visible, in the case of the gravitational interaction, even from a less sophisticated perspective than that offered by the field-theoretical model. When the gravitational action is described not in terms of a field theory but simply using the Newtonian force law, the distinction between mass as a property of matter that *causes* gravitation and mass as a reactive property of matter that *resists* the acceleration caused by a gravitational force is rather artificial. Indeed, it has long been known that all bodies fall with the same acceleration in a gravitational field whatever their mass (Galileo's principle). Within the context of the field-theoretical model this insight suggests far-going consequences for the understanding of an equation of motion.

In fact, since in classical mechanics the **CHARGE**-frame and the **SOURCE**-frame instantiate to the passive and active gravitational mass resp. mass density according to (VII) and (VIII), we may identify these two frames with each other and with the general **MASS**-frame:

$$\textbf{SOURCE} = \textbf{CHARGE} = \textbf{MASS} \qquad \textbf{(X)}$$

Recalling the relations that the **FORCE**-frame enters with the **ACCELERATION**-frame and the **FIELD**-frame according to (III) and (IV), our symbolic equations entail

$$\textbf{ACCELERATION} = \textbf{FIELD}. \qquad \textbf{(XI)}$$

This symbolic equation translates Galileo's principle to the assertion that in a gravitational field theory the local acceleration actually represents the gravitational field. This makes it possible to interpret the effect of gravitation, namely **ACCELERATION**, directly as a representation of the local force, i.e. as a **FIELD**, independently of the properties of the object exposed to it. We emphasize that we introduced our symbolic notation in order to be able to represent this kind of inference which can be made largely on the level of qualitative physical reasoning independent of any concrete representation. It also expresses the fact that the identification of the **ACCELERATION**-frame and the **FIELD**-frame is a general relation between two frames that is not tied to the concrete conceptualization of the gravitational interaction. It may hence guide the physical reasoning also in situations where new ways of mathematical representation or else new conceptual relations within a gravitational theory are being explored.

Eq. (XI) no longer contains **FORCE**. This insight crucial for the development of general relativity. It suggests that it should be possible to set up a theory where field phenomena are equivalent to acceleration phenomena. This, of course, is exactly the idea at the core of Einstein's equivalence principle.[36] In such a theory Galileo's principle would find the conceptual justification it lacked in classical mechanics, where it appeared as a mere empirical coincidence.

The insight that in a gravitational field theory the acceleration is directly equivalent to the field, symbolically represented by eq. (XI), also suggests the formulation of an equation of motion in a gravitational field that does not make use of the intermediate concept of force. The idea of eliminating the concept of force was familiar from classical physics and had been elaborated in the context of the Lagrange formalism of analytical mechanics. In elementary situations of classical mechanics the Langrangian or Lagrange function at the center of this formalism is simply the difference between the kinetic and the potential energy of a physical system:

$$L = T - V. \tag{17}$$

The Lagrange formalism provides an alternative way of obtaining equations of motion. In this formalism the trajectory of a material body is selected from the set of all kinematically possible trajectories satisfying given constraints. The criterion for the selection is that the action, defined as the same integral of the Langrangian along a given trajectory is stationary, i.e., takes on either a maximum or a minimum value, for the actual trajectory. This criterion is known as Hamilton's principle. Saying that the action is stationary is the same as saying that its variation vanishes:

36 See the discussion below. Cf. in this context Einstein's use of the word "Beschleunigungsfeld" (acceleration field) in (Einstein 1912b): "the hypothesis that the "acceleration field" is a special case of the gravitational field [...]" ("die Hypothese, daß das 'Beschleunigungsfeld' ein Spezialfall des Gravitationsfeldes sei [...]" (p. 355).

$$\delta\left\{\int L dt\right\} = 0. \tag{18}$$

The situation is similar to the problem of finding the shortest path connecting two points on a curved surface. This problem can be solved by looking for an extremal value among the lengths of all possible paths connecting these points.

The Lagrange formalism yields the explicit equation of motion for a particle in the form of the so-called Euler-Lagrange equations, which follow from Hamilton's principle:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_i}\right) - \frac{\partial L}{\partial x_i} = 0. \tag{19}$$

Under appropriate circumstances, these equations may be assimilated to the relation

$$\boldsymbol{F} = d\boldsymbol{p}/dt \tag{20}$$

between a force F and the change of momentum p familiar from classical mechanics. We capture the qualitative physical content of equations (19) and (20) by the symbolic equation

DIFF(MOMENTUM) – FORCE = 0, (XII)

introducing, at the same time, a **MOMENTUM**-frame and a **DIFF**-frame, the latter being in the present case instantiated by time-derivatives. The advantage of the Lagrange formalism compared to the explicit specification and use of forces becomes clear if one considers motion under geometrical constraints, such as the motion of a point particle on the surface of a sphere. Using the acceleration-implies-force model such geometrical constraints are realized by constraining forces which are defined only by their effect, i.e., if a body moving under geometric constraints departs from uniform rectilinear (inertial) motion this deviation is assumed to be caused by the constraining forces. The precise magnitude and direction of these forces are generally unknown and hence cannot be explicitly specified in order to obtain an equation of motion by instantiating the acceleration-implies-force model. The Lagrangian formalism uses the fact that these constraining forces do not perform work. They play no role in the interplay between kinetic and potential energy as captured by the Lagrange function if the latter is expressed solely in terms of those generalized coordinates that describe possible motions under the given constraints, without losing information about the physical situation.

The significance of the Lagrange formalism for Einstein's research on gravitation was twofold. For one, it represented a generalizable formalism that was applicable in cases where a force or a momentum was not easily identified. It therefore also applied in a geometrized theory by expressing the Lagrange function in terms of a geometry adapted to the physical situation at hand.[37] Describing motion in a gravitational field with the help of the Lagrange formalism, as we shall see, naturally suggests to conceive of gravitation as a consequence of geometry, rather than of force. And this conception of gravitation in turn suggests adopting the Lagrange formalism as the natural framework for formulating the equation of motion in a gravitational field.

3. THE ELEMENTS OF EINSTEIN'S HEURISTICS

In the course of the first period of Einstein's research on gravitation—between 1907 and 1912—specific components of the heuristics had crystallized as relatively stable structures which would guide his search for the gravitational field equation in the second period—documented in the Zurich Notebook. To some of them Einstein even attached labels, such as the "equivalence hypothesis,"[38] making their outstanding role for his heuristics evident, while other heuristic requirements were so obviously interwoven with canonical expectations from classical physics, such as the requirement of energy conservation, that they did not receive a special name. For ease of reference, we shall nevertheless introduce standard names.

Einstein's *equivalence principle*, which in the first period mainly served to find properties of special cases of gravitational fields, became in the second period a standard criterion for checking whether or not candidates for a general gravitational field equation incorporated his earlier insights about the intimate relation between gravitation and inertia. The *generalized relativity principle*, a closely related result of Einstein's research in the first period, was applied in the second period either as a starting point in the context of the mathematical strategy for choosing appropriate candidates for the gravitational field equation or as a validation criterion by which a candidate constructed in the context of the physical strategy was examined. The *conservation principle*, inherited from classical physics, played a crucial role in developing the theory of the static gravitational field and was similarly used in the second period both as a touch stone and as a building block. This was also the case for what we will call the *correspondence principle*. This principle represents the demand to incorporate in a new theory of gravitation the knowledge about Newtonian gravitation by requiring that the basic relations of the latter be recovered from the former in some approximation or as some special case. Its implementation as a component of Einstein's heuristics took clues from the relation between electrostatics and electrodynamics. Einstein thus expected that the generalized theory should be connected to the Newtonian theory via the intermediate case of the weak and static gravitational field.

---

37 Its significance for expressing the equation of motion in special relativity was realized by Max Planck (Planck 1906, 1907).

38 Although Einstein referred to the equivalence principle as a "hypothesis" in (Einstein 1907) and in (Einstein 1911), the terms "Aequivalenzhypothese" and "Aequivalenzprinzip" were used for the first time in (Einstein 1912b).

3.1 The Equivalence Principle and the Generalized Relativity Principle

According to Einstein's equivalence principle the effects of a homogeneous static gravitational field are equivalent to those in a uniformly and linearly accelerated reference frame. The equivalence principle, which establishes a connection between the gravitational field and inertial forces, is closely related to Galileo's principle that all bodies fall with the same acceleration in a gravitational field, independent of their constitution. While neither Galileo's principle nor the equivalence principle are part of the foundational structure of classical physics, they are part of the knowledge contained in it, as expressed by our symbolic equation (XI). Einstein established a meaningful connection between acceleration and the gravitational field by integrating two mental models of classical physics which originally belonged to different domains of knowledge, the mental model of a system with a homogeneous static gravitational field, familiar from everyday physics in local terrestrial laboratories, and the model of a system in uniformly accelerated motion (Einstein's famous elevator experiment), which was analyzable using standard tools of classical mechanics. The indistinguishability of motions in these two systems makes it possible to identify the terminals of these models and thus to establish an equivalence between gravitational and inertial forces as well as between an accelerated frame of reference and an inertial frame. These identifications turned out to have far-going consequences for the organization of physical knowledge. Such consequences can be spelled out if further elements of the knowledge of classical and special relativistic physics are taken into account and are combined, for instance, with simple mental models of ray optics leading to the conclusion that light is curved in a gravitational field.[39]

Einstein's "elevator model," admits an extension to a more general class of gravitational fields and accelerated motions. Such an extension was suggested, in particular, by the Machian idea to interpret the inertial forces occurring within a uniformly rotating system as due to the interaction with distant masses rather than due to "absolute space." Mach had compared an accelerated system—Newton's famous rotating bucket—with a system at rest in which an interaction with distant masses, the stars revolving around the bucket, accounts for the same physical phenomena as are produced by the inertial forces in the accelerated system. This thought experiment provided a blueprint for the elevator-thought-experiment, which is at the heart of Einstein's "principle of equivalence." In analogy to the "elevator model," a "bucket model" could thus be conceived as one in which the inertial forces occurring in a rotating reference frame are interpreted as the effects of a generalized gravitational field.

The elevator and the bucket models may both be considered as special cases of a general "gravito-inertial model" in which inertial forces resulting from arbitrarily accelerated motions are interpreted as coming from a "dynamic" gravitational field. This gravito-inertial model made it plausible to assume that inertial frames of reference play no privileged role in a theory that adequately describes such a generalized

39 See the discussion in "Classical Physics in Disarray …" (in this volume).

gravitational field. It also suggested that the generic properties of gravitational fields that can be thought of as resulting from accelerated motions are shared by arbitrary gravitational fields. It suggests, for instance, that the laws governing the motion of bodies are the same in both types of fields. More generally, the gravito-inertial model made it plausible that physical interactions taking place in a gravitational field are essentially equivalent to those taking place in a gravitation-free system that is described from the point of view of an accelerated observer. In hindsight, the equivalence principle—and the gravito-inertial model structuring the reasoning on which this principle is based—thus introduced four more or less distinct requirements into the search for a theory of general relativity:

– the theory should satisfy a "generalized principle of relativity" and eliminate as much as possible the privileged *a priori* structures which in the classical theory are associated with such notions as absolute space and inertial frames of reference;
– the theory should describe motion in a gravitational field as a "free fall" independent of the structure of the moving body;
– the theory should treat gravitation and inertia as aspects of one more general interaction; and
– the theory should describe non-gravitational physical interactions essentially in the same way as special relativity if an appropriate reference frame (local inertial frame) is chosen for that description.

These requirements are directly related to general relativity as we know it today. Historically, the impact of the equivalence principle on the search for a new theory of gravitation was much less straightforward than it may appear in hindsight. A number of conceptual and technical problems had to be resolved or at least disentangled before such a clear relation could emerge.[40] In particular, Einstein was convinced that the demand for a generalized relativity principle could be satisfied by requiring the equations of his theory to be generally covariant (Norton 1994, 1999). He lacked the modern notion of spacetime symmetries. Similarly, the description of motion in a gravitational field as "free fall" along a geodesic trajectory is closely related today to the understanding of the affine structure of spacetime. But Einstein did not have the concept of affine connection at his disposal and still saw the need to interpret the equation of motion in terms of a classical gravitational force.[41]

The equivalence principle and the generalized relativity principle did not give rise to requirements which the new theory had to satisfy as a set of fixed axioms; they acted in a more general and diffuse way as heuristic guiding principles which, in different contexts, had a variety of concrete implications not necessarily covered by their modern counterparts. The generalized principle of relativity, in particular, motivated Einstein to consider the absolute differential calculus as the appropriate lan-

40 See (Norton 1985), "Classical Physics in Disarray …" and "The First Two Acts" (both in this volume).
41 See "The Story of Newstein …" (in vol. 4 of this series).

guage for his new theory of gravitation but also to construct mathematical objects which are covariant merely under much more limited classes of coordinate transformations. The equivalence principle led him to identify qualitative consequences of general relativity such as light deflection even before the formulation of the definitive theory, albeit with numerically different results. It led Einstein to adopt the geodesic equation of motion as the law of motion appropriate for general gravitational fields but also to systematically check whether candidate field equations are covariant at least under transformations to linearly accelerated systems and to uniformly rotating systems.

In the context of Einstein's systematic search for the gravitational field equation documented in the Zurich Notebook the adoption of the generalized relativity principles amounted to a check of the covariance properties of a candidate field equation. But even the way in which this check was implemented—by the introduction of a generally-covariant differential operator along the mathematical strategy or by explicitly checking the behavior of a candidate under coordinate transformations along the physical strategy—depended on the specific perspective guiding the implementation. At the beginning of Einstein's search it was not at all clear whether he would eventually succeed in finding a generally-covariant field equation of gravitation incorporating the equivalence principle. From the outset it was unclear whether the ambitious aim of a generalized relativity principle and perhaps even the equivalence principle would be realizable or whether these postulates had to be restricted or modified in order to be able to satisfy other requirements to be imposed on such a field equation, such as the conservation principle.

3.2 The Conservation Principle

According to the conservation principle as it functioned in Einstein's heuristics, it should be possible to establish a balance of energy and momentum in a gravitational field, resulting in a conservation law if all contributions to the balance, including that of the gravitational field itself, are taken into account. This expectation was motivated by the experience of classical physics where such a balance of energy and momentum could indeed be obtained for all physical processes if only appropriate concepts of energy and momentum were identified for all relevant subdomains, such as mechanics, thermodynamics, and electrodynamics. This expectation had been both amplified and modified by the advent of special relativity, and in particular that of special relativistic continuum physics, which had shown that several distinct conservation laws of classical physics, such as those of mass, energy, and momentum actually had to be integrated into a single all-encompassing conservation law referring to a complex new entity, the stress-energy or energy-momentum tensor. Against this background, the conservation principle, understood as part of the heritage of classical and special-relativistic physics, introduced three more or less distinct requirements into the search for a theory of general relativity:

- the theory should take into account the close relation between mass and energy established by special relativity and consider not just mass but more generally mass and energy as embodied in the energy-momentum tensor (or some entity derived from it) as the source of the gravitational field;
- the theory should contain some generalization of the special-relativistic law for the conservation of energy and momentum; and, in particular,
- the gravitational field equation should be compatible with this generalized requirement of energy and momentum conservation.

From the perspective of today's understanding of general relativity, these requirements considerably restrict the choice of an acceptable gravitational field equation. But historically, just as with the generalized relativity principle, Einstein's heuristic expectations could not simply be turned into iron-clad axioms for the formulation of his new theory. Precisely because the requirements listed above were rooted in the knowledge of classical and special-relativistic physics, they were still embedded in a conceptual framework that was eventually overturned by general relativity. Furthermore, there were, at the outset of his search, still numerous possibilities for instantiating the general relations suggested by Einstein's classical expectations. It was, for instance, conceivable that not the energy-momentum tensor itself but its trace acts as the source of the gravitational field. For some time, Einstein assumed that he had to find a generally-covariant energy-momentum tensor of the gravitational field in analogy to the one for matter, while such a tensor does not exist according to the final theory. He also assumed that the conservation principle would play the role of an additional postulate of the theory, whereas it is implied by the correct gravitational field equations. Such conceptual novelties of general relativity could not have been anticipated on the basis of the knowledge of classical physics informing Einstein's heuristics. They were the eventual outcome of his heuristic schemes in the course of concrete and often futile attempts to identify a gravitational field equation compatible with criteria such as the conservation principle.

The effect on Einstein's search of the requirements here summarized under the label "conservation principle" depended on the specific questions he pursued and on the level of sophistication of the techniques at his disposal. At one point he erroneously convinced himself, for instance, that a gravitational theory based on a single scalar potential was incompatible with the conservation principle but then had to retract that argument in the light of a closer analysis of such a scalar theory.[42] The clear-cut function which the conservation principle eventually assumed as a compatibility requirement for an acceptable field equation in his search for such an equation documented in the Zurich Notebook was the result of his learning experience with the theory for static gravitational fields in 1912.[43] This experience demonstrated to Einstein the crucial significance of the conservation principle for his search. He became aware step by step of the full scope of the network of relations it implies. In the

42 See (Norton 1992a).
43 See "The First Two Acts" (in this volume) and the discussion below.

course of his research documented in the Zurich Notebook, involving the mathematically much more complex tensorial formalism, these relations combined to form a set of standard expectations for a field equation that had to be systematically checked for each candidate. Only towards the very end of this phase of his research did Einstein recognize the possibility of turning this network into a recipe for constructing a gravitational field equation satisfying the conservation principle—albeit as a requirement essentially still conceived within a classical framework.

3.3 The Correspondence Principle

The correspondence principle requires that the new relativistic theory of gravitation incorporate the empirically well-founded knowledge about gravitation contained in the classical Newtonian theory. Ideally, it should be possible to obtain the Newtonian theory as a limiting or special case from the new theory under appropriate conditions, such as low velocities and weak fields.[44] In contrast to the generalized relativity principle of which it was not clear at the outset to what extent it could be implemented in the new theory, the correspondence principle was a much less negotiable, if not absolutely necessary requirement for any acceptable theory of gravitation. It also seemed clear from the beginning how this principle would have to be implemented in concrete attempts to create a relativistic theory of gravitation. The classical theory offered a model for a gravitational field equation, the Poisson equation, even if this model does not take into account the relativistic demand of a finite speed of propagation of the gravitational action as would a field equation based on the d'Alembertian operator as in (6). But the Poisson equation did not only serve as a model for the structure of the new field equation. Einstein also expected it to emerge from a limiting process by which a relativistic field equation should touch base, via the intermediate case of a special-relativistic field equations based on the d'Alembertian operator, with the classical Newtonian theory. Einstein's theory of the static gravitational field provided another such base-line. Since it represents an intermediate situation between the full relativistic theory and the Newtonian case, he expected that the general theory would, under appropriate limiting conditions, first reproduce the results of the static special case and then, under further constraints, those of the Newtonian theory. A relativistic theory with this limiting behavior clearly would cover the full range of physical knowledge covered by the more specialized theories. Since the constraints imposed by the correspondence principle were embodied not just in abstract requirements but in well-developed theories, it follows that this heuristic principle could act not only as a compatibility condition for an acceptable gravitational field equation but also as a starting point for its construction.

---

44 For a discussion of the Newtonian limit of general relativity from a modern point of view, see (Kuenzle 1976, Ehlers 1981, 1986). For a discussion of the relation between Newtonian gravitation theory and general relativity from an axiomatic point of view as a case of reduction, see (Scheibe 1997, 1999, esp. ch. VIII).

As with the other criteria, there is a perspective from which the correspondence principle, together with a few other conditions, singles out general relativity as the only acceptable solution to Einstein's problem. But as we saw with the other heuristic principles, this hindsight-perspective tends to obscure rather than clarify the actual role of the correspondence principle in the creation of general relativity. This process involved conceptual innovations that could not have been anticipated on the basis of classical physics. From hindsight, we would rather have to say it could not even be anticipated *that* (let alone *how*) the definitive solution of his problem would yield the Newtonian theory since the classical limit of the final theory—which in some sense must exist for the reasons pointed out above—might not resemble the familiar Newtonian formulation of the classical knowledge about gravitation. Vice versa, the classical expectations concerning the relation between Newtonian and relativistic theory might impose restrictions on the choice of admissible candidates that could effectively rule out a satisfactory realization of Einstein's other heuristic requirements, in particular, of the generalized principle of relativity. The dilemma, in short, was that the correspondence principle represented, in view of its roots in the classical knowledge about gravitation, the most weighty of Einstein's heuristic principles but also the one most likely to be entangled with physical assumptions that would have to be given up if the new, relativistic theory of gravitation were to challenge those classical roots.

This dilemma could hardly be avoided. At the beginning of his search, Einstein sought to extrapolate the classical knowledge about gravitation into the new territory of a relativistic field theory. From his perspective, that territory, fortunately, was mapped out nicely by the implications of the Lorentz model. As we have seen, his model also determined the conceptualization of the relation between a generic field theory and the special case of a static field. In Maxwell's theory of the electromagnetic field that relation was well understood, so it could serve as a guide for exploring the analogous relation in the case of the relativistic gravitational field. The mental model of a field theory and the knowledge of classical physics it incorporates had governed Einstein's seemingly inductive procedure all along in examining special cases such as that of the static field.[45] It was clear to him from the outset that the static gravitational field corresponds to the electrostatic field while the field of a rotating reference frame corresponds to the magnetostatic field. The theory of electromagnetism also suggested that and how a many-component tensorial object representing the field in general turns into a much simpler object for the special case of a static field, which can be derived from a scalar potential. The fact that in classical physics both the electrostatic and the gravitational potential are represented by a scalar potential lent support to the assumption that a reduction to a scalar potential also takes place in a relativistic theory of gravitation, at least in the limit of weak static fields. Although this assumption eventually turned out to be wrong, it was backed by a long tradition in classical field theory to which no alternative was known and it initially prevented Einstein from accepting the Einstein tensor as a viable candidate for the left-hand side of the field equation.

Just as Einstein's other heuristic principles the correspondence principle did not act as an isolated axiom which in the end turned out to be either compatible or not with general relativity as we know it today. It was not an isolated statement at all but part of a network of arguments, affecting his heuristics in the context of a variety of considerations. The correspondence principle comprised, in particular, the demands that:

– the differential operator on the left-hand side of the gravitational field equation should, for weak fields, reduce to the d'Alembertian operator as in (6);
– the field equation, for weak static fields, should reduce to the Poisson equation for the scalar potential of classical physics;
– the same scalar potential should determine the behavior of a particle in a gravitational field, via the equation of motion.

The correspondence principle was also subject to modifications as Einstein's experience with attempts to implement this principle in concrete candidate field equations grew. The paradoxical fluid yet firm character of Einstein's qualitative reasoning on the level of his heuristic principles, which we have tried to grasp by describing it in terms of mental models and frames, allows it to first exclude and then support the correct field equations of general relativity. The correspondence principle thus left room for learning experiences as when Einstein found out that it was possible to meet the requirements of this principle with the help of additional constraints on the choice of the coordinate system.

While the technicalities of its implementation were subject to reconsideration and improvement, the basic structure of Einstein's understanding of the correspondence principle was stabilized by a wider context of arguments rooted in classical physics.

---

45 Compare the following equations from Einstein's correspondence: "I finished the investigations on the statics of gravitation (point mechanics electromagnetics gravitostatics) and am very satisfied with them. I really believe that I discovered a piece of truth. Now I ponder the dynamic case, going again from the more special to the more general." ("Die Untersuchungen über die Statik der Gravitation (Punktmechanik Elektromagnetik Gravitostatik) sind fertig und befriedigen mich sehr. Ich glaube wirklich, ein Stück Wahrheit gefunden zu haben. Nun denke ich über den dynamischen Fall nach, auch wieder vom spezielleren zum Allgemeineren übergehend.") Einstein to Ehrenfest, 10. March 1912, (CPAE 5, Doc. 369); "Lately I have been working like mad on the gravitation problem. Now I have gotten to the stage where I am finished with the statics. I do not know anything yet about the dynamic field, that will come only now. [...] You see that I am still far from being able to conceive of rotation as rest! Each step is devilishly difficult, and what I have derived so far is certainly still the simplest of all." ("In der letzten Zeit arbeitete ich rasend am Gravitationsproblem. Nun ist es soweit, dass ich mit der Statik fertig bin. Von dem dynamischen Feld weiss ich noch gar nichts, das soll erst jetzt folgen. [...] Du siehst, dass ich noch weit davon entfernt bin, die Drehung als Ruhe auffassen zu können! Jeder Schritt ist verteufelt schwierig, und das bis jetzt abgeleitete gewiss noch das einfachste.") Einstein to Michele Besso, 26. March 1912, (CPAE 5, Doc. 377); "My case corresponds to the electrostatic field in the theory of electricity, whereas the more general static case would also include the analog of the static magnetic field." ("Mein Fall entspricht in der Elektrizitätstheorie dem elektrostatischen Felde, wogegen der allgemeinere statische Fall noch das Analogon des statischen Magnetfeldes mit einschliessen würde.") Albert Einstein to Paul Ehrenfest, Prague, before 20 June 1912, (CPAE 5, Doc. 409).

Precisely because of this wider context, the modifications of Einstein's understanding in the course of his search for the field equation had the potential of challenging not only the technical aspects but also the conceptual framework of his heuristics.

3.4 Einstein's Heuristic Principles and his Double Strategy

Einstein's heuristic principles, as we have seen, did not constitute a set of axioms from which a theory of gravitation could be derived in a straightforward way. These principles yielded both too much and too little knowledge to find a new theory of gravitation—too little, because they were not sufficient to determine the new theory uniquely, too much, because they imposed requirements on the new theory that could not be maintained all at once. As we mentioned in the introduction, these principles initially even acted as competing approaches toward a relativistic theory of gravitation. In addition, their interpretation in concrete attempts to realize such a theory depended on the specific formalism applied and on the form other requirements took within that formalism. In the course of Einstein's work documented in the Zurich Notebook, these principles nonetheless developed together to become elements of a heuristic double strategy. Earlier research on the problem of a relativistic theory of gravitation, Einstein's own as well as that of others, had not only suggested the mathematical tools to be employed but had also circumscribed the requirements such a theory had to satisfy. As a consequence, the problem of identifying an acceptable gravitational field equation had become the task of constructing, as if in a theoretical laboratory, a more or less well-defined but never-tried device from a set of given building blocks.

Each of Einstein's heuristic principles against which constructions would have to be checked could be used either as a construction principle or as a criterion for their validity. The sequence in which the heuristic principles were used essentially determined their function. The approach we have labelled the "physical strategy" starts from the correspondence principle, i.e., from a candidate field equation which by inspection is seen to yield the Newtonian limit in the expected way. Such a candidate field equation is thus firmly rooted in classical physics. Typically, only mathematical knowledge familiar from the context of classical and special-relativistic physics was used in its construction. The compatibility of such a "physical candidate" with other criteria was, as a rule, less obvious and needed to be checked explicitly. If the primary goal was to stay as close to the familiar territory of classical physics as possible, the first thing to check was the conservation principle, which could turn out to be satisfied, give rise to modifications, or lead to the rejection of the candidate altogether. If the candidate survived this test, it was to be explored to what extent it complied with the generalized relativity principle, i.e., under how broad a class of coordinate transformations it would retain its mathematical form. For candidates that were not generally covariant, it had to be determined under which class of transformations the candidate was covariant and whether or not the restriction of this class was acceptable on physical grounds. In particular, it made sense to check whether at least the situations at the core of the equivalence principle, i.e., transformations to reference frames in uniform linear and rotational accelerated motion, were included in this class.

The approach we have labelled the "mathematical strategy" starts from the generalized relativity principle, i.e. from a candidate field equation which by inspection is seen to be covariant under a broad enough class of coordinate transformations. Since such a general principle of relativity was not part of classical physics, it was much less obvious than in the case of the correspondence principle what "by inspection" meant in this case. The expert mathematical knowledge of the time, however, provided him with a certain reservoir of suitable objects. Their relation to any meaningful physics was much less obvious than for a candidate of the physical strategy. It had to be checked explicitly whether such a "mathematical candidate" could be brought into agreement with the requirements of the correspondence principle. Failure to comply with the correspondence principle could lead to immediate rejection of the candidate, or generate additional conditions amounting to a restriction of the relativity principle. It could even trigger the discovery of a new way to obtain the Newtonian limit. It could also suggest how a given candidate was to be modified in order to pass the test. The situation was similar for the conservation principle, which represented another necessary condition for a physically meaningful theory. Since both the correspondence and the conservation principles could neither be circumvented nor substantially weakened, they tended, in turn, to impose restrictions on the generalized relativity principle or suggest modifications of the candidates.

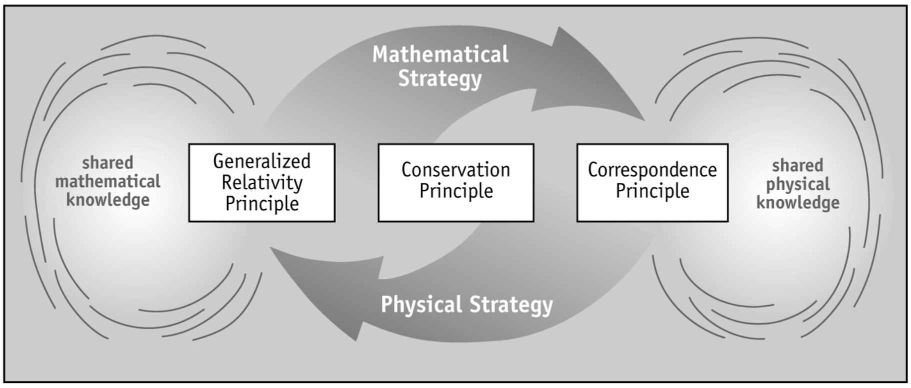

Figure 1: Einstein's double strategy arose from the different roles of the heuristic requirements of Generalized Relativity, Conservation, and Correspondence.

The two strategies are illustrated in Fig. 1 above. The physical and the mathematical strategy work with the same heuristic principles, draw on the same knowledge base of classical and special-relativistic physics, and essentially use the same mathematical representations. Why did they nevertheless produce, as we will see, different results in the course of Einstein's research? An answer is suggested by noting that the candidate solutions he examined are not determined directly by the two strategies but only through the concrete representations in which he attempted to embody his heuristic criteria. The two strategies did not act as an algorithm for producing solutions but rather as different channels for filling the Lorentz model with concrete mathematical and physical content. In other words, the two strategies constituted alternative ways for bringing to bear the available physical and mathematical knowledge on the problem of finding a gravitational field equation.

The notions of mental models and frames are helpful, we believe, for describing this process of knowledge assimilation. Because of its character as a mental model, the Lorentz model does not just represent an abstract scheme, but carries with it the experience of previous implementations. This prior experience includes the model's default settings enabling it to generate concrete candidate field equations even in the absence of sufficient knowledge about the properties of a relativistic gravitational field. The default settings make it possible to deal with the problem of insufficient knowledge by supplementing missing information drawn from prior experience. Moreover, since the experience of classical and special-relativistic physics entered the Lorentz model in the form of default settings, Einstein could give up prior assumptions in the course of his research without shattering his entire heuristic framework.

One and the same mental model may come with different sets of default settings, depending on prior experience, applications, knowledge resources, and higher-order models in which it is embedded. Default-settings depend on knowledge contexts. Classical field theory, the knowledge about Newtonian gravitation, the insights opened up by the elevator and the bucket models, the Machian interpretation of classical mechanics—all constitute different knowledge contexts relevant to the default assumptions of the Lorentz model when implemented in attempts to create a relativistic field theory of gravitation. The same is true for the mathematical resources of Gaussian surface theory, vector calculus, the theory of invariant forms, and the absolute differential calculus. Einstein's double strategy can be understood as a way of dealing with this problem of overabundant knowledge by consciously selecting alternate knowledge contexts dominating the default settings of the model. In this sense, the physical strategy, in particular, starts not just from the correspondence principle but from candidates embodying the classical knowledge about gravitation. The mathematical strategy likewise starts, not just from the generalized relativity principle but from candidates embodying the prior mathematical knowledge, in particular about generally-covariant, second-rank tensors of second order in the derivatives of the metric.

The selection of such different approaches dominating the default settings of the Lorentz model occurred initially, of course, in the hope that one or the other knowledge context would be more relevant or turn out to be more suitable to yield a full solution of the problem. Effectively, however, the alternation between different knowledge contexts led to a systematic exploration of resources that could not have been assimilated to the model all at once. The double strategy was not an astute plan for attacking the problem of finding field equations from two sides, the physical and the mathematical side. It emerged only gradually as a result of learning more about the implications of the field-theoretical model for gravitational field equations by varying its default settings. To understand Einstein's search for the gravitational field equation, it is therefore not enough to examine his heuristic principles as we have done in this chapter. We also have to reconstruct the default settings for the Lorentz model in different contexts of this search.

4. DEFAULT SETTINGS AND OPEN SLOTS IN THE LORENTZ MODEL FOR A GRAVITATIONAL FIELD EQUATION IN 1912

In this section we will introduce the principal entities figuring in Einstein's search for the gravitational field equation in the period documented by the Zurich Notebook. His research in this period focused on the problem of formulating a field equation for gravito-inertial phenomena, which had to satisfy all heuristic requirements, both those embodied in the mental model of a field equation and those that had emerged from his work between 1907 and 1912. The experience of these years had largely shaped the default assumptions that formed the starting point of Einstein's exploration of the mental model represented by the symbolic equation **OP(POT) = SOURCE**, cf. (I).

In particular, the metric tensor, which we will represent by the frame **METRIC**, was adopted as the representation of the gravitational potential and became the canonical instantiation of **POT** in the Lorentz model:

$$\mathbf{POT} =_{\mathbf{DEFT}} \mathbf{METRIC}, \qquad \mathbf{(XIII)}$$

where "$=_{\mathbf{DEFT}}$" is meant to express that the right-hand side of the equation represents the default-setting of the left-hand side. Similarly, we will refer to the energy-momentum tensor of matter and of the electromagnetic field, by the frame **ENEMO**, which became the new standard setting for **SOURCE**:

$$\mathbf{SOURCE} =_{\mathbf{DEFT}} \mathbf{ENEMO}. \qquad \mathbf{(XIV)}$$

These two key components of the gravitational field equation were generally-covariant tensors and thereby nurtured the expectation that the field equation itself would take the form of a generally-covariant tensorial equation, thus allowing Einstein to realize his ambition of creating a generalized relativity theory.

For the third component of the Lorentz model, the differential operator **OP**, the situation was more complicated. At the beginning of his search, Einstein was largely ignorant of the mathematical techniques necessary for constructing suitable candidates. The many requirements to be imposed on acceptable candidates prevented the selection of an obvious default assumption for the differential operator **OP** compatible with all these requirements.

We summarize the situation in the following figure which we will further elaborate in the following sections:

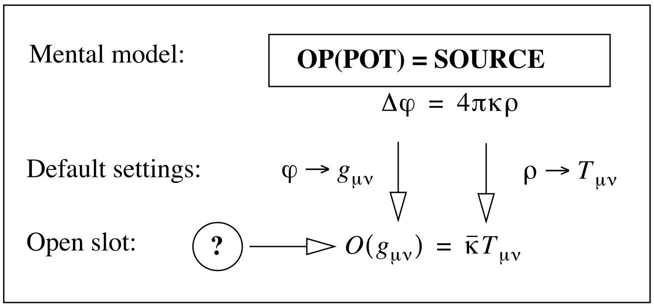

Figure 2: The Lorentz model evolved by changing the default settings for the ***POT****- and the* ***SOURCE****-slots, leaving the question of an appropriate instantiation of the* ***OP****-slot.*

4.1 The Metric as the Potential in the Gravitational Field Equation

In the middle of 1912 Einstein introduced the metric tensor as the new default setting for the **POT**- slot of the Lorentz model. This step affected both the field equation and the equation of motion. The grounds for this move had been prepared by his earlier attempts to set up a theory for the static gravitational field and his awareness that such a theory could only represent a special case within a wider framework suggested by the model. In these attempts Einstein had also learnt that Minkowski's spacetime framework for special relativity could be useful but had to be generalized for use within this larger context. This had been suggested, in particular, by Einstein's controversy with Abraham, pointing to the need for a generalization of the so-called "line element" used in the Minkowski framework, as well as by Einstein's insight into the geometrical consequences of applying special relativity to an accelerated system such as a rotating disk, pointing to the need for non-Euclidean geometry when describing gravitation.

An appropriate generalization of Minkowski's framework was found on the basis of the mathematical work of Gauss, Riemann, Christoffel, Ricci, and Levi-Civita. This led Einstein and Grossmann to the consideration of curvilinear coordinates and the introduction of a metric tensor $g_{\mu\nu}$ for a four-dimensional generalization of Gauss' theory of curved surfaces. Curvilinear coordinates are given by four functions x^{μ} with $\mu = 1, \ldots, 4$, mapping a point of spacetime to four numbers representing its coordinates similar to the use of coordinates in Gaussian surface theory. The generalized line element ds, giving the distance between two neighboring points in spacetime separated by coordinate differentials dx^{μ}, expresses a generalization of the Pythagorean theorem:

$$ds^2 = \sum_{\mu\nu} g_{\mu\nu} dx^{\mu} dx^{\nu}. \tag{21}$$

In the usual representation of Minkowski spacetime in Cartesian coordinates, this expression reduces to the four-dimensional form of the Pythagorean theorem in which the metric tensor $g_{\mu\nu}$ is given by the four-by-four matrix (c being the speed of light):

$$g_{\mu\nu} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & c^2 \end{bmatrix}. \tag{22}$$

According to the gravito-inertial model inertial forces resulting from arbitrarily accelerated motions can be interpreted as being equivalent to effects of a "dynamic" gravitational field. In the generalized formalism, this model then suggests a natural default setting for the **POT** slot of the frame as well as a natural candidate for the equation of motion in a given gravitational field. The generalized principle of relativity finds a natural expression in terms of the admissibility of arbitrary (smooth) curvilinear coordinate systems representing accelerated reference frames. The inertial motion of a particle in such a reference frame can, on the basis of the gravito-inertial model, be interpreted as motion in a special kind of gravitational field. In the generalized Minkowski formalism such a motion can be described by a geodesic curve, in complete analogy to Gaussian surface theory where geodesic curves represent the natural generalization of straight lines in Euclidean geometry. Combining these two perspectives, it becomes plausible to assume that the motion of a particle under the influence of *any* gravitational field is represented by a geodesic line in a curved spacetime.

Mathematically, a geodesic line can be described as an extremal curve in spacetime determined by a given metric tensor:

$$\delta\left\{\int ds\right\} = 0. \tag{23}$$

From a physical perspective, this equation can be seen as Hamilton's principle (cf. eq. (18)) for the Lagrangian of a free particle of mass m:

$$L = -m\frac{ds}{dt}. \tag{24}$$

The Euler-Lagrange equations (cf. eq. (19)) then suggest to consider the metric tensor $g_{\mu\nu}$ as representing the gravitational potential, i.e. **POT =$_{\text{DEFT}}$ METRIC**, as in eq. (XIII). The combination of the default setting (XIII) and the equation of motion (23) was compatible with the special case of Minkowski spacetime of special relativity where the metric tensor is given by eq. (22) and where the equation of motion of the form of eq. (23) had been developed well before Einstein had begun to work on the problem of gravitation. It was also supported by the special case a static gravitational field, as developed by Einstein in 1912, which could be integrated into the generalized Minkowski formalism in a special and seemingly natural way. Thus one could say that the **POT**-frame specializes to a $\mathbf{POT_{STAT}}$-frame in the context of Einstein's theory of static gravitation:

$$\mathbf{POT} \xrightarrow[\text{gravitation}]{\text{static}} \mathbf{POT_{STAT}}\ , \qquad \mathbf{(XV)}$$

and that the default setting for the $\mathbf{POT_{STAT}}$-frame was given by the following metric:

$$\mathbf{POT_{STAT}} =_{\mathbf{DEFT}} g_{\mu\nu} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & c^2(x, y, z) \end{bmatrix}, \qquad (25)$$

Here $c(x, y, z)$ is the gravitational potential in Einstein's static theory. In the following we shall refer to this metric as the "canonical metric for a static field." It represents Einstein's default setting for representing a static gravitational field, This default setting mediated between a generic field and the Newtonian case, and was crucial to the heuristics of the correspondence principle.

4.2 The Source-Term in the Gravitational Field Equation

In classical and special-relativistic physics, the relation between force and acceleration (cf. eq. (III)) is not the only way to characterize the effect of a force on a physical system. The effect can also be described in terms of a change in the momentum and in the energy of the system. In classical physics, the force is equal to the rate of change in *time* of the momentum, which can be symbolically expressed as (cf. eq. (XII)):

FORCE = DIFF(MOMENTUM). (XVI)

But the force is also equal to the rate of change in *space* of the energy, which can be symbolically expressed as (cf. eq. (V)):

FORCE = – GRAD(ENERGY). (XVII)

These relations also express that whenever a system gains or loses momentum and energy this must be due to the action of an external force. Note that we have here again introduced new frames **GRAD** and **ENERGY**. As with the previous examples, one could discuss different instantiations of these frames in the context of, say, classical point mechanics, special-relativistic point mechanics, or Maxwellian electrodynamics. But from this point on, we will introduce and make use of our symbolic notation in a more roundabout and indirect way, relying on an intuitive understanding that we hope is conveyed by our choice of names for our symbolic notation (**ENERGY** for energy, **ENEMO** for energy-momentum, etc.). At crucial junctures, however, we will explicitly discuss the concrete instantiations of these frames in Einstein's research and thus provide a general argument for the impact of heuristic reasoning on a qualitative level for Einstein's concrete explorations of pathways out of classical physics.

In special relativity, the concepts of energy and momentum are integrated into a single new concept, the 10-component energy-momentum or stress-energy tensor, which we symbolically represent by the frame **ENEMO** so that the relation between force, energy, and momentum can now be written as:[46]

$$\mathbf{FORCE = -\,DIV(ENEMO).} \qquad \mathbf{(XVIII)}$$

In his search for a relativistic gravitational field equation Einstein quickly realized that the source-term, i.e. the instantiation of **SOURCE** in the Lorentz model, had to be the energy-momentum tensor **ENEMO**. In our symbolic notation (cf. eq. (XIV)):

$$\mathbf{SOURCE =_{DEFT} ENEMO.}$$

Two lines of arguments, a mathematical and a physical one, made this default setting almost inescapable. From a mathematical point of view—or alternatively, from the point of view of filling the slots of the Lorentz model—something more complex than the scalar mass density was required for **SOURCE** because the gravitational potential is represented by a tensorial object. The slots on both sides of the field equation have to be filled by analogous mathematical objects. While it was in principle conceivable to construct a scalar object out of the metric tensor, e.g. by forming its determinant, and hence to have a scalar field equation, it was more plausible to Einstein that the 10-components of the metric tensor enter into some many-component field equation, just as with the many-component object representing the electromagnetic field.[47]

From a physical point of view—or alternatively from the point of view of the default settings of the Lorentz model based on prior research experience—the energy-momentum tensor had turned out to be the appropriate generalization of the concept of mass in a four-dimensional spacetime setting, i.e.:

$$\mathbf{MASS =_{DEFT} ENEMO.} \qquad \mathbf{(XIX)}$$

The elaboration of four-dimensional relativistic electrodynamics and hydrodynamics had shown that the introduction of this tensor was necessary in order to adequately describe the energetic and inertial behavior of an extended physical system.[48] In view of Einstein's expectation that, in his relativistic theory of gravitation, energy and

---

46 For the sign compare (CPAE 4, Doc. 1, 92) and "Einstein's Zurich Notebook" 05R (in this volume). The significance of the sign becomes clear when considering the energy-momentum gained or lost by a physical system, for instance in the case of a system with electromagnetic interactions. The divergence of the energy-momentum tensor of the electromagnetic field at a point describes the increase of the energy-momentum of the field at that point, which corresponds to the flow of energy-momentum from the charges to the field. This in turn equals the negative flow of energy-momentum from the field to the charges, which is given by the negative of the Lorentz force.

47 For attempts to build a scalar theory, see John Norton's discussion of Nordström's theory and Einstein's objections in "Einstein, Nordström, and the Early Demise of Scalar, Lorentz Covariant Theories of Gravitation" (in vol. 3 of this series).

mass, as well as gravitational and inertial mass, would essentially be equivalent (cf. eq. (IX)), the energy-momentum tensor was the natural candidate for the right-hand side of the field equation. To sum up, the default setting for the source slot of the Lorentz model of a field equation, eq. (XIV), was inherited from the default setting for the mass slot, eq. (XIX), resulting from the special-relativistic generalization of the mass concept of classical physics.

In order to make the choice of the default setting for **SOURCE** acceptable from the broader point of view provided by eq. (XIX), it was necessary to check whether **ENEMO** also satisfies further properties of **MASS** in classical and special-relativistic physics. The field-theoretical model suggested using the same instantiation of **MASS** both in the field equation and in the equation of motion.

The structure of an equation of motion in a gravitational field involving the **ENEMO**-frame was suggested by the special-relativistic relation between force and energy-momentum represented by eq. (XVIII). Combining this equation with the relation between force and potential, eq. (V), and the appropriate default setting for **CHARGE**, (see eqs. (VII) and (IX) one obtains:

$$\mathbf{GRAD(POT)\ x\ ENEMO = DIV(ENEMO).} \qquad \mathbf{(XX)}$$

Initially, this structural relation provided merely a heuristic hint of what a general equation of motion involving the energy-momentum tensor would look like. To validate this hint, Einstein used a default-setting for the energy-momentum tensor which allowed him to establish a connection between the proper realm of the stress-energy-momentum tensor, i.e. continuum mechanics, and the mechanics of point particles, for which an equation of motion was well established (eq. (23)).[49] In this way, he built a bridge between the knowledge embodied in eq. (XX) and the knowledge that the trajectory of a particle in a gravitational field is a geodesic.

The instantiation **ENEMO** that Einstein used to build this bridge and which, in fact, became its default setting, was the energy-momentum tensor for a swarm of independent particles ("dust"). In our symbolic notation:

$$\mathbf{ENEMO =_{DEFT} DUST,} \qquad \mathbf{(XXI)}$$

where the energy-momentum tensor for **DUST** is mathematically represented by (cf. eq. (4)):

---

48 See the discussion in "Einstein, Nordström, and the Early Demise of Scalar, Lorentz Covariant Theories of Gravitation" (in vol. 3 of this series). At some point in 1913 Einstein even convinced himself that this consideration would altogether rule out a scalar theory of gravitation, which he believed to be incompatible with the conservation laws, but then had to acknowledge that his choice of a tensorial theory with the energy-momentum tensor as a source term was reasonable but not unavoidable. His own subsequent exploration of a relativistic scalar theory of gravitation made it clear, however, that such a theory was based on *a priori* assumptions about the geometry of spacetime, which Einstein was not willing to accept. Hence even this apparently far-fetched consideration, based on Mach's critique of Newton's concept of space, contributed to stabilizing Einstein's choice of the energy-momentum tensor as the default-setting for **SOURCE**.

49 "Einstein's Zurich Notebook" 05R, p. 43R (in this volume).

$$T^{\mu\nu} = \rho \frac{dx^{\mu}}{ds}\frac{dx^{\nu}}{ds}. \tag{26}$$

In the Lagrangian formalism for a point particle in a gravitational potential given by the metric tensor (cf. eq. (24)), Einstein obtained expressions for the momentum and energy of a particle. He then applied these expressions to the energy-momentum tensor for dust and interpreted the resulting terms. In this way he arrived at an equation corresponding to the structural relation eq. (XX):[50]

$$\frac{1}{2}\sqrt{-g}\frac{\partial g_{\mu\nu}}{\partial x^{\sigma}}T^{\mu\nu} = \frac{\partial}{\partial x^{\nu}}(\sqrt{-g}g_{\sigma\mu}T^{\mu\nu}). \tag{27}$$

Introducing mixed tensor densities $\mathfrak{T}_{\sigma}^{\nu} = \sqrt{-g}g_{\sigma\mu}T^{\mu\nu}$, we can write this equation more compactly:

$$\frac{1}{2}g^{\alpha\mu}\frac{\partial g_{\mu\nu}}{\partial x^{\sigma}}\mathfrak{T}_{\alpha}^{\nu} = \frac{\partial \mathfrak{T}_{\sigma}^{\nu}}{\partial x^{\nu}}. \tag{28}$$

The agreement between this concrete result for the default setting **DUST** with the general relation eq. (XX) supported the underlying physical heuristics and suggested that this equation holds for an arbitrary symmetric energy-momentum tensor.[51]

Equation (27) turned out to be very important for Einstein's further research. First, it supported the choice of the energy-momentum tensor as the source of a gravitational field equation and stabilized this instantiation of both **SOURCE** and **MASS**. Second, it provided Einstein with one of the fundamental components for the Lorentz

---

50 "Einstein's Zurich Notebook" 05R, (in this volume). In this form eq. (27) is valid only for a symmetric tensor $T^{\mu\nu}$. See also (CPAE 6, Doc. 9, 95).

51 The terms in eq. (27) are interpreted in (Einstein and Grossmann 1913) in the following way: "We ascribe to equation (10) [i.e. our (27)] a validity range that goes far beyond the special case of the flow of incoherent masses. The equation represents in general the energy balance between the gravitational field and a arbitrary material process; one has only to substitute for [$T^{\mu\nu}$] the stress-energy tensor corresponding to the material system under consideration. The first sum in the equation contains the space derivatives of the stresses or of the density of the energy flow, and the time derivatives of the momentum density or of the energy density; the second sum is an expression for the effects exerted by the gravitational field on the material process." ("Der Gleichung [(27)] schreiben wir einen Gültigkeitsbereich zu, der über den speziellen Fall der Strömung inkohärenter Massen weit hinausgeht. Die Gleichung stellt allgemein die Energiebilanz zwischen dem Gravitationsfelde und einem beliebigen materiellen Vorgang dar; nur ist für [$T^{\mu\nu}$] der dem jeweilen betrachteten materiellen System entsprechende Spannungs-Energietensor einzusetzen. Die erste Summe in der Gleichung enthält die örtlichen Ableitungen der Impuls- bzw. Energiedichte; die zweite Summe ist ein Ausdruck für die Wirkungen, welche vom Schwerefelde auf den materiellen Vorgang übertragen werden." p.11). Indeed, except for the original derivation of eq. (27) in (Einstein and Grossmann 1913), in all later publications up to 1916, this relation appears in a form where the two conceptually distinct terms of the left-hand side are set equal, rather than in the form of eq. (27) which asserts the vanishing of a generally-covariant object that only happens to be represented by the algebraic difference of two terms as in eq. (27).

model, a general equation of motion which describes how material processes are affected by the gravitational field. Third, this equation became, as we shall see, the starting point for the formulation of the requirement of energy-momentum conservation that had to be satisfied by any candidate for the left-hand side of the gravitational field equation. Fourth, its left-hand side suggested, in connection with the relation between **FIELD** and **POT** in eq.(II), an instantiation of **FIELD**:

$$\textbf{FIELD} = -\,\textbf{GRAD(POT)} =_{\textbf{DEFT}} -\frac{1}{2}g^{\alpha\mu}\frac{\partial g_{\mu\nu}}{\partial x^{\sigma}} \equiv \tilde{\Gamma}^{\alpha}_{\sigma\nu}. \qquad \textbf{(XXII)}$$

This choice was plausible but not without alternatives. Eq. (28) can also be written as:

$$-\Gamma^{\alpha}_{\sigma\nu}\mathfrak{T}^{\nu}_{\alpha} = \frac{\partial \mathfrak{T}^{\nu}_{\sigma}}{\partial x^{\nu}}, \qquad (29)$$

with $\Gamma^{\alpha}_{\sigma\nu}$ —in the following symbolically represented as **CHRIST**—defined as minus the so-called Christoffel symbols (of the second kind):

$$\Gamma^{\alpha}_{\sigma\nu} = -\left\{ \begin{matrix} \alpha \\ \sigma\nu \end{matrix} \right\} = -\frac{1}{2}g^{\alpha\mu}\left(\frac{\partial g_{\mu\nu}}{\partial x_{\sigma}} + \frac{\partial g_{\mu\sigma}}{\partial x_{\nu}} - \frac{\partial g_{\sigma\nu}}{\partial x_{\mu}}\right). \qquad (30)$$

As a consequence, one obtains an alternative instantiation of **FIELD**:

$$\textbf{FIELD} =_{\textbf{DEFT}} -\,\textbf{CHRIST} =_{\textbf{DEFT}} \Gamma^{\alpha}_{\sigma\nu}. \qquad \textbf{(XXIII)}$$

The familiar form of the relation between field and potential in classical field theory made eq. (XXII) the natural first choice, and eq. (XXIII) only came into play when this first choice turned out to lead to difficulties.

Equation (27) had one final important implication. Written in the form:

$$\frac{\partial}{\partial x^{\nu}}(\sqrt{-g}g_{\sigma\mu}T^{\mu\nu}) - \frac{1}{2}\sqrt{-g}\frac{\partial g_{\mu\nu}}{\partial x^{\sigma}}T^{\mu\nu} = 0, \qquad (31)$$

its left-hand side could be conceived as a generic, generally-covariant differential operator known as "covariant divergence," here symbolically represented as $\textbf{DIV}_{\textbf{COV}}(\textbf{.})$, so that eq. (XX) can also be written as:[52]

$$\textbf{DIV}_{\textbf{COV}}\textbf{(ENEMO)} = \textbf{DIV(ENEMO)} - \textbf{GRAD(POT) x ENEMO} = \textbf{0}. \qquad \textbf{(XXIV)}$$

Although Einstein interpreted eq. (31) primarily from a physical point of view i.e. as a representation of the structure (XX), as we have seen, he knew, probably even before he became acquainted with the absolute differential calculus, that this equa-

52 Note that the embodiment eq. (31) of the symbolic eq. (XXIV) holds only for symmetric tensors.

tion involves a generic tensor operation which is generally covariant.[53] He had thus recognized the covariant divergence as a mathematical ingredient of his new theory that was meaningful in its own right and could in principle be used for other purposes. The formulation of eq. (31) is a prime example of how Einstein's physical strategy produced a result that turned out to be independent of the specifics of its derivation, such as the choice of **DUST** for **ENEMO**. Einstein even attempted to use the covariant divergence as a constituent of a candidate for the left-hand side of the gravitational field equation but failed because it vanishes when applied to the metric tensor.[54] The fact that the equation of motion expressed in terms of **ENEMO** turned out to be generally covariant must, in any case, have been an important confirmation of his program to establish a generally-relativistic theory of gravitation, suggesting that the other major constituent of the Lorentz model, the field equation, should also have this property.

4.3 The Differential Operator in the Gravitational Field Equation

For the differential operator acting as **OP** in eq. (I), Einstein did not have an immediately satisfactory candidate or even a heuristic shortcut for finding one. Substituting the metric tensor for the scalar gravitational potential quickly drove him out of any familiar mathematical terrain. He had to find a second-order differential operator acting on the metric tensor by relying either on attempts to directly construct such an operator or on the mathematical literature in order to find suitable starting points.

One of Einstein's earliest attempts[55] to construct a differential operator **OP** was to mimic the way in which the classical Laplace operator was formed, that is, by compounding the differential operations divergence and gradient familiar from three-dimensional vector calculus. In this way he obtained a first, natural instantiation for the differential operator on the left-hand side of the gravitational field equation:[56]

$$\mathbf{OP} =_{\mathbf{DEFT}} \mathbf{LAP} = \mathbf{DIV\ (GRAD)} \qquad \mathbf{(XXV)}$$

Applying **LAP** to the default setting for **POT**, we obtain what we will call the *core operator*:[57]

---

53 Einstein's remark "I have now found the most general equations." ("Ich habe nun die allgemeinsten Gleichungen gefunden.") in a letter to Ludwig Hopf, dated 16 August 1912 (CPAE 5, Doc. 416) in all probability refers to this insight, cf. the editorial note "Einstein on Gravitation and Relativity: The Collaboration with Marcel Grossmann" (CPAE 4, 294–301). The covariance of this equation was demonstrated in terms of the absolute differential calculus of Ricci and Levi-Civita by showing that it represents the covariant divergence of the (symmetric) contravariant stress-energy-tensor in Grossmann's "mathematical part" of (Einstein and Grossmann 1913, 32).

54 See p. 05R of "Einstein's Zurich Notebook" (in this volume).

55 The following discussion relies heavily on the analysis of Einstein's research notes contained in the Zurich Notebook. Since the actual historical path will be discussed in chapter 6, we will here only refer to the relevant pages of this notebook, without any further comments.

56 Cf. pp. 07R and 08L of "Einstein's Zurich Notebook" (in this volume).

57 Cf., e.g., p.07L of "Einstein's Zurich Notebook" (in this volume).

$$\mathbf{LAP(POT)} =_{\mathbf{DEFT}} \sum_{\alpha\beta} \frac{\partial}{\partial x^{\alpha}}\left(g^{\alpha\beta}\frac{\partial g^{\mu\nu}}{\partial x^{\beta}}\right). \qquad \mathbf{(XXVI)}$$

It has to be noted that this equation is only one of several possible instantiations of the frame **LAP**. Alternative instantiations typically involve additional factors of the determinant of the metric which typically affect the transformation behavior of the particular version of the core operator under consideration.

As a symbolic equation, the resulting tentative field equation reads:

$$\mathbf{LAP(POT) = ENEMO.} \qquad \mathbf{(XXVII)}$$

This instantiation of **OP** was supported by several arguments based, in particular, on the correspondence principle. The straightforward generalization of the Laplace operator was also plausible against the background of the field equation Einstein had developed for static gravitational fields. This field equation resulted from a simple instantiation of the Lorentz model obtained essentially by replacing the Newtonian potential in the classical Poisson equation by the variable speed of light, a move suggested by the equivalence principle:[58]

$$\Delta c = kc\rho. \qquad (32)$$

The constant k is related to the gravitational constant κ of the Poisson equation through $k = (4\pi\kappa/c^2)$.[59]

In the following, we discuss the implications of Einstein's heuristic framework for choosing and modifying the instantiations for the gravitational differential operator in his field equation: We examine the implications coming from the correspondence principle, the conservation principle, the generalized principle of relativity, and examine the Lagrangian formalism, respectively. This discussion is not meant as a substitute for a detailed account of Einstein's pathway, but as preparation for such an account by identifying the constraints under which it was pursued. These constraints

---

58 The explicit justification for this equation was follows. After noting that the variable velocity of light fulfills the Laplace equation for the matter-free case, Einstein continues: "It is easy to establish the presumably valid equation that corresponds to Poisson's equation. For it follows immediately from the meaning of c that c is determined only up to a constant factor that depends on the constitution of the clock with which one measures [the time] t at the origin of [the accelerated coordinate system] K. Hence the equation corresponding to Poisson's equation must be homogeneous in c. The simplest equation of this kind is the linear equation [eq. (32)] where k denotes the (universal) gravitational constant, and ρ the matter density." ("Es ist leicht diejenige vermutliche Gleichung aufzustellen, welche derjenigen von Poisson entspricht. Es folgt nämlich aus der Bedeutung von c unmittelbar, daß c nur bis auf einen konstanten Faktor bestimmt ist, der davon abhängt, mit einer wie beschaffenen Uhr man t im Anfangspunkte von K mißt. Die der Poissonschen Gleichung entsprechende muß also in c homogen sein. Die einfachste Gleichung dieser Art ist die lineare Gleichung [eq. (32)], wenn unter k die (universelle) Gravitationskonstante, unter ρ die Dichte der Materie verstanden wird.") (Einstein 1912b, 360)

59 The relation is obtained by identifying $c^2/2$ with the Newtonian potential φ and neglecting terms of order $(\partial_i c)^2$, cf. (Einstein 1912a, 362).

were rooted in the knowledge of classical physics, which provided the default settings for the frames with which Einstein operated. That these default settings often led to conflicting results necessitating their modification or replacement lies in the nature of Einstein's search, whose outcome could not be anticipated.

4.4 Implications of the Correspondence Principle

A gravitational field equation based on the core operator as given by eq. (XXVI) is in accordance with the correspondence principle, thus strengthening the role of this operator as an instantiation for the left-hand side of the field equation. For weak fields this differential operator reduces to the d'Alembertian operator, the default-setting for **OP** in the weak-field limit. The transition to this limiting case can be represented symbolically as:[60]

LIM(OP(.)) $=_{\mathbf{DEFT}}$ LIM(LAP(.)) $=_{\mathbf{DEFT}}$

$$\Box(\;) = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right). \tag{XXVIII}$$

The weak-field equation thus takes on the canonical form:

$$\Box g_{\mu\nu} = \kappa T_{\mu\nu}. \tag{33}$$

This equation can also be written as:

$$\Box h_{\mu\nu} = \kappa T_{\mu\nu}, \tag{34}$$

where $h_{\mu\nu}$ is defined by

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \tag{35}$$

with $|h_{\mu\nu}| \ll 1$ denoting small deviations from the Minkowski metric:

$$\eta_{\mu\nu} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & c^2 \end{bmatrix}. \tag{36}$$

If the source is taken to first order as a pressureless, static cloud of dust of density ρ (compare eq. (4)), one can neglect all terms of the energy-momentum tensor on the right-hand side of eq. (33) except for the T_{44} - term, which can be identified with the

60 Cf. (Einstein and Grossmann 1913, 13). Note that in contrast to the *Entwurf* operator, the core operator reduces to the Laplacian for a static metric of the form (25) for *strong* static fields as well. Einstein never seems to have considered this case.

gravitating mass density appearing in the classical Poisson equation. The neglected terms in the energy-momentum tensor involve the velocity of the gravitating matter which, in the Newtonian case, will be small compared to the velocity of light. If one now considers the case of a *static* weak field, introducing $\mathbf{LIM_{STAT}}$ and using the static metric of the canonical form (25) on the left-hand side of the weak-field equation, one has:

$$\mathbf{LIM_{STAT}(OP(POT)) =_{DEFT} LIM(LAP(POT_{STAT}))}. \qquad \mathbf{(XXIX)}$$

This expression reduces to the Laplace operator acting on a single component of the metric. Eq. (33) thus reduces to the familiar Poisson equation:

$$\Delta g_{44} = \kappa\rho \text{ or equivalently } \Delta h_{44} = \kappa\rho. \qquad (37)$$

The equation of motion in the Newtonian limit can be obtained from eq. (31) under similar assumptions, i.e., small velocities and a weak static field. The result is:

$$\frac{d^2 x_i}{dt^2} = -\frac{1}{2}\frac{\partial g_{44}}{\partial x_i} = -\frac{1}{2}\frac{\partial h_{44}}{\partial x_i} \qquad (i = 1, 2, 3). \qquad (38)$$

This equation shows that $g_{44}/2$ resp. $h_{44}/2$ plays the role of the Newtonian gravitational potential.

The assumption that the left-hand side of the field equation has the form of eq. (XXVIII) is not independent from the assumption that the metric tensor for weak static fields has the form (25). Under appropriate circumstances, a weak-field equation of this form gives rise to solutions precisely of this canonical form.[61] In other words, the most natural assumption for the form of a weak-field equation and the most natural assumption for the metric of a static field supported each other. A further argument supporting Einstein's understanding of the correspondence principle as implying a canonical metric of the form (25) was independent from the field equation but also related to the roots of this principle in the framework of classical physics.[62] This argument is based on Galileo's principle, that is, the requirement that all bodies fall with the same acceleration in a given gravitational field, and makes use of the basic relations between force, momentum, energy, and acceleration as understood in classical physics, with some additional ingredients from special relativity such as the equivalence of mass and energy. Einstein argued that particles with different energy, and hence different inertial mass, fall with different accelerations in a static gravitational field, *unless* such a field is represented by a metric tensor of the canonical form (25). As a criterion for the validity of Galileo's principle he used the requirement that the ratio of the force acting on a particle and its energy depend neither on the particle's mass nor on its velocity.

---

61 See (Norton 1984, 120–121).

62 The following argument is based on a reconstruction of p. 21R of "Einstein's Zurich Notebook" (in this volume).

4.5 Implications of the Conservation Principle

The fact that the core operator was firmly anchored in knowledge about the familiar cases of static and Newtonian gravitation made it the natural starting point for Einstein's "physical strategy." The core operator, however, had to pass a number of further checks, which could result in modifications. In particular, it remained to be seen how the operator could be brought into agreement with the conservation principle and the generalized relativity principle.

An acceptable field equation (I) had to be compatible with the equation of motion and the related structural insight into energy-momentum conservation represented by eq. (XX). This compatibility could be checked by replacing **ENEMO** in eq. (XX) by the left-hand side of the field equation, i.e. by **OP**:

$$\mathbf{GRAD(POT)\ x\ OP(POT) = DIV(OP(POT))}, \qquad \mathbf{(XXX)}$$

or, in the notation of eq. (XXIV):

$$\mathbf{DIV_{COV}(OP(POT)) = 0}. \qquad \mathbf{(XXXI)}$$

It was necessary to check whether this "conservation compatibility check" could be satisfied for a given candidate field equation if need be by imposing extra conditions, in addition to the field equation.

In the course of Einstein's research documented in the Zurich Notebook it became clear that the conservation compatibility check fails for a field equation based on the core operator:

$$\mathbf{GRAD(POT)\ x\ LAP(POT) \neq DIV(LAP(POT))}. \qquad \mathbf{(XXXII)}$$

This problem may not have surprised Einstein as it was already familiar to him from his theory of static gravitational fields. There he had also encountered the difficulty that the first choice of a field equation for the static field (eq. (32)) turned out to be incompatible with momentum conservation.[63] To demonstrate this conflict, Ein-

63 Cf. Einstein's second thoughts about the paper expressed in a letter to the editor of the *Annalen der Physik*, Wilhelm Wien: "I asked you this morning to return my manuscript, and now I am asking you to keep it after all. To be sure, not everything in the paper is tenable. But I think I should let the thing stand as it is, so that those interested in the problem can see how I arrived at the formulas." ("Heute Morgen bat ich Sie, mir mein Manuskript zurückzusenden und nun bitte ich Sie es doch zu behalten. Es ist zwar nicht alles haltbar, was in der Arbeit steht. Aber ich glaube die Sache doch so lassen zu sollen, damit diejenigen, welche sich für das Problem interessieren, sehen, wie ich zu den Formeln gekommen bin.") Einstein to Wien, 11 March 1912, (CPAE 5, Doc. 371). Since (Einstein 1912b), which contains eq. (32) was received by the *Annalen* on February 26, and (Einstein 1912c) where the problem with this equation is discussed, was received four weeks later, on March 23, the problem Einstein refers to in the letter to Wien is most probably the incompatibility with the conservation principle discussed in the following. To the published discussion of the potential equation for c in (Einstein 1912b) Einstein added a footnote reading: "A soon to be published paper will show that equation (5a) and (5b) cannot yet be exactly right. However, they will be provisionally used in the present paper." ("In einer in kurzem nachfolgender Arbeit wird gezeigt werden, daß die Gleichungen [$\Delta c = 0$] und [eq. (32)] noch nicht exakt richtig sein können. In dieser Arbeit sollen sie vorläufig benutzt werden.")

stein considered an assembly of masses fixed to a rigid, massless frame and showed that this assembly of masses would set itself in motion if the field equation were assumed to be (32). Following the logic underlying eq. (XXX), he substituted the left-hand side of the field equation for the mass density ρ in the expression for the force-density (compare eq. (V)):

$$\boldsymbol{F} = -\rho \operatorname{grad} c. \tag{39}$$

The integral of this expression over space (under the assumption that c is constant at infinity) should vanish on account of momentum conservation. However, the expression

$$\boldsymbol{F} = -\frac{1}{k}\frac{\Delta c}{c} \operatorname{grad} c, \tag{40}$$

resulting from this substitution cannot be transformed into a divergence expression, and momentum conservation is violated. The rigid massless frame would start to move, in contradiction with Newton's principle *actio* = *reactio*.

It is easily seen that the Poisson equation of classical mechanics and electrostatics does not present this problem. In a later paper Einstein himself explained how this can be shown in a way that suggests a generalization of the argument to the case of a relativistic gravitational field theory.[64] In electrostatics the ν th component of the momentum conferred to matter per unit volume and time (or the force density, compare (V)) is:

$$-\frac{\partial \varphi}{\partial x_\nu}\rho,$$

where φ represents the potential and ρ the density of the electrical charge. It can then be demonstrated that a field equation of the form (cf. eq. (5)):

$$\sum_\nu \frac{\partial^2 \varphi}{\partial x_\nu^2} = -\rho$$

satisfies the requirement of momentum conservation. This is done by showing that the rate of change of momentum:

$$-\frac{\partial \varphi}{\partial x_\nu}\cdot\rho = \frac{\partial \varphi}{\partial x_\nu}\sum_\mu \frac{\partial^2 \varphi}{\partial x_\mu^2}$$

can be transformed into a divergence expression, i.e., an expression with the property that the integral over a closed system vanishes so that the total momentum is conserved.

The challenge resulting from the problem with Einstein's first static theory was to find an expression for the force, the momentum transferred from the gravitational

64 See (Einstein and Grossmann 1913, part 1, §5).

field to material processes that can be written as a divergence. Let us try to capture the heuristic behind his reasoning in our symbolic notation. We conceive of the **FORCE**-frame as a divergence of some **FIELDMASS**-frame in an equation of the form:

FORCE = DIV(FIELDMASS), **(XXXIII)**

where **FIELDMASS** represents, in the three-dimensional case, the momentum (or, alternatively, the energy) and, in the four-dimensional case, the energy-momentum of the gravitational field. Such a force expression had to be extracted from a revised field equation in which the default setting **LAP** is replaced by a modified frame, let us call it **GRAV** for **OP**:

OP(POT) $=_{\mathbf{DEFT}}$ GRAV(POT) = LAP(POT) + CORR(POT). **(XXXIV)**

The correction term **CORR** introduced in the new choice for **OP** in eq. (XXXIV) had to be compatible, of course, with the correspondence principle and in particular with the default setting eq. (XXVIII) so that the condition

LIM(CORR(POT)) = 0 **(XXXV)**

follows. The correction term has to make sure that both eq. (XXX), the conservation compatibility check, and eq. (XXXIII), the equivalent divergence condition for the gravitational force, are satisfied:

FORCE =

– GRAD(POT) X GRAV = – DIV(GRAV) = DIV(FIELDMASS). **(XXXVI)**

In view of the definition of **GRAV** as a sum of **LAP** and **CORR** (see eq. (XXXIV)) one thus obtains the following symbolic equation:

LAP(POT) x GRAD(POT) =

– DIV(FIELDMASS) – CORR(POT)) x GRAD(POT). **(XXXVII)**

The crucial result is that this relation suggests a generic operational procedure for identifying the desired correction term, regardless of specific instantiations of the frames involved. From the force expression for Einstein's first field equation for static fields, eq. (40), it follows that the term which serves as the starting point for such a procedure, corresponding to **LAP(POT) x GRAD(POT)**, is:[65]

$$\frac{1}{c}\partial_i\partial_i c\partial_k c, \tag{41}$$

By repeated application of the Leibniz rule for the differentiation of products, one obtains an equation of the form (XXXVII):

65 In the following we assume summation over repeated (spatial) indices.

$$\frac{1}{c}\partial_i\partial_i c\partial_k c = \partial_i T_{ik} + \frac{1}{2c^2}\partial_i c\partial_i c\partial_k c \tag{42}$$

with

$$T_{ik} = \frac{1}{c}\partial_i c\partial_k c - \frac{\delta_{ik}}{2c}\partial_j c\partial_j c, \tag{43}$$

so that

$$\partial_i T_{ik} = \left(\frac{1}{c}\partial_i\partial_i c - \frac{1}{2c^2}\partial_i c\partial_i c\right)\partial_k c \tag{44}$$

is a divergence term and corresponds to **– DIV(FIELDMASS)**, while

$$\frac{1}{2c^2}\partial_i c\partial_i c\partial_k c \tag{45}$$

corresponds to **– CORR(POT)) x GRAD(POT)**. Eq. (45) therefore gives the correction term necessary to satisfy the conservation principle. In other words, this principle not only served to refute the first static field equation (32), it also provided Einstein with a procedure for constructing a modified field equation complying with this principle.

Einstein thus arrived at a new field equation (Einstein 1912b), the core of his so-called "second theory:"[66]

$$\frac{\Delta c}{c} - \frac{\text{grad}^2 c}{2c^2} = k\sigma. \tag{46}$$

Since this revised equation no longer represents a direct analogue of the Poisson equation, Einstein faced the challenge to find a plausible physical interpretation of it. He had to reexamine both the equivalence principle and the role of energy and momentum conservation. A remarkable feature of eq. (46) is that the first derivative of the gravitational potential enters in a non-linear way so that the left-hand side of eq. (46) may be symbolically expressed with the help of eqs. (II) and (XXXIV) as:

GRAV =$_{\text{DEFT}}$ DIV(FIELD) + FIELD^2 (XXXVIII)

The second term had not been encountered before in working with the mental model of a gravitational field theory. It also threatened one of Einstein's key heuristic assumptions, the principle of equivalence, which could only be upheld for infinitesimally small fields.[67] This restriction made it all the more pressing to provide a plausible physical justification for the correction term. Einstein found such a justification in implications of both field theory and special relativity, i.e. in the fact that a field may

66 The theory advanced in this paper is commonly referred to as Einstein's second theory of static gravitation. Its main difference pertains to the amended field equation.

carry energy and that any kind of energy, being equivalent to mass, should act as a source of the gravitational field.[68]

The physical interpretation of the modified field equation (46) is brought out more clearly by rewriting it as:

$$\Delta c = k\left(c\sigma + \frac{1}{2ck}\text{grad}^2 c\right). \tag{47}$$

This form of the equation suggests that the term

$$\frac{1}{2ck}\text{grad}^2 c,$$

which appears on the right-hand side on the same footing as the material source, be interpreted as the energy density of the gravitational field acting as its own source.

This physical interpretation also supported the conclusion that the general field equation would be non-linear, a conclusion which, after this experience with the special case of the static field, became a standard expectation in Einstein's further search. In terms of our symbolic equations, the revised form of the generic field equation could either be expressed with the help of the **GRAV** and **CORR**-frames (see eq. (XXXIV)) or with the help of the **FIELDMASS**-frame, representing in the general, four-dimensional case the energy-momentum of the gravitational field, as:[69]

NORM(POT) = ENEMO + FIELDMASS. (XXXIX)

NORM(POT) thus represents a new setting of the differential operator slot **OP** allowing the field equation (I) to be written in the "normal" form of eq. (XXXIX); we thus define

OP $=_{\mathbf{DEFT}}$ NORM, (XL)

with a corresponding new setting for **SOURCE**:

---

67 "Thus, it seems that the only way to avoid a contradiction with the reaction principle is to replace equations (3) and (3a) with other equations homogeneous in c for which the reaction principle is satisfied when the force postulate (4) is applied. I hesitate to take this step because by doing so I am leaving the territory of the unconditional equivalence principle. It seems that the latter can be maintained for infinitely small fields only." ("Eine Beseitigung des genannten Widerspruches gegen das Reaktionsprinzip scheint also nur dadurch möglich zu sein, daß man die Gleichungen [$\Delta c = 0$] und [eq. (32)] durch andere in c homogene Gleichungen ersetzt, für welche das Reaktionsprinzip bei Anwendung des Kraftansatzes [39] erfüllt ist. Zu diesem Schritt entschließe ich mich deshalb schwer, weil ich mit ihm den Boden des unbedingten Äquivalenzprinzips verlasse. Es scheint, daß sich letzteres nur für unendlich kleine Felder aufrechterhalten läßt.") (Einstein 1912c, 455–456)

68 The discussion referred to in the following is introduced in (Einstein 1912c) by the phrase: "The term added in equation (3b) [$c\Delta c = \frac{1}{2}(\text{grad } c)^2 = kc^2\sigma$] in order to satisfy the reaction principle wins our confidence thanks to the following argument." ("Das in Gleichung (3b) zur Befriedigung des Reaktionsprinzipes hinzugesetzte Glied gewinnt unser Vertrauen durch die folgenden Überlegungen."), p. 456–7.

69 For the following, see "Untying the Knot ..." (in vol. 2 of this series), sec. 3.

$$\textbf{SOURCE} =_{\textbf{DEFT}} \textbf{ENEMO + FIELDMASS.} \qquad \textbf{(XLI)}$$

This form of the field equation clearly brings out the parallelism between the energy-momentum of matter and the energy-momentum of the gravitational field. Eq. (XXXIX) is the symbolic expression of what eventually became Einstein's standard or normal expectation for the form of a field equation with the property that it is compatible with the conservation principle and with the requirement that gravitational energy and momentum enter the field equation on the same footing as the energy and momentum of matter. With this normal form the conservation principle takes on a particularly simple form. From the last equality in eq. (XXXVI) and the field equation it follows that:

$$\textbf{DIV(ENEMO) + DIV(FIELDMASS) =}$$

$$\textbf{DIV(ENEMO + FIELDMASS) = 0.} \qquad \textbf{(XLII)}$$

This symbolic equation expresses the expectation that the conservation laws should hold for gravitation and matter taken together. Accordingly, the conservation compatibility check for **NORM(POT)** becomes

$$\textbf{DIV(NORM) = 0,} \qquad \textbf{(XLIII)}$$

(cf. eq. (XXXI))

It was natural to expect that **NORM** would take on the classical form of a divergence of the field, generated both by material processes and the energy-momentum of the gravitational field itself. The field operator might be brought into such a simple form, resembling the familiar structure from electromagnetic field theory by some appropriate mathematical manipulation, involving the source-term of the field equation as well. In other words, one would have the revised settings:

$$\textbf{OP(POT)}=_{\textbf{DEFT}}\textbf{NORM(POT)}_{\textbf{CLASS}} \textbf{ = DIV(FIELD),} \qquad \textbf{(XLIV)}$$

with a corresponding setting for **SOURCE**:

$$\textbf{SOURCE} =_{\textbf{DEFT}} \textbf{ENEMO + FIELDMASS.} \qquad \textbf{(XLV)}$$

Note, however, that the requirement expressed by eq. (XLII) may not be compatible with the requirement expressed by eq. (XXXIX) if the particular form eq. (XLIV) for the left-hand side of the field equation is imposed.[70]

In summary, Einstein's experiences with implementing the conservation principle in his theory for static gravitational fields turned out to be of crucial significance for his further research, shaping the expectation for the differential operator in the generic gravitational field equation. Reflecting on these experiences, he could conclude, in particular, that

– the field equation would probably be non-linear and contain a term representing the gravitational field acting as its own source;

70 See the discussion in "Untying the Knot ..." (in vol. 2 of this series), secs. 3, §3.

– just as with the static field equation, the field equation might have to be found in two steps with a first step involving a linear second-order differential operator and a second step involving the non-linear, first-order correction terms;
– the correction term might be identified by trying to establish an energy-momentum balance, beginning with a linear, second-order differential operator as a first step.

4.6 Implications of the Generalized Relativity Principle

When starting from an instantiation of the left-hand side of the mental model of a gravitational field equation rooted in physical knowledge such as the core operator, the most challenging problem was to identify its transformation properties and to find out whether or not they allow the implementation of a generalized principle of relativity. Alternatively, one could start from an instantiation rooted in mathematical knowledge. While the physical strategy automatically takes care of the correspondence principle, the mathematical strategy automatically takes care of the generalized relativity principle. In the latter case the main challenge was the implementation of the correspondence and conservation principles, including a check of their mutual compatibility. In the course of his research, Einstein developed a strategy for addressing this challenge. This strategy involved replacing one immediately given default setting for **OP** by a more sophisticated one, better adapted to the purpose at hand. In this respect, the strategy resembles the strategy discussed above for adapting a setting suggested by the correspondence principle to the necessities implied by the conservation principle, i.e., for the transition from **LAP(POT)** to **GRAV(POT)**.

Instantiations for **OP** suggested by the mathematical strategy typically have well-defined transformation properties (e.g. are generally covariant). As his research proceeded Einstein familiarized himself with the relevant mathematical literature, in collaboration with his mathematician friend Marcel Grossmann.[71] While the mathematical horizon enlarged it came to include more and more sophisticated mathematical objects. At the beginning, the mathematical instrumentarium was limited to that of linear vector and tensor analysis in four dimensions as developed by Minkowski, Sommerfeld, and Laue.[72] After a number of unsuccessful attempts to employ these techniques in the construction of a suitable differential operator,[73] the core operator emerged as the most satisfactory candidate which could be obtained at this level of mathematical sophistication. The core operator, however, was covariant only under linear transformations and did thus not lead to a substantial generalization of the relativity principle.

---

71 For a discussion of the Grossmann's role in the search for and reception of pertinent mathematical literature, see (Pais 1982, chap.12c; Norton 1992b, appendix; Reich 1994, chap. 5.3; CPAE 4, 294).

72 Cf. note 30.

73 Cf. pp. 39L–40L of "Einstein's Zurich Notebook" (in this volume).

Einstein subsequently became familiar with the so-called Beltrami invariants.[74] These mathematical objects, in particular the second Beltrami invariant, could be seen as a generalization of the ordinary Laplace operator and must have looked promising. They are generally covariant and thus provide a good starting point for pursuing the mathematical strategy. It was difficult, however, to see how the second Beltrami invariant, defined only for scalar functions, could be applied to a gravitational potential represented by the metric tensor. Einstein thus had two plausible but mutually incompatible default settings, the second Beltrami invariant for the differential operator, and the metric tensor for the gravitational potential. Einstein's dilemma at this point is illustrated in Fig. 3.

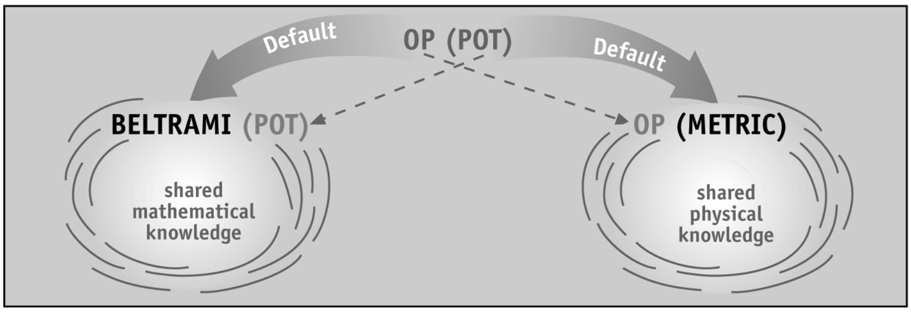

Figure 3: The incompatibility of instantiating the operator slot of Einstein's mental model of a gravitational field equation with the Beltrami invariant, and the potential slot with the metric produced posed a dilemma for Einstein.

The breakthrough for the mathematical strategy came when Einstein got acquainted with the Riemann tensor and its potential to produce suitable candidates for the differential operator in the gravitational field equation. The Riemann tensor represents a second-order differential operator on the metric and is generally covariant. Moreover, by a general theorem any generally-covariant differential operator, which consists of the metric components and its derivatives, and contains no higher than second-order derivatives and is linear in those, can be constructed from the Riemann tensor by tensor-algebraic operations.[75] It must have been clear to Einstein from the outset that the Riemann tensor itself could not play the role of an instantiation for **OP**. First, since the energy-momentum tensor appearing on the right-hand side of the field equation is a second-rank tensor with two indices, the differential operator on the left-hand side was required to have the same property. The Riemann tensor, however, is a fourth-rank tensor, with four indices. Second, a field equation with the Riemann tensor on the left-hand side would be much too restrictive. It would require that, outside the sources, the metric would be strictly Minkowskian so no

74 Cf. pp. 06L–07L of "Einstein's Zurich Notebook" (in this volume).
75 See (Einstein and Grossmann 1913, part II, §4). See also (Bianchi 1910).

non-trivial gravitational potential could exist, a conclusion, for instance, manifestly wrong for the field of a point mass.

A second-rank tensor serving as a natural candidate for **OP**, however, could be extracted from the Riemann tensor in various ways. Let us designate the frame of such a candidate by **RIEM**:

$$\mathbf{OP(POT) =_{DEFT} RIEM(POT).} \qquad \mathbf{(XLVI)}$$

As was pointed out above, such candidates come with assurances about their behavior under coordinate transformations (here designated as **TRAFO**). They inherit these transformation properties from their progenitor, the fourth-rank Riemann tensor. The default setting for this property is general covariance (here designated as **GCOVARIANT**):

$$\mathbf{TRAFO(RIEM) =_{DEFT} GCOVARIANT.} \qquad \mathbf{(XLVII)}$$

When relating the frame **RIEM** to the default settings **GRAV** or **NORM** for **OP** suggested by the correspondence and conservation principles (cf. eq. (XXXIV) and eq. (XL)), one typically finds a relation of the form:

$$\mathbf{RIEM(POT) = GRAV(POT) + DIST(POT),} \qquad \mathbf{(XLVIII)}$$

where **DIST(POT)** represents “disturbing” terms incompatible with the requirements of the correspondence and conservation principles. To obtain from **RIEM(POT)** a “reduced” candidate satisfying these principles one has to impose the revised default setting for the left-hand side of the gravitational field equation:

$$\mathbf{OP(POT) =_{DEFT} RIEM_{RED}(POT) = RIEM(POT) - DIST(POT),} \qquad \mathbf{(XLIX)}$$

which can be obtained from **RIEM** either by requiring that

$$\mathbf{DIST(POT) = 0.} \qquad \mathbf{(L)}$$

or by requiring that **DIST(POT)** behaves as a tensor under some group of coordinate transformations, in which case it can be subtracted leaving a reduced candidate invariant under this now restricted group of transformations:

$$\mathbf{TRAFO(RIEM_{RED}) => TRAFO(DIST),} \qquad \mathbf{(LI)}$$

Conditions such as (XLVII) can typically be derived from first-order conditions on the metric tensor, corresponding to a restriction of the admissible coordinate systems (here designated as **COORD(POT)**):

$$\mathbf{COORD(POT) = 0 => DIST(POT) = 0.} \qquad \mathbf{(LII)}$$

Such a coordinate restriction comes in turn with its own transformation behavior, but typically is at least covariant at least under linear transformations:

$$\mathbf{TRAFO(COORD) =_{DEFT} LINEAR.} \qquad \mathbf{(LIII)}$$

Coordinate systems selected in this way assumed for Einstein the role of privileged reference frames, similar to the distinguished role of inertial reference systems in classical physics. It is in these preferred coordinate systems that the physical laws are supposedly valid in their usual form. The condition **COORD(POT) = 0** represented for him a true limitation of the generalized relativity principle and is therefore referred to here as a “coordinate restriction.”

From a modern perspective, the relation between a generally-covariant candidate for the left-hand side of the field equation and the condition expressed by eq. (LII) can be interpreted in an entirely different way: Since the Newtonian theory clearly does not hold in arbitrary coordinate systems, while generally-covariant field equations do, special coordinates have to be introduced to obtain the Newtonian limit. A *coordinate condition* in the modern sense, however, does not have the meaning of an overall restriction on the choice of admissible coordinates; it is only a tool adapted for this specific purpose. This tool in no way imposes a restriction on the covariance of the field equation, but is available precisely *because* of it. For the Einstein of the Zurich Notebook, however, it was more natural to think of eq. (LII) as a coordinate *restriction,* valid not only in the context of a special situation such as that of the Newtonian limit but necessary in general to ensure that the candidate gravitation tensor takes on the canonical form of eq. (XXXIV). The transformation properties of $\mathbf{RIEM_{RED}(POT)}$ are thus constrained by those of the coordinate restriction, a relation we can express as:

$$\mathbf{TRAFO(RIEM_{RED}) => TRAFO(COORD).} \qquad \mathbf{(LIV)}$$

An additional restriction of the generalized relativity principle typically follows from the conservation principle, given that its mathematical implementation (e.g., by eq. (XLIII)) does, in general, not lead to a generally-covariant equation:

$$\mathbf{TRAFO(DIV(NORM)) \neq GCOVARIANT.} \qquad \mathbf{(LV)}$$

Just as with the correspondence principle (cf. eq. (LII)), the condition **DIV(NORM) = 0** may be inferred from a simpler, possibly first-order condition representing the restriction to coordinate systems in which the conservation principle holds:

$$\mathbf{ENERG(POT) = 0 => DIV(NORM) = 0.} \qquad \mathbf{(LVI)}$$

For the transformation properties of the gravitational field equation we thus have similarly:

$$\mathbf{TRAFO(NORM) => TRAFO(ENERG),} \qquad \mathbf{(LVII)}$$

or, taken together with relation (LIV), replacing $\mathbf{RIEM_{RED}}$ and **NORM** by **GRAV**:

$$\mathbf{TRAFO(GRAV) => TRAFO(COORD) + TRAFO(ENERG).} \qquad \mathbf{(LVIII)}$$

This relation expresses that the transformation properties of the left-hand side of the gravitational field equation are restricted by the needs of the correspondence and the conservation principles taken together. In summary:

$$\mathbf{TRAFO(GRAV) <=>}$$

$$\mathbf{TRAFO(RIEM) + TRAFO(COORD) + TRAFO(ENERG).} \qquad \mathbf{(LIX)}$$

What this symbolic equation says is that the transformation properties of the field equation are known if those of the original default setting rooted in mathematical knowledge are given together with those of the coordinate restrictions imposed to satisfy the correspondence and the conservation principles.

The above considerations raise the more general problem of the compatibility between the mathematical implementations of the correspondence and conservation principles (the corresponding compatibility condition is designated here as **CC-COMP(GRAV)**). While this question could typically be dealt with at the level of the compatibility of the respective coordinate restrictions eq. (LII) and eq. (LVI), it was conceivable that the compatibility requirement gave rise to new conditions with implications not only for the transformation properties of the field equation but for other questions as well, including the question of whether the given default setting for **GRAV** was acceptable at all:

CC–COMP(GRAV) => (COORD = 0) + (ENERG = 0). **(LX)**

It was also conceivable that a conflict between a candidate for **GRAV** and the correspondence and conservation principles arose because the default setting for the metric of static gravitational fields eq. (25) was incompatible with one of the coordinate restrictions following from these principles:

COORD(POT$_{\mathbf{STAT}}$) ≠ 0, **(LXI)**

ENERG(POT$_{\mathbf{STAT}}$) ≠ 0. **(LXII)**

4.7 Implications of the Lagrange Formalism

At some point in his research, Einstein realized the significance of the Lagrange formalism not only for formulating the equation of motion but also for deriving the field equation.[76] Because of its earlier application in the context of classical electromagnetic field theory this formalism came with its own default-settings, which played an important role in Einstein's search for the gravitational field equation. Classical field theory suggested, in particular, to choose a Lagrangian quadratic in the field:

LAGRANGE =$_{\mathbf{DEFT}}$ FIELD$^{\mathbf{2}}$. **(LXIII)**

The Lagrangian for the free Maxwell field for instance is of this form (cf. eq. (11)):

$$L = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{48}$$

The use of the Lagrange formalism had two immediate advantages for Einstein. First, when following the physical strategy, he could focus on a scalar object, the Lagrangian, to explore the transformation properties of his theory rather than on the

76 The first paper in which he made use of the Lagrangian formalism for this purpose is (Einstein and Grossmann 1914). This approach was fully developed in (Einstein 1914a).

more complex tensorial objects representing candidates for the left-hand side of the field equation:[77]

TRAFO(GRAV) <=> TRAFO(LAGRANGE). **(LXIV)**

Second, when following the mathematical strategy, he could rely on an expression for **FIELDMASS** directly delivered by this formalism in terms of the Lagrangian to explore the validity of the conservation principle. The formalism produces a field equation which can easily be brought into a form corresponding to the default settings eq. (XLIV) and eq. (XLV):

DIV(FIELD) = ENEMO + FIELDMASS. **(LXV)**

It remains, of course, to be checked in each concrete case whether the resulting expression for **FIELDMASS** is compatible with the expectation for such an expression following from the conservation principle and, in particular, with eq. (XLII).[78]

The introduction of the Lagrange formalism had one further consequence for Einstein's search, which eventually turned out to be decisive for identifying the gravitational field equation of general relativity. Due to the default setting eq. (LXIII), the Lagrange formalism helped to highlight the importance of the **FIELD**-frame, pointing to the alternative between eq. (XXII) and eq. (XXIII), one leading to the non-covariant *Entwurf* theory, the other to an essentially generally-covariant theory which quickly opened up the pathway toward the field equation of general relativity.

5. TESTING THE CANDIDATES: EINSTEIN'S CHECK LIST FOR GRAVITATION TENSORS

The reservoir of candidates for the left-hand side of the field equation **OP** available to Einstein was determined by the mathematical knowledge available to him. Roughly three levels of knowledge can be distinguished, each coming with its own set of candidates as shown in Fig. 4 below. Not all candidates played the same prominent role in Einstein's research. The four most important ones were the *Entwurf* operator, the Ricci tensor, the Einstein tensor, and the November tensor (Einstein 1915a).

Einstein examined these four differential operators twice in the course of two exploratory phases of his work. He first confronted them with his heuristic requirements in the period documented by the Zurich Notebook dating from the winter 1912–1913, and then once more in the fall of 1915, as documented by publications and correspondence. He came to different conclusions in these two stages of his work. Before we discuss in detail in which way his research experience led him to

77 Einstein's point of view was, however, criticized by the mathematician Tullio Levi-Civita, who contested that the Euler-Lagrange equations have the same covariance group as the Lagrangian in the case of the *Entwurf* theory. See, e.g., Tullio Levi-Civita to Einstein, 28 March 1915 (CPAE 8, Doc. 67).

78 For the detailed mathematical considerations, see "Untying the Knot ..." (in vol. 2 of this series), secs. 3.1 and 3.2.

these different views, we shall systematically examine how the various candidates fare when confronted with his heuristic requirements and give an overview of the results of his checks. In this way, we shall be able to establish the potential of these candidates independently of their actual role in the dramatic history of Einstein's search for the field equation. As a consequence, the twists and turns of this search will become understandable as reactions to the epistemic constraints and potentials inherent in the knowledge resources available to Einstein. These constraints are largely embodied in the mental models and frames guiding his research, as well as in their default settings and the instantiations of their open slots. But the conflicting implications of these default settings and instantiations were only revealed in the course of Einstein's elaboration of his theory on the level of concrete mathematical

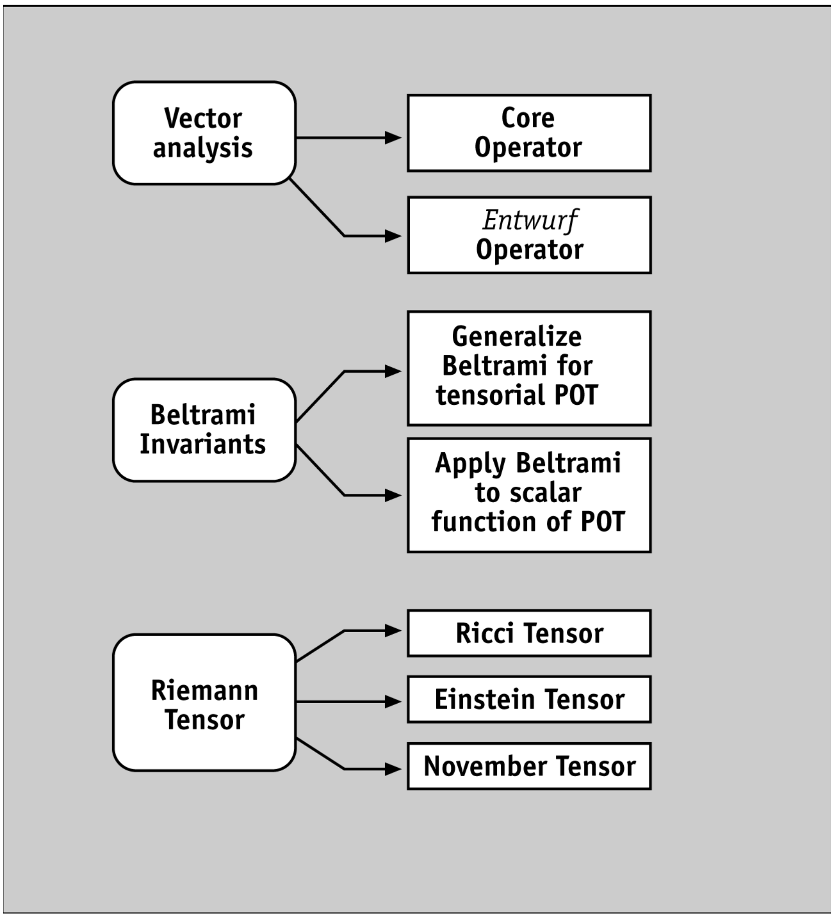

Figure 4: A list of Einstein's most important candidates for differential operators to fill the operator slot of his mental model of a gravitational field equation.

representation. As a matter of fact, the fate of a candidate not only depended on the structural constraints of his heuristics but also on the *order* in which these structures were implemented, on the *exploration depth* with which they were treated, and on the *perspective* under which Einstein examined the answers to his questions. As we shall show in more detail in the next section, in all three of these *performative dimensions* of his evaluation of candidates, the situation of the winter of 1912–1913 was very different from that of October and November 1915 when he once more examined these candidates.

5.1 The Entwurf Operator and the Correspondence Principle in the Winter of 1912–1913

The *Entwurf* operator, first written down in the Zurich Notebook and then published by Einstein and Grossmann in the Spring of 1913, gives rise to the field equations:

$$\mathbf{ENTWURF} =_{\mathbf{DEFT}} \frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^{\alpha}}\left(g^{\alpha\beta}\sqrt{-g}\frac{\partial g^{\mu\nu}}{\partial x^{\beta}}\right) - g^{\alpha\beta}g_{\tau\rho}\frac{\partial g^{\mu\tau}}{\partial x^{\alpha}}\frac{\partial g^{\nu\rho}}{\partial x^{\beta}} = \kappa(T^{\mu\nu} + t^{\mu\nu}) \quad (49)$$

with the following expression for the gravitational energy-momentum:

$$\mathbf{FIELDMASS} =_{\mathbf{DEFT}} -\kappa t^{\mu\nu} = \frac{1}{2}g^{\mu\alpha}g^{\nu\beta}\frac{\partial g^{\tau\rho}}{\partial x^{\beta}}\frac{\partial g_{\tau\rho}}{\partial x^{\alpha}} - \frac{1}{4}g^{\mu\nu}g^{\alpha\beta}\frac{\partial g^{\tau\rho}}{\partial x^{\alpha}}\frac{\partial g_{\tau\rho}}{\partial x^{\beta}}. \quad (50)$$

The *Entwurf* field equations satisfy the correspondence principle just as the core operator does since the correction terms which distinguish the two vanish in the limiting procedure for obtaining the Newtonian theory. One has, in particular, (cf. eq. (XXIX)):

$$\mathbf{LIM_{STAT}(ENTWURF) = LIM(LAP(POT_{STAT}))}. \quad \mathbf{(LXVI)}$$

5.2 The Entwurf Operator and the Conservation Principle in the Winter of 1912–1913

By their very construction, the *Entwurf* field equations satisfy the conservation principle since the correction terms distinguishing them from the core operator are generated in such a way that an identity of type (XXXVII) holds. Evidently, the field equations (49) are of the form (XXXIX), while an equation of the form (XLII) expresses the conservation principle:

$$\frac{\partial}{\partial x^{\nu}}(\boldsymbol{T}^{\mu\nu} + \boldsymbol{t}^{\mu\nu}) = 0. \quad (51)$$

5.3 The Entwurf Operator and the Generalized Relativity Principle in the Winter of 1912–1913

The principal challenge for the *Entwurf* theory was the question of the transformation properties of the *Entwurf* field equations and hence of the extent to which the theory satisfies the generalized relativity principle. The *Entwurf* operator had *not* been obtained from a generally-covariant object along the mathematical strategy, (cf. eq. (XLIX)). By construction, the *Entwurf* operator is covariant only under linear transformations (cf. eq. (LIII)):

$$\textbf{TRAFO(ENTWURF)} =_{\textbf{DEFT}} \textbf{LINEAR.} \qquad \textbf{(LXVII)}$$

In different stages of Einstein's work during the reign of the *Entwurf* theory, i.e., between the winter of 1912–1913 and the fall of 1915, he took different positions on the question of whether or not the theory admits a wider class of coordinate transformations. These positions ranged from the acceptance that the *Entwurf* theory is covariant only under linear transformations to the belief that it fully complies with the demands of a generalized relativity principle. Einstein at first believed that the issue of the transformation properties of the *Entwurf* equations was wide open and could be settled only by an extensive mathematical investigation. In the summer of 1913, however, he came to the conclusion that a mere inspection of the form of eq. (51) was sufficient to resolve the problem in favor of the claim that the *Entwurf* theory could *only* be covariant under linear transformations.[79] He thus accepted that the conservation principle requires a severe limitation of the generalized relativity principle.

5.4 The Entwurf Operator and the Correspondence Principle in the Fall of 1915

In the course of his elaboration of the *Entwurf* theory, Einstein succeeded in deriving the field equations from a Lagrange formalism with the default setting for the field given by eq. (XXII). The field then enters the Lagrangian in the form of eq. (LXIII). After an initial attempt to select this default setting for the field with the help of a consistency argument involving the conservation principle (see below), he returned to the correspondence principle as the main argument for choosing, among several options to specify the field variable, the default setting eq. (XXII), giving rise to the familiar *Entwurf* field equation.

5.5 The Entwurf Operator and the Conservation Principle in the Fall of 1915

In the course of his elaboration of the *Entwurf* theory, Einstein succeeded in bringing its field equation into the canonical form described by eqs. (XLIV), (XLV) with the condition (XLIII), a form that was expected on the basis of classical field theory:

---

79 Cf. Einstein to H.A. Lorentz, 16 August 1913, (CPAE 5, Doc. 470, Norton 1984, 126).

$$\textbf{ENTWURF} =_{\textbf{DEFT}} \frac{\partial}{\partial x^\alpha}(\sqrt{-g}g^{\alpha\beta}\tilde{\Gamma}^\lambda_{\mu\beta}) = \kappa(T^\lambda_\mu + t^\lambda_\mu), \qquad (52)$$

with

$$\textbf{FIELDMASS} =_{\textbf{DEFT}} \kappa t^\lambda_\mu = \sqrt{-g}\Big(g^{\lambda\rho}\tilde{\Gamma}^\alpha_{\tau\mu}\tilde{\Gamma}^\tau_{\alpha\rho} - \frac{1}{2}\delta^\lambda_\mu g^{\rho\tau}\tilde{\Gamma}^\alpha_{\beta\rho}\tilde{\Gamma}^\beta_{\alpha\tau}\Big). \qquad (53)$$

Here $\tilde{\Gamma}^\alpha_{\sigma\nu}$ represents the default setting for the field as given by eq. (XXII). In 1914 Einstein erroneously believed that a compatibility requirement resulting from the conservation principle and the generalized relativity principle (cf. eq. (87) below) would uniquely fix the *Entwurf* Lagrangian. However, this requirement merely corresponds to demanding the compatibility between **FIELDMASS** in the sense of eq. (XLV) and **FIELDMASS** in the sense of eq. (XLII) and does not substantially restrict the choice of possible gravitation tensors.[80]

5.6 The Entwurf Operator and the Generalized Relativity Principle in the Fall of 1915

Einstein quickly discovered that his argument based on the form of eq. (51) was fallacious since the energy-momentum expression of the gravitational field does not represent a generally-covariant tensor.[81] But he soon found another, seemingly powerful argument in order to justify the *Entwurf* theory's lack of general covariance, the so-called hole argument.[82] To identify the covariance group of the *Entwurf* field equations compatible with this argument, Einstein again made use of eq. (51) but now in a different way which corresponds to the conservation compatibility check as represented by eq. (XLIII), i.e. he combined energy-momentum conservation with the gravitational field equation in order to derive a condition for the class of admissible coordinate systems. By exploring the transformation properties of the Lagrangian (cf. eq. (LXIV), Einstein and Grossmann (1914) claimed to have shown that this condition is both necessary and sufficient (cf. eqs. (LV) and (XLIV)):

$$\textbf{TRAFO}(\textbf{NORM}_{\textbf{CLASS}}) \Leftrightarrow \textbf{TRAFO}(\textbf{DIV}(\textbf{NORM}_{\textbf{CLASS}})) \qquad \textbf{(LXVIII)}$$

with

$$\textbf{DIV}(\textbf{NORM}_{\textbf{CLASS}}) =_{\textbf{DEFT}} B_\mu = \frac{\partial^2}{\partial x^\nu \partial x^\alpha}(\sqrt{-g}g^{\alpha\beta}\tilde{\Gamma}^\nu_{\sigma\beta}) = 0. \qquad (54)$$

These four third-order differential equations for the metric tensor complement the ten gravitational field equations and embody the conditions enforcing the restriction of general covariance characteristic of the theory. They were understood by Einstein and

80 See "Untying the Knot ..." (in vol. 2 of this series), sec. 3.

81 Cf. the footnote in (Einstein and Grossmann 1914, 218).

82 For historical discussion, see (Earman and Norton 1987, Stachel 1989b) and "What Did Einstein Know ..." (in vol.2 of this series) as well as further references cited therein.

Grossmann as determining the coordinate systems "adapted" to the *Entwurf* theory.[83] It was difficult to see exactly which transformations to accelerated coordinate systems are admitted by these conditions.

5.7 The Ricci Tensor and the Correspondence Principle in the Winter of 1912–1913

The generally-covariant Ricci tensor, first taken into consideration by Einstein in the Zurich Notebook (see p. 22R), can be expressed in terms of the Christoffel symbols (cf. eqs. (30) and (XLVI)) as:

$$\textbf{RIEM} =_{\textbf{DEFT}} \textbf{RICCI} =_{\textbf{DEFT}} R_{\mu\nu} = \frac{\partial\Gamma^{\alpha}_{\mu\nu}}{\partial x^{\alpha}} - \frac{\partial\Gamma^{\alpha}_{\mu\alpha}}{\partial x^{\nu}} + \Gamma^{\beta}_{\nu\beta}\Gamma^{\alpha}_{\mu\alpha} - \Gamma^{\beta}_{\mu\nu}\Gamma^{\alpha}_{\beta\alpha}. \qquad (55)$$

The validity of the correspondence principle could be examined by bringing **RICCI** into the form (cf. eqs. (XXXIV) and (XLIX))

$$\textbf{RIEM} =_{\textbf{DEFT}} \textbf{RICCI}_{\textbf{RED}} = \textbf{LAP(POT)} + \textbf{CORR(POT)}, \qquad \textbf{(LXIX)}$$

and by checking whether (cf. eq. (L))

$$\textbf{DIST(POT)} = \textbf{0}.$$

In the Zurich Notebook Einstein identified the relevant terms as:

$$\textbf{DIST(POT)} =_{\textbf{DEFT}} \frac{\partial^2 g_{\alpha\alpha}}{\partial x_{\mu}\partial x_{\nu}} - \frac{\partial^2 g_{\mu\alpha}}{\partial x_{\alpha}\partial x_{\nu}} - \frac{\partial^2 g_{\nu\alpha}}{\partial x_{\alpha}\partial x_{\mu}}. \qquad (56)$$

The vanishing of these disturbing terms can be achieved by imposing a set of four *first-order* partial differential equations for the metric tensor which is given by (cf. eq. (LII)):

$$\textbf{COORD(POT)} =_{\textbf{DEFT}} \textbf{COORD}_{\textbf{HARM}}\textbf{(POT)} =_{\textbf{DEFT}} g^{\lambda\kappa}\Gamma^{\mu}_{\lambda\kappa} = 0. \qquad (57)$$

---

83 This expression is chosen to resolve the ambiguity of German expressions which may be translated by "condition" as well as by "restriction." Cf. the formulations in (Einstein and Grossmann 1914): "understood [...], that an acceptable theory of gravitation implies necessarily a specialization of the coordinate system." ("eingesehen [...], daß eine brauchbare Gravitationstheorie notwendig einer Spezialisierung des Koordinatensystems bedarf [...]", p. 218); "restriction" ("Einschränkung", p. 218, note); "true condition" ("wirkliche Bedingung", p. 219); "conditions [...], by which we restricted the coordinate systems" ("Bedingungen [...], durch die wir die Koordinatensysteme eingeschränkt haben.", p. 225). In a letter to Michele Besso, ca. 10 March 1914, (CPAE 5, Doc. 514), Einstein comments on eq. (54): "These are 4 third-order equations for the [...] or [...], which can be conceived as the conditions for the special choice of the reference system." ("Dies sind vier Gleichungen dritter Ordnung für die $g_{\mu\nu}$ [...], welche man als die Bedingungen für die spezielle Wahl des Bezugssystems auffassen kann.")

These equations, representing the "harmonic" coordinate restriction, were interpreted by Einstein as singling out a particular class of coordinate systems that were then called "isothermal" and are now referred to as "harmonic" coordinates.

The reduced Ricci tensor **RICCI$_{\mathbf{RED}}$** suffered from yet another problem related to Einstein's understanding of the correspondence principle. It consists in a conflict between the harmonic coordinate restriction and the canonical metric for a static gravitational field **POT$_{\mathbf{STAT}}$** (cf. eq. (25)):

$$\mathbf{COORD_{HARM}(POT_{STAT}) \neq 0.} \qquad \mathbf{(LXX)}$$

However, as far as the available evidence from the Zurich Notebook and other contemporary sources show, this argument played no role in evaluating the reduced Ricci tensor.[84]

5.8 The Ricci Tensor and the Conservation Principle in the Winter of 1912–1913

As far as the conservation principle is concerned, the exploration depth reached in the Zurich Notebook was characterized by the fact that Einstein examined only the weak-field equation following from a gravitational field equation based on the Ricci tensor. He considered, in other words, an equation of the type of eq. (33). For such a weak-field equation in which the source is given by pressureless dust (cf. eq. (4)), Einstein succeeded in representing the force exerted by the gravitational field as a divergence expression in the sense of eq. (XXXVI)):

$$\mathbf{-\,GRAD(POT)\ X\ LIM(RICCI) = DIV(LIM(FIELDMASS)),} \qquad \mathbf{(LXXI)}$$

which in his notation reads:[85]

$$\sum_{\kappa i m}\gamma_{\kappa\kappa}\frac{\partial^2 g_{im}}{\partial x_\kappa^2}\frac{\partial g_{im}}{\partial x_\sigma} = \sum_{\kappa i m}\gamma_{\kappa\kappa}\left[\frac{\partial}{\partial x_\kappa}\left(\frac{\partial g_{im}}{\partial x_\kappa}\frac{\partial g_{im}}{\partial x_\sigma}\right) - \frac{1}{2}\frac{\partial}{\partial x_\sigma}\left(\frac{\partial^2 g_{im}{}^2}{\partial x_\kappa}\right)\right]. \qquad (58)$$

The conservation compatibility check similarly takes on a simpler form if considered for the weak field case. In first-order approximation the covariant derivative in eq. (XXXI) can be replaced by an ordinary derivative and **OP(POT)** by **LAP(POT)** with its default setting according to eq. (XXVIII) so that this condition can be written, in Einstein's notation, as:

$$\mathbf{LIM(DIV_{COV}(OP)) = DIV(LIM(LAP)= 0.} \qquad \mathbf{(LXXII)}$$

Interchanging the two differential operations,

---

84 See "Untying the Knot ..." (in vol. 2 of this series), fn. 12 for further discussion.

85 See p. 19R of "Einstein's Zurich Notebook" and sec. 5.4.2 of the "Commentary" (in vol. 2 of this series), fn. 10 for further discussion. Einstein's notation, which is somewhat sloppy, is explained in detail in the commentary; note that he used an imaginary time coordinate and that the terms $g_{\mu\nu}$ here stand for the small deviations $h_{\mu\nu}$ from the covariant Minkowski metric.

DIV(LIM(LAP(POT)) = LIM(LAP(DIV(POT))), **(LXXIII)**

which, in Einstein's notation amounts to:

$$\frac{\partial}{\partial x_m}(\Box g_{im}) = \Box\left(\frac{\partial g_{im}}{\partial x_m}\right) = 0, \tag{59}$$

it becomes clear that the conservation compatibility check is satisfied at the weak-field level if an appropriate set of first-order conditions hold in the sense of eq. (LVI):

LIM(ENERG) =$_{\mathbf{DEFT}}$ DIV(POT) = 0 => DIV(LIM(LAP)) = 0. **(LXXIV)**

More specifically, the conservation compatibility check works out in the weak field limit if the condition:

$$\textbf{LIM(ENERG)} =_{\textbf{DEFT}} \textbf{COORD}_{\textbf{HERTZ}} = \textbf{DIV(POT)} =_{\textbf{DEFT}} \frac{\partial g_{im}}{\partial x_m} = 0 \tag{60}$$

is fulfilled. This condition was mentioned by Einstein in a letter to Paul Hertz from 22 August 1915[86] and will therefore be called the "Hertz condition" or the "Hertz restriction" depending on the context. In the case at hand, it is appropriately referred to as the "Hertz restriction" since it represents a restriction of the admissible coordinates required by the conservation principle.

As it turned out, the combination of the two coordinate restrictions eq. (57) and eq. (60), resulting from the correspondence and the conservation principle, and the weak-field field equation imposed a restriction which Einstein considered to be unacceptable. According to this condition, the trace of the source term has to vanish, which can be expressed in terms of eq. (LX) as:

TRACE(SOURCE) = 0 => CC-COMP(LIM(LAP)). **(LXXV)**

This condition was indeed incompatible with the default setting for the source term of the gravitational field equation, pressureless dust (cf. eq. (XXI)). Combining restrictions eq. (57) and eq. (60) furthermore implies that the trace of the potential must be constant which is obviously in conflict with the default setting for the metric of a static field eq. (25).[87]

5.9 The Ricci Tensor and the Generalized Relativity Principle in the Winter of 1912–1913

Given the compatibility problem just described, the transformation properties of the reduced Ricci tensor remained unexplored.

---

86 For a detailed discussion of this letter, see (Howard and Norton 1993).

87 See "Commentary ..." (in vol. 2 of this series), sec. 5.4.3 for detailed discussion.

5.10 The Ricci Tensor and the Correspondence Principle in the Fall of 1915

When Einstein returned to the Ricci tensor in November 1915 both the perspective and the exploration depth of his investigation had changed. He then considered the Ricci tensor in coordinate systems with:

$$\sqrt{-g} = 1 \tag{61}$$

in which it takes on the simpler form:

$$\mathbf{RICCI} =_{\mathbf{DEFT}} \frac{\partial \Gamma^{\alpha}_{\mu\nu}}{\partial x^{\alpha}} + \Gamma^{\beta}_{\mu\alpha}\Gamma^{\alpha}_{\nu\beta}. \tag{62}$$

Within this framework, the Ricci tensor could be brought into the appropriate weak-field form eq. (XXXIV) by assuming the Hertz condition:

$$\mathbf{COORD_{HERTZ}} = \mathbf{DIV(POT)} =_{\mathbf{DEFT}} \frac{\partial g_{im}}{\partial x_m} = 0. \tag{63}$$

In November 1915, Einstein was aware of the fact that it was sufficient for satisfying the correspondence principle to use such an equation (cf. eq. (LII)) in the modern sense of a coordinate *condition* that simply makes use of the freedom within a generally-covariant framework to pick appropriate coordinate frames—without imposing an overall restriction. In this sense, Einstein's understanding of the correspondence principle had been substantially enhanced by a greater exploration depth of his formalism.[88]

In contrast to eq. (LXX) we now have:

$$\mathbf{COORD_{HERTZ}(POT_{STAT}) = 0,} \tag{LXXVI}$$

so that the conflict between the coordinate condition and the canonical metric for a static field is apparently removed. This is, in any case, what Einstein at first must have believed when he published, in November 1915, a gravitational field equation based on the Ricci tensor. What he seems to have overlooked, however, was the fact that his canonical metric given by eq. (25) was incompatible with the condition eq. (61) on which his entire framework, including the coordinate condition eq. (63), crucially depended. In other words, the available evidence suggests that Einstein had first published his theory based on the Ricci tensor although it actually violates the correspondence principle as he then conceived it.

He only realized the challenge represented by the choice of the Ricci tensor for his understanding of the correspondence principle when he examined the implication of this choice for the explanation of Mercury's perihelion motion, an examination that gave him a nearly perfect match with the observational data.[89] Einstein at first

88 See "Untying the Knot ..." (in vol. 2 of this series), secs. 1.5 and 6 for detailed discussion.

interpreted this agreement as evidence in favor of his hypothesis of an electromagnetic theory of matter which had made the proposal of a field equation with the Ricci tensor as its left-hand side acceptable to him (see below).

5.11 The Ricci Tensor and the Conservation Principle in the Fall of 1915

By the fall of 1915, the exploration depth of Einstein's investigation had been increased, in particular, by the development of a technique allowing him to derive a gravitational energy-momentum expression **FIELDMASS** for the full field equation from the Lagrange formalism, if the coordinate condition eq. (61) is assumed and the default setting for the field is given by eq. (XXIII). He was thus able to bring a field equation based on the Ricci tensor into a form corresponding to eq. (XXXIX) with the conservation equation (XLII):

$$\textbf{NORM(POT)} =_{\textbf{DEFT}} (g^{\nu\lambda}\Gamma^{\alpha}_{\mu\nu})_{,\alpha} - \frac{1}{2}\delta^{\lambda}_{\mu}g^{\rho\sigma}\Gamma^{\alpha}_{\beta\rho}\Gamma^{\beta}_{\alpha\sigma} = -\kappa(T^{\lambda}_{\mu} + t^{\lambda}_{\mu}), \quad (64)$$

with

$$\textbf{FIELDMASS} =_{\textbf{DEFT}} \kappa t^{\lambda}_{\sigma} = \frac{1}{2}\delta^{\lambda}_{\sigma}g^{\mu\nu}\Gamma^{\alpha}_{\beta\mu}\Gamma^{\beta}_{\alpha\nu} - g^{\mu\nu}\Gamma^{\alpha}_{\mu\sigma}(\Gamma_{\alpha\nu},)^{\lambda} \quad (65)$$

and

$$\textbf{DIV(ENEMO + FIELDMASS)} =_{\textbf{DEFT}} (T^{\lambda}_{\mu} + t^{\lambda}_{\mu})_{,\lambda} = 0. \quad (66)$$

Einstein, however, did not manage to comply with the requirement expressed by the default setting eq. (XLIV). Bringing the left-hand side of the field equation into the form of eq. (XLIV) would result in a formulation in which the right-hand side no longer satisfies the default setting eq. (XLI) for **SOURCE**:

$$\textbf{NORM(POT)}_{\textbf{CLASS}} = \textbf{DIV(FIELD)} =_{\textbf{DEFT}} (g^{\nu\lambda}\Gamma^{\alpha}_{\mu\nu})_{,\alpha} = -\kappa\left(T^{\lambda}_{\mu} + t^{\lambda}_{\mu} - \frac{1}{2}\delta^{\lambda}_{\mu}t\right) \quad (67)$$

with

$$\textbf{SOURCE} =_{\textbf{DEFT}} \textbf{ENEMO + FIELDMASS} \neq_{\textbf{DEFT}} -\kappa\left(T^{\lambda}_{\mu} + t^{\lambda}_{\mu} - \frac{1}{2}\delta^{\lambda}_{\mu}t\right). \quad (68)$$

Equation (66) could be used to perform the conservation compatibility check in a straightforward manner. Einstein succeeded in showing that this check turned out successful if the trace of the energy-momentum tensor vanishes (cf. eq. (LXXV))—without imposing any further conditions on the admissible coordinate systems:

TRACE(SOURCE) = 0 => DIV(ENEMO + FIELDMASS) = 0. (LXXVII)

The odd assumption of a vanishing trace, violating the default assumption eq. (XXI), was now acceptable to Einstein since both the exploration depth of his investi-

89 See (Einstein 1915b) and for historical discussion, (Earman and Janssen 1993).

gation and his perspective had changed. He now reexamined the Ricci tensor from the perspective of an electromagnetic theory of matter in which this condition was fulfilled from the outset, given that the trace of the electromagnetic energy-momentum tensor vanishes.[90]

5.12 The Ricci Tensor and the Generalized Relativity Principle in the Fall of 1915

The field equations based on the Ricci tensor as formulated by Einstein in the fall of 1915 represents, according to his heuristic criteria, a complete implementation of the generalized principle of relativity. The conservation compatibility check for these field equations had given Einstein, as we have seen, merely a condition on the trace of the energy-momentum tensor which does not imply any restriction on the choice of coordinate systems. As a consequence, there no longer was any conflict between the conservation and the correspondence principles as he had encountered it in the winter of 1912–1913. It was thus possible to impose either the harmonic coordinate condition eq. (57) or the combination of eqs. (61) and (63) in order to reduce the Ricci tensor to the canonical weak field form eq. (LXIX) from which the Newtonian limit could be obtained—at least if the objection resulting from eq. (LXX) could be solved or circumvented.

5.13 The Einstein Tensor and the Correspondence Principle in the Winter of 1912–1913

The generally-covariant Einstein tensor, first taken into consideration, albeit only in the weak-field approximation, in the Zurich Notebook, can be expressed in terms of the Ricci tensor $R_{\mu\nu}$ and its trace R (cf. eq. (55)) as:

$$\textbf{RIEM} =_{\textbf{DEFT}} \textbf{Einstein} =_{\textbf{DEFT}} E_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R. \tag{69}$$

A field equation based on the Einstein tensor may also be written by shifting the trace term to the right-hand side by a simple mathematical argument. The equation then reads

$$R_{\mu\nu} = -\kappa\left(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T\right), \tag{70}$$

where T is the trace of the energy-momentum tensor $T_{\mu\nu}$. Here the Ricci tensor again appears on the left-hand side as the differential operator acting on the metric tensor.

---

90 See (Einstein 1915d) where the consequence is called “introducing an admittedly bold additional hypothesis on the structure of matter.” (“Einführung einer allerdings kühnen zusätzlichen Hypothese über die Struktur der Materie”, p. 799).

The exploration level of the first examination of the Einstein tensor in the winter of 1912–1913 was, just as that of Einstein's analysis of the Ricci tensor, characterized by a focus on the weak-field equations and the assumption that the correspondence principle could only be satisfied by a coordinate restriction. Given that the Einstein tensor results from a modification of the Ricci tensor according to eq. (69), it was natural to presuppose the harmonic coordinate restriction $\mathbf{COORD_{HARM}(POT) = 0}$ (cf. eq. (57)). As a matter of fact, in the winter of 1912–1913 the Einstein tensor was obtained directly by an ad hoc modification of the weak-field form of the gravitational field equation eq. (33), resulting in:[91]

$$\Box\left(g_{i\kappa} - \frac{1}{2}\delta_{i\kappa}U\right) = T_{i\kappa} \tag{71}$$

with the trace term:

$$U = \sum g_{\kappa\kappa}, \tag{72}$$

or alternatively as:

$$\Box g_{ij} = T_{ij} - \frac{1}{2}\delta_{ij}(\sum T_{\kappa\kappa}). \tag{73}$$

It surely would have been possible for Einstein to carry out the corresponding modification on the level of the original Ricci tensor, turning it into what we now call the Einstein tensor, by the subtraction of a trace term.

On closer inspection, however, the harmonic coordinate restriction does not achieve the desired reduction of the field equation to the required standard form in the sense of eq. (LII). Indeed, if the left-hand side is brought into the canonical form eq. (XXVIII) so that eq. (73) is obtained, the right-hand side does obviously not represent the default setting for **SOURCE** as given by eq. (XIV):

$$\mathbf{SOURCE} \neq_{\mathbf{DEFT}} T_{ij} - \frac{1}{2}\delta_{ij}(\sum T_{\kappa\kappa}). \tag{LXXVIII}$$

Instead an additional trace term appears which in general is not constant. If one examines, in particular, a static mass distribution as the default setting for **SOURCE** so that the 44 component is the only non-vanishing term of the energy-momentum tensor, it follows from the weak-field equation (73) that *all* diagonal components of the metric tensor will be variable so that one has in general:

$$g_{ii} \neq \text{const for } i = 1, \ldots, 4. \tag{74}$$

As a consequence, the weak-field equation (73) no longer admits the canonical metric $\mathbf{POT_{STAT}}$ defined by eq. (25) as a solution. At that point in time, Einstein saw no way to avoid this default setting for the potential, and he rejected the Einstein tensor—lin-

91 See "Commentary" (in vol. 2 of this series), sec. 5.4.3.

earized and reduced by the harmonic coordinate restriction—as a candidate for the left-hand side of the gravitational field equation.[92]

5.14 The Einstein Tensor and the Conservation Principle in the Winter of 1912–1913

At the weak-field level, the results of Einstein's check of the conservation principle turned out to be promising. In spite of the additional trace term it was possible to write the gravitational force density in the required form of a divergence of the gravitational energy-momentum density, which in Einstein's notation reads (cf. eqs. (XXXVI) and (LXXI)):

$$\begin{aligned}&\sum_{\iota\kappa\nu}\frac{\partial}{\partial x_\nu}\left(\frac{\partial g_{\iota\kappa}}{\partial x_\nu}\frac{\partial g_{\iota\kappa}}{\partial x_\sigma}\right)-\frac{1}{2}\sum_{\iota\kappa\nu}\frac{\partial}{\partial x_\sigma}\left(\left(\frac{\partial g_{\iota\kappa}}{\partial x_\nu}\right)^2\right)\\&-\frac{1}{2}\sum_{\nu}\frac{\partial}{\partial x_\nu}\left(\frac{\partial U}{\partial x_\nu}\frac{\partial U}{\partial x_\sigma}\right)+\frac{1}{4}\sum\frac{\partial}{\partial x_\sigma}\left(\left(\frac{\partial U}{\partial x_\nu}\right)^2\right).\end{aligned} \tag{75}$$

At the level of the weak-field equation it was also immediately clear that the conservation compatibility check is no longer in conflict with the correspondence principle, in contrast to what he had found before for the Ricci tensor. In analogy with eq. (59) one now obtains:

$$\frac{\partial}{\partial x_\kappa}\left(\Box\left(g_{\iota\kappa}-\frac{1}{2}\delta_{\iota\kappa}U\right)\right) = \Box\left(\frac{\partial}{\partial x_\kappa}\left(g_{\iota\kappa}-\frac{1}{2}\delta_{\iota\kappa}U\right)\right) = 0, \tag{76}$$

which, in symbolic notation, corresponds to (cf. eqs. (LXXII), (LXXIV)):

DIV(LIM(EINSTEIN)) = LIM(LAP(COORD$_{\mathbf{HARM}}$)) = 0. **(LXXIX)**

In other words, the conservation compatibility check is, in the weak-field limit, satisfied *because* of the harmonic coordinate restriction eq. (57) required by the correspondence principle—without imposing any restriction on the trace of the energy-momentum tensor.

5.15 The Einstein Tensor and the Generalized Relativity Principle in the Winter of 1912–1913

Whether or not the generalized relativity principle was satisfied would, according to Einstein's understanding in the winter of 1912–1913, depend on whether the coordinate restrictions necessary to fulfill his other heuristic criteria would leave him enough covariance. In view of the clash between the Einstein tensor and the corre-

92 See "Commentary" (in vol. 2 of this series), sec. 5.4.6.

spondence principle (see eq. (LXXVIII)), this issue remained unexplored at this point in time.

5.16 The Einstein Tensor and the Correspondence Principle in the Fall of 1915

When Einstein returned to the Einstein tensor in November 1915 he focused on coordinate systems with:

$$\sqrt{-g} = 1. \tag{77}$$

The field equations based on the Einstein tensor then take on the form (cf. eqs (70) and (62)):

$$\Gamma^{\alpha}_{\mu\nu,\alpha} + \Gamma^{\alpha}_{\beta\mu}\Gamma^{\beta}_{\alpha\nu} = -\kappa\left(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T\right). \tag{78}$$

In view of eqs. (69) and (70), one could use, just as in the case of the Ricci tensor, the harmonic condition eq. (57)[93] to bring the left-hand side of the field equation into the required form. This condition was now understood as a coordinate condition in the modern sense. Proceeding in this way, one reduces the Einstein field equation for weak fields to an equation of the form (73). The usual transition to the Newtonian theory could now proceed by taking the energy-momentum tensor of dustlike matter as the source and neglecting all terms except the T_{44}-term, which can be identified with the gravitating mass density ρ appearing in the classical Poisson equation (cf. eq. (37)).

What remained to be shown was that the canonical metric for a static field was compatible with the non-standard form of the right-hand side of the weak-field equations (73). Even in 1915 this conflict remained, in a sense, unresolved. The additional trace term on the right-hand side made it impossible to accept the canonical metric for static fields as a solution of the weak field equations since the 11 ... 33 components of the source term had to be retained in the transition to the Newtonian case (cf. eq. (74)). Therefore the correction term in the Einstein tensor made the transition to the Newtonian case *a fortiori* impossible following the procedure suggested by the correspondence principle. All this had been known to Einstein in 1912[94] and remained, of course, true also in 1915, when he took up the Einstein tensor a second time (Einstein 1915d).

But now Einstein was able to circumvent this problem. Even though the field equation failed to satisfy the correspondence principle as hitherto understood, this did not affect the equation of motion. In the weak-field limit of the equation of motion, the non-standard character of the weak-field Einstein equation plays no role. For weak static gravitational fields and for velocities negligible in comparison with that of light, the general equation of motion (31) reduces, as we have seen, to eq. (38).

---

93 For a comment on the role of the Hertz condition in this context, see Albert Einstein to Karl Schwarzschild, Berlin, 19 February 1916, (CPAE 8, Doc. 194).

94 Cf. pp. 20L–21R of “Einstein’s Zurich Notebook” (in this volume).

This equation now implies that, under the conditions assumed, a gravitational field equation based on the Einstein tensor is actually compatible with the experimental data on gravitation that are adequately described by Newton's theory if $g_{44}/2$ is, as usual, identified with the Newtonian potential, while the other components of the metric tensor play no role at this level of the weak-field limit of the Einstein equation.

5.17 The Einstein Tensor and the Conservation Principle in the Fall of 1915

When Einstein returned to the Einstein tensor in late 1915, the greater exploration depth of his investigation made it possible to establish an energy-momentum expression for the gravitational field of the required form. Also the question of the conservation compatibility check could now be addressed in a straightforward manner. He succeeded in bringing the field equation into a form corresponding to eq. (XXXIX) with a conservation equation of the form of eq. (XLII):

$$\textbf{NORM(POT)} =_{\textbf{DEFT}} (g^{\nu\lambda}\Gamma^{\alpha}_{\mu\nu})_{,\alpha} - \frac{1}{2}\delta^{\lambda}_{\mu}\kappa(t+T) = -\kappa(T^{\lambda}_{\mu} + t^{\lambda}_{\mu}). \tag{79}$$

The satisfaction of the conservation compatibility check (cf. eq. (XLIII)) now no longer imposes any additional conditions interfering with the field equation as was the case for the tensor where this check implied that the trace of both sides of the field equation has to vanish.

In the field equation based on the Einstein tensor, the trace terms of the energy-momentum of matter and of the gravitational field enter, in contrast to what happens for the Ricci tensor (cf. eq. (67)), in complete parallel to each other. As a matter of fact, the introduction of these trace terms corresponds to changing the default setting eq. (XLI) for **SOURCE** into:

SOURCE =$_{\textbf{DEFT}}$

(ENEMO – 1/2 TRACE(ENEMO)) +

(FIELDMASS – 1/2 TRACE(FIELDMASS)). (LXXX)

With this new instantiation for the source term, Einstein now also managed to comply with the expectation for the left-hand side of the field equation expressed by the default setting eq. (XLIV):

$$\textbf{NORM(POT)}_{\textbf{CLASS}} = \textbf{DIV(FIELD)} =_{\textbf{DEFT}} (g^{\nu\lambda}\Gamma^{\alpha}_{\mu\nu})_{,\alpha} \tag{80}$$

with

$$\textbf{SOURCE} =_{\textbf{DEFT}} -\kappa\left(\left(T^{\lambda}_{\mu} - \frac{1}{2}\delta^{\lambda}_{\mu}T\right) + \left(t^{\lambda}_{\mu} - \frac{1}{2}\delta^{\lambda}_{\mu}t\right)\right). \tag{81}$$

Note, however, that with this redefinition of the source-term the field equation no longer corresponds with the canonical expectation for its right-hand side expressed by the default setting eq. (XLII) suggested by the conservation principle. Even for a field equation based on the Einstein tensor it is simply impossible to satisfy all expectations raised by the experience of classical field theory!

5.18 The Einstein Tensor and the Generalized Relativity Principle in the Fall of 1915

Since neither the correspondence nor the conservation principle imposed any further restrictions, field equations based on the Einstein tensor fully implement the generalized relativity principle.

5.19 The November Tensor and the Correspondence Principle in the Winter of 1912–1913

The November tensor, first considered in the Zurich Notebook, can be obtained from the Ricci tensor (cf. eqs. (55) and (62)) by restricting the covariance group to unimodular transformations and then splitting off a term:

$$\mathbf{RIEM} =_{\mathbf{DEFT}} \mathbf{NOVEMBER} =_{\mathbf{DEFT}} N_{\mu\nu} = \frac{\partial \Gamma^{\alpha}_{\mu\nu}}{\partial x^{\alpha}} + \Gamma^{\beta}_{\mu\alpha}\Gamma^{\alpha}_{\nu\beta}. \tag{82}$$

The exploration level of the November tensor in the winter of 1912–1913 was, in general, characterized by a limitation to the weak-field equation and the expectation that the implementation of the correspondence and conservation principles requires a coordinate restriction. The correspondence principle, in particular, can be satisfied if the Hertz restriction eq. (60) is imposed, bringing **NOVEMBER** into the form (cf. eqs. (XXXIV) and (XLIX)):

$$\mathbf{RIEM} =_{\mathbf{DEFT}} \mathbf{NOVEMBER_{RED}} = \mathbf{LAP(POT)} + \mathbf{CORR(POT)}, \qquad \mathbf{(LXXXI)}$$

so that also the weak-field equation takes on the canonical form of eq. (33) which can be solved by the canonical metric for a static field given by eq. (25). One now also has (cf. eqs. (LXX) and (LXXVI)):

$$\mathbf{COORD_{HERTZ}(POT_{STAT}) = 0.} \qquad \mathbf{(LXXXII)}$$

5.20 The November Tensor and the Conservation Principle in the Winter of 1912–1913

The weak-field equation for the November tensor has the same form as that obtained from a field equation based on the Ricci tensor since (cf. eq. (XXVIII)):

$$\mathbf{LIM(RICCI) = LIM(NOVEMBER) = LIM(LAP).} \qquad \mathbf{(LXXXIII)}$$

It is clear therefore that the conservation principle holds, at least in the weak-field limit. It is possible to form a divergence expression such as that given by eq. (LXXI) and to satisfy the conservation compatibility check as represented by eq. (LXXII) if the Hertz restriction eq. (60) is imposed. But contrary to the case of the Ricci tensor, the coordinate restrictions required by the correspondence and the conservation principles, respectively, now coincide (cf. eq. (LX)):

$$\mathbf{LIM(ENERG) = COORD_{HERTZ} = 0.} \qquad \textbf{(LXXXIV)}$$

5.21 The November Tensor and the Generalized Relativity Principle in the Winter of 1912–1913

The check of the generalized relativity principle was eased by the fact that the transformation behavior of the reduced November tensor (cf. eq. (XXXIV)) could be inferred from the transformation properties of the restriction distinguishing it from the original November tensor, the Hertz restriction. Indeed, if the Hertz restriction remains covariant under a given unimodular coordinate transformation so must the reduced November tensor (cf. eq. (LIX)). In the winter of 1912–1913, Einstein examined this transformation behavior for the two cases central to the heuristics governed by the equivalence principle, the case of uniform acceleration ("the elevator") and the case of rotation ("the bucket"). To simplify matters, he considered the case of infinitesimal transformations and found that, while the Hertz restriction is satisfied by infinitesimal rotations, it is not by infinitesimal transformations to a uniformly accelerated system.[95] At least as far as the exploration level of his calculations (limited to the weak-field case) allowed, Einstein could conclude that the reduced November tensor clashes with the equivalence principle, even in the case of infinitesimal transformations. He may well have found that transformations to finite rotations are incompatible with the Hertz restriction as well.

5.22 The November Tensor and the Correspondence Principle in the Fall of 1915

When Einstein returned to the November tensor in 1915, he could make use of the results he had established earlier, in particular with regard to the correspondence principle and how to satisfy that principle by imposing the Hertz restriction (cf. eq. (LXXXI)). His reexamination was, on the other hand, characterized by an increased exploration depth, which allowed him to treat this restriction as a coordinate condition in the modern sense. As we shall see, the conservation principle again leads to a coordinate restriction following from **DIV(NORM) = 0** (cf. eq. (XLIII)) which made it necessary to recheck the compatibility of this condition with the correspondence principle. As it turned out, the conservation compatibility check only gives rise to a weak scalar condition in this case, which in the weak-field limit, can easily be satisfied if the Hertz condition is fulfilled (cf. eq. (LVI)):

---

95 See "Commentary" (in vol. 2 of this series), sec. 5.5.3 and secs. 4.5.2–4.5.3.

$$\mathbf{LIM(ENERG(POT))} =_{\mathbf{DEFT}} \mathbf{COORD}_{\mathbf{HERTZ}} = 0 \Rightarrow$$

$$\mathbf{LIM(DIV(NORM))} = 0. \qquad \mathbf{(LXXXV)}$$

Under these circumstances, the Hertz condition can thus be considered as a strengthening of the restriction **DIV(NORM) = 0** following from the conservation principle. But as this sharpening turned out to be necessary only for the purpose of implementing the correspondence principle by choosing a class of coordinate systems suitable for this purpose, the Hertz condition could now indeed be interpreted, for the first time, as a *coordinate condition* in the modern sense.

5.23 The November Tensor and the Conservation Principle in the Fall of 1915

In the fall of 1915, Einstein succeeded in deriving a gravitational energy-momentum expression **FIELDMASS** for the full field equation based on the November tensor from a Lagrange formalism in which the default setting for the field is given by eq. (XXIII). He brought the field equation into a form corresponding to eq. (XXXIX), thus obtaining eqs. (64), (65), and (66), familiar from our discussion of the Ricci tensor.

What remained was the check of compatibility with the conservation principle and the question of which coordinate transformations it allowed. This question could now be addressed not just on the weak-field level—where the transformation properties of the Hertz restriction had led to a disappointing answer—but on the level of the full field equation. Einstein succeeded in expressing the conservation compatibility check in terms of an equation of the form of eq. (XLIII) which now, however, has the remarkable property that it represents not four equations but rather follows from a single scalar condition (cf. eq. (LVI)):[96]

$$\mathbf{ENERG(POT)} =_{\mathbf{DEFT}} \mathbf{DIV(SCALAR(POT))} = 0 \Rightarrow \mathbf{DIV(NORM)} = 0, \mathbf{(LXXXVI)}$$

where:

$$\mathbf{SCALAR(POT)} =_{\mathbf{DEFT}} \sum_{\alpha\beta} \frac{\partial^2 g^{\alpha\beta}}{\partial x^\alpha \partial x^\beta} - \sum_{\sigma\tau\alpha\beta} g^{\sigma\tau}\Gamma^\alpha_{\sigma\beta}\Gamma^\beta_{\tau\alpha} = 0. \qquad (83)$$

This condition clearly is much less restrictive than the Hertz restriction. As mentioned above, the Hertz restriction could therefore be reinterpreted as a coordinate condition, obtained by strengthening the weak-field version of this scalar condition.

5.24 The November Tensor and the Generalized Relativity Principle in the Fall of 1915

The November tensor was obtained from the generally-covariant Ricci tensor by imposing a restriction to unimodular coordinate transformations. The conservation compatibility check (cf. eq. (83)) gave rise to a further restriction of the choice of

96 See "Untying the Knot ..." (in vol. 2 of this series), sec. 6, eqs. (75)–(78), for detailed discussion.

admissible coordinate systems, the "November restriction," as it might be called. Combining the trace of the full field equation with eq. (83), the following scalar equation results:[97]

$$\frac{\partial}{\partial x^{\alpha}}\left(g^{\alpha\beta}\frac{\partial}{\partial x^{\beta}}\log\sqrt{-g}\right) = -\kappa T. \tag{84}$$

This additional coordinate restriction requires, in particular, that the coordinate system cannot be chosen in such a way that $\sqrt{-g} = 1$ since this would imply the physically implausible consequence that the trace of the energy-momentum tensor vanish.[98]

Since the November restriction was much weaker than the Hertz restriction, it offered a way to overcome the latter's fatal implications for the equivalence principle. In particular, transformations of a given coordinate system to a rotating system or a system whose origin moves in any given way were now allowed so that the generalized principle of relativity is amply, but not fully satisfied. Indeed if a given coordinate system, for instance the usual representation of Minkowski space in Cartesian coordinates, satisfies this coordinate restriction, any other system resulting from the given one by a unimodular transformation must also fulfill this restriction, which is covariant under unimodular transformations.

6. CHANGING HORSES: EINSTEIN'S CHOICE OF GRAVITATIONAL TENSORS FROM 1912–1913

The checklist for candidates for the left-hand side of the field equations that we used in the preceding section was based on the heuristic criteria that Einstein had essentially established by the end of 1912. The decision as to which candidate fares best given these heuristic criteria depends on the state of elaboration of the various mathematical and physical consequences associated with that candidate. The relative arbitrariness of elaborating the consequences of a physical theory along various conceivable pathways, which from the outset can never be overlooked in their totality, therefore entails an element of historical contingency. As the comparison between the *Entwurf* theory, maintained by Einstein essentially for three years, and his final theory of general relativity shows, this contingency may take the form of different physical theories with different empirical consequences, which, at the time, were open to debate.

---

97 See (Einstein 1915a, p.785). Cf. "Untying the Knot ..." (in vol. 2 of this series), sec. 6, eqs. (79)–(82).

98 "In writing the previous paper, I was not yet aware that the hypothesis $\sum T^{\mu}_{\mu} = 0$ is, in principle, admissible." ("Bei Niederschrift der früheren Mitteilung [Einstein 1915a] war mir die prinzipielle Zulässigkeit der Hypothese $\sum T^{\mu}_{\mu} = 0$ noch nicht zu Bewußtsein gekommen.") (Einstein 1915b, 800).

Furthermore, even if the more developed state of elaboration reached by Einstein by the fall of 1915 is taken into account, it is as we have seen in the previous section the November tensor rather than the Einstein tensor which fits Einstein's original heuristic criteria best. The November tensor had passed all tests of Einstein's checklist with only a minor adjustment of the generalized relativity principle while the Einstein tensor had failed the test of the correspondence principle as originally conceived by Einstein. This was all the worse for the Einstein tensor since the generalized relativity principle was an ambitious and idiosyncratic goal which was not shared by many of Einstein's contemporaries, while the correspondence principle had all the support of classical physics and special relativity. That it was the Einstein tensor that in the end won the race can only be understood by taking into account another aspect of the historical process, which we have so far neglected, changes in the heuristic criteria themselves as well as in their relative importance. We therefore need to take a closer look at the actual development of Einstein's thinking.

Why exactly did he turn from one candidate to the other? How did his judgement of candidates evolve? What made him come back eventually to previously discarded candidates after spending almost three years working out a more or less satisfactory relativistic theory of gravitation based on one of them? These questions are the focus of this and the next chapter dealing with what one might describe as Einstein's discovery process or better, as his "investigative pathway."[99] As we have argued, the eventual success of Einstein's research was based on applying shared knowledge resources to the problem of gravitation. The actual mechanism of these applications has so far been considered only from a single perspective, that of assimilating physical and mathematical resources to the basic model of a field equation. In the following, we shall argue that focusing on the exploitation of these resources not only allows us to understand the basic pattern of Einstein's search, the alternation between physical and mathematical strategy. It also allows us to reconstruct, to a surprising extent, the actual course of his search, if we take into account an additional cognitive process as well. While the assimilation of physical and mathematical knowledge to the Lorentz model of a gravitational field equation is basically a top-down process that is guided by the relatively stable high-level cognitive structures at the core of Einstein's heuristic criteria, a reflection on the experiences resulting from such an assimilation, including its failures, could trigger a corresponding bottom-up process of accommodating these high-level structures, including the very mental model itself, to the outcome of these experiences. These two complementary processes were mediated by the external representation of the mental model in terms of mathematical language. The combination of these processes produced conclusions that evolved with the elaboration of the formalism and with the accumulation of Einstein's experience. In order to substantiate this schematic account, we shall, in the following, review his pathways, first in the period documented by the Zurich Notebook and then—in the next chapter—in the period between 1913 and 1915. Relying heavily on the joint

99 See (Renn, Damerow and Rieger 2001; Holmes, Renn and Rheinberger 2003).

work presented in this volume,[100] we shall interpret these pathways as being governed by an interplay between assimilation and accommodation, mediated by the mathematical formalism.

6.1 The Tinkering Phase in the Zurich Notebook

The earliest notes on gravitation in the Zurich Notebook represent a stage of Einstein's search for the field equation in which he had few sophisticated mathematical tools at hand that would allow him to construct candidates fitting the framework provided by the Lorentz model. Even his knowledge of the metric tensor and its properties was still rudimentary. Only gradually did he find ways of exploiting his knowledge of vector analysis for his search. Eventually he familiarized himself with the scalar Beltrami invariants as another instrument that allowed him to investigate the few building blocks at his disposal, that is, the metric as a representation of the gravitational potential, the four-dimensional Minkowski formalism, and his theory of the static gravitational field. In spite of the lack of mathematical sophistication characterizing this early tinkering phase, not to mention the failure to produce promising candidate field equations, it is in this period that Einstein acquired essential insights shaping his research in subsequent phases of work.

These insights consisted, first of all, in a number of concrete results that later turned out to be useful, such as the identification of the core operator (cf. eq. (XXVI)), the establishment of a repertoire of techniques for dealing with coordinate transformations, results on the transformation properties of the Hertz restriction, and, most importantly, the successful implementation of the Lorentz model of an equation of motion in a generally relativistic framework (cf. eq. (XX)). The most far-reaching insights of this period were, however, of a different nature. They consisted in more general ideas resulting from a reflection on the experiences in the tinkering phase, ideas that were largely independent from the concrete mathematical material to which they were applied. Here we encounter a second function of reflection in this context, beyond that of modifying one or the other of Einstein's heuristic principles: reflection could also result in higher-level structures operating on a strategic level, that is, guiding the implementation of these heuristic principles. The most important example is certainly the idea to first impose a coordinate restriction on an object of broad covariance in order to satisfy the correspondence principle and then to explore the transformation properties of this coordinate restriction in order to check the extent to which the generalized principle of relativity is satisfied as well (cf. eq. (LIV)). Even the alternation between more physically and more mathematically motivated approaches emerged as a distinct pattern in this period, again with far-reaching implications for Einstein's subsequent research. The reflection on the experiences of this tinkering phase thus led to what one might describe as a "chunking" of Einstein's

100 See, in particular, "Commentary" for this section and "Untying the Knot ..." for the next section (both in vol. 2 of this series).

heuristic principles in terms of procedures that interconnected them in such a way as to ease their implementation as a whole.[101] Such procedures could involve the subsequent translation of these principles into well-defined mathematical requirements on a gravitational field theory (e.g., the choice of a generally-covariant candidate, followed by the stipulation of a coordinate restriction) or they could consist in alternating between physical and mathematical default settings.

6.2 Assimilating Knowledge about the Static Gravitational Field to a Metric Formalism (39L–39R)

When Einstein began systematically to explore a metric theory of gravitation, he was confronted with the problem that the knowledge resources available to him for constructing such a theory presented themselves as more or less isolated building blocks that could not easily be fitted together. On the page of the Zurich Notebook which documents the point of departure of his exploration (p. 39L), he therefore started his investigation simply by listing three such building blocks, the line element in terms of the metric tensor representing the gravitational potential, the four-dimensional Minkowski formalism, and his theory of the static gravitational field. How could they be brought into relation to each other?

The principal challenge was to assimilate the knowledge about the special case of a scalar, static gravitational potential to a tensorial formalism. If such an assimilation were successful, the mental model of a field equation for a tensorial gravitational potential would acquire a physically meaningful instantiation. Einstein's first consideration of the problem of gravitation that is recorded in the Zurich Notebook is precisely such an attempt to assimilate the static case to a metric formalism, concentrating on two of the slots of the Lorentz model for a field equation, that for the differential operator and that for the gravitational potential. For reasons that we have discussed earlier, the default setting for the latter slot was given by the canonical metric for a static field (cf. eq. (25)). Brought into proper mathematical form, Einstein's scalar field equation for the static gravitational field could therefore be conceived as a second-order partial differential equation for the one variable component of this special metric tensor, expressed in a special coordinate frame in which the metric takes on its canonical form. Exploiting mathematical knowledge about the behavior of a tensorial field equation under coordinate transformations, one should then be able to generalize this equation for one component to a field equation for the full metric tensor.

By transforming the equation for the static field into a more general coordinate system, Einstein made an observation that suggested a new pathway to him. He found that, under linear coordinate transformations, the metric tensor behaves exactly the same way as the second-order partial derivatives of a scalar function. This observation opened up a new possibility for drawing on hitherto unexploited mathematical

101 For the concept of "chunking" in cognitive science, see (Minski 1987).

resources and thus for identifying a suitable differential operator acting on the metric tensor.

6.3 Assimilating Knowledge about Scalar Differential Invariants to a Metric Formalism (40L–41L)

Einstein's key problem was that the default-settings for two of the slots of the mental model for a gravitational field equation, suggested by his earlier experiences with implementations of this model, could not be matched to each other (see Fig. 3, p. 173). While the default setting for the gravitational potential was represented by the canonical metric, the default setting for the differential operator was, at this point, an object like the Laplace operator, applicable only to scalar functions and covariant only under linear transformations. Was there a way of bridging this gap between a scalar differential operator and a tensorial potential? Einstein's insight into the analogy between the transformational properties of the metric tensor and those of the second-order partial derivatives of a scalar function offered such a bridge, allowing him to bring to bear on this problem mathematical knowledge about scalar differential operators. It suggested the possibility of building some higher-order differential operator acting on a scalar function, which could then be translated into a differential operator acting on the metric tensor. All that was needed for such a translation was the replacement of a second-order partial derivative term by the corresponding components of the metric tensor; the remaining partial derivatives could then be considered as a differential operator acting on the metric.

What could be gained by such a roundabout procedure? If the scalar differential operators involved are just linearly covariant, like the Laplace operator, relatively little. If, however, scalar differential invariants are taken as the building blocks of such a construction, it could lead to the formulation of a generally-covariant differential operator for the metric tensor. There is some indication in the Zurich Notebook that this may have been Einstein's hope. In any case, he systematically checked whether various higher-order scalar differential operators would yield, after translation, a suitable candidate for the left-hand side of the gravitational field equations. But apparently he was unable to single out a candidate promising to fulfill his other heuristic criteria as well, and did not pursue this investigation for the time being. As is clear from later pages of the notebook, however, Einstein did not consider the potential of scalar differential invariants for his project to be exhausted. The purpose of a somewhat obscure calculation on the immediately following pages (pp. 40R-41L), dealing with linear transformations of an algebraic quadratic form, might well have been to learn more about such invariants and their properties.

6.4 Implementing the Lorentz Model of the Equation of Motion (05R)

At some later point, Einstein made a new beginning in his research on a theory of gravitation. He now turned to the other element of the field-theoretical model, the equation of motion. He probably had realized by this time that the equation of motion for a point particle in a gravitational field corresponds to the equation for a geodesic curve in a four-dimensional curved spacetime (cf. eq. (23)).[102] But he probably also had realized that an equation of motion in this sense was not quite the match of the gravitational field equation for which he was looking. The default-setting for the source-slot of the Lorentz model for the field equation was not a point particle but the energy-momentum tensor (cf. eq. (XIV)). The mental model of a field equation together with special relativistic continuum theory now suggested what such an equation should look like in terms of the energy-momentum tensor (cf. eq. (XX)).[103] Such an equation would provide, at the same time, an expression for energy-momentum balance in the presence of a gravitational field.

When Einstein studied the equation of motion problem, he was confronted with the challenge of how to link his general expectations concerning the structure of such an equation with his concrete knowledge about the motion of point particles in a gravitational field. To bridge this gap he made use, as we have discussed before, of a particular model of matter, which allowed him to link point mechanics and continuum mechanics, i.e. the model of "dust" (cf. eq. (XXI)). At a mathematical level, the bridge was built with the help of the Lagrangian formalism (cf. eq. (19)). The dust model allowed Einstein to generalize the equation of motion derived within the Lagrange formalism into a relation between components of the energy-momentum tensor. This relation suggested, in turn, what the full tensorial equation of motion in gravitational field should look like, if it was supplemented by both mathematical and physical default-assumptions provided by the corresponding special relativistic equation.[104] As discussed above, eq. (XX) expresses the energy-momentum balance in a gravitational field, i.e. the generalization of the special relativistic relation between force, energy, and momentum (cf. eq. (XVIII)). Einstein also realized that, from a mathematical point of view, it corresponds to the covariant divergence of the energy-momentum tensor (cf. eq. (XXIV)). This remarkable convergence of physical and mathematical perspectives must have confirmed the expectation that his result also applies to other kinds of sources and turned Einstein's equation into the default-setting for the equation of motion in the Lorentz model and for the energy-momentum balance in a gravitational field.

---

102 He reproduced the proof that the trajectory of a force-free motion constrained to a two-dimensional surface is a geodesic on a page of the notebook immediately following the consideration of quadratic invariants mentioned in the previous subsection (see p. 41R).

103 Einstein emphasized the central role of the energy-momentum tensors and the importance of special-relativistic continuum mechanics in an article he wrote in 1912 but never published, see (CPAE 4, Doc. 1, 63).

104 See the discussion in (Norton 2000, Appendix C).

6.5 A Mathematical Toy Model as a New Starting Point (6L–7L)

The mismatch between the instantiations for two of the slots of the mental model of a field equation, that for the differential operator and that for the gravitational potential, left Einstein with two principal options as to how to proceed. He could continue trying to build an appropriate differential operator applicable to the metric tensor or he could tentatively explore substitutions of the default-setting for the gravitational potential, thus creating "toy-models" in the sense of obviously unrealistic instantiations of the model. Even if that meant temporarily suspending the insight that the gravitational potential is represented by the metric tensor, it might still be possible to gain knowledge from exploring such toy-models that could be helpful in constructing a more realistic candidate field equation.

When Einstein became familiar with the generally-covariant Beltrami invariants as a generalization of scalar differential operators, they must have appealed to him as a promising starting point for his search for a relativistic gravitational field equation. A field equation based on those invariants would automatically satisfy the heuristic requirement of the generalized principle of relativity. A first attempt to construct a differential operator for the metric out of operators acting on a scalar function had, as we have seen, turned out to be too speculative. It was hence worth trying to explore a generalization of the scalar Poisson equation in a generally-covariant setting by using—instead of the Laplace operator—the second Beltrami invariant applied to a scalar function. While such a generally-covariant scalar field equation was only a toy model, it confronted Einstein with a serious problem, viz. that of reconciling a mathematically satisfactory candidate with the physical knowledge of his theory of static gravitational fields, (see Fig. 3, p. 173). In a sense, a scalar field equation formulated in terms of the second Beltrami invariant represents the counterpart of the scalar field equation of Einstein's static theory: while the latter constitutes an initial, physically plausible instantiation for the field-theoretical model, the former represents an equally plausible initial instantiation rooted in mathematical knowledge. In both cases, the resulting field equations were merely starting points for further investigations that had to make contact with knowledge not yet embodied in these first default-settings.

It therefore comes as no surprise that Einstein tried to find out under which conditions a generally-covariant scalar field equation formulated in terms of the second Beltrami invariant reduces to the ordinary Poisson equation. Such a reduction must be possible if the candidate (or rather toy) field equation is to comply with the correspondence principle. It turned out that the implementation of this heuristic principle in this concrete case requires an additional constraint on the choice of the coordinates, supplementing the field equation. Essentially by inspection, Einstein could identify the harmonic coordinate restriction (cf. eq. (57)) as a condition that would make sure that the Beltrami field equation reduces to the ordinary Poisson equation for weak gravitational fields. In other words, the exploration of a toy field equation taught Einstein that a candidate field equation obtained from a mathematical default-setting may require an additional coordinate restriction to be viable from a physical point of view as well; it also familiarized him with a specific example of such restriction, which later turned out to be useful when studying the Ricci tensor.

How could the toy field equation be turned into a real candidate field equation? If the Beltrami field equation is considered as a mathematically reasonable structure to which physical knowledge should now be assimilated, such as the insight that the gravitational potential is actually represented by the metric tensor, it made sense to try to bring this knowledge into an appropriate mathematical form. If unimodular coordinate transformations are assumed, the determinant of the metric transforms as a scalar and can be used to fill the potential-slot of a scalar field equation. The next question was whether the resulting field equation, for the special case of a static field, could be related to the familiar static field equation. Einstein tried to extend this approach by taking into account different versions of a Beltrami-type field equation but failed to integrate the mathematical and the physical knowledge in this way.

6.6 A Physical Toy Model as a New Starting Point (7L–8R)

Einstein's first exploration of the Beltrami invariant had not answered the question as to how to get from a mathematically plausible scalar differential equation to a tensorial field equation that is both mathematically and physically plausible. Reflecting on this gap, Einstein may well have considered the possibility of dividing this transition into two steps. The first would be to construct a tensorial field equation that, even if its mathematical properties were unclear at the outset, made good sense physically. The second step would take him, relying on mathematical tools, from such a physically-plausible toy field equation to the final equation.

In any case, instead of taking a simplified instantiation for the potential-slot of Lorentz model for a field equation to explore a mathematical toy model, Einstein now chose a simplified instantiation for the differential operator slot, while keeping the realistic setting for the potential slot, i.e. the metric tensor. His experience with vector calculus and its use in physics allowed him to write down a straightforward translation of the ordinary Laplacian operator into a differential operator acting on the metric tensor, the core operator. Einstein's experience with the Beltrami invariants must have made it clear that the core operator could hardly represent a generally-covariant object. From the way in which it was constructed, however, it was equally clear that a field equation based on the core operator satisfies the correspondence principle. For this reason, the core operator (cf. eq. (XXVI)) became the default-setting for all of Einstein's subsequent attempts to implement this principle.

This candidate now had to be checked against the other heuristic requirements and, in particular, its behavior under coordinate transformations needed to be explored. This could be done in two distinct ways: either by directly checking the transformational behavior of the core operator, or by considering it—in the sense indicated above—an intermediate step towards the final field equation. Einstein began with the first option. To get beyond linear transformations, however, he used a special kind of coordinate transformations, which explicitly depend on the metric tensor, and which he later called “non-autonomous transformations”.[105] The behavior of the core operator under such transformations is determined by differential equations for the transformation matrices involving the metric tensor and its derivatives. Einstein succeeded in writing down, at least for infinitesimal transformations, the essential term in such a differential equation. But probably in view of the complexity that this condition would take on for finite transformations and, more generally, in view of the unfamiliar character of these non-autonomous transformations, he abandoned this approach and turned instead to the more familiar territory of ordinary coordinate transformations.

In that case the only way to go beyond linear transformations was to generalize the core operator. Einstein developed an ingenious method for doing so. First of all, he considered the two differential operators constituting the core operator separately, the divergence and the gradient (or exterior derivative, cf. eq. (XXV)). He then took the familiar form of these operators applied to some second-rank tensor in Minkowski spacetime with (pseudo-)Cartesian coordinates as his starting point. Einstein now made the assumption that these operators actually transform as tensors under arbitrary coordinate transformations. Under this assumption, a coordinate transformation carrying these operators from their special form in Cartesian coordinates to arbitrary coordinates should reveal their generic form. The idea was similar to that of obtaining a generalization of the line element of Minkowski spacetime to that of a generic curved spacetime by passing from pseudo-Cartesian to arbitrary coordinates in Minkowski spacetime. In both cases one simply had to assume that an equation obtained for the Minkowski metric in arbitrary coordinates is actually valid for the metric of a generic spacetime. Although Einstein did not see his calculations through to the end, he essentially succeeded in finding covariant generalizations of the constituents of the core operator. Eventually he must have realized, however, that this success amounted to no more than a Pyrrhic victory since these generally-covariant differential operators give zero when applied to the metric tensor. In other words, a generalized core operator built from these covariant differential operators is not suitable as a candidate for the left-hand side of the gravitational field equation. Eventually, this failure forced Einstein to take the peculiar non-autonomous coordinate transformations of his first approach much more seriously than he had probably intended when he first encountered them.

6.7 Identifying the Core Operator as the Target of the Mathematical Strategy (8R–9R)

In his next attempt Einstein, reflecting on his earlier failures and insights, combined his prior experiences to develop a procedure for constructing candidate field equations that he would repeatedly use in the notebook (cf. eq. (XLVIII)). The genesis of

---

105 See Einstein to H. A. Lorentz, 14 August 1913, (CPAE 5, Doc. 467). For discussion, see “Commentary ...” (in vol. 2 of this series), sec. 4.3.

this procedure as the result of an oscillation between a more mathematically and a more physically motivated attempt illustrates Einstein's learning experience in the course of his search, which therefore cannot be seen simply as the successive elimination of unsatisfactory alternative candidates.

The attempt to conceive of the physically plausible core operator as the representation of a more general covariant object in specific coordinates had failed because of the degeneracy of the corresponding differential operations when applied to the metric. It made therefore sense to return to the earlier direct exploration of the transformation properties of the core operator. This pathway had not definitively failed yet but turned out to be too rough. Considered from a higher level of reflection, the core operator could not just serve as a physically plausible starting point but also as the possible target of a strategy starting from a mathematically well-defined object. At this point, the only such mathematically well-defined objects that Einstein had at his disposal were the Beltrami invariants. It therefore was natural to deal with them once again, but now not with the theory of the static gravitational field but with the core operator as the more promising physically meaningful target. This approach came with a new challenge, the task to extract a tensorial object, the core operator, from a scalar invariant. This challenge turned out to be manageable.

In short, the idea was to once more start from a mathematically motivated instantiation, the second Beltrami invariant, trying to exploit its familiar mathematical properties in order to determine the transformational behavior of the physically plausible core operator. The necessary bridges between tensorial and scalar objects were readily at hand. From his earlier experience, Einstein knew that he could use the determinant of the metric tensor in the second Beltrami invariant if he considered only unimodular transformations. Now he realized that he could, in turn, try to extract a tensor from a scalar by conceiving the latter as a contraction between two tensors, in this case of the metric tensor and the core operator.

The concrete implementation of this approach confronted Einstein with a number of problems, minor and major. There was, first of all, the need for a restriction to unimodular coordinate transformations. More importantly, when trying to extract the core operator from the second Beltrami invariant applied to the determinant of the metric tensor, he encountered an additional first-order term that required further consideration. Einstein's understanding of the conservation principle, and in particular his experience with his second theory of the static gravitational field, must have immediately suggested to him that this first-order term might be related to an expression for gravitational energy-momentum (cf. eq. (XXXIV)).

In the end, however, Einstein did not succeed in establishing a convincing bridge between core operator and Beltrami invariants. As a consequence, he failed to clarify the transformational properties of the core operator or of a suitably amended candidate gravitational tensor constructed from it. At this point, he took up once more the direct exploration of the transformational properties of the core operator earlier abandoned because of the intricacy of the non-autonomous transformations involved in it. This resort to an earlier approach was, however, no return to square one. Einstein could benefit from the insights he had made in the meantime, in particular from the breaking down of the original problem into simpler ones that the introduction of the Beltrami invariants had made possible. The new first-order term posed a problem analogous to the one Einstein had first encountered when comparing a mathematical toy model based on the second Beltrami invariant with the ordinary Laplace operator and thus suggested the remedy of introducing a coordinate restriction as an additional hypothesis under which a mathematically acceptable expression reduces to a physically plausible one. One could limit then the direct exploration of transformational properties to the remainder term distinguishing the second Beltrami invariant from the contraction of the core operator with the metric tensor. This task was simpler than the original one given the structure of the remainder term. If the class of non-autonomous transformations leaving this term invariant could be determined, one would thereby have found the class of transformations leaving the principal term, i.e., the contraction of the core operator with the metric, invariant as well (cf. eq. (LI)). In this way, a bridge would have been built between the transformational behavior of the mathematically well-defined second Beltrami invariant and that of the physically plausible core operator. In spite of this simplification with respect to Einstein's original attempt to determine the transformational behavior of the core operator, even this reduced task still turned out to be too cumbersome to carry out.

Although this entire episode was fraught with frustrations of reasonable hope, it gave Einstein strategic insights well beyond the concrete mathematical material at hand. There was, first of all, the recognition of the canonical form for the left-hand side of a gravitational field equation, which would have to consist of a core operator plus first-order correction terms somehow related to gravitational energy-momentum (cf. eq. (XXXIV)). Second, the experiences of this episode brought the mathematical strategy into a form that was to dominate much of the subsequent work documented in the notebook. The general idea now was to start from an object with well-defined mathematical properties, in particular with a broad enough covariance group to meet the demands of the generalized principle of relativity (cf. eq. (XLVI)). The next step was to extract from it a candidate gravitation tensor with well-defined physical behavior, more specifically the core operator possibly with correction terms not invalidating the correspondence principle (cf. eq. (XLVIII)). The extent to which the generalized principle of relativity was actually fulfilled could be determined by checking the transformational behavior of the term distinguishing the candidate gravitation tensor from the mathematical starting point (cf. eq. (LI)). This term could be eliminated by imposing the appropriate coordinate restriction (cf. eqs. (L), (LII)). It is remarkable that this strategy, crucial to Einstein's exploration of the Riemann tensor, was in place before he had even seen a single realistic candidate gravitation tensor.

6.8 Subjecting the Core Operator to a Piecemeal Approach (10L–12R, 41L–R)

While in the last episode Einstein had developed an overall strategy for solving his problem and failed, he now took a more piecemeal approach. He focused on a mathematically much simpler object to avoid the complexity of the non-autonomous transformations he had considered so far. On the whole, this phase of his work was characterized by the attempt to break down his main problem, the identification of appropriate field equations, into smaller, more manageable pieces in the hope of identifying reliable building blocks that could then be used to put the puzzle together.

This line of pursuit was largely shaped by the options and constraints that had emerged in the course of Einstein's preceding experience. In particular, while he had just established a paradigm for what was to become his mathematical strategy, for the time being this strategy was powerless for want of mathematical objects other than the Beltrami invariants that could serve as input. The core operator, on the other hand, inspired confidence as a solid achievement that would be a physically meaningful starting point were it not for the difficulties of determining its transformation properties. Yet in view of the absence of other mathematical resources, the use of non-autonomous transformations may have seemed unavoidable. And if hope was to remain of connecting possible results to the Beltrami invariants, the only advanced mathematical objects at Einstein's disposal, the transformations should be unimodular as well.

It is against this backdrop that the emergence of the main idea guiding Einstein's work in this episode becomes understandable. This work may have sprung from the idea to consider simpler mathematical objects that would make a direct approach to the examination of their transformation properties feasible. In any case, at some point it must have become clear to Einstein that the full problem, the determination of the transformation properties of the core operator, could actually be broken down into the study of such simpler objects, if possible vectors or even scalars (cf. eq. (XXV)). The preceding experience with the attempt to extract the core operator from the scalar Beltrami invariant may have triggered this idea. It therefore made sense to carefully check the covariance of these simpler objects in particular under the transformations relevant to the implementation of the elevator and the bucket models, i.e. transformations in Minkowski spacetime to uniformly accelerated or to rotating systems.

The realization of the idea just described brought Einstein to study the transformation properties of the Hertz restriction (cf. eq. (60)). If this restriction was imposed, the core operator reduces to a simpler object whose transformation properties can then be determined separately. But in spite of the greater simplicity of these objects, it turned out to be necessary to introduce a further simplification and to limit the analysis to infinitesimal transformations. With these presuppositions in place, Einstein was able to obtain some specific results even if these were not all that encouraging. In particular, when attempting to implement the elevator model and satisfy the equivalence principle, he found that compatibility with the covariance properties of the Hertz restriction required a modification of the transformation to a uniformly accelerated system, which turned out to be unacceptable for physical reasons. In short, Einstein found it difficult to establish a match between the transformational properties of the objects under study and his physical expectations. How much of the room opened up by his main idea, that of splitting the core operator into two simpler pieces, did he actually investigate in the course of his calculations? It is clear from his notes that he was aware of variants of this operator and hence of alternative splits (involving e.g. the harmonic restriction instead of the Hertz restriction), but he left this option unexplored.

In the end, Einstein once again assembled a number of isolated results that later became useful. In addition, he gained strategic insights governing the subsequent course of his research. As far as his specific results are concerned, he established, for instance, that the Hertz restriction is covariant under infinitesimal non-autonomous transformations to a rotating system in Minkowski spacetime but, as mentioned above, not under a transformation to a uniformly accelerated system. He also found that the core operator is covariant under antisymmetric non-autonomous transformations in Minkowski spacetime, without however being able to associate physical meaning with this result. In the course of his work, he gradually shifted the emphasis of his quest from the transformation properties of the constituents of the core operator to a careful reexamination of the physically relevant transformations themselves. He thereby again accumulated some useful findings such as the derivation of the metric for Minkowski spacetime in rotating coordinates from the Lagrangian formalism. Eventually he made a fresh start, taking unimodular transformations as a starting point for implementing the generalized principle of relativity. He once more tried to match them with transformations to a uniformly accelerating system, again without success. He then abandoned this attempt to incorporate the equivalence principle, alongside with his entire endeavor to deal with the covariance properties of the core operator on the basis of his piecemeal strategy. What remained, apart from specific achievements, was the experience that non-autonomous transformations could be handled after all, at least when applied to sufficiently simple objects. But this was an insight that Einstein would be able to put to good use only much later, when he explored the transformation properties of the finished *Entwurf* theory, in particular in its Lagrangian formulation (cf. eqs. (LXIII), (LXIV)). Of more immediate impact was his realization, fostered by the disappointments produced even by his piecemeal strategy, that it might be prudent to put the pursuit of an audacious interpretation of the generalized principle of relativity on hold, turning instead to the physical requirements embodied in the conservation principle.

6.9 Using the Core Operator as the Starting Point for the Physical Strategy (13L–13R)

Einstein's next move was to look at his problem from a different angle, bracketing the intricate problems raised by the generalized principle of relativity and making sure that what he had achieved so far was at least sound in other respects. And even after the disappointing yield of his piecemeal approach, the identification of the core operator as a candidate compatible with the correspondence principle remained such a sound result. Einstein now tried to address the seemingly intractable aspects of the exploration of the core operator by reducing the ambitious goals imposed by the generalized principle of relativity. For this purpose, he set up a more manageable framework for dealing with the other central, but as yet unexamined aspect of the validity of the core operator as a candidate gravitation tensor, its compatibility with the conservation principle.

His previous research had already suggested that perhaps the core operator needed to be supplemented by additional first-order terms representing the energy-momentum of the gravitational field in the gravitational field equation. But at that point the problem of energy-momentum conservation had occurred only as a marginal aspect of the relativity problem at the center of Einstein's attention. The impasse of his work on this problem provided a natural occasion to return to the issue.

Einstein created a manageable framework by restricting all considerations to unimodular, linear transformations. The requirement of linearity would secure the tensorial character of the core operator, while the requirement of unimodularity kept the door open for establishing contact with the Beltrami invariants later. By setting up a systematic framework for generating vectors and tensors involving the metric Einstein could hope, first of all, to reduce the ambiguities of his approach and second, to gain solid ground for examining the relation between core operator and conservation principle without the interference of the relativity problem. Such an examination might, in particular, help to find the correction terms that he had earlier tried to obtain from the first Beltrami invariant.

Einstein constructed a framework for generating tensorial objects involving the metric with well-defined transformation properties, beginning with the Hertz expression (cf. eq. (60)). He set up a survey of the first-order objects and then stopped, either because these were the objects in which he was mainly interested with a view to the correction terms needed for the core operator or because even this limited overview dashed any hopes he might have had for a reduction of the space of possible candidates. Whatever the case may be, it was at this point that he once more took up the core operator directly, checking its compatibility with energy-momentum conservation. As it turned out, his network of results had become dense enough to allow for such a check which, even if it failed, would probably still provide hints about what was still needed to enforce compatibility. Einstein combined the left-hand side of a field equation based on the core operator with the expression for the energy-momentum balance he had established earlier (cf. eq. (XXX)). He thus produced an expression corresponding to (cf. eq. (XXXII)):

GRAD(POT) X LAP – DIV(LAP) **(LXXXVII)**

which *a priori* could be expected to be of third differential order. But if one now assumes that a field equation of the form (cf. eq. (XXXIX)):

LAP = ENEMO + FIELDMASS **(LXXXVIII)**

holds, then compatibility with the conservation principle requires that the above expression reduces to the second-order expression (cf. eqs. (XXXI), (XXIV)):

GRAD(POT) x FIELDMASS – DIV(FIELDMASS) **(LXXXIX)**

with an appropriate explicit form of **FIELDMASS**. Einstein found indeed that imposing the Hertz condition implies that no third-order terms appear. But, unfortunately, he failed to arrange the terms in the resulting expression in a way that would have allowed him the extraction of an explicit form of **FIELDMASS**.

In summary, Einstein succeeded neither in identifying the conditions under which the core operator is compatible with the conservation principle nor in finding the correction terms that could possibly help establishing such compatibility. Even for such a simple object as the core operator the differential equation resulting from his compatibility check appeared to be too complicated. His calculations and the reflections stimulated by them had nonetheless laid the groundwork for the global approach we have called his "physical strategy," all elements of which were now assembled. The core operator provided him with the starting point for this approach and the calculations just considered constituted the conservation compatibility check for this candidate (cf. eq. (XXXI)). The restriction to linear transformations made it possible to postpone a check of the extent to which the generalized relativity principle was satisfied. What was still lacking was a procedure for guessing or generating suitable correction terms to be added to the core operator to turn it into a viable candidate. Einstein had arrived at a dead end, but not with empty hands. Not only had he accumulated a reservoir of insights and tools that would be useful for his further search, but he had developed two overall strategies, each capable of guiding this search. For the time being, however, both strategies were doomed to be abeyant as long as certain elements that could trigger their application were missing. But as soon as an appropriate incentive was provided, either of them could be activated. For the physical strategy to become productive, all that was needed was a way to generate plausible correction terms to the core operator. For the mathematical strategy to become productive, all that was required were tensors with second-order derivatives of the metric and a well-defined transformational behavior. As it turned out, the latter option was realized first.

6.10 The Systematic Search Phase in the Zurich Notebook

The raw material needed to set the mathematical strategy in motion was evidently delivered by Marcel Grossmann whose name appears next to the first occurrence of the Riemann tensor in the notebook. With this entry, the first phase of Einstein's research was over and a phase of systematic searching for suitable gravitational field equations began. The Riemann tensor represented something like a raw diamond for Einstein to which he could now apply the various extraction schemes that he had elaborated earlier as well additional stratagems he developed in the course of his search. Among these schemes was the contraction of the fourth-rank Riemann tensor to yield a second-rank candidate gravitation tensor, the extraction of such a candidate from a scalar object, the stipulation of coordinate restrictions, and the possibility of modifying candidate field equations by adding or subtracting terms. The products to which these extraction processes gave rise had an impact on Einstein's search procedure going well beyond their immediate evaluation as being either refinements or debris, as is particularly evident from the identification and subsequent rejection of the Einstein tensor.

Einstein's procedure was guided throughout by the Lorentz model, which suggested that candidates for the left-hand side of the field equations have the form of a core operator plus correction terms (cf. eq. (XXXIV)). His prior experience with extracting such candidates from the second Beltrami invariant furthermore gave him guidance on how to handle those terms not fitting his expectations, i.e., how to eliminate them with the help of a coordinate restriction. Since he started from objects with well-defined transformation properties, the main heuristic criteria to be checked were the correspondence and conservation principles. From the point of view of the mathematical strategy, both criteria posed similar challenges and thus seemed to call for similar responses, viz. coordinate restrictions to be imposed in addition to the field equations. This parallel strengthened Einstein's expectation that the stipulation of these heuristic principles required a restriction of the covariance of the object used as the starting point of the mathematical strategy. On the weak-field level, the two restrictions, one resulting from the correspondence, the other from the conservation principle could easily be compared with each other; their compatibility or rather the lack thereof was an important driving force in the search for field equations (cf. eq. (LX)).

The severe restriction on the generalized relativity principle that seemed to be the almost unavoidable consequence of Einstein's procedure made it all the more urgent to check whether or not at least the most essential requirements associated with this principle were satisfied and, in particular, whether the important special case of rotation was included. Not surprisingly, Einstein more than once reexamined this special case during his search.

The difficulties Einstein encountered in the course of his attempts to enforce his heuristic criteria within the formalism he was weaving around the Riemann tensor naturally provoked a reflection on the validity, the physical meaning, and the mathematical implementation of these criteria. After all, they may just have been prejudiced. Does the conservation principle really require the covariant divergence of the stress-energy tensor to vanish (cf. eq. (XXIV))? Does the correspondence principle really demand a static gravitational field to be represented by a spatially flat metric (cf. eq. (25))? Can the generalized relativity principle perhaps be satisfied for rotation by metric tensors other than the one obtained from a coordinate transformation of the standard Minkowski metric? Such questioning of his original heuristic criteria and the default settings suggested by them would eventually pave the way for the breakthrough of 1915. But in the winter of 1912–1913, the answers that Einstein found to these questions confirmed his original conceptions and solidified them by extending the network of inferences in which they were embedded. Ironically, it was precisely the lack of a candidate field equation complying with his heuristic criteria and worthy of further elaboration that also prevented, for the time being, the construction of an even wider network of inferences that would allow these criteria to be overcome.

While Einstein's search would eventually turn up just such a candidate, the *Entwurf* field equation, this candidate was no longer the result of an extraction from the Riemann tensor.

What was overturned in the course of the research documented by the notebook was not Einstein's reliance on his heuristic criteria but the way in which he tried to meet them following his mathematical strategy. Even when he managed to find a candidate for which the coordinate restrictions implied by the correspondence and conservation principles, respectively, could be matched, at least on the weak-field level, the ensuing restriction of the generalized relativity principle and the question of how to satisfy the conservation principle for the full equation made the entire attempt look futile. The appeal of the generalized principle of relativity thus gradually faded away, and the conservation principle gradually emerged as the major stumbling block of the mathematical strategy and, at the same time, as the key stone for a new approach corresponding to a successful implementation of the physical strategy.

Einstein's checks of the conservation principle in the context of the mathematical strategy were limited to weak-field equations. Accordingly, all of his results concerning candidate gravitation tensors—positive as well as negative—were provisional only. For the time being, the limited exploration level of the conservation principle could not be overcome in the context of the mathematical strategy. First, Einstein had no systematic mathematical technique at his disposal for implementing this principle beyond the weak-field level. He would acquire such a technique only much later when developing a variational formalism for the *Entwurf* theory in 1914. Second, the ad-hoc strategies he used to implement the conservation principle beyond the weak-field level necessitated substantial modifications of the candidate field equations serving as the starting point of the mathematical strategy, modifications that made the transformation properties of the proposed field equations intractable despite their origin in the generally-covariant Riemann tensor.

Einstein's experiences with extracting candidate gravitation tensors from the Riemann tensor thus displayed a remarkable parallelism to his prior experiences with the Beltrami invariants. In both cases, the advantages gained by starting from an invariant or generally-covariant object had to be gradually given up in favor of satisfying the other heuristic requirements rooted in the knowledge of classical physics until finally nothing was left of the covariance properties that recommended these objects in the first place. But it was not only this twofold experience of failure that ultimately triggered a switch from the mathematical to the physical strategy. It was precisely the main weakness of Einstein's attempts to come to terms with the conservation principle, i.e., the limitation to the weak-field level, that indicated a way out of the impasse. If the implementation of the conservation principle at the weak-field level could not be the final word, it made sense to take energy-momentum conservation for the weak-field equations as a starting point for identifying those additional terms that were needed to turn the core operator into a viable candidate complying with this heuristic requirement, irrespective of the generalized principle of relativity. Einstein thus found a way of solving the problem that had blocked the pursuit of the physical strategy before, viz. the lack of a procedure for generating plausible correction terms to the core operator. His difficulties with the mathematical strategy suggested a procedure, whose first elements were found before they turned into a systematic mechanism. Once more the essential pattern governing the next step of Einstein's research, the derivation of the *Entwurf* field equations, had been prepared by reflecting on the blocked pathways encountered in the previous episode.

The failure of Einstein's pursuit of the mathematical strategy in the Zurich Notebook resulted in the derivation of the *Entwurf* field equations along the physical strategy. The establishment of these field equations, compatible with both the correspondence and the conservation principles, ended, for the time being, his systematic search for gravitational field equations. What remained from his efforts in the winter of 1912–1913, however, was more than yet another and, as it eventually turned out, unsatisfactory candidate that would eventually be discarded. There was the November tensor, which did not immediately fall victim to any knock-out argument derived from Einstein's heuristic checklist, and which was dropped, not in favor of a better candidate, but in favor of a seemingly better strategy. There were the Ricci tensor and the linearized Einstein tensor, which had been explored only at the weak-field level. From this perspective, their later revival is not surprising. But apart from candidates that he would consider again in late 1915, Einstein's search for field equations in the winter of 1912–1913 also left its mark at the strategic level, both in his subsequent attempts to consolidate the *Entwurf* theory and in his renewed search for field equations at the end of 1915. In fact, even when he focused exclusively on the *Entwurf* theory, he never abandoned the expectation, grounded in the experience documented by the Zurich Notebook, that it should be possible to arrive at the same field equations using either the physical or the mathematical strategy. It was this persistence, perhaps more than the potential of any not yet fully explored candidate, that prevented Einstein from ceasing his quest before he had reached his goal of a generally-relativistic theory of gravitation in late 1915.

6.11 Fitting the Riemann Tensor to the Lorentz Model (14L–18R)

When Marcel Grossmann introduced Einstein to the Riemann tensor, this new mathematical resource fell on ground that was well-prepared by Einstein's previous investigations. The expectations with which he approached the new object, however, sent him in a direction very different to where our modern expectations would take us, viz. the derivation of the Einstein field equation from the Riemann tensor. For Einstein, the Lorentz model essentially prescribed the steps to take to evaluate the new candidate. His prior attempts to implement this model had led him, in particular, to expect a field equation with a left-hand side of the form (XXXIV), i.e., a left-hand side of the form 'core operator plus correction terms,' which is incompatible with what we now take to be the correct field equations.

The central role of the Riemann tensor within the absolute differential calculus as the wellspring of all other "differential tensors" and "differential invariants"—a role of which Grossmann was certainly aware (Einstein and Grossmann 1913, 35)—and its unexplored status in Einstein's investigations must initially have nourished high hopes for the project of extracting from it a suitable left-hand side of the field equations. Einstein may even have expected that the direct pathway from the Riemann tensor to an object fitting the Lorentz model would produce the desired result, without any of the moves and tricks that had been necessary in the earlier attempts based on more pedestrian mathematics. If needed, however, by now such auxiliary schemes were available to Einstein should difficulties arise. In any case, the fourth-rank Riemann tensor had to be turned into a second-rank tensor that could serve as the left-hand side of gravitational field equations whose right-hand side was the second-rank stress-energy tensor (cf. eq. (XIV)). This was a straightforward mathematical operation, which Einstein carried out as soon as he had been handed the Riemann tensor. Unfortunately, the result of this operation, the second-rank Ricci tensor (cf. eq. (55)) did not fit to instantiate the open operator slot of the Lorentz model for the field equation but contained additional, unwanted second-order terms invalidating the correspondence principle (cf. eq. (56)). Einstein's first attempt to assimilate the Riemann tensor to his mental model thus resulted in the condition that these disturbing terms would have to vanish. The appearance of such an additional condition is reminiscent of similar hindrances he had encountered exploring the Beltrami invariants.

Now that the direct approach had failed, Einstein was forced to exploit the tricks and tools he had assembled before. The most obvious way to connect his new predicament with his earlier experiences was the construction of a scalar object from the Riemann tensor, the Ricci curvature scalar. This scalar object could be subjected to exactly the same procedure as the scalar Beltrami invariant. Einstein thus attempted to extract a tensorial object from it in analogy to his earlier treatment of the second Beltrami invariant, i.e., by conceiving the scalar as the contraction of this new, contravariant tensorial object and the covariant metric tensor. Also in analogy with his earlier work on the Beltrami invariants, he set the determinant of the metric equal to unity to simplify his calculations, thereby imposing a restriction to unimodular coordinate transformations.

The hope was that the new second-rank tensor extracted in this way from the curvature scalar would represent a suitable candidate for the left-hand side of the field equations, meeting the requirements of the Lorentz model. Unfortunately, the considerable calculational effort required to pursue this option failed to produce more acceptable results than the direct approach. Einstein even briefly considered introducing an additional condition on the metric tensor—a weaker form of the Hertz restriction—but apparently gave up this idea because it did not seem to promise an easy way out either. He then tried to make some progress by comparing the two unsatisfactory candidates he had extracted from the Riemann tensor in both their contravariant and covariant forms. This procedure also followed the example set by his experiments with the Beltrami invariants and may similarly have been driven by a concern for uniqueness and the hope to learn from combining different pathways. While the procedure was given up without reaching a definite conclusion, it gave an insight that quickly proved to be important. Using techniques familiar from his Beltrami experiments, Einstein found that the constancy of the determinant of the metric could be used to replace one of three disturbing second-order terms occurring in the Ricci tensor by a first-order expression.

At the same time, it must have been clear to him that disturbing second-order terms of some sort were there to stay, and, consequently, that at least some aspects of what might initially have appeared to be mere stop gaps were there to remain as well. Precisely because of the original promise of the Riemann tensor, it was clear that the problem could no longer be the lack of mathematical resources and that no amount of calculational sophistication would suffice to turn the Riemann tensor into an acceptable candidate gravitation tensor without introducing further hypotheses, in all likelihood with serious physical repercussions. The need for further hypotheses was also suggested by the fact that the conservation principle had not played any role in the analysis of the Riemann tensor so far. It was to be expected that this heuristic requirement would exact its price as soon as the physical consequences of a gravitation theory based on the Riemann tensor were pursued any further.

6.12 Establishing a Contradiction between the Correspondence and the Conservation Principles (19L–19R)

When Einstein had tried to match the second Beltrami invariant to the correspondence principle, he had hit upon the harmonic coordinate restriction as a suitable auxiliary hypothesis. Trying to match the Ricci tensor to the correspondence principle, he found that the same hypothesis could be used to eliminate *all* disturbing second-order terms. This first-order condition on the metric tensor was suggested by the condition following from the restriction to unimodular coordinates. This immediately gave it a similar status, i.e., that of a global coordinate restriction. Against the background of his earlier experience with the Beltrami invariants, the introduction of such an auxiliary hypothesis was clearly an application of what we have called the mathematical strategy. The natural next step would thus have been to explore the transformational properties of this additional restriction (cf. eq. (LIV)).

However, Einstein's earlier experience had also involved wrestling with the conservation principle. He had come to realize that this principle might entail further restrictions, affecting the covariance properties of the theory. Knowing that exploring the transformational properties of such extra conditions could become quite involved, he first tackled the issue of conservation. It made sense to collect all necessary restrictions first, and establish the transformational properties of overall restriction later (cf. eq. (LIX)). To explore the emerging network of conditions, Einstein simplified his framework, focusing on a first-order, weak-field approximation. He thereby effectively introduced another toy model, now with the goal to explore the entanglement of correspondence and conservation principles.

In weak-field approximation, the harmonic coordinate restriction coming from the correspondence principle could easily be related to the restriction coming from the requirement of compatibility between field equation and energy-momentum conservation. It was immediately clear that the field equations in first-order approximation satisfy the divergence condition (LXXI). The conservation compatibility check (LXXIV) gave rise to an additional restriction which could also be brought into a first-order form, for comparison with the harmonic restriction. The combination of the resulting Hertz restriction with the harmonic restriction implies that the trace of the metric tensor must be constant. This implication was inacceptable to Einstein on physical grounds. It was incompatible not only with the default-setting for the metric tensor of a weak static gravitational field (25) but also, via the field equations (cf. eq. (LXXV)) with the default-setting for the stress-energy of matter as given by eq. (XXI). In view of this discrepancy between the mathematical consequences of his heuristic principles and his physical expectations, it is not surprising that Einstein at this point reexamined the legitimacy of the conservation compatibility check which had evidently triggered this conflict. A crucial implication with physical significance was the vanishing of the covariant divergence of the energy-momentum tensor (cf. eq. (XXIV)). Within his weak-field approximation, Einstein therefore rederived this relation from first principles, i.e., from the continuity equation and the equation of motion. In this way, he not only extended his network of arguments to include the latter results but, more importantly, he firmly established the existence of a contradiction within this network, with no simple escape by adjusting his heuristic principles.

In summary, Einstein's exploration of the Ricci tensor as a candidate for the left-hand side of the gravitational field equations had ended in an impasse. At the same time, this exploration had helped him to further extend his strategic resources. They now included, in particular, the consideration of a weak-field equation. Furthermore, the mathematical strategy was amplified by adding as a routine a compatibility check of the restrictions resulting from the correspondence and the conservation principles, respectively. As a result, the notion of coordinate restrictions as a virtually unavoidable consequence of combining a generalized relativity principle with other physical requirements was solidified. Perhaps the most important result of Einstein's exploration of the Ricci tensor was, however, the establishment of a sharp contradiction in the argumentative framework. The identification of this contradiction offered a range of fairly clear options of how to avoid it. Among the alternative pathways to explore was the option of changing the physical default settings entering his argument, in particular those for the metric tensor of a static field and for the stress-energy or energy-momentum tensor. Another option was to reconsider the implementation of the correspondence principle with the help of the harmonic coordinate restriction, e.g., by extracting a new candidate from the Riemann tensor with the help of a different coordinate restriction. Probing a different implementation of the correspondence principle probably looked like the more sensible option given that Einstein's reconsideration of the conservation principle had strongly confirmed *its* implications. In the course of his research, Einstein eventually pursued all of these options. The option he chose to explore first came courtesy of the new toy model he had introduced, the weak-field equation. Why should it not be possible to tinker with the field equations themselves, within the weak-field framework, in order to find out whether there really was no way to satisfy all requirements on the table, including the harmonic coordinate restriction?

6.13 Matching the Riemann Tensor and the Correspondence Principle: the Failure of the Linearized Einstein Tensor (20L–21R)

The preceding considerations had shown Einstein that the contradiction between the coordinate restrictions implied by the correspondence and conservation principles, respectively, had to be taken seriously enough to entertain even a modification of the form of the field equations. His starting point had been a weak-field equation obtained from the Ricci tensor by imposing the harmonic coordinate restriction to satisfy the correspondence principle. The most obvious conflict was that between the implication of the conservation principle that the trace of the stress-energy tensor of matter must vanish, on the one hand, and the default-setting for this tensor (XXI), on the other hand. Einstein's earlier experience with the adjustment of his original theory of the static gravitational field to the requirements of the conservation principle helped to make a modification of the field equation acceptable as a possible way out of this dilemma. In addition, the weak-field equations made the exploration of possible modifications easier by making it possible to study the interplay between the various constraints in a mathematically simplified form. While the overall logic of this exploration was dominated by the mathematical strategy, the challenges it produced for the various physical default-settings of Einstein's search for the gravitational field made it necessary to reflect on his heuristic presuppositions as well and to go back once more to the physical principles guiding his search such as the equivalence principle and even to the more secure part of his theory in the making, the equation of motion.

Since neither the conservation principle nor the default-setting for the stress-energy tensor of matter could be given up easily, the conflict between them first turned Einstein's attention to the source slot of the field equation, or rather on its default-setting, the energy-momentum tensor of matter according to eq. (XIV). If this default setting could be changed, the default setting **DUST** (XXI) for the stress-energy might well be retained without leading to a conflict with the conservation principle. By replacing the default-setting eq. (XIV) with a traceless quantity, Einstein was indeed able to avoid the conclusion that the trace of the stress-energy tensor has to vanish if the trace of the field equation vanishes, as it would have to as a result of combining harmonic and Hertz restrictions, as we have seen.

This remarkable achievement did not provide an entirely satisfactory solution to the compatibility problem of the correspondence and conservation principles. Einstein had resolved the conflict between the combined coordinate restrictions following from these principles and the default-setting for the energy-momentum tensor (XXI), the discrepancy between the combined coordinate restrictions and the default-setting for the metric tensor of a weak static gravitational field (25) still existed. The preceding experience had taught Einstein how modifying the field equation could help in dealing with disturbing coordinate restrictions. If that method had worked to get rid of the unwanted trace condition, why not try to use it again to get rid of the Hertz restriction altogether rather than to make it compatible with the harmonic coordinate restriction?

Once again, an unsuccessful line of thought had thus paved the way for an important strategic insight, which, in this case, gave Einstein the harmonically reduced and linearized Einstein tensor as a candidate for the left-hand side of the field equations. Instead of giving up and replacing the default-setting for the right-hand side of the field equations, changing the way in which the source-term enters the equation, he modified the way in which the gravitational potential enters the left-hand side of the equations, i.e. the default-setting for the weak-field version of **LAP** (XXVIII), the d'Alembert operator. In a sense, this may have appeared to Einstein as the more conservative approach because it interfered less with the canonical form of the field equation. More specifically, Einstein changed the left-hand side of the field equations by adding a trace term in such a way that the object on which **LAP** operates becomes equal to the left-hand side of the harmonic coordinate restriction if the divergence of this left-hand side is taken (cf. eq. (LXXIX)). In this way, the vanishing divergence of the energy-momentum tensor, which is required by the conservation principle and which originally resulted in the Herz restriction, is now implied by the harmonic coordinate restriction alone—without imposing an additional constraint (cf. eq. (76)).

Now that the correspondence principle and the conservation compatibility check in its weak-field form had been taken care of, the next step was to make sure that the conservation principle was satisfied in all of its facets. For the modified field equations, Einstein needed to check, in particular, whether the gravitational force could be represented as the divergence of a gravitational stress-energy expression. The weak-field equations passed this test without any problem, in spite of the additional trace term they involve (cf. eq. (75)). It was less obvious how this success could be extended to the full version of the equations. A half-hearted attempt to solve this problem was, apparently, enough for Einstein to see that this extension represented a major challenge and that it might even bring back additional coordinate restrictions.

After this preliminary exploration of the conservation issue beyond the weak-field case, Einstein returned to the correspondence principle and discovered that another conflict between the modified field equations and the default-settings of his search was still unresolved. The canonical metric for a static field (25) is no longer a solution of the modified field equations. Since the weak-field equation with the added trace term had otherwise fared fairly well in comparison to earlier candidates, it made sense to carefully reexamine the legitimacy of the one obstacle that remained, the default-setting for the weak static gravitational field. He tested its justification by physical knowledge in the same way in which he had earlier checked the legitimacy of the Hertz restriction when it proved to be an obstacle. He turned to the more solid ground provided by the equation of motion. Since the entire theory of the static field had, in a sense, originated from the equation of motion with the help of the equivalence principle, a check of the default assumption about the static gravitational field with the help of this principle was the most natural option to pursue. From this perspective, the crucial question was whether the default-setting for the weak static gravitational field was actually inescapable given the equivalence principle. What does Galileo's principle of equal acceleration in a gravitation field, on which the equivalence principle hinges, imply about the form of a metric tensor for a weak static gravitational field?

All Einstein had to do to address this question was to formulate Galileo's principle in terms of his metric formalism. The conceptual framework within which his question was formulated suggested to do so by trying to identify the elements of Newton's equation (cf. eq. (III)) within this formalism, in particular the force term and the mass (or energy) term. His earlier work on the equation of motion and his experience with the Lagrange formalism gave him the tools for writing down the required quantities. Since their interpretation was governed by the conceptual framework of Newtonian physics, Einstein could draw the conclusion that, if the force was to vary as the energy, so as to ensure the validity of Galileo's principle, the metric for a static field must take on its canonical form. As a consequence of this inference, based on combining physical with mathematical elements, in a way that in hindsight can be recognized as problematic, the default-setting for the weak static gravitational field became even more firmly rooted in Einstein's heuristic framework, making its clash with the linearized Einstein tensor so much the worse for the latter.

In summary, Einstein's attempt to match the Riemann tensor first with the correspondence principle and then with the conservation principle by setting up a field equation for which only the harmonic coordinate restriction was needed as a supplementary condition had left him in the end without a viable candidate to pursue. The promising candidate he had found in the process had to be rejected because it seemed to be irreconcilable with the equivalence principle. Thus, in this dramatic episode of the search for gravitational field equations, the Einstein tensor of general relativity was, albeit only in a weak-field approximation and for harmonic coordinates, identified and discarded. Clearly, the criteria that led to its rejection had to be changed before it could be accepted. In particular, the default-setting for the metric of a weak static field had to be given up, in spite of its support by the canonical form of the weak-field equation and the—in hindsight—spurious argument based on the equivalence principle. The rejection of the equations in the winter of 1912–1913 was a matter of heuristic criteria that were still rooted in classical physics and that were incompatible with general relativity as we know it today. It was also a matter of a network of arguments that were still too loosely woven to produce a contradiction between any candidate field equation and these classical criteria, that could seriously challenge the latter rather than leading only to the rejection of the former.

Again, the failure to establish an acceptable candidate field equation in this preceding episode strengthened Einstein's vision and generated new strategic insights. The mathematical strategy was now fully operative, from the extraction of a candidate from the Riemann tensor, via the introduction of a weak-field approximation, to the matching of the coordinate restrictions following from the correspondence and the conservation principle, respectively. Among the new insights may have been an appreciation of the difficulty in passing from a weak-field to a full-fledged implementation of the conservation principle. But before this problem could even be addressed another candidate equation was needed. The options for avoiding the original conflict between the coordinate restrictions resulting from the correspondence and the conservation principle, respectively, suggested a different implementation of the correspondence principle. The pathway toward such a different implementation and thus to a new candidate was, in a sense, suggested by the original conflict itself. So far, Einstein had tried to get rid of the Hertz restriction in order to follow the path indicated by the harmonic restriction. Since this path looked like a dead end, it made sense to abandon the harmonic restriction, retaining the Hertz restriction instead.

6.14 Matching the Riemann Tensor and the Conservation Principle: the Failure of the November Tensor (22L–25R)

The Ricci tensor and the Einstein tensor do not exhaust the potential represented by the Riemann tensor and the mathematical strategy for producing candidates for the gravitational field equations. As mentioned above, the direction in which to proceed was indicated by the as yet unresolved conflict between the correspondence and the conservation principles, which was embodied in the clash between two coordinate restrictions, the harmonic restriction and the Hertz restriction, respectively. Since the constraints imposed by the conservation principle appeared to be unavoidable, and since Einstein's earlier attempt to suppress the need for the Hertz condition had failed, he now explored the possibility of realizing the correspondence principle in a new way, without the help of the harmonic restriction.

It was once again Marcel Grossmann who prepared the ground for pursuing this other possibility. At the price of a restriction to unimodular transformations, the Ricci tensor could be split into two parts, each part individually transforming as a tensor under unimodular transformations. One of those two parts was a promising new candidate for the left-hand side of the field equations, the "November tensor" (cf. eq. (82)).

The November tensor has a surprisingly elegant form: the divergence of a Christoffel symbol plus a quadratic expression in the Christoffel symbols. If the Christoffel symbols were taken to represent the gravitational field (cf. eq. (XXIII)), the candidate would have the canonical form of eq. (XXXVIII). Such an interpretation, however, was in conflict with Einstein's heuristics at this stage, which demanded the implementation of the correspondence principle first by imposing an appropriate coordinate restriction; furthermore the default setting for the gravitational fields was given by eq. (XXII). At this point, Einstein only looked for an interpretation of a candidate in terms of field components once he had found the *reduced* field equations, i.e., once he had imposed a coordinate restriction to meet the demands of the correspondence principle (cf. eq. (LXXXI)).

As he had done before, Einstein expanded the Christoffel symbols in terms of derivatives of the metric to identify the disturbing second-order terms preventing the implementation of the correspondence principle. These disturbing second-order terms, it turned out, could be eliminated with the help of the Hertz restriction so that the harmonic restriction was no longer needed. Since the Hertz restriction also guarantees the vanishing of the divergence of the linearized stress-energy tensor, the conflict between correspondence and conservation principle was thus resolved, at least at the weak-field level.

Now that this major conflict was settled, new problems arose, among them the question of the transformations allowed by the reduced field equations and the question of the implementation of the conservation principle for the full field equations. Einstein first addressed the issue of covariance which, given the known transformational behavior of the November tensor and following a strategy first established in the context of the Beltrami invariants (cf. eq. (LIX)), could be addressed by exploring the transformation properties of the Hertz restriction. Einstein could also build on an earlier analysis of the Hertz restriction which seemed to indicate that transformations in Minkowski space to a linearly accelerated frame presented a problem, but that transformations to rotating frames did not.

A complete clarification of the transformation properties of the Hertz restriction could be obtained by a larger effort dealing with non-autonomous transformations. Before undertaking such an effort, Einstein preferred, it seems, to turn once more to the conservation issue. How could he extend his results concerning the conservation principle from the weak-field level to the full field equations? He must have been aware of the crucial role of the first-order correction terms to the core operator. Considering the reduced November tensor, i.e. the terms left of the November tensor after imposing the Hertz restriction, Einstein was confronted with a number of such first-order terms, destroying the simple structure which the new candidate displayed when written in terms of the Christoffel symbols. Einstein tried to reintroduce the Christoffel symbols. While this allowed him to group certain terms more effectively, the resulting expression became even more opaque, mixing as it did first-order derivatives of the metric and Christoffel symbols. It was difficult to see, on the basis of this expression, how the conservation principle for the full field equation could be satisfied.

At this point Einstein had an idea that may seem ingenious but whose grounds were prepared by the contrast between the simple and elegant original structure of the November tensor expressed in terms of the Christoffel symbols and its confusing complexity when written in terms of derivatives of the metric. What was needed was a preservation of the original structure in terms of what, in Einstein's understanding, would be the true representation of the gravitational field, viz. the first-order derivatives of the metric (cf. eq. (XXII)). The resulting candidate gravitation tensor would then be of the canonical form (XXXVIII) and, in all likelihood, comply with both the correspondence and the conservation principle. The idea was to impose a new coordinate restriction that would effectively allow Einstein to replace the Christoffel symbols by first-order derivatives of the metric. This was somewhat more difficult than the above sketch would suggest, as it required in particular the introduction of an indirectly defined coordinate restriction amounting to the stipulation that an object we have designated as the "theta expression" behaves as a tensor. It nevertheless proved to be fairly successful. Not only did Einstein manage to obtain a "theta-reduced November tensor" of the desired canonical form but, along the way he also found out that he no longer needed the Hertz restriction as an additional condition to recover the Newtonian theory.

The next challenge was to determine the covariance properties of this theta-reduced November tensor implied by the somewhat strange new coordinate restriction. Einstein first derived a general condition for the infinitesimal non-autonomous transformations leaving the theta expression invariant, which, however, was just as formidable as the conditions of this kind that he had encountered earlier. He then turned to a special case and tried to identify the class of transformations in Minkowski space that preserve the theta condition. In doing so, he found a puzzling result: among the metric tensors satisfying the theta coordinate restriction was a metric corresponding to Minkowski space in rotating coordinates but with interchanged covariant and contravariant components, an object we shall call the "theta rotation metric," or simply the "theta metric." This curious result continued to concern Einstein almost until the end of the research period covered by the Zurich Notebook.

What did this strange finding actually mean? Was rotation covered by the theta restriction or was it not? To answer this question, Einstein had to find a physical interpretation of the curious theta rotation metric, exploring whether or not it was possible to connect it to the dynamics of rotation. He did so in various ways. First he rederived the equations of motion with the help of the Lagrange formalism in order to identify Coriolis and centrifugal forces. He abandoned this approach because it became too involved. Then he switched covariant and contravariant components in the theta condition, since this reformulated condition would obviously admit the ordinary rotation metric as a solution. Finally, in yet another attempt to come to terms with the physical interpretation of the theta condition, Einstein took recourse to the law of energy-momentum conservation, reformulating it in terms of the covariant rather than the contravariant stress-energy tensor and trying to extract from the reformulated law and from the theta metric the correct expression for the centrifugal force. Due to an error, Einstein at first convinced himself that this was actually possible but then appears to have developed doubts.

While the physical meaning of the theta rotation metric remained obscure, its exploration, nonetheless, had two consequences for Einstein's subsequent work: First, it proved increasingly difficult to reach a comprehensive implementation of the generalized relativity principle, and rotation increasingly became something of a litmus test, the one case of accelerated motion that Einstein expected his theory to cover to comply with his original heuristic mission. Second, checking the theta metric with respect to the dynamics of rotation may well have directed Einstein's attention once again to the significance of the force expression as a clue to viable field equations.

Let us try to reconstruct such a clue by means of our symbolic expressions. It was indeed possible to derive a force expression from the linearized field equation, expressing it as the divergence of the gravitational stress-energy density (cf. eqs. (XXXIII), (XXVIII)):

LIM(FORCE) = DIV(FIELDMASS). (XC)

If such an expression offers a physically meaningful starting point, for instance because it vanishes for rotation, it might serve as a criterion for picking a suitable gravitation tensor instead of just serving as an indirect consistency check by way of a particular solution such as the theta metric. In that case, the force expression could perhaps be reinterpreted as representing an exact quantity even though it was obtained in linear approximation. The force expression could thus become, in a way similar to the transition from Einstein's first to his second theory of the static gravitational field, the starting point for extracting a suitably corrected full gravitation tensor from it (cf. eq. (XXXVI)):

FORCE = GRAV x FIELD (XCI)

A gravitation tensor **GRAV** constructed in this way would automatically satisfy the conservation principle and looked promising with respect to the generalized relativity principle, at least as far as rotation was concerned. After all, it fulfilled a necessary condition for being compatible with the relativity of rotation, the vanishing of the corresponding force expression in the case of rotation.

If Einstein were in fact trying to implement such ideas, he ran into a number of difficulties, caused in part by calculational errors. First of all, Einstein did at first not systematically construct a candidate gravitation tensor **GRAV** but seems to have merely guessed it. Second, the gravitation tensor he extracted from the force equation does not vanish for rotation as he had hoped, but then he found that this extraction itself involved errors whose elimination might well yield the desired result after all. Third, he must have realized that by postulating a physically meaningful force expression as his new starting point he effectively abandoned the link with the November tensor with its well-defined transformation properties. It therefore made sense to interpret the candidate field equation extracted from the force expression not as the definitive result of a physical strategy but rather as the new preliminary target, itself subject to further corrections, of the mathematical strategy starting from the November tensor. In this way, the advantage of well-defined transformational properties might be combined with that of a physically meaningful force expression ensuring the satisfaction of the conservation principle and perhaps even covering the generalized relativity principle for the case of rotation.

In a sense, Einstein may have reached once again reached the constellation he had reached earlier when establishing the core operator as the physically meaningful target (modulo correction terms) of a mathematical strategy taking the second Beltrami invariant as its starting point. Now the place of the Beltrami invariant was taken by the November tensor and that of the core operator by a candidate gravitational field equation that received its physical meaning not just from the correspondence principle but from the conservation principle as well. In the end, however, Einstein once again was unable to build a convincing bridge between his mathematical starting point and his physically meaningful target.

Unable to build a bridge between the two, Einstein put the November tensor to one side for the time being and explored the field equation suggested by his force expression. However, at this bifurcation point of his research, he does not seem to have had, perhaps for the first time, any promising idea about how to proceed. Neither his overall heuristics nor his remarkable ability to draw strategic lessons from failure suggested a plausible next step. The entries in the notebook at this point do not seem to follow any coherent and well-defined strategy. As noted above, the field equation suggested by the force expression vanishing for rotation does itself not vanish for the rotation metric. Does it perhaps vanish for the curious theta rotation metric obtained by interchanging covariant and contravariant components? This question may sound absurd but was nevertheless pursued by Einstein. He even convinced himself—arbitrarily adjusting a coefficient—that his candidate field equation does indeed vanish for the theta metric, a conclusion that is in fact erroneous. This specious result encouraged him to resort to an earlier trick: if the candidate field equation vanishes for the theta metric, a new candidate field equation could be constructed by interchanging contravariant and covariant components that would vanish for the ordinary rotation metric. The new candidate resulting from this crude operation was mathematically ill-defined. Nevertheless, Einstein explored it. It covered, or so he may have believed, the case of rotation, was compatible with the correspondence principle, and looked promising as far as the conservation principle was concerned.

This last issue called for a closer examination and brought Einstein back to the starting point of this phase of his search, the expression for the force density. The new candidate would be compatible with the conservation principle if it gave rise to a force density that can be represented as the divergence of a stress-energy expression for the gravitational field. Now that he had found an apparently viable candidate complying with the rotation criterion by merely formal manipulations, it made sense to repeat the procedure that originally brought him to the expression for the force density and that had been the point of departure of this whole line of reasoning. Following this procedure, Einstein began to write down the force density for the linearized version of the new candidate field equation, which he then tried to make exact. If everything worked out as expected, his procedure should correspond to the transition from eq. (XC) to eq. (XCI) so that he should be able to reconstruct his full candidate in this way, that is, essentially from inserting the right candidate for **LAP** into the expression for the force expression **LIM(FORCE)** and then generating the correction terms yielding **GRAV**.

Unfortunately, things did not work out in the end. Einstein managed to extract terms from **LIM(FORCE)** that had the required form of a divergence or that could be put on the left-hand side of the field equation as correction terms, but he also encountered a term that could not be treated in either of these two ways. But he got close as only one disturbing term remained. Remarkably, the terms that were put on the left-hand side of the field equation not only induced correction terms of the form **CORR(POT) x FIELD** but also a term of the form **LAP(POT) x FIELD**. This suggested that Einstein's expansion of **LIM(FORCE)** might actually produce an identity if only the right expression for **LAP(POT)** was taken as the starting point. Eventually, Einstein nonetheless abandoned the entire calculation, probably not only because he failed to establish the compatibility of his candidate with the conservation principle but also because he may have realized at some point that this candidate did not make good sense mathematically in the first place.

In summary, Einstein had extracted yet another candidate from the Riemann tensor in addition to the Ricci, the harmonically reduced Ricci, and the harmonically reduced Einstein tensors: the November tensor. In the end, this candidate was judged to be just as unsatisfactory as its predecessors. Its original appeal gradually waned because of the problems Einstein ran into when he tried to turn the November tensor into a viable candidate by adding appropriate coordinate restrictions. The discouraging result was that it hardly made any difference whether a candidate resulted from reducing a covariant object by additional coordinate restrictions or whether it was merely constructed *ad hoc*. Either way, the main challenges, the compatibility with rotation and the satisfaction of the conservation principle, had to be addressed directly, by explicit construction. As a consequence, Einstein's mathematical strategy lost its appeal and gave way to another tinkering phase.

In this tinkering phase Einstein focused on the expression for the gravitational force which had the advantage of having a clear physical interpretation. Such an expression had already played a key role in the transition from his first to his second theory of the static gravitational field. If the force can be written as a divergence, the conservation principle is satisfied automatically. And if the expression for the force happened to vanish for rotation, there was at least a chance of meeting some of the demands of the generalized relativity principle as well. Einstein's problem was that he had yet to find a way of systematically extracting a candidate gravitation tensor from such a force expression. His attempts to construct or guess candidate gravitation tensors along this line tended to destroy the promise of his initial ansatz. Given a candidate consisting of some version of the core operator plus correction terms suggested by a force expression, it still had to be checked against the conservation principle. This in turn meant forming a force expression from the linearized field equation which then was to be rewritten as a divergence, possibly with the help of introducing new correction terms to the original ansatz. Having tried this procedure once if not twice without being able to reproduce his original ansatz, Einstein noticed that he could actually begin simply with the core operator and *use the conservation principle as a means for producing correction terms* to it. The realization of this possibility was the birth of the *Entwurf* strategy and, as far as the research documented in the Zurich Notebook is concerned, the end of the mathematical strategy.

6.15 Matching Correspondence and Conservation Principles: The Emergence of the Entwurf Equations (25R–26R)

As the November tensor gradually dropped out of sight, the mathematical strategy launched with the introduction of the Riemann tensor into Einstein's research fell victim to the attrition associated with the efforts to realize Einstein's physically motivated heuristics when starting from a generally-covariant object. His principal instrument for implementing the correspondence and the conservation principles was to impose coordinate restrictions, which, first of all, had to be brought into agreement with one another and then tended to consume the original benefit provided by a generally-covariant starting point. Meanwhile, he had developed more concise ideas about what a viable candidate satisfying both the correspondence and the conservation principles should look like. In the course of his research he had even encountered candidates that seemed close to satisfying these heuristic criteria. Einstein, however, never succeeded in building a bridge between a mathematically viable starting point such as the November tensor and a candidate that looked promising from a physical point of view. While the November tensor, in particular, was never quite refuted, it just became more and more uninteresting.

On a heuristic level, Einstein's difficulties in implementing simultaneously the correspondence and the conservation principles counteracted the potential advantage of starting from a candidate satisfying the generalized relativity principle. Even when the battle was won at the weak-field level, the mathematical strategy failed to provide any hint for winning it at the level of the full equations. Instead, such a hint came from mere formal manipulations of the force expression that had already guided Einstein's pathway from his first to his second theory of the static field. Against the background of his prior experience with a physical strategy and the inadequacy of the mathematical strategy to cope with the conservation principle, this hint prepared the ground for the derivation of the *Entwurf* field equation.

The gist of this derivation consists in starting from the force expression for the core operator which is then transformed into a divergence expression plus terms which are identified as correction terms (cf. eq. (XXXVII)). The resulting identity then yields both the correction terms and the gravitational stress-energy expression whose divergence corresponds to the force expression for the definitive gravitation tensor. The gravitation tensor produced in this fashion even happened to involve the gravitational stress-energy expression in such a way that the field equation could be written in the canonical form of eq. (XXXIX), with the energy-momentum of the gravitational field entering the field equations on the same footing as the energy-momentum tensor of matter (cf. eqs. (49) and (50)). This strategy was the result of Einstein's reflection on his earlier attempts to generalize the representation of the force as a divergence expression from the weak-field to the general case. Rather than guessing the right correction terms, he had now found a systematic construction procedure, which seemed to uniquely identify a candidate gravitation tensor compatible with both the correspondence and the conservation principle.

The match between these two heuristic principles was achieved at the expense of the generalized principle of relativity. All that could be known in that respect about a candidate gravitation tensor produced in this way was its covariance under linear transformations (cf. eq. (LXVII)). Einstein was ready to accept this consequence. He had already turned his attention to a simplified approach encompassing only linear transformations once before so that he could get a better handle on the conservation principle. This was when his earlier attempts to determine the covariance properties of the core operator with or without the help of the Beltrami invariants had run into similar difficulties as his efforts involving to the Riemann tensor. While, at that point, the restriction to linear transformations was merely a presupposition for formulating the conservation problem, Einstein may now have felt that this was the price to pay for its solution.

In the course of Einstein's pursuit of the mathematical strategy, the conservation principle had emerged as the major challenge for his search for the gravitational field equation. This challenge triggered the switch to a physical strategy, judiciously incorporating results found while pursuing the mathematical strategy, from the form of the field equation to the role of coordinate restrictions. With the establishment of the *Entwurf* field equations with the help of this physical strategy Einstein had succeeded, for the first time in the course of his research documented by the notebook, to satisfy the conservation principle without restriction to the weak-field level. The major challenge for Einstein's research now was the generalized principle of relativity. How far could the covariance properties of the *Entwurf* field equation be extended or was their covariance really restricted to linear transformations only? Why were the physically satisfying *Entwurf* field equations not generally-covariant to begin with? These were the questions delineating Einstein's research program for the further exploration of the *Entwurf* equation. Its compliance with the heuristic principles rooted in classical physics, the correspondence and the conservation principle, made it possible to consider these questions not as incentives for continuing the search for gravitational field equations but as remaining puzzles within an established conceptual framework, that of the *Entwurf* theory. For the time being, Einstein's search for the gravitational field equations was over—even if this meant turning his back on a reservoir of possible further candidates.

7. PROGRESS IN A LOOP: EINSTEIN'S GENERAL RELATIVITY AS A TRIUMPH OF THE "*ENTWURF*" THEORY IN THE PERIOD FROM 1913 TO 1915

7.1 Consolidation, Elaboration, and Reflection

This chapter focuses on what has traditionally been seen as the most uneventful period of Einstein's search for a generalized theory of relativity, the time between spring of 1913 and fall of 1915, in which he clung to the erroneous *Entwurf* theory, which he published together with Marcel Grossmann before the end of June 1913. According to the dramatic narratives of the emergence of general relativity, this period was one of stagnation, it was the calm interval between two major thunderstorms, Einstein's tragic struggle with and eventual rejection of generally-covariant field equations in the winter of 1912–1913 in favor of a theory with only limited covariance properties and the sudden revelation of errors in the *Entwurf* theory, which led immediately to its demise and then to a triumphant, if gradual, return to generally-covariant field equations in the fall of 1915. The assumption of a "pitfall" in 1912–1913 and of a "breakthrough" in late 1915 constitutes the traditional explanation for the most peculiar feature of the genesis of general relativity, Einstein's double discovery of generally-covariant gravitational field equations, first formulated in 1912 and then rediscovered in 1915. How else, if not by the introduction and later elimination of errors, can this closed loop be explained?[106]

From the perspective of an historical epistemology, the supposed period of stagnation between 1913 and 1915 can be considered a period in which new knowledge was assimilated to a conceptual structure still rooted in classical physics. As a result of this assimilation of knowledge, this conceptual structure became richer, both in terms of an ever more extended network of conclusions that it made possible, and in terms of new opportunities for ambiguities and internal conflicts within this network. It was this gradual process of enrichment that eventually created the preconditions for a reflection on the accumulated knowledge which, in turn, induced a reorganization of the original knowledge structure. The enrichment of a given conceptual structure by the assimilation of new knowledge and the subsequent reflective reorganization of the enriched structure are the two fundamental cognitive processes which explain the apparent paradox that the preconditions for the formulation of general relativity matured under the guidance of a theory that is actually incompatible with it.

As we will argue in the following, the results achieved on the basis of the *Entwurf* theory should not be understood as so many steps in the wrong direction, whereupon it appears that their only function was to make the deviation from the truth evident, but rather as instruments, first for accumulating knowledge and then for rearranging it in a new order. Both these processes are essential to the development of scientific knowledge. The second half of this chapter covers Einstein's papers of November 1915, with the intention to demonstrate the role in these papers of insights and techniques developed in the period before.

When elaborating the *Entwurf* Theory, Einstein still pursued the same heuristics that had shaped his search for a gravitational field equation in the winter of 1912–1913 as documented by the Zurich Notebook, although the heuristics were now governed by the perspective of consolidation rather than by that of exploration of alternatives. In particular, the unresolved tensions between Einstein's heuristic principles guided his attempts to consolidate the *Entwurf* theory. These attempts were characterized by two complementary approaches. Following a defensive approach, he attempted to justify the restricted covariance of the *Entwurf* field equations by arguments based on the knowledge of classical physics. Following a bold approach, he attempted to look for a generalization of the relativity principle even in the framework of the *Entwurf* theory. The outcome of these efforts was, as we shall see, that he eventually succeeded in both, the defensive and the bold approach.

In the first phase of the consolidation period of the *Entwurf* theory, lasting roughly from spring to summer 1913, Einstein formulated two problematic argu-

106 This section relies heavily on (Renn 2005c) and "Untying the Knot ..." (in vol. 2 of this series).

ments, an argument based on the consideration of energy-momentum conservation and the notorious hole argument, both justifying the restricted covariance properties of the theory. Although these arguments later turned out to be erroneous, they were nevertheless significant in bringing out decisive conceptual peculiarities of general relativity that distinguished it from classical theories, including the *Entwurf* theory. The further development, refutation or clarification of these arguments revealed the non-locality of energy-momentum conservation, as expressed by the nonexistence of a local energy-momentum tensor for the gravitational field, the impossibility of ascribing physical significance to single spacetime points independent of the metric, and the fundamental connection between conservation laws and symmetries of spacetime structure later explicated in the Noether theorems.

In the second phase of the consolidation of the *Entwurf* theory, roughly lasting from fall 1913 to the end of 1914, Einstein elaborated this theory, which he originally found along the physical strategy, from the point of view of the complementary mathematical strategy. Guided by this heuristic strategy, Einstein found a new derivation of the *Entwurf* field equations, which he completed by the fall of 1914. The second phase of the consolidation period of the *Entwurf* theory had, like the first phase, a paradoxical character. On the one hand, Einstein's findings stabilized the *Entwurf* theory, on the other hand they provided instruments for overcoming the objections that had earlier prevented him from accepting candidate gravitational field equations found along the lines of the mathematical strategy. In particular, the variational techniques explored in the context of the new derivation of the *Entwurf* theory made it possible for Einstein to solve one of the crucial problems associated with the candidates for a gravitational field equation that he had discarded in the winter of 1912–1913, the establishment of energy-momentum conservation. That he did not immediately draw this consequence was partly a matter of perspective since, from the point of view governing the consolidation period, there was no reason for reexamining the earlier candidates.

Having discovered flaws in the *Entwurf* theory and its derivation along a modified mathematical strategy, Einstein eventually abandoned the consolidation phase and subsequently returned to a new exploratory phase, searching once more for the correct gravitational field equation. In hindsight, he gave three reasons for his rejection of the *Entwurf* theory: It could not explain the perihelion shift of Mercury; it did not allow the interpretation of a rotating system as being equivalent to the state of rest, and hence did not satisfy his Machian expectations, and finally, the derivation of its field equations along a mathematical strategy involved an unjustified assumption. For a short time, the theory survived *all* of these problems. Even the last problem, the discovery of a flaw in the derivation of the field equations, did not lead to a refutation of the Entwurf theory but only to a successful attempt of repairing it on a technical level. But the discovery of this problem had nevertheless fargoing consequences on the level of Einstein's reflection on the results he had achieved. By its very nature this discovery had a double effect:

- It showed that the adaptation of the mathematical strategy to the *Entwurf* theory failed and forced Einstein to return to the arguments at the core of the physical strategy as the only possible justification of the *Entwurf* theory.
- It showed that the mathematical strategy adapted to the *Entwurf* theory did not single out this theory but rather opened up the possibility of examining other candidate field equations. And he *needed* new equations because of the problem of rotation.

Together with the other short-comings found earlier, the discovery of the error in the derivation of the field equations, after a period of reflection, caused Einstein to drop his attempts to consolidate the *Entwurf* theory and eventually brought him back to an exploratory phase.

The second part of this chapter deals with the short period of three weeks before Einstein presented the definitive field equations of general relativity to the Prussian Academy on 25 November 1915. This period began when Einstein started to check whether the *Entwurf* field equations are necessarily the only solution to his problem and thus returned to his 1912–1913 attempts to search for a solution by examining candidate field equations familiar from his pursuit of the mathematical strategy, the November tensor, the Ricci tensor, and the Einstein tensor. By focusing on the impact of Einstein's achievements under the reign of the *Entwurf* theory, it is possible to answer the question of why in 1915 he could accept field equations that he had rejected in 1912–1913. In a note Einstein submitted to the Prussian Academy on 4 November 1915, he proposed a new gravitation theory based on the November tensor, considered earlier in the Zurich Notebook. In contrast to the situation in 1912–1913, he was now able to demonstrate that the new theory complies with the conservation principle. Just as he had done in the *Entwurf* theory, Einstein continued to interpret the conservation principle as implying a restriction of the admissible coordinate systems which now, however, turned out to be much less restrictive than the condition he had earlier found on the basis of his examination of the weak field equation (cf. eq. (LXXXVI)). He thus reached a decoupling of the coordinate restrictions implied by the conservation and the correspondence principles, respectively. Reflecting on this decoupling, Einstein was now able, for the first time, to conceive the choice of coordinates required for implementing the correspondence principle as a coordinate condition in the modern sense.

In an addendum to the note published on 4 November, Einstein reinterpreted another already familiar candidate in a new context, the Ricci tensor. This new context was provided by a speculative electromagnetic theory of matter, probably stimulated by the contemporary work of David Hilbert on such a theory. Due to this new context, Einstein shifted the restriction on the choice of coordinates, which he had found for the theory based on the November tensor, to a restriction of the choice of a particular kind of matter acting as the source of gravitational fields. Einstein thus arrived at a generally-covariant theory based on the Ricci tensor, which he considered as being merely a reinterpretation of the theory based on the November tensor so that he could take over essential results such as those concerning energy-momentum conservation.

He thus partly resolved—on the basis of the *Entwurf* theory—and partly circumvented—on the basis of an electromagnetic theory of matter—the objections he had earlier encountered against such a theory when he first considered it in 1912–1913.

In a paper submitted on 18 November 1915 Einstein calculated the perihelion shift of Mercury, claiming to provide support for the hypothesis of an electromagnetic nature of matter on which his new theory of gravitation was based. In a sense, the Mercury problem now offered a theoretical laboratory for the Ricci Tensor. Einstein's paper is largely based on techniques he had developed jointly with Besso in 1913 in the context of the *Entwurf* theory. It also includes the crucial insight that the determination of the Newtonian limit for a gravitational field equation involves, in general, more complex considerations than originally envisioned along the physical strategy, and that were used earlier to object to the harmonically reduced Ricci tensor in the Zurich Notebook. This insight into the complex nature of the correspondence principle had already been attained in 1913 as well, in the context of the *Entwurf* theory (at least by Besso) but was then of no relevance as Einstein and Besso did their original calculations in the consolidation phase of this theory.

Einstein's more sophisticated understanding of the Newtonian limit had, in the context of the renewed exploratory phase at the end of 1915, decisive consequences: It made it possible for him, in his final paper of that period, to base a theory of gravitation on the Einstein tensor, whose harmonically reduced and linearized form had been rejected in 1912 because of its apparent incompatibility with the correspondence principle. Also the status of energy-momentum conservation changed in the new theory. The insight into its different status was a consequence and not a presupposition of the establishment of the definitive version of general relativity, which on Einstein's part was established entirely in the conceptual framework of the *Entwurf* theory. The *Entwurf* theory and general relativity were initially not separated by a conceptual gulf, but merely by technical insights on the one hand, and a change of perspective on the other. Remarkably, these were both the result of the same process, the elaboration of the *Entwurf* theory during the supposed period of stagnation. Since the technical achievements attained in this period could still have been, in principle, assimilated to the theory that had given rise to them, taken by themselves they would have induced only a linear progress, thus yielding an increasingly sophisticated and increasingly complex *Entwurf* theory. It thus becomes clear that in light of the new technical achievements of the consolidation phase of the *Entwurf* theory, Einstein's reflection on his earlier knowledge, including previously discarded candidate gravitation tensors, was the crucial process that made the establishment of general relativity the result of progress in a loop.

7.2 The First Phase of the Consolidation Period of the Entwurf Theory: The Defensive and the Bold Approach

With the formulation of the *Entwurf* field theory and its publication by Einstein and Grossmann in the spring of 1913, the search for the field equations, as documented by the Zurich Notebook, had manifestly come to an end. Einstein no longer examined different candidates by comparing them with his heuristic expectations. Instead, he used his growing mastery of the mathematical representation to develop the one most promising candidate he had found. He thus entered the consolidation period of the *Entwurf* theory. A consolidation of the *Entwurf* theory was necessary in view of the main problem left open by the *Entwurf* paper of 1913, the determination of the covariance properties of the field equations and thus of the extent to which the new theory realized the generalized principle of relativity. The covariance was more restricted than that of the candidate gravitation tensors derived from the generally-covariant Riemann tensor that had formed the points of departure of Einstein's mathematical strategy. It therefore made sense to address this problem by trying to explain and justify the restricted covariance of the *Entwurf* equations and to explore these covariance properties in the hope of generalizing the relativity principle as much as possible.

Einstein's probing of these two approaches came to a first conclusion in August 1913. All his bold efforts up to that point to identify by explicit calculations non-linear transformations under which the *Entwurf* field equations might be covariant had failed, including an attempt to show that the metric for Minkowski spacetime in rotating coordinates is a solution of these equations. At the same time, his defensive efforts had led to a first result. Not surprisingly, this result was based on the conservation principle which had earlier motivated a restriction of the generalized principle of relativity on several occasions in the Zurich Notebook. An interpretation of the expression for energy-momentum conservation in the *Entwurf* theory, following the model of classical and special-relativistic physics, was now taken by Einstein to indicate that the *Entwurf* theory is covariant only under linear transformations. Both results, the failure of his attempts to identify non-linear transformations and the conservation argument, as we shall call it, thus pointed in the same direction and encouraged Einstein to look for further arguments along the defensive.

By the end of August 1913, he found, quite possibly in discussion with Michele Besso,[107] another argument against general covariance, the so-called "hole argument," which is based on the assumption, again motivated by classical and special-relativistic physics, that points in spacetime can be identified by means of coordinate systems, independently from any physical processes. In a formulation by Besso, the argument merely seeks to express the non-uniqueness of the metric tensor in terms of two distinct sets of functions which solve the same set of differential equations with given boundary values. Einstein elaborated this argument to a construction of two distinct solutions for the metric tensor considered within one and the same coordinate system. This more sophisticated version of the hole argument makes use of the idea, in modern parlance, to drag values of the metric tensor from one spacetime point to the other and later raised the important question of which aspects of a generally-covariant theory are physically meaningful.

---

107 See "What Did Einstein Know ..." (in vol. 2 of this series).

By means of the hole argument, Einstein convinced himself that generally-covariant gravitational field equations, together with boundary values, do not determine uniquely the metric tensor representing the gravitational field. Having thus identified an apparently fundamental reason for rejecting general covariance, he interpreted his earlier argument from the conservation principle as providing the necessary specialization of the reference frames to be used within the theory. With these results, the *Entwurf* theory had come to a certain closure, ending the first phase of its consolidation period.

7.3 The Failure of the Generalized Principle of Relativity: A Conflict Between Formalism and Physical Intuition

Einstein's decision to settle for the non-generally-covariant field equations of the 1913 *Entwurf* paper was the consequence of his failure to find generally-covariant equations and not of a conviction or an insight that such equations could not exist. In the spring of 1913, he could not be sure that he had just failed to find generally-covariant equations that would fulfill his hopes for fully implementing the generalized principle of relativity. In the notebook, he had considered several candidates for generally or at least unimodularly covariant field equations and found them defective. But the fact that these candidates failed for *different* reasons must have made it difficult for Einstein to accept that generally-covariant field equations did not exist since these different reasons did not include a clear hint as to why a full implementation of the generalized relativity principle could not exist. Either the conflict with the realization of the Newtonian limit as required by the correspondence principle, or with the demonstration of energy-momentum conservation as required by the conservation principle, or both, led to the rejection of a promising candidate, but these conflicts by themselves did not provide any counter-argument against the possibility of the generalized principle of relativity. The failure to find generally-covariant field equations was, after all, merely a technical result, incompatible with the physical intuition incorporated in the generalized principle of relativity.[108] The conflict between formalism and physical understanding motivated Einstein's further elaboration of the *Entwurf* theory.

If Einstein, at the time of the notebook or of the publication of the *Entwurf* paper, had seen any reason to modify or restrict this principle, he might have done so explicitly in order to justify his failure to achieve its full implementation. In a letter to Ehrenfest from May 1913, in which he announced the forthcoming publication of the *Entwurf* paper, he asserts his firm belief in a generalized principle of relativity, but points out that he had been unable to realize this principle on the level of the theory's formalism:

---

108 For discussion of the problematic relation between the physical intuition incorporated in the generalized principle of relativity and general covariance, see (Janssen 2005).

> I slowly convinced myself that *privileged coordinate systems do not exist at all*. However, I succeeded only partly in arriving at this position also from a formal point of view.[109]

In the *Entwurf* paper itself, Einstein refers to the conflicts between the generalized relativity principle and his other heuristic principles in order to justify the new theory's lack of general covariance (Einstein and Grossmann 1913, 11). He emphasized, in particular, the difficulties he had found in realizing the correspondence principle, suggesting a second-order field equation, as a justification for his failure to achieve a generally-covariant field equation. He also admitted that his introduction of the *Entwurf* field equations was merely based on plausible assumptions and not on a strict derivation from postulates such as a generalized principle of relativity.

The failure to find generally-covariant field equations was most evident from an intrinsic asymmetry of the *Entwurf* theory, between the non-generally-covariant field equation and the generally-covariant equation for material processes in a gravitational field, i.e., the equation for energy-momentum conservation. This asymmetry is also emphasized in the *Entwurf* paper itself. Einstein attempted to interpret it as a clue for justifying the failure to establish a generally-covariant field equation, pointing at the different ways in which the metric tensor enters into the equation for energy-momentum conservation, on the one hand, and the field equation, on the other:

> This exceptional position of the gravitational equations in this respect, as compared with all of the other systems, has to do, in my opinion, with the fact that only the former can contain second derivatives of the components of the fundamental tensor.[110]

In a sense, he simply turned the description of the problem into its solution at this point. Einstein's problem remained that this assertion was merely speculative and, in the final account, based on nothing but his failure to find appropriate, generally-covariant, second-order gravitational field equations.

7.4 The Failure of Einstein's Search for Non-Linear Transformations

While in the spring of 1913 Einstein made his first attempts at justifying the lack of general covariance of the *Entwurf* field equation, he tried, at the same time, to overcome this problem. In fact, it could not be excluded that, even though all derivations in the *Entwurf* theory merely involve the assumption of linear covariance, the field equations would turn out to be covariant under a wider class of transformations. In the *Entwurf* paper, this question is singled out as the most important one left to be resolved (Einstein and Grossmann 1913, 18).

---

109 "Die Überzeugung, zu der ich mich langsam durchgerungen habe, ist die, *dass es bevorzugte Koordinatensysteme überhaupt nicht gibt*. Doch ist es mir nur te[i]lweise gelungen, auch formal bis zu diesem Standpunkt vorzudringen." Einstein to Paul Ehrenfest, 28 May 1913, (CPAE 5, Doc. 441).

110 "Die diesbezügliche Ausnahmestellung der Gravitationsgleichungen gegenüber allen anderen Systemen hängt nach meiner Meinung damit zusammen, daß nur erstere zweite Ableitungen der Komponenten des Fundamentaltensors enthalten dürften." (Einstein and Grossmann 1913, 18)

There are indications that Einstein attempted to find out by calculation whether the *Entwurf* field equation transforms also under a wider class of transformations. From his Zurich Notebook, he was familiar with techniques for exploring the transformational behavior of field equations. But the *Entwurf* field operator confronted him with a case that was far more complex than any other candidate in the notebook for which he had attempted to determine the transformational behavior by directly subjecting it to coordinate transformations.

One indication for Einstein's explorative attempts and their failure comes from a couple of pages of the so-called Einstein-Besso manuscript, pages that were probably written around June 1913 and that deal specifically with the problem of rotation.[111] Another indication comes from a letter Einstein wrote on 14 August 1913 to Hendrik A. Lorentz. This letter marks the preliminary end of Einstein's search for non-linear transformations of the *Entwurf* field equation and provides a succinct resume of the situation of the *Entwurf* theory just on the verge of the renouncement of the bold approach. The letter begins with the confession that the lack of general covariance represents a profound dilemma for the new theory of gravitation:

> And now to gravitation. I am delighted that you so warmly espouse our investigation. But, unfortunately, there are still such major snags in the thing that my confidence in the admissibility of the theory is still shaky. So far the "Entwurf" is satisfactory insofar as it concerns the effect of the gravitational field on other physical processes. For the absolute differential calculus permits the setting up of equations here that are covariant with respect to arbitrary substitutions. The gravitational field ($g_{\mu\nu}$) seems to be the skeleton, so to speak, on which everything hangs. *But, unfortunately, the gravitation equations themselves do not possess the property of general covariance*. Only their covariance with respect to *linear* transformations is certain. But all of our confidence in the theory rests on the conviction that an acceleration of the reference system is equivalent to a gravitational field. Hence, if not all of the equation systems of the theory, and thus also equations (18), permit other than linear transformations, then the theory refutes its own starting point; then it has no foundation whatsoever.[112]

The letter to Lorentz continues with a description of Einstein's unsuccessful attempts to find non-linear transformations under which the *Entwurf* field equation remains covariant, discussing two types of transformations, autonomous and non-autonomous ones.[113] In the Zurich Notebook, he had attempted on several occasions to find the transformational properties of a physically plausible candidate by deriving differential equations for the transformation coefficients involving the metric tensor. But he had never found a simple solution to the problem posed in this way. In view of the many reasons in favor of the *Entwurf* theory, he must have applied this technique to it with even greater persistence. Einstein's letter to Lorentz shows, however, that these efforts remained as unsuccessful as they had been in the Zurich Notebook. He was, in fact, ready to give up the bold approach to solve the most fundamental problem of the *Entwurf* theory and turn to the defensive approach, searching for more substantial

111 See (CPAE 4, Doc 14 [pp. 41–42]) and the discussion in (Janssen 1999).

arguments to justify the lack of covariance of the field equations than those adduced in the *Entwurf* paper.

7.5 Einstein's Reinterpretation of the Conservation Principle

For the candidates which Einstein had encountered pursuing the mathematical strategy, the conservation principle typically implied a restriction of the generalized principle of relativity. Given this experience, it must have been plausible for him to examine whether an explanation for the restricted covariance of the *Entwurf* field theory could perhaps also be found in the context of energy-momentum conservation. Only a day after Einstein sent the letter to Lorentz quoted above, on 15 August 1913, he indeed found a way to justify the limited covariance of the *Entwurf* theory on the basis of the conservation principle. A crucial heuristic ingredient of his argument was the parallelism between gravitational energy and other forms of energy, represented in the *Entwurf* theory by eq. (51). In another letter to Lorentz, written on 16 August 1913, Einstein wrote:

> Furthermore, yesterday I found out to my greatest delight that the doubts regarding the gravitation theory, which I expressed in my last letter as well as in the paper, are not appropriate. The solution of the matter seems to me to be as follows: The expression for the energy principle for matter & gravitational field taken together is an equation of the form (19), i.e., of the form $\sum \frac{\partial F_{\mu\nu}}{\partial x_{\nu}} = 0$; starting out from this assumption, I set up

112 "Nun zur Gravitation. Ich bin beglückt darüber, dass Sie mit solcher Wärme sich unserer Untersuchung annehmen. Aber leider hat die Sache doch noch so grosse Haken, dass mein Vertrauen in die Zulässigkeit der Theorie noch ein schwankendes ist. Befriedigend ist der Entwurf bis jetzt, soweit es sich um die Einwirkung des Gravitationsfeldes auf andere physikalische[e] Vorgänge handelt. Denn der absolute Differenzialkalkül erlaubt hier die Aufstellung von Gleichungen, die beliebigen Substitutionen gegenüber kovariant sind. Das Gravitationsfeld ($g_{\mu\nu}$) erscheint sozusagen als das Gerippe an dem alles hängt. *Aber die Gravitationsgleichungen selbst haben die Eigenschaft der allgemeinen Kovarianz leider nicht*. Nur deren Kovarianz *linearen* Transformationen gegenüber ist gesichert. Nun beruht aber das ganze Vertrauen auf die Theorie auf der Überzeugung, dass Beschleunigung des Bezugssystems einem Schwerefeld äquivalent sei. Wenn also nicht alle Gleichungssysteme der Theorie, also auch Gleichungen (18) [i.e. the gravitational field equations] ausser den linearen noch andere Transformationen zulassen, so widerlegt die Theorie ihren eigenen Ausgangspunkt; sie steht dann in der Luft." Einstein to Hendrik A. Lorentz, 14 August 1913, (CPAE 5, Doc. 467). In view of the later development, in which the question of whether the manifold with its coordinate systems or the metric tensor is the "skeleton" on which all physical processes depend acquired a certain significance, it is remarkable that in the above formulation Einstein singled out the metric tensor as the crucial object. As we will see below when discussing the hole argument, in defending the lack of general covariance of his field equations Einstein for a while assumed that the points of the manifold identified by certain sets of coordinates actually have a reality and physical significance by themselves, that is, also independently from the metric tensor.

113 This fact also suggests that Einstein at this point did not assume that the *Entwurf* field equations remain covariant under rotations.

> equations (18). But a consideration of the general differential operators of the absolute differential calculus shows that an equation so constructed is never absolutely covariant. Thus, by postulating the existence of such an equation, we tacitly specialized the choice of the reference system. We restricted ourselves to the use of such reference systems with respect to which the law of momentum and energy conservation holds in this form. It turns out that if one privileges such reference systems, then only more general linear transformations remain as the only ones that are justified.[114]

With this insight, the conservation principle was no longer merely a technical impediment to the full implication of the generalized relativity principle but provided the concrete physical reason for the restriction to specific, well-defined transformations. Accordingly Einstein continued in his letter to Lorentz:

> Thus, in a word; *By postulating the conservation law, one arrives at a highly determined choice of the reference system and the admissible substitutions*. Only now, after this ugly dark spot seems to have been eliminated, does the theory give me pleasure.[115]

Einstein's argument presupposes, however, that the objects appearing in his equation for energy-momentum conservation eq. (51) behave themselves as tensors. This is true for the stress-energy tensor of matter, but not for the stress-energy expression for the gravitational field, as Einstein came to realize a few months later.[116] But if this quantity fails to behave as a tensor, the transformational properties of eq. (51) cannot be read off by inspection as Einstein claimed to be able to do in his letter to Lorentz.

Einstein immediately incorporated the argument for linear covariance found on 15 August 1913 and exposed to Lorentz a day later into his manuscript for a lecture he was invited to hold on 23 September 1913 in Vienna.[117] In § 6 of this lecture, enti-

114 "Ferner fand ich gestern zu meiner grossen Freude, dass die gegenüber der Gravitationstheorie in meinem letzten Briefe, sowie in der Arbeit geäusserten Bedenken nicht angezeigt sind. Die Sache scheint sich mir folgendermassen zu lösen. Ausdruck des Energieprinzips für Materie & Gravitationsfeld zusammen ist eine Gleichung von der Form (19) d.h. von der Form $\sum \frac{\partial F_{\mu\nu}}{\partial x_\nu} = 0$; von dieser Voraussetzung ausgehend stellte ich die Gleichungen (18) [i.e. the field equations] auf. Nun zeigt aber eine Betrachtung der allgemeinen Differenzialoperatoren des absoluten Differenzialkalküls, dass eine so gebaute Gleichung niemals absolut kovariant ist. Indem wir also die Existenz einer solchen Gleichung postulierten, spezialisierten wir stillschweigend die Wahl des Bezugssystems. Wir beschränkten uns auf den Gebrauch solcher Bezugssysteme, inbezug auf welche der Erhaltungssatz des Impulses und der Energie in dieser Form gilt. Es zeigt si[ch], dass bei der Bevorzugung solcher Bezugssysteme nur mehr allgemeine lineare Transformationen als allein berechtigt übrig bleiben." Einstein to Hendrik A. Lorentz, 16 August 1913, (CPAE 5, Doc. 470).

115 "Also kurz gesagt: *Durch die Postulierung des Erhaltungssatzes gelangt man zu einer in hohem Masse bestimmten Wahl des Bezugssystems und der zuzulassenden Substitutionen*. Erst jetzt macht mir die Theorie Vergnügen, nachdem dieser hässliche dunkle Fleck beseitigt zu sein scheint."

116 Einstein found the fallacy of this argument in the period between ca. 20 January 1914 and ca. 10 March 1914, see Einstein to Heinrich Zangger, ca. 20 January 1914, (CPAE 5, Doc. 507) and Einstein to Paul Ehrenfest, before 10 March 1914, (CPAE 5, Doc. 512), as well as Einstein to Heinrich Zangger, ca. 10 March 1914, (CPAE 5, Doc. 513). The last two letters mention, for the first time after a long period of intermission (to which the first letter to Zangger evidently still belongs), progress in the work on the gravitation problem related to the covariance properties of the *Entwurf* equation.

tled “Bemerkungen über die mathematische Methode,” Einstein summarized some of the mathematical properties of the *Entwurf* theory. In the previous paragraph he had dealt with equations describing the effect of the gravitational field on other physical processes, and in particular with energy-momentum conservation, in the next he was going to develop the gravitational field equation. Towards the end of § 6, he took the occasion to comment on the remarkable asymmetry between the transformational behavior of these two types of equations (Einstein 1913, 1257). He explained that the arguments in favor of a generalized principle of relativity, on the one hand, and his argument in favor of its restriction, on the other, are located on different levels, one on the general level of the spacetime structure, the other on the level of concrete physical requirements. Whereas the first level remains the relevant one for all equations dealing with the effects of the gravitational field on other physical processes, the second level becomes relevant only for the gravitational field equation itself. The introduction of a specialization of the reference system only “after the fact” may well have motivated him to continue his search for an explanation that makes the restriction of the generalized relativity principle understandable also on the level of the spacetime structure. Furthermore, the asymmetry between the transformational behavior of the gravitational field equations and that of all other equations of physics hinted at an explanation that is related to the mathematical nature of the field equations. As we shall now see, Einstein eventually found such an explanation in the hole argument.

7.6 The Construction of the Hole Argument[118]

The use of the conservation principle to justify the limited covariance of the *Entwurf* field equation was a relief but contributed, in effect, little to *understanding* the restriction of the generalized principle of relativity. In a first attempt to come to terms with this unsatisfactory situation, Einstein reformulated the Machian heuristics that had originally suggested the introduction of this principle. This attempt is visible already in his earliest statements concerning the conservation argument described above, but is most clearly expressed in a letter Einstein wrote to Ernst Mach in the second half of December 1913:

> It seems to me absurd to ascribe physical properties to “space.” The totality of masses produces the $g_{\mu\nu}$ field (gravitational field), which in turn governs the course of all processes, including the propagation of light rays and the behavior of measuring rods and clocks. First of all, everything that happens is referred to four completely arbitrary spacetime variables. If the principles of momentum and energy conservation are to be satisfied, these variables must then be specialized in such a way that only (completely) linear substitutions shall lead from one justified reference system to another. The reference sys-

117 Einstein completed the manuscript for this lecture about a month earlier, see (CPAE 4, note 17) and “What Did Einstein Know ...” note 19 (in vol. 2 of this series).

118 See also “What Did Einstein Know ...” (in vol. 2 of this series) and further references cited there.

> tem is, so to speak, tailored to the existing world with the help of the energy principle, and loses its nebulous aprioristic existence.[119]

In this reformulated argument against absolute space, the role of the cosmic masses and their relations in constituting space is now taken over by the conservation principle which is claimed to provide the physical justification of preferred reference frames. But Einstein's revised Machian heuristics nevertheless failed to completely ban the plausibility of a generalization of the relativity principle. This is evident from the efforts by Einstein and his friend Michele Besso to find a deeper connection between the limited transformation properties of the *Entwurf* field equation and the structure of spacetime.

These efforts are documented, in particular, by a manuscript in the hand of Michele Besso.[120] In a group of pages, the first of which carries the dateline 28 August 1913, Besso listed a number of problems which he had probably encountered while working jointly with Einstein on the problem of Mercury's perihelion shift in the context of the *Entwurf* theory.[121] These problems were evidently not intended as a program for the further development of the *Entwurf* theory but rather constituted the results of reflections on what had been achieved so far. In a note found in the part of the manuscript belonging to a later period, Besso modestly characterized himself as being merely an "*orrechiante*," that is, as an amateur who had the privilege of listening to a great master. But some of Besso's observations turned out to be most consequential, introducing a possibly naive but fresh perspective. It seems in fact that he went together with Einstein through his list of problems, some of them directly formulated as questions, and that they discussed them one by one; at some later point Besso then entered Einstein's responses.[122] Three of Besso's problems are relevant for the present discussion, the first concerning the issue of rotation, the second regarding the logical status of the restriction implied by the conservation principle, and the third concerning what later was to become the hole argument, Einstein's central argument to defend the restricted covariance of the *Entwurf* theory.

---

119 "Für mich ist es absurd, dem "Raum" physikalische Eigenschaften zuzuschreiben. Die Gesamtheit der Massen erzeugt ein $g_{\mu\nu}$-Feld (Gravitationsfeld), das seinerseits den Ablauf aller Vorgänge, auch die Ausbreitung der Lichtstrahlen und das Verhalten der Massstäbe und Uhren regiert. Das Geschehen wird zunächst auf vier *ganz willkürliche* raum-zeitliche Variable bezogen. Diese müssen dann, wenn den Erhaltungssätzen des Impulses und der Energie Genüge geleistet werden soll, derart spezialisiert werden, dass nur (ganz) *lineare* Substitutionen von einem berechtigten Bezugssystem zu einem anderen führen. Das Bezugssystem ist der bestehenden Welt mit Hilfe des Energiesatzes sozusagen angemessen und verliert seine nebulose apriorische Existenz." Einstein to Ernst Mach, second half of December 1913, (CPAE 5, Doc. 495).

120 For more detailed discussion of both the contents and the dating of this document, see "What Did Einstein Know ..." (in vol. 2 of this series).

121 For a facsimile reproduction of these pages, see (Renn 2005a, 126–130).

122 This pattern is made evident also by the arrangement of problems and answers on these pages, the answers being often written in a slightly different hand than the questions and squeezed in between the lines or at the margin.

In his notes on the first problem, Besso summarized the failure of Einstein's hopes to fully implement his Machian heuristics within the *Entwurf* theory. Besso noted that it was impossible to conceive of rotation as equivalent to a state of rest in a gravitational field that is a solution of the *Entwurf* field equations.[123] This represented a challenge either for the *Entwurf* theory or the generalized principle of relativity. In the consolidation period of the *Entwurf* theory the decision was made in favor of the theory. That was bound to change only in the context of Einstein's renewed exploratory phase in the fall of 1915. In his notes Besso studied the question of whether the failure of the *Entwurf* theory to interpret rotation as a state of rest can possibly be explained by a failure of the conservation principle in a rotating system. If that were so, he would have succeeded in establishing the desired physical connection between the problem of implementing the generalized relativity principle and the restriction of the admissible coordinate systems required by the conservation principle. The remainder of the text, probably going back to Einstein's intervention, shows that this attempt of explaining the problem with rotation in terms of the conservation principle does not work.

In a second passage, Besso posed a more general question concerning the role of energy-momentum conservation for the selection of admissible coordinate systems:

> Is every system that satisfies the conservation laws a justified system?[124]

If the conservation principle is really the explanation for the restriction on the choice of coordinate systems, it should not only be a necessary but also a sufficient condition for this choice. Besso therefore wondered whether *any* coordinate system satisfying the constraints imposed by the conservation laws is compatible also with the covariance of the field equations. Since there is no note which may be traced back to a reaction by Einstein on this issue, it seems that he did, at first, not seriously consider Besso's suggestion. In fact, given Einstein's belief at that point that the conservation equation (51) is covariant under linear transformations only, the question may have held little interest for him since it offered no promise of generalizing the covariance properties of the *Entwurf* field equations beyond linearity. Eventually, however, when Einstein turned to the exploration of a mathematical strategy for the *Entwurf* theory, he did realize, as we shall see, the significance of Besso's question.

After writing down his second question, Besso sketched an idea of how the failure of realizing general covariance on the level of the gravitational field equation might be explained, namely as a problem of the uniqueness of its solutions. The text of his third problem reads:

> The requirement of [general] covariance of the gravitational equations under arbitrary transformations cannot be imposed: if all matter [is given] were contained in one part of space and for this part of space a coordinate system [is given], then outside of it the coordinate system could still [essentially] except for boundary conditions be chosen arbi-

123 See "What Did Einstein Know ..." sec. 3 (in vol. 2 of this series).
124 "Ist jedes System, welches den Erhaltungssätzen genügt, ein berechtigtes System?"

> trarily, [through which the g arbitrarily] so that a unique determinability of the g's cannot be obtained.
>
> It is, however, not necessary that the g themselves are determined uniquely, only the observable phenomena in the gravitation space, e.g., the motion of a material point, must be.[125]

In this passage Besso imagines a central mass to be surrounded by empty space and wonders whether the solution for the metric tensor is, in this case, determined uniquely for the empty region. His mental model appears to be the inverse of Einstein's famous hole argument where matter may be anywhere outside an empty hole for which the problem of the ambiguity of solutions then supposedly arises. In Besso's argument the ambiguity of solutions is conceived as being due to the arbitrary choice of the coordinate system in the empty region, giving rise to arbitrary coordinate expressions for the metric tensor (which have to satisfy, however, the boundary conditions). Apart from the inversion of hole and matter, Besso's "proto-hole argument" thus corresponds to the primitive version of the hole argument that was traditionally ascribed to Einstein, charging him with the naivety of being unaware that different coordinate representations of the metric tensor do not correspond to different solutions of the field equations. It is therefore remarkable that even Besso immediately realized the flaw of this naive version since he added, in the second paragraph of the above text, that only observable phenomena, such as the motion of a particle, should be determined uniquely.

How did Besso's idea emerge and how was it transformed into the hole-argument familiar from Einstein's later publications? As to the first question, it seems plausible that Besso related, in the context of his reflections on the *Entwurf* theory, the problem of the restriction of general covariance to other problems that had arisen for this theory, in particular in the course of his joint research with Einstein. One such problem was rotation, as we have just seen. Another problem was the perihelion shift of Mercury, the central subject of a paper Besso planned to write.

In 1913 Besso had in fact encountered the problem of uniqueness when he worked on the perihelion problem in the context of the *Entwurf* theory. In fact, a note in the Einstein-Besso manuscript explicitly refers to the question of uniqueness in connection with the ansatz used for solving the *Entwurf* field equation by an approximation procedure. For the first step of that iterative procedure Einstein and Besso had used a metric with only one variable component, the same spatially flat metric (25)

---

125 "Die Anforderung der [allgemeinen] Covarianz der Gravitationsgleichungen für beliebige Transformationen kann nicht aufgestellt werden: wenn in einem Teile des Raumes alle Materie [gegeben ist] enthalten wäre und für diesen Teil ein Coordinatensystem, so könnte doch ausserhalb desselben das Coordinatensystem noch, [im wesentlichen] abgesehen von den Grenzbedingungen, beliebig gewählt werden, [wodurch die g beliebig eine] so dass eine eindeutige Bestimmbarkeit der g s nicht eintreten könne.
Es ist nun allerdings nicht nötig, dass die g selbst eindeutig bestimmt sind, sondern nur die im Gravitationsraum beobachtbaren Erscheinungen, z.B. die Bewegung des materiellen Punktes, müssen es sein."

that is also crucial for obtaining the Newtonian limit of the *Entwurf* theory. Besso then wondered whether that choice of a metric was sufficiently general for recovering all possible solutions of the field equation:

> Is the static gravitational field in § 1 $g_{\mu\nu} = 1$, 1 to 3 $g_{44} = f(x, y, z)$ a particular solution? Or is it the general solution expressed in particular coordinates?[126]

It may have been this question arising in the context of the perihelion calculation that suggested to Besso that covariant field equations suffer, in general, from a problem of uniqueness. In fact, the physical model of Besso's proto-hole argument is strikingly similar to that of the perihelion problem, a central mass surrounded by empty space. And when Besso reminded himself that it was not the expression for the metric tensor that mattered but physically observable phenomena, he chose the motion of a material particle, such as that of Mercury around the sun, as an example.

How did Einstein react to Besso's consideration of the proto-hole argument and how did the definitive version of the hole argument emerge? Einstein's reaction is, it seems, preserved in a text written below Besso's note quoted earlier. It starts, just as Einstein's earlier remark concerning rotation and energy conservation, with a characteristic "Of no use" ("Nützt nichts"):

> Of no use, since with [the] a solution a motion is also fully given. If in coordinate system 1, there is a solution K_1, then this same construct is also a solution in 2, K_2; K_2, however, also a solution in 1.[127]

Remarkably, Einstein did not simply agree with Besso's conclusion that the ambiguity of the coordinate representation of the metric tensor was of no physical consequence. He apparently found Besso's idea to justify the lack of general covariance of the *Entwurf* field equations on the basis of a uniqueness argument intriguing and effectively reinterpreted it as an argument about the nature of space and time, and, in particular, about the role of coordinate systems in identifying points in space and time. In fact, a solution of the field equation in a particular coordinate system, expressed in terms of functions representing the components of the metric tensor, can be transformed to another coordinate system, producing a different set of functions representing the same solution. But if, as Einstein's argument suggests, this set of functions "this same construct" ("dieses selbe Gebilde") can somehow be related to the original coordinate system, it there represents a different metric which, however, solves the same field equation, provided that the right-hand side of these equations remains unchanged by the coordinate transformation, which is the case for an empty region where the stress-energy tensor of matter vanishes. To avoid the issue of additional boundary conditions, it turned out to be convenient for Einstein to reverse the

---

126 "Ist das stat Schwerefeld des § 1 $g_{\mu\nu} = 1$, 1 bis 3, $g_{44} = f(x, y, z)$ ein spezielles? oder ist es das allgemeine, auf spec. Coordinaten zurückgeführtes" (CPAE 4, Doc. 14, [p. 16]).

127 "Nützt nichts, denn durch eine Lösung ist auch eine Bewegung voll gegeben. Ist im Koordinatensystem 1 eine Lösung K_1, so ist dieses selbe Gebilde auch eine Lösung in 2, K_2; K_2 aber eine Lösung in 1." For a facsimile of this passage, see Fig. 2 on p. 300 of "What Did Einstein Know ..." (in vol. 2 of this series) and (Renn 2005a, 128).

physical model proposed by Besso and consider, instead of a void with a lump of mass, matter with a hole in it—the well-known configuration of the hole argument.[128]

Einstein's interpretation of a solution initially given in one coordinate system as referring to another coordinate system implicitly presupposes that coordinate systems have their own physical reality and allow to identify points in spacetime. The crucial but hidden point of this reinterpretation of Besso's proto-hole argument is therefore a reification of coordinate systems, which are conceived as part of the physical set-up constituting a solution and not only as a mathematical device for describing it. Only the later refutation of the hole argument made it eventually clear that it is not "motions" in the sense used here which constitute physically real events but rather spacetime coincidences for which a coordinate-independent description can be given.[129]

Our reconstruction suggests that the hole argument was, in spite of its philosophical appeal, not rooted in a metaphysical prejudice concerning the nature of space and time or the role of coordinate systems, preventing Einstein from accepting generally-covariant field equations. On the contrary, it was the necessity of justifying a non-generally-covariant field equation that led to the construction of this argument, triggering a peculiar interpretation of the physical significance of coordinate systems, an interpretation moreover that largely remained implicit in the initial formulation of the argument. The hole argument was just the kind of argument Einstein had been after in his earlier attempts to justify the failure of general covariance: a mathematical argument related to the structure of space and time. It was this peculiar perspective, shaped by the context of the consolidation period of the *Entwurf* theory, that probably led him to take Besso's naive point seriously and search for a physically significant interpretation of a mathematically trivial property, the coordinate dependence of expressions for the metric tensor. It is hardly surprising that to formulate such an interpretation, Einstein relied on the conceptual resources of classical physics, implicitly defining what a motion is in terms of the relation between a particle and a coordinate system. As a result, he found a way of relating the formalism of absolute differential calculus to a physical interpretation of coordinate systems that allowed him to justify the restricted covariance of the *Entwurf* field equations. In short, the necessity of interpreting a complex mathematical formalism under a peculiar perspective was crucial for the emergence of the hole argument. Only when Einstein eventually succeeded in formulating physically acceptable, generally-covariant field equations did he abandon this argument and revise the physical interpretation of coordinate systems as well as of space and time associated with it.[130] The deep conceptual insight into the crucial role of spacetime coincidences was thus no presuppo-

128 For detailed discussion, see sec. 4 of "What Did Einstein Know ..." (in vol. 2 of this series).

129 For selected references to the extensive literature on the hole argument, see "What Did Einstein Know ..." note 95 (in vol. 2 of this series).

130 For discussion of Einstein's later retraction of the hole argument, see "What Did Einstein Know ..." sec. 4 (in vol. 2 of this series) and (Janssen 2005, 73–74)

sition of general relativity but merely a consequence of its establishment—effectively implying a refutation of the hole argument.

With the advent of the hole argument, the conservation principle lost its role as the primary physical reason for the restriction of general covariance of the *Entwurf* field equations; the restriction to linear transformations now merely appeared as a concrete result in harmony with a more general insight. Einstein reinterpreted his first argument defending the restricted covariance of the *Entwurf* field equation accordingly as a specific physical complement to what he saw as a general "logical" principle.[131] That he considered the hole argument not merely as an addition but as the solution of a puzzle left unresolved by the earlier physical argument is confirmed by a letter he wrote in the beginning of November 1913 to Ehrenfest:

> The gravitation affair has been clarified to my *complete satisfaction* (namely the circumstance that the equations of the gr. field are covariant only with respect to *linear* transformations. For it can be proved that *generally covariant* equations that determine the field *completely* from the matter tensor cannot exist at all. Can there be anything more beautiful than this, that the necessary specialization follows from the conservation laws?[132]

7.7 The Second Phase of the Consolidation Period of the Entwurf Theory: A Mathematical Strategy for the Entwurf Theory

The insights Einstein had acquired pursuing the mathematical strategy in the Zurich Notebook continued to set standards for his further elaboration of the *Entwurf* theory developed along the lines of the physical strategy. In particular, the procedure at the core of the mathematical strategy by which non-generally-covariant field equations could be extracted from a generally-covariant object remained plausible. Even if generally-covariant field equations were excluded for a satisfactory relativistic theory of gravitation, the physical and the mathematical strategies should converge because only in this way was it possible to fully clarify and stabilize the relation between the physical and the mathematical knowledge expressed in the theory. Even if Einstein had good reasons for restricting the generalized principle of relativity, it still made sense for him to search for a derivation of the field equations of the *Entwurf* theory along the mathematical strategy, albeit now with the aim of confirming what had

---

131 In a later paper he formulated with regard to these two arguments: "But there are two weighty arguments that justify this step [i.e. the restriction of general covariance], one of them of logical, the other one of empirical provenance" ("Es gibt aber zwei gewichtige Argumente, welche diesen Schritt rechtfertigen, von denen das eine logischen, das andere empirischen Ursprungs ist", CPAE 4, Doc. 25, [178]).

132 "Die Gravitationsaffäre hat sich zu meiner *vollen Befriedigung* aufgeklärt (der Umstand nämlich, dass die Gleichungen des Gr. Feldes nur *linearen* Transformationen gegenüber kovariant sind. Es lässt sich nämlich beweisen, dass *allgemein kovariante* Gleichungen, die das Feld aus dem materiellen Tensor *vollständig* bestimmen, überhaupt nicht existieren können. Was kann es schöneres geben, als dies, dass jene nötige Spezialisierung aus den Erhaltungssätzen fliesst?" Einstein to Paul Ehrenfest, before 7 November 1913, (CPAE 5, Doc. 481).

already been found through the physical strategy. This approach is characteristic for what we call the "second phase" of the consolidation of the *Entwurf* theory.

It was probably in pursuing a mathematical strategy for the *Entwurf* theory that, in early 1914, Einstein discovered a flaw in his conservation argument for linear covariance. It turned out that the quantity representing the stress-energy of the gravitational field is not a tensor. Einstein thus realized that, while the conservation principle still requires a restriction on the admissible coordinate systems, this restriction was not as stringent as it had seemed before. The issue of the covariance properties of the *Entwurf* theory was therefore reopened since the hole argument only excluded general covariance but did not by itself prescribe a specific covariance group. It remained to be clarified, in particular, in which sense the transformational properties of the *Entwurf* field equation were restricted by the conservation principle, which now appeared as implying a "weak" restriction only taking full advantage perhaps of the leeway left by the hole argument.

A suggestion by the Zurich mathematician Paul Bernays made it possible for Einstein and Grossmann to return in early 1914 to the "bold" approach, once again exploring the transformational properties of the *Entwurf* field equation by direct calculation. Bernays suggested to derive the *Entwurf* field equation from a variational principle in order to be able to focus attention on a single scalar quantity, the Lagrangian, rather than on the complex tensorial objects constituting the field equation itself. Einstein and Grossmann succeeded indeed in finding a Lagrangian from which the *Entwurf* field equations could be derived. They found that this Lagrangian is invariant under transformations between coordinate systems specified solely by the requirement of energy-momentum conservation. The necessary restriction of covariance following from the conservation principle thus turned out to be also a sufficient one, just as Michele Besso had envisaged. This result seemed to be in perfect agreement also with the hole argument since the four additional equations resulting from the conservation principle were apparently just enough to remove the ambiguity in the metric field on which this argument turns.

With these results, attained by March 1914, the "defensive" and the "bold" approaches had converged and, once again, a sense of closure in the development of the *Entwurf* theory was reached, this time on a higher level than half a year earlier and more durable: it would last until October 1915. Einstein was convinced that he had obtained an optimal realization of the generalized relativity principle and that he had understood the profound reasons for the impossibility of general covariance. He even came to believe that the restriction of covariance imposed by the conservation principle in fact does not imply a restriction of the possible solutions to the field equation but merely a restriction of the possible coordinate systems in which these solutions can be expressed. Consequently, Einstein also became convinced that the *Entwurf* theory fully realized the equivalence principle and other heuristic ideas, such as a description of Minkowski spacetime in a rotating frame of reference as a special case of the gravitational field, in spite of the difficulties at the level of explicit calculations.[133]

The success of this exploration of the *Entwurf* theory along the lines of the mathematical strategy encouraged Einstein to undertake a new derivation of its field equations; he completed this derivation by the fall of 1914. By February 1914, he had abandoned his earlier conviction that the *Entwurf* field equation had no relation to the absolute differential calculus essential to the mathematical strategy. When Einstein took up the project of deriving the *Entwurf* field equation along the mathematical strategy, however, he did not start from the generally-covariant objects suggested by the original mathematical strategy as he had done in the Zurich Notebook. He rather generalized the variational derivation, developed together with Grossmann, into a mathematical formalism applicable not only to the *Entwurf* field equations but to other candidate field equations. He then searched for a mathematical reason to justify choosing the Lagrangian corresponding to the *Entwurf* equation and erroneously convinced himself that he had actually found such reasons.

7.8 A Prelude: The First Reawakening of the Mathematical Strategy

Einstein's failure to reach general covariance had been a target of criticism by his colleagues.[134] In January 1914 he wrote a paper in reply to such criticism (Einstein 1914b). Apart from presenting his arguments in favor of a restricted covariance of the *Entwurf* field equations such as the hole argument, he had to admit that the relation of this field equations to the generally-covariant objects of the absolute differential calculus was still an open problem. As a consequence, the relation between physical and mathematical strategies, which should have been just two different pathways to the same result, also remained unclear.

Defending the restricted covariance of the *Entwurf* theory, Einstein had to acknowledge that there are profound reasons why generally-covariant equations should exist which correspond to the *Entwurf* field equation (Einstein 1914b, 177–178). He argued that there must be, in modern terms, a coordinate-free representation of any meaningful mathematical relation between physical magnitudes. Ideally, the *Entwurf* equation should be derived from such a representation by a suitable specialization of the coordinate system. This would correspond to its derivation along the lines of the mathematical strategy. But as if to excuse himself for the failure to realize such a derivation, Einstein claimed that the hole argument and the argument from energy momentum conservation suggested that it would not be worthwhile to search for the generally-covariant counterpart of the *Entwurf* equation.[135]

In spite of this excuse Einstein embarked, at about the same time, on precisely such a search, albeit for another gravitation theory with restricted covariance properties serving as a toy model On 19 February 1914 he submitted a joint paper with

133 See "What Did Einstein Know ..." sec. 3 (in vol. 2 of this series).

134 See, e.g., (Abraham 1914, 25).

135 (CPAE 4, Doc. 25, [179]). See "Untying the Knot ...," (in vol. 2 of this series) note 57, for the relevant passage. See (Norton 1992a) for a historical discussion.

Adriaan Fokker on a generally-covariant reformulation of Nordström's special relativistic theory of gravitation (Einstein and Fokker 1914). They demonstrated that the field equation of this theory, in its original version only Lorentz covariant, can in fact be obtained from a generally-covariant equation. Just as in the mathematical strategy employed in the Zurich Notebook, a generally-covariant expression derived from the Riemann tensor served as the starting point of their approach, from which a suitable left-hand side of the gravitational field equation was then obtained by imposing additional conditions on the metric tensor. In the case of the Nordström theory, the additional condition amounted to the requirement of the constancy of the speed of light. This additional condition in turn led to a restriction on the admissible coordinate systems, in this case to those systems which are adapted to the principle of the constancy of the velocity of light (Einstein and Fokker 1914, 326).

As a consequence of the successful reformulation of Nordström's theory in generally-covariant terms, it was only natural to search for an analogous reformulation also of the *Entwurf* theory, in spite of the skepticism which Einstein had expressed in his earlier paper. That such a search made sense was precisely the conclusion which Einstein and Fokker drew at the end of their joint paper:

> Finally, the role that the Riemann-Christoffel differential tensor plays in the present investigation suggests that this tensor may also open the way for a derivation of the Einstein-Grossmann gravitation equations that is independent of physical assumptions. The proof of the existence or nonexistence of such a connection would represent an important theoretical advance.[136]

In a footnote to the above passage, they added:

> The argument in support of the nonexistence of such a connection, presented in §4, p. 36 of "Entwurfs einer verallgemeinerten Relativitätstheorie" ["Outline of a Generalized Theory of Relativity"], did not withstand closer scrutiny.[137]

On the cited page of the *Entwurf* paper, Einstein and Grossmann had simply claimed that, in the case of field equations with restricted covariance, it was understandable that no relation to generally-covariant tensors could be established (Einstein and Grossmann 1913, 36). But in view of Einstein's realization that a connection with a generally-covariant formulation must exist for any physically meaningful theory, the failure to discover such a connection could no longer be defended in this simple way.

In their paper, Einstein and Fokker used a terminology for the relation between generally-covariant equations and field equations with restricted covariance that

---

136 "Endlich legt die Rolle, welche bei der vorliegenden Untersuchung der Riemann-Christoffelsche Differentialtensor spielt, den Gedanken nahe, daß er auch für eine von physikalischen Annahmen unabhängige Ableitung der Einstein-Großmannschen Gravitationsgleichungen einen Weg öffnen würde. Der Beweis der Existenz oder Nichtexistenz eines derartigen Zusammenhanges würde einen wichtigen theoretischen Fortschritt bedeuten." (Einstein and Fokker 1914, 328).

137 "Die in §4, p. 36, des "Entwurfs einer verallgemeinerten Relativitätstheorie" angegebene Begründung für die Nichtexistenz eines derartigen Zusammenhanges hält einer genaueren Überlegung nicht stand."

would soon become standard in the further analysis of the covariance properties of the *Entwurf* theory. They spoke, in particular, of "preferred" ("*bevorzugt*") coordinate systems, "adapted" ("*angepasste*") to a certain physical situation.[138] Nordström's theory and the terminology developed for its treatment helped to further pursue the questions which had to be answered for a derivation of the *Entwurf* theory along the lines of the mathematical strategy to succeed: What were the "preferred" coordinate systems of the *Entwurf* theory? And what was the physical condition to which these coordinate systems are "adapted"? Although Einstein must have believed that he had answers to these questions, given his argument in favor of a restriction to linear transformations from energy-momentum conservation, it remained open how these answers could assist him in relating the *Entwurf* theory to its unknown generally-covariant counterpart. It was the experience gathered with Nordström's theory that eventually helped him to make progress in this regard—by challenging the answers that had seemingly settled the fate of the generalized relativity principle in the *Entwurf* theory.

7.9 A First Consequence of the Return to the Mathematical Strategy

In early March Einstein wrote to his friends about a breakthrough in his work on the *Entwurf* theory.[139] By this time he had not only recognized the fallacy of his argument for restricted covariance from energy-momentum conservation but had also investigated, jointly with Marcel Grossmann, the covariance properties of the theory in a new way. This new analysis, contained in a joint paper published on 29 May 1914 (Einstein and Grossmann 1914), was based on the use of variational techniques which allowed them to pursue the bold approach of exploring covariance properties by direct calculation.

Einstein's breakthrough was prepared by his reflection on the relation between non-covariant and covariant formulations of a theory, substantiated by his analysis of Nordström's theory. In light of these considerations, the *Entwurf* theory appears as a specialization of a generally-covariant theory to coordinate systems which are adapted to a certain physical condition. In the case of the *Entwurf* theory, this physical condition was the validity of energy-momentum conservation in the sense of eq. (51).

If generally-covariant field equations are expressed in coordinates adapted to this condition, they should take on the form of the *Entwurf* field equation eq. (52).

In analogy to the treatment of the Nordström theory, eq. (51) should be considered as a condition on the metric tensor $g_{\mu\nu}$, providing the necessary coordinate restriction. But in Einstein's original version of the argument for restricted covariance from energy-momentum conservation, this equation does not so much *provide* a con-

---

138 See (Einstein and Fokker 1914, 326).

139 See Einstein to Paul Ehrenfest, before 10 March 1914, (CPAE 5, Doc. 512) and Einstein to Heinrich Zangger, ca. 10 March 1914, (CPAE 5, Doc. 513).

dition on the metric tensor, but actually *presupposes* one. In fact, the argument that eq. (51) is only covariant under linear transformations only works, as Einstein was aware, if it is assumed that $t^{\mu\nu}$ has the same transformational behavior as $T^{\mu\nu}$, i.e. if it is a generally-covariant object. However, $t^{\mu\nu}$ is a coordinate-dependent expression.[140] While the assumption that it is generally covariant may have at first appeared plausible to Einstein in the light of his conviction that gravitational and other forms of energy should behave in the same way as sources of the gravitational field, this assumption becomes much less plausible once eq. (51) is seen as imposing a coordinate restriction on generally-covariant equations. That perspective requires in fact a much closer examination of its ingredients, since it is now the content rather than the form of the equation that matters. Apart from checking more closely the character of $t^{\mu\nu}$, Einstein's earlier experience with what we have called the conservation compatibility check in the sense of eq. (XLIII) suggested expressing $T^{\mu\nu} + t^{\mu\nu}$ in this equation by means of the field equations so that eq. (51) becomes a condition merely in terms of the metric tensor and its derivatives. One thus obtains eq. (54) as a condition for the class of admissible coordinate systems.

This equation played a central role in Einstein's new approach to the problem of the covariance properties of the *Entwurf* field equation. It first appeared in Einstein's and Grossmann's 1914 paper[141] and expresses in fact a physically motivated coordinate restriction in a sense that was quite familiar to him from his experiences along the mathematical strategy in the Zurich Notebook. In distinction from the original mathematical strategy, however, the generally-covariant equation from which the *Entwurf* field equation should be derivable by means of this coordinate restriction was unknown. But finding this generally-covariant equation may have been precisely Einstein's point in formulating eq. (54). In summary, a reconsideration from the perspective of the mathematical strategy of the argument for restricted covariance based on energy-momentum conservation could have led Einstein both to see the fallacy of his original argument and to cast it into a new form.

This reconstruction is supported by the timing of the transformation of the original argument for a linear covariance of the *Entwurf* equations into an argument about a coordinate restriction in the sense of the Zurich Notebook. In the manuscript of a popular exposition on his theory,[142] which Einstein completed by the end of January 1914,[143] he still included the argument in its original form. When he submitted the paper for publication by March, 21, 1914, that is, before he left Zurich, the passage arguing for the linear covariance of the *Entwurf* field equation was cancelled. By the beginning of March, Einstein had already achieved a breakthrough along the varia-

---

140 There are a number of arguments by which Einstein could have seen his fallacy: 1) There are no generally-covariant tensors involving only the metric and its first order derivatives. 2) In a suitably chosen coordinate system, the stress-energy complex of the gravitational field $t_{\mu\nu}$ can be made to vanish.

141 See (Einstein and Grossmann 1914, 218).

142 The published version is (Einstein 1914c). For references to and transcriptions of the manuscript version, see the annotations in (CPAE 4, 621–622).

143 See Einstein to Heinrich Zangger, ca. 20 January 1914, (CPAE 5, Doc. 507).

tional approach, as we know from his correspondence. The paper by Einstein and Fokker on Nordström's theory, on the other hand, in which, as we have also seen, the application of the mathematically strategy to the *Entwurf* theory is formulated as a program, was submitted on 19 February 1914. In other words, Einstein must have reformulated his argument based on the conservation principle at some point between the end of January and the beginning of March 1914, at the time or shortly after he was working on the application of the mathematical strategy to Nordström's theory.

With the discovery of the fallacy in the original argument, the question of the covariance properties of the *Entwurf* field equation was open again. While it was obvious that eq. (54) imposes a necessary condition on the coordinates systems "adapted" to this theory, it remained to be clarified whether this condition is also a sufficient one and how it related the *Entwurf* field equation to its generally-covariant counterpart. Einstein had thus arrived at a point where it made sense for him to take up the second point raised in Besso's notes:

> Is *every* system *that satisfies the conservation laws* a *justified* system?[144]

In the context of Einstein's reconsideration of the *Entwurf* field equation from the perspective of the mathematical strategy, the relation between conservation laws and "justified" coordinate systems must have assumed a new significance. Exploring whether or not this field equation actually retained its form under transformations between the "preferred" coordinate systems characterized by eq. (54) now became a pressing task. Unfortunately, the absolute differential calculus offered little help in addressing this task.

7.10 A New Turn for the Mathematical Strategy: Variational Calculus

When Einstein took up the mathematical strategy once again and adapted it to the *Entwurf* theory, he faced difficulties achieving concrete results along this strategy, and must have searched out mathematical advice. It is unclear at exactly which point Grossmann re-entered the story. Perhaps he was already instrumental in recognizing the fallacy of Einstein's original argument for restricted covariance from energy-momentum conservation. Perhaps he entered the picture only when Einstein needed help in exploring the consequences of the new coordinate restriction eq. (54). But Grossmann was, it seems, as little successful as Einstein in establishing relations between the *Entwurf* theory and absolute differential calculus. At some point they both turned to another Zurich mathematician colleague, Paul Bernays, for help.[145] Bernays advised Einstein and Grossmann to bring the field equation of the *Entwurf* theory into the form of a variational principle.[146]

The reformulation of the *Entwurf* theory in terms of a variational principle did not, however, provide any clue concerning the relation of this theory to absolute dif-

---

144 "Ist *jedes* System, *welches den Erhaltungssätzen genügt*, ein *berechtigtes* System?" For a facsimile of this passage, see Fig. 2 on p. 300 of "What Did Einstein Know ..." (vol. 2 of this series).

ferential calculus. The latter would have suggested, as it later did to Hilbert, to take the Ricci scalar as a starting point for such a reformulation. But, by analyzing the relation of this scalar to the B_μ of their coordinate restriction, Einstein and Grossmann convinced themselves that the B_μ do not form a generally-covariant vector and that the *Entwurf* field equation has nothing to do with the invariant Ricci scalar.[147] Nevertheless, a variational reformulation of the *Entwurf* had, for Einstein and Grossmann, one chief advantage: instead of having to study the covariance properties of a complex tensorial field equation, they could instead explore the invariance group of a single scalar object, the action integral (cf. eq. (LXIV)). Much later Einstein still considered the simplification due to the introduction of the more familiar scalar quantities the main advantage of the variational calculus.[148]

In early 1914, the suggestion to make use of the variational calculus brought Einstein and Grossmann back to the initial bold approach of exploring by direct calculation the covariance properties of the *Entwurf* field equation. In pursuing this approach they could rely on their experience from the Zurich Notebook where they had attempted to study the covariance properties of objects found along the physical strategy or of coordinate restrictions by means of infinitesimal transformations.

However, one crucial presupposition of the new approach had to be established first, the expression for the action integral from which the *Entwurf* field equation could be derived by means of the variational formalism. In their 1914 paper Einstein and Grossmann only give the end result, without mentioning what had motivated them to introduce a particular Langrangian, other than its successful employment in deriving the field equations.[149] Probably they found the Lagrangian of the *Entwurf* theory by starting from an expression quadratic in the fields in analogy to classical and special-relativistic physics according to the default setting eq. (LXIII). With Einstein's default setting for the components of the gravitational field, the Lagrangian required for deriving the *Entwurf* field equation was found to be:

$$L = g^{\mu\nu} g_{\beta\mu,\alpha} g_{\alpha\nu,\beta} = g^{\mu\nu} \tilde{\Gamma}^{\alpha}_{\beta\mu} \tilde{\Gamma}^{\beta}_{\alpha\nu}. \tag{85}$$

145 Bernays later became known for his work in mathematical logic and set theory, was in Zurich from 1912 to 1919 after completing a mathematical doctoral thesis on the analytic theory of binary quadratic forms under the supervision of the mathematician E. Landau. Before coming to Zurich, Bernays had spent two years in Göttingen studying mathematics and physics chiefly with Hilbert, Landau, Weyl, Klein, Voigt and Born. Bernays was at the time concerned with an extension of the special theory of relativity.

146 In their paper, Einstein and Grossmann acknowledge the stimulation received from Bernays in a footnote (Einstein and Grossmann 1914, 218).

147 See (Einstein and Grossmann 1914, 225).

148 Einstein to Lorentz, 19 January 1916 and Einstein to T. De Donder, 23 July 1916 (CPAE 8, Docs. 184 and 240).

149 See (Einstein and Grossmann 1914, 219).

The *Entwurf* theory was thus solidified by connecting its field equation in yet another way with the established knowledge of classical and relativistic physics, in this case about the canonical form of a Lagrangian.

The identification of both the Langrangian for the *Entwurf* field equation and of a physically motivated coordinate restriction now gave Einstein and Grossmann a clear definition of their next goal, the establishment of a relation between the coordinate restriction and the transformational properties of this Lagrangian. Does the coordinate restriction eq. (54) resulting from the conservation principle indeed constitute not only a necessary but also a sufficient condition for the covariance of the *Entwurf* Lagrangian? In that case, by establishing a connection between conservation and covariance, Einstein would have achieved a result effectively preparing the later Noether theorem.[150] The close connection between conservation and covariance, first suggested by the ill-fated argument for a restriction of the *Entwurf* theory to linear transformations, became a heuristic guiding principle for Einstein's further exploration and a criterion that he expected a satisfactory theory to fulfill.

The means to answer his question concerning the covariance properties of the *Entwurf* Lagrangian was provided by the infinitesimal coordinate transformations explored earlier in the Zurich Notebook. With their help, Einstein and Grossmann succeeded in establishing a connection between the transformational properties of the action integral for the *Entwurf* Lagrangian and the physically motivated coordinate restriction eq. (54). Their bold approach had finally given them what they had failed to achieve with the help of the absolute differential calculus—a link between the physical and the mathematical strategies.

7.11 Looking Back on a Breakthrough: The General Relativity of the Entwurf Theory

With their proof of the covariance properties of the *Entwurf* field equations, Einstein and Grossmann had finally closed the crucial gap in their 1913 publication. But in Einstein's view, they had achieved much more. In the time between the completion of this proof by early March 1914 and the discovery of a problem with transformations to rotating frames of reference in September 1915, he was convinced that he had finally reached not only a generalization of the relativity principle but a truly general theory of relativity. In early March he wrote to Paul Ehrenfest that the proof of the existence of "most general transformations" leaving the field equations covariant demonstrated the validity of the principle of equivalence as well:

> The work on gravitation progresses, but at the cost of extraordinary efforts; gravitation is coy and unyielding! The equivalence principle is valid after all in the sense that there exist highly general transformations that transform the gravitational equations into themselves. What has been found is simple, but the search is hell![151]

In a similar vein he expressed himself in a contemporary letter to Heinrich Zangger:

150 See "Untying the Knot ..." (in this volume).

> I was toiling again on the gravitation theory to the point of exhaustion, but this time with unheard-of success. That is to say that I succeeded in proving that the gravit. equations hold for arbitrarily moving reference systems, and thus that the hypothesis of the equivalence of acceleration and the gravitational field is absolutely correct, in the widest sense. Now the harmony of the mutual relationships in the theory is such that I no longer have the slightest doubt about its correctness.[152]

In these passages Einstein left it somewhat open what he meant by qualifying the covariance he had reached as "most general" or "in the widest sense." In another contemporary passage he showed himself convinced that the transformation to a rotating coordinate system was comprised by this covariance:

> By means of a *simple* calculation I have been able to prove *that the gravitation equations hold for every reference system that is adapted to this condition*. From this it follows that there exist acceleration transformations of the most varied kind that transform the equations to themselves (e.g., also rotation), so that the equivalence hypothesis is preserved in its original form, even to an unexpectedly large extent.[153]

The "simple calculation" to which Einstein refers must be the demonstration of the covariance properties of the *Entwurf* equation published jointly with Grossmann, since he emphasizes the crucial element of this demonstration, the condition for adapted coordinate systems. He obviously perceived the more specific properties of the field equation, such as its covariance under rotation, as being merely a trivial consequence of this proof. Einstein thus believed he had achieved a full implementation of the generalized principle of relativity.

Yet, the exact relation of the *Entwurf* field equation to the absolute differential calculus had not been clarified. It seems, however, that Einstein did not bother too much about this problem. When he learned that Grossmann had finally succeeded in establishing such a relation, Einstein viewed this result as a nice complement to what they had already achieved earlier but not more. In late March or early April 1914 he wrote to Ehrenfest:

---

151 "Die Gravitation macht Fortschritte, aber unter ausserordentlichen Anstrengungen; sie ist spröde! Das Aequivalenzprinzip gilt nun doch in dem Sinne, dass es höchst allgemeine Transformationen gibt, die die Gravitationsgleichungen in sich überführen. Das Gefundene ist einf[a]ch, aber das Suchen ganz verflucht." Einstein to Paul Ehrenfest, before 10 March 1914, (CPAE 5, Doc. 512).

152 "Ich habe mich wieder bis zur Erschöpfung geplagt mit der Gravitationstheorie, aber diesmal mit unerhörtem Erfolge. Es ist nämlich der Beweis gelungen, dass die Gravit. Gleichungen für beliebig bewegte Bezugssysteme gelten, dass also die Hypothese von der Aequivalenz der Beschleunigung und des Gravitationsfeldes durchaus richtig ist, im weitesten Sinne. Nun ist die Harmonie der gegenseitigen Beziehungen in der Theorie eine derartige, dass ich an der Richtigkeit nicht mehr im Geringsten zweifle." Einstein to Heinrich Zangger, 10 March 1914, (CPAE 5, Doc. 513).

153 "Ich habe beweisen können durch eine *einfache* Rechnung, *dass die Gleichungen der Gravitation für jedes Bezugssystem gelten, welches dieser Bedingung angepasst ist*. Hieraus geht hervor, dass es Beschleunigungstransformationen mannigfaltigster Art gibt, welche die Gleichungen in sich selbst transformieren (z.B. auch Rotation), sodass die Aequivalenzhypothese in ihrer ursprünglichen Form gewahrt ist. sogar in ungeahnt weitgehendem Masse." Einstein to Michele Besso, ca. 10 March 1914, (CPAE 5, Doc. 514), Einstein's emphasis. The "einfache Rechnung" probably refers to the covariance proof. For an alternative interpretation, see (Janssen 1999, n. 125).

> Grossmann wrote me that now he also is succeeding in deriving the gravitation equations from the general theory of covariants. This would be a nice addition to our examination.[154]

The more Einstein thought about the proof of the covariance properties of the *Entwurf* field equation he had jointly developed with Grossmann, the more he became convinced that what he had reached was general covariance. This is apparent from his ever more optimistic assessments of their result. In June 1914 Einstein wrote to Wien:

> In Zurich I had found the proof for covariance in the gravitation equations. Now the theory of relat[ivity] really is extended to arbitrarily moving systems.[155]

In July Einstein wrote to Planck, also claiming that he had now his theory covered every possible manifold and that the restriction was only one of the coordinate system:

> Then also a brief reply to a comment you made recently at the Academy in the welcoming speech. There is an essential difference between the reference system restriction introduced by classical mechanics for the theory of relativity and that which I apply in the theory of gravitation. For the latter can always be adopted no matter how the $g_{\mu\nu}$'s may be selected. To the contrary, the specialization introduced by the principle of the constancy of the velocity of light presupposes differential correlations between the $g_{\mu\nu}$'s, that is, correlations that ought to be very difficult to interpret physically. Satisfaction of these correlations cannot be forced by the appropriate choice of a reference system for any given manifold.[156]

According to the explanation given to Planck, Einstein considered the principal distinction between the specialization of the reference system in classical mechanics and in the special theory of relativity, on the one hand, and that which he had introduced in his new gravitation theory, on the other hand, to be the fact that in the latter case the specialization of the reference system refers only to the choice of the coordinate system in an otherwise arbitrarily given manifold. Einstein made this point particu-

---

154 "Grossmann schreibt mir, dass es ihm nun auch gelingt, die Gravitationsgleichungen aus der allgemeinen Kovariantentheorie abzuleiten. Es wäre dies eine hübsche Ergänzung zu unserer Untersuchung." Einstein to Paul Ehrenfest, 10 April 1914, (CPAE 8, Doc. 2).

155 "In Zürich fand ich noch den Nachweis der Kovarianz der Gravitationsgleichungen. Nun ist die Relat[ivitäts]theorie wirklich auf beliebig bewegte Systeme ausgedehnt." Einstein to Wilhelm Wien, 15 June 1914, (CPAE 8, Doc. 14).

156 "Sodann noch eine kurze Beantwortung einer Bemerkung, die Sie neulich in der Akademie in der Begrüssungsrede geäussert haben. Es gibt einen prinzipiellen Unterschied zwischen derjenigen Spezialisierung des Bezugssystems, welche die klassische Mechanik bezw. die Relativitätstheorie einführt und zwischen derjenigen, welche ich in der Gravitationstheorie anwende. Die letztere kann man nämlich stets einführen, wie auch die $g_{\mu\nu}$ gewählt werden mögen. Diese durch das Prinzip der Konstanz der Lichtgeschwindigkeit eingeführte Spezialisierung dagegen setzt Differenzialbeziehungen zwischen den $g_{\mu\nu}$ voraus, und zwar Beziehungen, deren physikalische Interpretation sehr schwierig sein dürfte. Das Erfülltsein dieser Beziehungen kann nicht für jede gegebene Mannigfaltigkeit durch passende Wahl des Bezugssystems erzwungen werden." Einstein to Max Planck, 7 July 1914, (CPAE 8, Doc. 18).

larly clear in a letter he wrote to Lorentz a few months later. In this letter Einstein explained in what sense the restriction to adapted coordinate systems in his understanding was compatible with the claim that the theory would be a "general" theory of relativity. He referred to an analogous situation in the Gaussian theory of surfaces:

> Although I prefer certain reference systems, the fundamental difference to the Galilean preference is, however, that my coordinate selection makes no physical assumptions about the world; let this be illustrated by a geometric comparison. I have a plane of unknown description which I want to subject to geometric analysis. If I require that a coordinate system (p, q) on the plane be selected in such a way that
>
> $$ds^2 = dp^2 + dq^2,$$
>
> I therefore assume that then the surface can be unfolded on to a [Euclidean] plane. Were I only to demand, however, that the coordinates be chosen in such a way that
>
> $$ds^2 = A(p, q)dp^2 + B(p, q)dq^2$$
>
> i.e., that the coordinates be orthogonal, then I am assuming nothing about the nature of the surface; this can be obtained on any surface.[157]

The analogy with Gaussian surface theory suggests a geometrical interpretation of the coordinate restriction introduced in Einstein's theory of gravitation. A letter Einstein wrote in 1915 to Paul Hertz shows that he had searched in vain for such an interpretation and that for elucidating the meaning of this coordinate restriction, he had little more to offer than the comparisons he mentioned in the letter to Lorentz.

> He who has wandered aimlessly for so long in the chaos of possibilities understands your trials very well. You do not have the faintest idea what I had to go through as a mathematical ignoramus before coming into this harbor. Incidentally, your idea is very natural and would by all means be worth following up, if it could be carried through at all, which, based upon my experiences gathered during my wayward wanderings, I doubt very much.
>
> Given an arbitrary manifold of 4 dimensions (given $g_{\mu\nu}(x_\sigma)$). How can one distinguish a coordinate system or a group of such? This appears not to be possible in any way simpler than the one chosen by me. I have groped around and tried all sorts of possibilities, e.g., required: The system must be chosen such that the equations
>
> $$\sum_\nu \frac{\partial g^{\mu\nu}}{\partial x_\nu} = 0 \qquad (\mu = 1-4)$$

157 "Ich bevorzuge zwar auch gewisse Bezugssysteme, aber der fundamentale Unterschied gegenüber der Galileischen Bevorzugung besteht darin, dass meine Koordinatenwahl nichts über die Welt voraussetzt; dies sei durch einen geometrischen Vergleich erläutert. Es liegt mir eine Fläche unbekannter Art vor, auf der ich geometrische Untersuchungen machen will. Verlange ich, es solle auf der Fläche ein Koordinatensystem (p, q) so gewählt werden, dass $ds^2 = dp^2 + dq^2$, [s]o setze ich damit voraus, dass die Fläche auf eine Ebene abwickelbar sei. Verlange ich aber nur, dass die Koordinaten so gewählt seien, dass $ds^2 = A(p, q)dp^2 + B(p, q)dq^2$ ist, d.h. dass die Koordinaten orthogonal seien, so setze ich damit über die Natur der Fläche nichts voraus; man kann dies auf jeder Fläche erzielen." Einstein to H.A. Lorentz, 23 January 1915, (CPAE 8, Doc. 47).

> are satisfied throughout.
> At least it seemed definite to me *a priori* that a transformation group exceeding the Lorentz group must exist, because those observations summed up in the words "relativity principle" and "equivalency principle" point to it.
> The coordinate limitation that was finally introduced deserves particular trust because it establishes a link between it and the postulate of the event's complete determination.
> A theoretical differential geometric interpretation of preferred systems would be of great value. The weakest point of the theory as it stands today consists precisely in this, that the group of justified transformations are by no means closely assessable. There is not even any *exact* proof that arbitrary motions can be transformed to motionlessness.[158]

The letter shows that Einstein saw all coordinate restrictions he had examined to function essentially on the same level, that is, to be generally imposed as conditions supplementary to the field equations; this is made clear by his formulation that he assumed what we have called the "Hertz restriction" eq. (60) to be satisfied "everywhere." He evidently treated the Hertz restriction on the same level as the condition eq. (54) for adapted coordinate systems, despite their different form. Both conditions were motivated, as we have seen, by the conservation principle. But, as Einstein points out in his letter, the condition for adapted systems could also be justified on a deeper level; the causality considerations were related to the hole argument, and therefore inspired more confidence. The letter to Hertz furthermore confirms that Einstein was convinced that this condition just implies a particular choice of the coordinate system without restricting the range of possible manifolds. He implicitly claimed that, in the *Entwurf* theory, all motions can be transferred to rest, although he admitted that he had been unable to demonstrate this "exactly."

---

158 "Wer selber im Chaos der Möglichkeiten sich so viel herumgetrieben hat, begreift Ihre Schicksale sehr gut. Sie haben ja keine blasse Ahnung, was ich als mathematischer Ignorant habe durchmachen müssen, bis ich in diesen Hafen eingelaufen bin. Übrigens ist Ihre Idee sehr natürlich und wäre auf jeden Fall ernster Verfolgung wert, wenn sie sich überhaupt durchführen liesse, was ich auf Grund meiner im Herumirren allmählich angesammelten Erfahrung sehr bezweifle.
Gegeben eine beliebige Mannigfaltigkeit von 4 Dimensionen ($g_{\mu\nu}(x_\sigma)$ gegeben). Wie kann man ein Koordinatensystem bezw. eine Gruppe von solchen auszeichnen? Es scheint dies auf einfacher als die von mir gewählte Art nicht möglich zu sein. Ich habe herum getastet und alles Mögliche versucht, z.B. verlangt: Das System soll so gewählt werden, dass überall die Gleichungen [eq.] erfüllt seien.
Immerhin schien es mir a priori sicher, dass eine über die Lorentz-gruppe hinausgehende Transformationsgruppe vorhanden sein müsse, da jene Erfahrungen, die mit den Worten Relativitätsprinzip, Aequivalenzprinzip zusammengefasst werden, darauf hinweisen.
Die schliesslich eingeführte Koordinatenbeschränkung verdient deshalb besonderes Vertrauen, weil sie sich mit dem Postulat der vollständigen Bedingtheit des Geschehens in Zusammenhang bringen lässt.
Eine flächentheoretische Interpretation der bevorzugten Systeme wäre von sehr grossem Werte. Der schwächste Punkt der Theorie bei ihrem heutigen Stande besteht nämlich gerade darin, dass man die Gruppe der berechtigten Transformationen durchaus nicht scharf übersieht. *Exakt* ist nicht einmal der Beweis geliefert, dass beliebige Bewegungen auf Ruhe transformiert werden können." Einstein to Paul Hertz, 22 August 1915, (CPAE 8, Doc. 111). For an extensive discussion of this letter, see (Howard and Norton, 1993).

In spite of such reservations, Einstein was nevertheless convinced that he had reached all his major original heuristic goals within the *Entwurf* theory. What remained were only some minor problems, such as the establishment of the connection of the methods used by Einstein and Grossmann to the absolute differential calculus, and a clarification of some other mathematical aspects of the theory. Einstein also had the impression that the crucial proof of the covariance properties of the *Entwurf* field equation still required improvement. This was the task he set himself in mid-1914 in the context of composing a major review article, finished by the end of October and providing a full exposition of the finally complete theory which now was called, for the first time, the "general theory of relativity" (Einstein 1914a).

7.12 The Revised Covariance Proof and the Definitive Formulation of the Hole Argument[159]

When at the end of 1914 Einstein looked back on his first review article on general relativity, entitled "The Formal Foundations of General Relativity" and submitted on 29 October 1914, the revision of the covariance proof appeared to him as the central achievement, as is suggested in a letter he wrote in December to Paul Ehrenfest:

> In recent months I reworked extremely carefully the basis of the general theory of rel. The covariance proof of last spring was not yet completely right. Otherwise, I have also been able to penetrate a few things more clearly. Now I am entirely satisfied with that matter. You will soon receive the paper; read it, you will find it very enjoyable.[160]

The proof of the covariance properties of the *Entwurf* field equation as originally conceived by Einstein and Grossmann was based on the idea that an infinitesimal, adapted coordinate transformation leaves the variation of the action integral invariant. The variation of the manifold giving rise to the variation of this integral had to be performed in two steps, an "adapted" variation, making it possible to vary the coordinate system along with the manifold so that it remains adapted to it, and a variation of the manifold that merely corresponds to the introduction of a new coordinate system, a "coordinate variation." The problem with the original proof was that the first of these two variations was not clearly defined. In fact, Einstein and Grossmann had obvious difficulties in arguing for the possibility of an appropriate variation of the adapted coordinate system "following" that of the manifold. This variation of the coordinate system was introduced rather as an afterthought to the variation of the manifold, an afterthought which left open exactly how the variation of the manifold is restricted by the condition that it must be possible to vary the adapted coordinate system along

---

159 See (Cattani and De Maria 1989b).

160 "In den letzten Monaten habe ich die Grundlage der allgemeinen Rel Theorie nochmals höchst sorgfältig bearbeitet. Der Kovarianzbeweis vom letzten Frühjahr war noch nicht ganz in Ordnung. Auch sonst habe ich manches klarer durchdringen können. Nun bin ich aber völlig zufrieden mit der Angelegenheit. Du erhälst bald die Arbeit, lies sie, Du wirst grosse Freude daran finden." Einstein to Paul Ehrenfest, December 1914 (CPAE 8, Doc. 39).

with it. The origin of Einstein's difficulties was his concept of a manifold being closely tied to its representation by the metric tensor and hence lacked the clear-cut distinction from the representation of a manifold in terms of coordinates.[161]

At the outset of his new approach, Einstein distinguished more clearly a variation of the manifold and a variation of the coordinate system. It was probably for the purposes of such a cleaner separation of the different kinds of variations that he treated coordinate systems—as suggested by the hole argument and in contrast to the modern understanding—as being essentially given independently from the manifold for whose description they serve. Einstein believed that in this way he could refer to two different manifolds, or rather one manifold before and, after the variation, to one and the same coordinate system.[162] The ensuing challenge to refer changes of the values of the metric tensor due to a coordinate transformation to one and the same coordinate system, given independently from the manifold, may well have induced him to formulate more clearly than he had done before the artifice of transposing values of the metric tensor characteristic of the hole argument.

Einstein's treatment of the hole argument in his 1914 review paper is in fact the first published version of this argument that makes plain how values of the metric tensor given at one point of the manifold are to be referred to another point, a notion implicit in its original formulation in late August 1913 but obscured in the subsequent published presentations. It is also the first version that introduces a distinct notation for the metric tensor and its representation in a particular coordinate system.[163] The mature and more elaborate formulation of the hole argument was hence closely associated with the reworking of the covariance proof. Revisiting, together with Marcel Grossmann, the covariance properties of the *Entwurf* field equation, Einstein arrived at a formulation of this argument that now pointed to philosophical questions concerning the mathematical representation of the physical properties of space and time.

7.13 A Shaky Mathematical Derivation and a Spin-off with Consequences

In the introduction to his 1914 review paper Einstein mentioned the peculiar combination of physical and mathematical arguments that led him to the *Entwurf* theory and announced a purely mathematical derivation:

> In recent years I have worked, in part together with my friend Grossmann, on a generalization of the theory of relativity. During these investigations, a kaleidoscopic mixture of postulates from physics and mathematics has been introduced and used as heuristical tools; as a consequence it is not easy to see through and characterize the theory from a formal mathematical point of view, that is, only based upon these papers. The primary

161 See (Norton 1992b).

162 See (Einstein 1914a, 1071–1073). Einstein conceived a variation of the metric tensor generated by a coordinate transformation, referring its result to the same original coordinate system. His transformation can thus not be an ordinary coordinate transformation, but must be the kind of transport of values of the metric tensor from one coordinate system to the other as it is essential to the hole argument.

163 See (Einstein 1914a, 1067).

> objective of the present paper is to close this gap. In particular, it has been possible to obtain the equations of the gravitational field in a purely covariance-theoretical manner ... .[164]

The basis for this derivation was provided by the variational formalism. In their paper of early 1914, Einstein and Grossmann had, as we have seen, solved the problem of identifying an appropriate Lagrangian from which the *Entwurf* equations were derived. Since their Lagrangian now represented the natural starting point for building up the entire edifice of Einstein's theory, the question arose whether this Lagrangian could be justified by reasons other than that of generating the desired field equation. In order to answer this question, Einstein generalized the formalism jointly developed with Grossmann to apply to an arbitrary Lagrangian. While this generalization was rather straightforward, it was a more challenging task to pinpoint the assumptions by which the resulting formalism could be specialized again so as to determine the Lagrangian appropriate for the *Entwurf* field equation. Einstein's approach was effectively guided by the mathematical strategy presupposing a generic mathematical object, which is then specialized in light of concrete physical requirements. In the generalized formalism of his 1914 review paper, such physical requirements had to be formulated as mathematical criteria serving to select the *Entwurf* Lagrangian.

It was a combination of two criteria that helped Einstein to achieve this goal, one derived from the conservation principle, the other from the generalized principle of relativity. He formulated both criteria in terms of differential conditions for the Lagrangian and concluded, erroneously as it later turned out, that the requirement of their compatibility singles out a particular candidate. In the *Entwurf* theory, the implementation of the conservation principle imposed, as we have seen, a coordinate restriction $B_\mu = 0$ (cf. eq. (54)). This condition played, as we have also seen, the double role of ensuring the satisfaction of the conservation principle and of determining the covariance properties of the field equation. The exact same equation $B_\mu = 0$ also played a role in Einstein's interpretation of the generalized theory, but here only in the context of analyzing its covariance properties. The formulation of the conservation principle within the generalized framework yielded a slightly different equation, now comprising two terms instead of one:

$$\sum_\nu \frac{\partial S_\mu^\nu}{\partial x_\nu} - B_\mu = 0. \tag{86}$$

164 "In den letzten Jahren habe ich, zum Teil zusammen mit meinem Freunde Grossmann, eine Verallgemeinerung der Relativitätstheorie ausgearbeitet. Als heuristische Hilfsmittel sind bei jenen Untersuchungen in bunter Mischung physikalische und mathematische Forderungen verwendet, so daß es nicht leicht ist, an Hand jener Arbeiten die Theorie vom formal mathematischen Standpunkte aus zu übersehen und zu charakterisieren. Diese Lücke habe ich durch die vorliegende Arbeit in erster Linie ausfüllen wollen. Es gelang insbesondere, die Gleichungen des Gravitationsfeldes auf einem rein kovarianten-theoretischem Wege zu gewinnen ... ." (Einstein 1914a, 1030)

The existence of two similar but not identical conditions could be turned into a compatibility argument identifying the *Entwurf* theory as a special case of the generalized formalism. He thus demanded:

$$S_\mu^\nu \equiv 0, \tag{87}$$

and claimed that this condition, together with the additional requirement that the Lagrangian be a homogeneous function of second degree in the gravitational fields, determines uniquely the *Entwurf* Lagrangian. Einstein's additional "mathematical" requirement has, as is the case for his other constraints, also physical aspects, here the analogy with the Lagrangian for a free electromagnetic field (cf. the default setting eq. (LXIII)). In his paper Einstein did not explicitly prove his claim. It may well have been his long-held conviction that the conservation principle determines uniquely the *Entwurf* field equation, that simply correlated with his belief that the *Entwurf* theory can be uniquely characterized with the help of eq. (87).

It turned out later that Einstein's reasoning was flawed. A more careful analysis of his formalism later showed him that eq. (87) did not actually impose a strong additional selective criterion, but could be easily fulfilled by simply requiring that the Lagrangian be an invariant under linear transformations, a criterion that does not help to single out the *Entwurf* theory. Einstein had imposed this requirement implicitly in the context of his analysis of the covariance properties and had effectively suppressed the condition involving S_μ^ν in this context, thus coming up with what appeared to be two different sets of conditions, one derived from the generalized principle of relativity, the other from the conservation principle. A deeper exploration of his formalism, first achieved about a year later, eventually offered him the insight that the analysis of the covariance and the conservation aspects actually implied the same set of conditions, an important step towards what later became Noether's theorem. This step was prepared by Einstein's seemingly successful attempt to derive the *Entwurf* theory along a mathematical strategy in which, alongside the conservation principle, covariance considerations had assumed the role of the correspondence principle in restricting the admissible candidate Lagrangians. More than anything else, it was the supposed achievement of being able to renounce the correspondence principle as part of Einstein's derivation that gave it the appearance of being largely independent of specific physical knowledge about gravitation (Einstein 1914a, 1076).

Nevertheless, what Einstein had achieved was satisfactory also from a physical point of view. In the course of his mathematical elaboration of the *Entwurf* theory, he had brought its field equation into a form satisfying all structural requirements following from the conservation principle. In particular, he succeeded in identifying, even for a generic Lagrangian, an expression for the stress-energy tensor of the gravitational field, i.e. for **FIELDMASS**. Furthermore, he was able to write the field equation in a form corresponding to eqs. (XLIV) and (XLV), thus demonstrating the structural analogy with classical field theory as well as the parallelism between the stress-energy tensor of matter and of the gravitational field as sources of the field. In this way, Einstein had built up a mathematical framework lending itself to direct physical interpretation:

> The system of equations (81) allows for a simple physical interpretation in spite of its complicated form. The left-hand side represents a kind of divergence of the gravitational field. As the right-hand side shows, this is caused by the components of the total energy tensor. A very important aspect of this is the result that the energy tensor of the gravitational field itself acts field-generatingly, just as does the energy tensor of matter.[165]

7.14 From Consolidation to Exploration

7.14.1 Living with the Less than Perfect

If considered in hindsight of general relativity, the *Entwurf* theory has considerable flaws: it does not comply with the only astronomical test available before 1919[166] for a relativistic gravitation theory, the anomalous perihelion advance of Mercury by ca. 43" per century, which is inexplicable in terms of Kepler's laws; it does not include the Minkowski metric in rotating coordinates as a solution and hence disappointed Einstein's Machian expectations; and the mathematical derivation from general principles was based on an error. Einstein's discovery of these flaws in the *Entwurf* theory may appear to constitute a step-by-step refutation, clearing the way for a new approach. However, uncovering these flaws did not immediately shatter the *Entwurf* theory. As was shown above, the *Entwurf* theory had emerged as a theory firmly grounded in the knowledge of classical physics, incorporating, in particular, both the correspondence and the conservation principles. At the same time, the theory allowed for a limited extension of the generalized relativity principle to at least general linear transformations, this limitation being, however, justified by both physical and mathematical arguments. Whatever Einstein achieved in the second phase beyond this state—in terms of an astronomical confirmation of the theory, of a further generalization of the relativity principle, or in terms of its mathematical elaboration—was not necessary to support the theory. The successes and failures beyond the core established in the consolidation period concerned the more ambitious part of Einstein's heuristics, in particular the extension of the generalized principle of relativity, which from the beginning was a less stringent criterion for the validity of his new gravitation theory than its relation to the knowledge of classical physics.

In this section, we shall briefly assess the impact of the discovery of flaws in the *Entwurf* theory on Einstein's attitude with respect to his theory. It demonstrates his

---

165 "Das Gleichungssystem (81) [cf. eq. (52)] läßt trotz seiner Kompliziertheit eine einfache physikalische Interpretation zu. Die linke Seite drückt eine Art Divergenz des Gravitationsfeldes aus. Diese wird—wie die rechte Seite zeigt—bedingt durch die Komponenten des totalen Energietensors. Sehr wichtig ist dabei das Ergebnis, daß der Energietensor des Gravitationsfeldes selbst in gleicher Weise felderregend wirksam ist wie der Energietensor der Materie." (Einstein 1914b, 1077)

166 For a discussion of the status of the other two classical tests, gravitational light bending and gravitational redshift, by 1919, see the introduction to (CPAE 9).

ability to live with the less than perfect or, more specifically, his resistance to abandoning an elaborate edifice because of damage it suffered on one of its floors.

7.14.2 The Mercury Problem

In temporal order, Einstein's discovery of the failure of the *Entwurf* theory to yield the correct perihelion shift of Mercury came first; it was made as early as the summer of 1913. The extensive research notes by Einstein and Besso, which document their joint effort to calculate Mercury's perihelion motion, show that this endeavor was actually part of a broader program that included not only the *Entwurf* theory, but also Nordström's gravitation theory, and not only the perihelion shift of Mercury, but also other possible checks of a non-Newtonian gravitation theory, such as its compatibility with the effects anticipated on the basis of Einstein's Machian heuristics. This broader perspective may have shaped Einstein's reaction to finding that the *Entwurf* theory could not account for the astronomically observed value of the perihelion shift. First of all, this anomaly could not be explained by other contemporary gravitation theories; second, there might have been a purely astronomical explanation for it; and third, there was a range of other possible checks of the *Entwurf* theory, such as the deflection of light in a gravitational field and gravitational redshift. In view of this situation, the negative finding on Mercury's perihelion shift was not a result of immediate significance for the validity of the *Entwurf* theory. It had required some effort to perform the perihelion calculation, but from the beginning it must have been at best a hope that a relativistic gravitation theory could actually account for this effect. Einstein himself did not publish his negative result. He encouraged Besso to complete a paper offering a comparative evaluation of contemporary gravitation theories both on empirical and epistemological grounds.[167] In his contemporary letters, he appeared more convinced of or worried by, as the case might be, the theory's internal consistency or lack thereof.

The failure of the perihelion calculation was not mentioned in Einstein's publications and hardly ever in his contemporary correspondence. It only played a role in Einstein's later justifications of his abandonment of the *Entwurf* theory. If his and Besso's extensive manuscript notes had not survived, one would not have known how much effort they had invested into this calculation. And yet, this calculation had a profound impact on the genesis of general relativity, which is discussed more extensively below, by affecting the speed with which Einstein could calculate the perihelion shift on the basis of his later generally-covariant theory.[168] This was possible because the formalism he had developed for the *Entwurf* theory turned out to be more generally applicable and hardly required any modification when used in the context of another gravitation theory. But the perihelion calculation also had more subtle effects which, as we shall see, later turned out to be beneficial to Einstein's renewed

---

167 Einstein to Michele Besso, after 1 January 1914 (CPAE 5, Doc. 499).
168 For detailed historical discussion, see (Earman and Janssen 1993).

exploration of generally-covariant candidate field theories. It led, in particular, to an improved understanding of the Newtonian limit.

7.14.3 The Rotation Problem

The method developed by Einstein and Besso for calculating the perihelion advance was based on an iterative procedure for finding approximate solutions of the field equation. It also turned out to be applicable to the investigation of another question of great heuristic significance for Einstein's attempt to generalize the relativity principle.[169] As we have discussed earlier, this attempt was guided, from the beginning, by the idea of conceiving rotation as being equivalent to a state of rest, interpreting the inertial forces arising in a rotating frame of reference as a special gravitational field. If the *Entwurf* theory were actually compatible with this heuristics, the Minkowski metric in rotating coordinates should be a solution of its field equations.

The inertial forces arising in a rotating frame, centrifugal and Coriolis forces, are of a different order in the angular velocity, the centrifugal force depending on its square, the Coriolis forces depending linearly on this velocity. Einstein and Besso's approximation scheme could thus be used to check whether one can obtain from a first-order approximation of a Minkowski metric in rotating coordinates, containing only the components relevant for the Coriolis forces, the correct second order term relevant for the centrifugal forces. The result of this calculation could then be compared with that of the direct transformation of the Minkowski metric in rotating coordinates.

In a scratch notebook Einstein first wrote down the one component relevant for the centrifugal force and then two components relevant for the Coriolis force. Underneath he wrote:

> Is the first equation [concerning the centrifugal force] a consequence of the other two [concerning the Coriolis force] on the basis of the theory?[170]

In a page of the bundle of manuscripts used jointly by Einstein and Besso for their calculations on the perihelion shift, Einstein actually performed this check (CPAE 4, Doc. 14, [41–42]). Although the approximation scheme applied to the *Entwurf* theory does not yield the correct value for the 4–4 component of a Minkowski metric in rotating coordinates, he at first came to the erroneous conclusion that it actually does, and ended his calculation with the remark "stimmt" (CPAE 4, Doc. 14, [41]).

There is, however, as early as 1913 evidence that this was not Einstein's last word. In the draft for his paper on contemporary gravitation theories, Besso listed the failure of the *Entwurf* theory to yield the correct combination of centrifugal and Coriolis forces, in other words, its failure to include the Minkowski metric in rotating coordi-

---

169 For a discussion of this procedure, see the editorial note on the Einstein-Besso manuscript in (CPAE 4, 346–349), as well as (Earman and Janssen 1993, 142–143).

170 "Ist die erste Gleichung Folge der beiden letzten auf Grund der Theorie?" (CPAE 3, Appendix A, [66]).

nates as a solution, as problems to be thought about and to be discussed with Einstein. From these notes it seems that at some point by the end of August 1913, Besso was aware of this problem. Einstein's contemporary correspondence suggests that he as well had realized by mid-August that the positive result mentioned above was actually based on an error.[171]

The problem resurfaced only when Einstein had convinced himself, after the discovery of a flaw in his original conservation argument (cf. subsection "Einstein's Reinterpretation of the Conservation Principle," p. 235, that there actually exists a large variety of transformation to accelerated reference systems under which the *Entwurf* theory was covariant. On 20 March 1914, Michele Besso wrote to Einstein, after the latter had reported his progress in analyzing the covariance properties of the *Entwurf* field equations:

> Does the result obtained also give a clue, perhaps, for a more complete treatment of the rotation problem, so that one can get the correct value of the centrifugal force? Unfortunately, my brain, at least the way it has been trained, is much too feeble for me to answer this question myself, or even to guess from what side it could be attacked. For reasons already discussed, it seems to me that it (?) is of importance for the astronomical problem as well (for until now it at least seemed that a system in which no Coriolis forces flit about could still be a seat of centrifugal forces, or the reverse).[172]

The passage clearly confirms that Besso was aware by spring 1914—and also assumed Einstein to be aware—that the "incomplete treatment" of the problem of rotation (probably referring to the use of an approximation procedure) did not yield the correct result for the Coriolis force. Besso also claimed that the solution to this problem might be relevant for the calculation of the perihelion shift of Mercury (possibly the "astronomical problem" to which he alluded). Einstein did not, however, at this point check the compatibility of his general insights into the covariance properties of the *Entwurf* field equations with concrete calculations on the level of his approximation scheme.

It was only in September 1915 that Einstein rediscovered, to his surprise, the result that the iterative solution of the *Entwurf* field equations does not yield the correct Minkowski metric in rotating coordinates. This is known from a letter he wrote

171 See, e.g., Einstein to H. A. Lorentz, 14 August 1913, (CPAE 5, Doc. 467). Einstein did not allude to anything like having established the Minkowski metric in rotating coordinates as a solution of the *Entwurf* equations (at least in second-order approximation) in his letters to Lorentz nor in those he wrote to other colleagues and friends, who would have been interested in the issue, such as Ernst Mach, Erwin Freundlich, Heinrich Zangger, Paul Ehrenfest, and Michele Besso. See also the extended discussion in "What Did Einstein Know ..." (in vol. 2 of this series).

172 "Gibt das erreichte Resultat vielleicht auch einen Wink für eine vollständigere Behandlung des Drehungsproblems, so dass man den richtigen Wert der Centrifugalkraft bekommen kann? Leider ist mein Kopf, wenigstens so wie er einmal erzogen ist, viel zu schwach, um mir die Frage selbst zu beantworten, oder auch nur zu ahnen, wo man sie angreifen könnte. Aus schon besprochenen Gründen scheint sie (?) mir auch für das astronomische Problem von Bedeutung (weil es früher wenigstens so aussah, ein System in welchem keine Corioliskräfte huschen, doch Sitz von Centrifugalkräften sein könnte, oder umgekehrt)." Michele Besso to Einstein, 20 March 1914, (CPAE 5, Doc. 516, 606).

on 30 September 1915 to Erwin Freundlich, in which he now also connected this finding with the perihelion problem, just as Besso had done in the letter quoted above. Evidently Einstein was quite concerned by his finding:

> I am writing you now about a scientific matter that electrifies me enormously. For I have come upon a logical contradiction of a quantitative nature in the theory of gravitation, which proves to me that there must be a calculational error somewhere within my framework. [...]
>
> Either the equations are already numerically incorrect (numerical coefficients), or I am applying the equations in a principally incorrect way. I do not believe that I myself am in the position to find the error, because my mind follows the same old rut too much in this matter. Rather, I must depend on a fellow human being with unspoiled brain matter to find the error. If you have time, do not fail to study the topic.[173]

The letter leaves open in which context Einstein redid the earlier calculation. It is plausible to assume that it was once more Besso who stimulated the reconsideration of this problem. In fact, Besso and Einstein probably discussed the Mercury as well as the rotation problem during the latter's stay in Switzerland in September 1915.[174] The letter to Freundlich was sent only a week after Einstein's return to Berlin.[175] It was probably written as a reaction to a request for political support and represented one of the first occasions for Einstein to present the revived rotation problem to a colleague who must have been interested in it because of its implication for the understanding of the Mercury problem.[176]

Evidently, this time Einstein found the rotation problem much more alarming than he did in the summer of 1913. In his letter to Freundlich he still did not substantially question the *Entwurf* field equation but merely took into consideration that he applied

173 "Ich schreibe Ihnen jetzt in einer wissenschaftlichen Angelegenheit, die mich ungeheuer elektrisiert. Ich bin nämlich in der Gravitationstheorie auf einen logischen Widerspruch quantitativer Art gestossen, der mir beweist, dass in meinem Gebäude irgendwo eine rechnerische Unrichtigkeit stecken muss. [...] Entweder sind die Gleichungen schon numerisch unrichtig (Zahlenkoeffizienten) oder ich wende die Gleichungen prinzipiell falsch an. Ich glaube nicht, dass ich selbst imstande bin, den Fehler zu finden, da mein Geist in dieser Sache zu ausgefahrene Gleise hat. Ich muss mich vielmehr darauf verlassen, dass ein Nebenmensch mit unverdorbener Gehirnmasse den Fehler findet. Versäumen Sie nicht, wenn Sie Zeit haben, sich mit dem Gegenstande zu beschäftigen." Einstein to Erwin Freundlich, 30 September 1915 (CPAE 8, Doc. 123), extensively discussed in (Janssen 1999), here just a summary,

174 He wrote to Elsa Einstein from Lucerne: "In Zurich I was together with Besso very often; my stay in Zurich was very much improved by it, but thus I neglected my duties to others." ("In Zürich war ich sehr viel mit Besso zusammen, mein Aufenthalt wurde dadurch sehr verschönert, doch vernachlässigte ich so meine Pflicht gegen andere.") Einstein to Elsa Einstein, 11 September 1915, (CPAE 8, Doc. 116).

175 See Calender (CPAE 8, 998).

176 Einstein had written to Freundlich in March of the same year on the perihelion problem, see Einstein to Erwin Freundlich, 19 March 1915 (CPAE 8, Doc. 63). He had written a letter to Lorentz a day after his return from Switzerland, Einstein to H. A. Lorentz, 23 September 1915, (ibid., Doc. 122), in which he did not mention this problem, probably because he was ashamed.

it incorrectly—probably a reference to the unclear status of the approximation procedure—or that some numerical coefficients were wrong.

That the discovery of this flaw, taken by itself, did not amount to a refutation of the *Entwurf* theory is made evident also by Einstein's immediate reaction to the rotation problem. Apparently encouraged by his general results on the covariance properties of the *Entwurf* theory, which, as we have seen, amounted for him to the claim that there was no physical restriction of the generalized relativity principle but only on the choice of admissible coordinates, he attempted to show that the *Entwurf* field equation could be solved by a rotating system in a different set of coordinates; but this attempt failed as well.[177] Shortly afterwards, he developed a new derivation of the *Entwurf* field equation, to which we will turn below, effectively confirming its immunity with regard to the rotation problem. It was only after his return to the November tensor, that he listed the problem of rotation as one of the three flaws which undermined his trust in the *Entwurf* theory.

Einstein's diverse reactions over the course of time to the same problem of the *Entwurf* theory are correlated with his changing perspectives during the elaboration of this theory. When he first believed that his rotation calculation worked, it seemed like progress on the bold approach, without providing him with a general insight into the covariance properties of the *Entwurf* theory. When it turned out that it does not actually work, this negative result became irrelevant because Einstein successfully developed his defensive approach with the supposed consequence that the *Entwurf* field equation is covariant only under linear transformations. When Einstein then, in the second phase of the consolidation period of the *Entwurf* theory, achieved more general insights into its covariance properties, these insights seemed to make a check on the level of concrete calculations superfluous. Eventually, Einstein took it for granted that rotation did not present a problem for the *Entwurf* theory. Only when he reviewed the problem in September 1915, possibly at Besso's prompting, he finally connected his general considerations with his concrete calculations—and rediscovered the problem. This result now questioned not only his earlier conviction concerning rotation, but also more generally the significance of his insights into the theory's covariance properties. Still, the discovery of this failure implied little more than a return to the status of the *Entwurf* theory at the end of the first phase of the consolidation period, its covariance being guaranteed only for general linear transformations.

7.14.4 The Failure of the Covariance Proof

We now turn to the last flaw that Einstein discovered, probably in early October, some weeks before he gave up the *Entwurf* theory. This flaw concerns Einstein's attempt to derive the *Entwurf* field equation along the mathematical strategy. As our earlier discussion of this endeavor has made clear, one of its problems was the neces-

177 See the calculations on the back of the draft of letter Einstein wrote to Otto Naumann after 1 October 1915, (CPAE 8, Doc. 124).

sity to adapt general tools, such as variational calculus, to the requirement of restricted covariance. This aspect had been at the center of Einstein's discussion in early 1915 with the Italian mathematician Tullio Levi-Civita, which apparently was triggered by a letter from Max Abraham.

Abraham may well have been one of the first to discover a problem with Einstein's derivation of the *Entwurf* field equation from a Lagrangian function. On 23 February 1915 he wrote to Levi-Civita:

> Really I did not understand on which hypotheses his new demonstration is based. Among all possible invariants that could be used to construct the [Lagrangian] function H he chooses very arbitrarily the one that yields his field equations.[178]

Abraham thus succinctly summarized the crucial weakness of Einstein's proof.

But Levi-Civita's exchange with Einstein did not touch upon this crucial problem. Levi-Civita focused instead on a specific technical problem in Einstein's derivation; his proof of the claim that the candidate for the left-hand side of the field equations is a tensor.[179] He produced a counter-example which Einstein, however, declared irrelevant by pointing to the fact that Levi-Civita's example does not satisfy the condition of being covariant under the linear transformations that he had explicitly stipulated.[180] As we shall see, he only later would realize the questionable role of this condition in his proof. Levi-Civita, in any case, did not insist on this aspect. Einstein had more difficulties in responding to other problems in his proof to which Levi-Civita drew his attention. In spite of his attempts to rebut the latter's criticism, Einstein eventually had to admit that his derivation was incomplete, without, however, losing faith in it actually fulfilling its purpose in yielding the *Entwurf* field equations.[181] On the contrary, as Einstein wrote to Levi-Civita during their controversy:

> I must even admit that, through the in-depth considerations to which your interesting letters have led me, I have become only more firmly convinced that the proof of the tensor character of $\mathfrak{E}_{\mu\nu}/\sqrt{-g}$ is correct in principle.[182]

Further objections by Levi-Civita did not shatter this conviction. Nevertheless Einstein and Levi-Civita agreed upon a shortcoming of Einstein's proof; eventually when Levi-Civita proposed an alternative gravitation theory involving a scalar gravitational potential, Einstein lost interest in the exchange and broke it off.[183]

Einstein's discovery of the crucial flaw in his proof was not stimulated by Levi-Civita's criticism but by a reconsideration of this proof in light of a paper by Lorentz about six months later. When Einstein returned from Switzerland on 22 September

---

178 Quoted after (Cattani and De Maria 1989b, 184–185).

179 For an extensive discussion, see (Cattani and De Maria 1989b).

180 See Einstein to Tullio Levi-Civita, 5 March 1915 (CPAE 8, Doc. 60).

181 See Einstein to Tullio Levi-Civita, 5 May 1915, (CPAE 8, Doc. 80).

182 "Ich muss sogar gestehen, dass ich durch die tieferen Überlegungen, zu denen mich Ihre interessanten Briefe brachte, noch fester in der Überzeugung wurde, dass der Beweis vom Tensorcharakter von $\mathfrak{E}_{\mu\nu}/\sqrt{-g}$ im Prinzip richtig ist." Einstein to Tullio Levi-Civita, 8 April 1915 (CPAE 8, Doc. 71)

183 See Einstein to Tullio Levi-Civita, 5 May 1915, (CPAE 8, Doc. 80).

1915, he found Lorentz's recently published paper on Hamilton's principle in the theory of gravitation, including a treatment of electromagnetic fields. In a letter to Lorentz dated the following day Einstein wrote:

> Your article delighted me. I have also found a proof for the validity of the [relativistic] energy-momentum conservation principle for the electromagnet. field taking gravitation into consideration, as well as a simplified covariant theoretical representation of the vacuum equations, in which the "dual" six tensor [*Sechserverktor*] concept proves unessential. At the moment I am occupied with studying your paper.[184]

Lorentz's paper on generally-covariant Maxwell theory introduced a generic Hamiltonian principle without deriving Einstein's specific choice from it.[185] In a first reaction to this paper, Einstein attempted to convince Lorentz that the theory of invariants actually leads to such a specific choice. Although the letter in which Einstein expressed this conviction is not preserved, this much can be concluded from a subsequent letter in which Einstein revoked his claim:

> Subsequent reflections on the last letter I sent you have revealed that I made erroneous assertions in that letter. In actual fact the invariant theory method does not yield more than Hamilton's principle when determining your function $Q(= H\sqrt{-g})$.[186]

Evidently, it was the thorough comparison with Lorentz's approach that directed Einstein's attention to a flaw in his derivation of the *Entwurf* field equations.

On reexamining his 1914 derivation, Einstein found that the condition of linearity, which had apparently entered his argument as an unproblematic default setting, was less innocent than it first appeared to him:

> The reason why I did not notice this last year is that on p. 1069 of my article I had frivolously introduced the condition that H was invariant against linear transformation.[187]

As we discussed earlier, it was by introducing this condition that Einstein had found the condition $B_\mu = 0$ for the choice of an adapted coordinate system, while he had derived the condition $\sum_\nu \partial S_\mu^\nu / \partial x_\nu - B_\mu = 0$ as a consequence of energy-momentum conservation—without taking into account the linear covariance of the Lagrangian.

---

184 "Über Ihre Abhandlung habe ich mich sehr gefreut. Ich habe auch einen Beweis für die Gültigkeit des Impuls Energ[ie]satzes des elektromagnet. Feldes mit Berücksichtigung der Gravitation gefunden sowie eine kovariantentheoretisch vereinfachte Darstellung der Vakuumgleichungen, indem sich der Begriff des "dualen" Sechservektors als entbehrlich erweist. Ich bin gerade mit dem Studium Ihrer Arbeit beschäftigt." Einstein to H. A. Lorentz, 23 September 1915, (CPAE 8, Doc. 122).

185 See (Lorentz 1915).

186 "Nachträgliche Überlegungen zu dem letzten Briefe, den ich an Sie richtete, haben gezeigt, dass ich in diesem Briefe Unrichtiges behauptete. Thatsächlich liefert die invariantentheoretische Methode nicht mehr als das Hamilton'sche Prinzip wenn es sich um die Bestimmung der Ihrer Funktion $Q(= H\sqrt{-g})$ handelt." Einstein to H. A. Lorentz, 12 October 1915, (CPAE 8, Doc. 129).

187 "Dass ich dies letztes Jahr nicht merkte liegt daran, dass ich auf Seite 1069 meiner Abhandlung leichtsinnig die Voraussetzung einführte, H sei eine Invariante bezüglich *linearer* Transformationen." (CPAE 8, Doc. 129)

The basic error which Einstein discovered thus consisted in requiring the compatibility of two conditions derived under different assumptions, once with linearity and once without. An acceptable derivation of the *Entwurf* field equation from the Lagrangian formalism had therefore to be based on additional assumptions. In his letter to Lorentz, Einstein reintroduced the correspondence principle in order to justify the selection of the *Entwurf* Lagrangian among the lengthy list of candidates given in his 1914 paper:

> That $Q/\sqrt{-g}$ had been set by me as equal to the fourth expression given there can be justified by the fact that only with this choice does the theory contain Newton's in approximation. That I believed it possible to support this selection on the equation S_μ^λ was based on error.[188]

Einstein's derivation along a mathematical strategy was thus reduced, in its substance, to that of the original 1913 *Entwurf* paper. He had come back to his starting point—but with one important difference: In spite of the failure of his hope to achieve a derivation essentially from covariant theory only, he had effectively found a derivation in which mathematical principles came first and were supplemented by the physical requirements of energy-momentum conservation and Newtonian limit embodied in the default settings of his field equation. In other words, Einstein had established a derivation which follows precisely the pattern of his attempted derivations along the mathematical strategy in the Zurich Notebook. But instead of taking a generally-covariant object suggested by absolute differential calculus as a starting point, Einstein's point of departure was now the variational calculus he had developed himself.

From Einstein's correspondence it becomes clear that he did not yet consider the state of affairs just described as a reason for abandoning the *Entwurf* theory. There is no trace of this in his letter to Lorentz. Also in a letter written to Zangger a few days later,[189] Einstein treated gravitation as one among several topics, clearly considering his current work on it as business as usual:

> It has unfortunately become clear to me now that the "new stars" have nothing to do with the "lens effect," moreover that, taking into account the stellar densities existing in the sky, the latter must be such an incredibly rare phenomenon that it would probably be futile to expect one of the like.[190]

---

188 "Dass $Q/\sqrt{-g}$ von mir gleich dem vierten der dort angegebenen Ausdrücke gesetzt werde, lässt sich dadurch rechtfertigen, dass die Theorie nur bei dieser Wahl die Newton'sche als Näherung enthält. Dass ich glaube, diese Auswahl auf die Gleichung S_μ^λ stützen zu können, beruhte auf einem Irrtum." (CPAE 8, Doc. 129)

189 The letter was dated by the editors of (CPAE 8) as 15 October 1915, but since it is explicitly dated only as "Friday" and other indications leave a window between 30 September and 21 October, it may well have been written on 8 October, i.e. before the letter to Lorentz.

190 "Seit ich hier bin, habe ich sehr fest auf meiner Bude gearbeitet. Es ist mir nun leider klar geworden, dass die "neuen Sterne" nichts mit der "Linsenwirkung" zu thun haben, das ferner letztere mit Rücksicht auf die am Himmel vorhandenen Sterndichten ein so ungeheuer seltenes Phänomen sein muss, dass man wohl vergeblich ein solches erwarten würde. Ich schrieb eine ergänzende Arbeit zu meiner letztjährigen Untersuchung über die allgemeine Relativität. Gegenwärtig arbeite ich etwas in Wärmetheorie." Einstein to Heinrich Zangger, 15 October 1915, (CPAE 8, Doc. 130).

Einstein evidently considered what he believed to be the impossibility of relating the observation of nova stars to the gravitational lensing effect he had predicted in 1912[191] to be more important than the problem he had discovered in the mathematical proof of the *Entwurf* theory. Even the paper that he mentions as being supplementary to his 1914 review is probably not a reference to a planned revocation of his proof[192] but to a paper mentioned in his earlier letter to Lorentz on generally-covariant electrodynamics, which was eventually published in 1916 (Einstein 1916b). Thus, also the third flaw which Einstein discovered in the *Entwurf* theory did not immediately lead to its refutation.[193] It nevertheless must have had a subversive effect on his belief in this theory as will be discussed in the next section.

In summary, we have seen that the endeavor to develop a mathematical strategy for deriving the *Entwurf* field equation had to face tensions between the non-covariance of the theory and the properties of a mathematical apparatus naturally tuned to generally-covariant objects. These tensions implied the necessity to repeatedly rework the mathematical analysis of the covariance properties of the theory and the derivation of the field equation based on it; they became particularly evident in the criticism by Abraham and Levi Civita. The latter's critique, however, never questioned the goal of Einstein's derivation, to show the supposed uniqueness of the *Entwurf* field equation. Einstein's exposition of his theory to the Göttingen mathematicians and physicists in a series of six Wolfskehl lectures, delivered in late June and early July 1915 at the invitation of David Hilbert,[194] may also have induced a renewed reflection on the justification of the *Entwurf* theory along the mathematical strategy.

The problematic character of Einstein's derivation of the *Entwurf* field equation from a Lagrangian formalism may have emerged, as we have also seen, in the context of applying the formalism to a different context, that of a generally-covariant Maxwell theory. It was here that the formalism developed specifically for the *Entwurf* theory first proved its greater generality. Against this background Einstein's discovery of a flaw in his derivation could have had a double effect: It pointed out the familiar physical arguments for deriving the *Entwurf* theory, and it suggested taking up once more the mathematical strategy of trying out the newly empowered techniques on different candidates. Initially Einstein chose the first option, but his rederivation of the *Entwurf* theory from a variational formalism plus one extra physical condition may

---

191 See (Renn and Sauer 2003b).

192 As conjectured in (Janssen 1999, note 51).

193 We leave aside here the question of a possible, direct or indirect, influence by David Hilbert, who had also found a flaw in Einstein's reasoning around this time. It is, however, unclear whether and if so when and what Einstein may have learned about Hilbert's work prior to 7 November 1915 when their extant correspondence on this issue begins. See (Corry 2004, ch. 7) and further references cited therein for a discussion of the interaction between Einstein and Hilbert in this period. See also the discussion below in sec. 7.18.1

194 Fragments of an auditor's notes taken during one of the lectures are published in (CPAE 6, Appendix B). For further historical discussion of Einstein's Göttingen visit in June and July 1915, see (Corry 2004, pp. 320–329).

have made it clear to him how much more general this formalism was, particularly since the conservation principle no longer presented a major obstacle.

7.15 Einstein's November Revolution: the Restoration of an Old Candidate

7.15.1 Looking Back in Anger and Hope

With his publication of 11 November 1915, submitted on the 4th of November, Einstein made his definite rejection of the *Entwurf* theory public:

> My efforts in recent years were directed toward basing a general theory of relativity, also for nonuniform motion, upon the supposition of relativity. I believed indeed to have found the only law of gravitation that complies with a reasonably formulated solution in a paper that appeared last year in the *Sitzungsberichte*.
> Renewed criticism showed to me that this truth is absolutely impossible to show in the manner suggested. That this seemed to be the case was based upon a misjudgment.[195]

Among the three major flaws he had meanwhile found in the *Entwurf* theory, the Mercury failure, the rotation failure, and the breakdown of its mathematical derivation, the latter was publicly the most visible, documented as it was by Einstein's lengthy 1914 review paper. It was, in any case, the only failure explicitly mentioned in his first 1915 article:

> The postulate of relativity—*as far as I demanded it there*—is always satisfied if the Hamiltonian principle is chosen as a basis. But in reality, it provides no tool to establish the Hamiltonian function H of the gravitational field.[196]

As we have seen, none of the problems of the *Entwurf* theory, taken by themselves or together, resulted in an immediate rejection of this theory. Therefore it is not self-evident that Einstein, in view of these problems, finally decided to give up the *Entwurf* theory, as he pointed out in the 1915 paper:

> For these reasons I lost trust in the field equations I had derived, and instead looked for a way to limit the possibilities in a natural manner.[197]

---

195 "In den letzten Jahren war ich bemüht, auf die Voraussetzung der Relativität auch nicht gleichförmiger Bewegungen eine allgemeine Relativitätstheorie zu gründen. Ich glaubte in der Tat, das einzige Gravitationsgesetz gefunden zu haben, das dem sinngemäß gefaßten, allgemeinen Relativitätspostulate entspricht, und suchte die Notwendigkeit gerade dieser Lösung in einer im vorigen Jahre in diesen Sitzungsberichten erschienenen Arbeit darzutun.
Eine erneute Kritik zeigte mir, daß sich jene Notwendigkeit auf dem dort eingeschlagenen Wege absolut nicht erweisen läßt; daß dies doch der Fall zu sein schien, beruhte auf Irrtum." (Einstein 1915c, 778)

196 "Das Postulat der Relativität, *soweit ich es dort gefordert habe*, ist stets erfüllt, wenn man das Hamiltonsche Prinzip zugrunde legt; es liefert aber in Wahrheit keine Handhabe für eine Ermittelung der Hamiltonschen Funktion H des Gravitationsfeldes." (Einstein 1915c, 778)

197 "Aus diesen Gründen verlor ich das Vertrauen zu den von mir aufgestellten Feldgleichungen vollständig und suchte nach einem Wege, der die Möglichkeiten in einer natürlichen Weise einschränkte." (Einstein 1915c, 778)

This step cannot be exclusively accounted for, we believe, on the basis of the failures of that theory but only becomes plausible in view of the unexploited resources that he still had at his disposal from his earlier work in the Zurich Notebook. This is also suggested by the sentences following immediately in the paper:

> In this pursuit I arrived at the demand of general covariance, a demand from which I parted, though with a heavy heart, three years ago when I worked together with my friend Grossmann. As a matter of fact, we were then quite close to that solution of the problem, which will be given in the following.[198]

In fact, after abandoning the *Entwurf* theory Einstein returned to one of the mathematical objects he had encountered along the mathematical strategy three years ago in the Zurich Notebook, the November tensor (cf. eq. (82)). As the above passage suggests, he did not insist on the specific requirement of general covariance, but merely required a "more general covariance of the field equations." It was thus, above all, a return to the mathematical strategy applied to the absolute differential calculus that marked the turning point of 11 November 1915, rather than a radical break with his earlier experiences concerning the restriction of covariance if it turned out to be necessary. Indeed, Einstein now emphatically embraced the absolute differential calculus:

> Nobody who really grasped it can escape from its charm, because it signifies a real triumph of the general differential calculus as founded by Gauss, Riemann, Christoffel, Ricci, and Levi-Civita.[199]

In the context of our account, the crucial questions for understanding Einstein's shift in late 1915 are:

1. What eventually convinced him to give up the *Entwurf* theory and return to the mathematical strategy applied to the absolute differential calculus?
2. How could *any* of the tensors explored and discarded in the course of Einstein's work on the Zurich Notebook now again represent a resource for a renewed exploration?
3. What made the November tensor particularly suitable for such a renewed exploration? And, finally:
4. How exactly did Einstein find his way back to the November tensor?

The answer to the first question follows from our analysis of the third flaw Einstein discovered in the *Entwurf* theory, the erroneous derivation. His last derivation of the *Entwurf* field equation had effectively reinstalled the mathematical strategy in its

---

198 "So gelangte ich zu der Forderung einer allgemeineren Kovarianz der Feldgleichungen zurück, von der ich vor drei Jahren, als ich zusammen mit meinem Freunde Grossmann arbeitete, nur mit schwerem Herzen abgegangen war. In der Tat waren wir damals der im nachfolgenden gegebenen Lösung des Problems bereits ganz nahe gekommen." (Einstein 1915c, 778)

199 "Dem Zauber dieser Theorie wird sich kaum jemand entziehen können, der sie wirklich erfaßt hat; sie bedeutet einen wahren Triumph der durch Gauss, Riemann, Christoffel, Ricci und Levi-Civiter begründeten Methode des allgemeinen Differentialkalküls." (Einstein 1915c, 779)

original sense, starting from a generally-covariant object which is then checked and, if necessary, modified according to physical criteria. However, both the starting point and the check list of criteria now looked somewhat different from how they did in 1912–1913. If Einstein's starting point had then been a second-rank tensor representing a candidate for the left-hand side of the field equation, it was now a scalar representing the Lagrangian of the theory. And if the conservation principle was then a criterion that had to be laboriously checked for each single candidate, it was now automatically fulfilled for any candidate fitting into the general framework. The correspondence principle therefore remained the crucial criterion for choosing the right candidate. If Einstein at any point after his letter to Lorentz of 12 October 1915 decided to actually check his claim that the *Entwurf* theory was determined uniquely by this criterion, his search would have been governed by the renewed mathematical strategy. In a word, checking the *Entwurf* theory and pursuing the mathematical strategy simply coincided in the end.

To answer the second question of why, in general, tensors earlier discarded could now be considered worthy of further examination, we must turn once again to the Zurich Notebook. Einstein's examination in the notebook of candidate gravitation tensors extracted from the Riemann tensor was restricted to the weak-field form of the field equation. In the notebook, Einstein mastered energy-momentum conservation only for weak fields—with the exception of the *Entwurf* operator at the end of the notebook. Therefore, all candidates extracted from the Riemann tensor were left only partially explored in the notebook when Einstein decided to move on to the next candidate. This unexplored potential of the candidates encountered along the mathematical strategy was one of the essential reasons why he considered it worthwhile to reexamine them in 1915. Another crucial reason for such progress in a loop—or by reflection—was the fact that Einstein's renewed mathematical strategy now drew on more resources, in particular, the variational calculus as applied for the *Entwurf* theory.

In answering the third question of what made the November tensor particularly suitable for a renewed exploration, we again must look at the Zurich Notebook. There Einstein, with Grossmann's help, had derived the November tensor from the Ricci tensor under the stipulation of unimodular coordinate transformations. Contrary to the Ricci tensor, the November tensor satisfies Einstein's physically motivated criteria because it could be reduced to a form suitable for obtaining the Newtonian limit by assuming a coordinate restriction (the Hertz restriction) that was the same as the restriction required by energy-momentum conservation in the weak field limit (cf. eq. (LXXXIV)). Contrary to the Einstein tensor, the harmonization of these two restrictions for the November tensor did not require a change in the view of how the Newtonian limit was to be achieved. Clearly, the November tensor offered a natural starting point for a renewal of the mathematical strategy, since in its case the check of the weak field equation with regard to the conservation and correspondence principles produced a positive outcome—shadowed only by the restricted covariance properties of the candidate that resulted from the November tensor by imposing the required coordinate restriction. The questions that remained were whether or not this result could be extended to the full field equation as well, and which restrictions of covariance were implied for the full equation by the conservation principle. These questions could now be addressed with the help of an improved mathematical apparatus.

The answer to the fourth question, concerning the actual path that Einstein took in rediscovering the November tensor in the fall of 1915 is suggested by several comments pointing to the crucial role of the default setting for the gravitational field eq. (XXII) which, in hindsight, played the role of a "fateful prejudice" with its substitution by the default setting eq. (XXIII) being the "key to the solution."[200] The path leading from the *Entwurf* field equation to a field equation based on the November tensor can be reconstructed with fair confidence in view of the default setting for the Lagrangian in terms of the field eq. (LXIII). It was, as we have discussed, this default setting, rooted in classical field theory, which had also made the *Entwurf* Lagrangian look particularly promising.

It must have been tempting for Einstein to look for other Lagrangians that could be interpreted in this way as involving a "square" of the gravitational fields, experimenting with the definition of the gravitational field. In fact, the internal logic of the mathematical representation exerted a pressure on Einstein's interpretation, since in that representation the connection coefficients, i.e. the Christoffel symbols, have a central importance in, e.g. the concept of a covariant derivative, the geodesic equations, or the definition of the Riemann and Ricci tensors. This role of the connection coefficients was at odds with the significance that Einstein attached to the coordinate derivatives of the metric. Thus, when Einstein took the mathematical tradition more seriously again, the mathematical knowledge that was accumulated in the representation, forced him to reconsider his physical prejudices. And it was indeed not far fetched to reinterpret the Christoffel symbol as representing the field, a choice that almost immediately leads to the Lagrangian of the November theory. Einstein's earlier experience, documented in the Zurich Notebook, might have helped find this path from the *Entwurf* to the November theory because at that time he had already explored the relation between gravitation tensors expressed by Christoffel symbols and their expression in terms of the derivatives of the metric, e.g. in the context of studying the so-called "theta-condition."[201]

In summary, in the fall of 1915 Einstein succeeded in combining insights from his earlier mathematical strategy and canonical mathematical knowledge with the achievements of the physically motivated *Entwurf* theory. He had probably omitted the November tensor from the Zurich Notebook because he lacked the mathematical means to build a full-scale theory around it, in particular, with regard to the implementation of the conservation principle. His unsuccessful attempt at deriving the *Entwurf* theory from a mathematical strategy had laid just those means in his hands. The failure of the mathematical derivation of the *Entwurf* theory left Einstein with a formalism that initially seemed tailor-made for this very purpose, but then turned out to

200 See "Untying the Knot ..." (in vol. 2 of this series).
201 See the "Commentary …" (in vol. 2 of this series).

be much more generally applicable. An attempt to rederive the *Entwurf* field equation within this formalism turned almost automatically into a renewal of his search along the mathematical strategy. Einstein's physical expectations, not only of the correspondence and the conservation principles, but also of the role of the gravitational field in the Lagrangian and in the equation of motion, must have quickly led him to identify the November tensor as the most appropriate candidate, which not only was probably the easiest to handle given Einstein's propensity for unimodular coordinates but for which the implications of the mathematical and the physical strategies seemed to coincide.

7.15.2 Removing an Old Stumbling Block and Encountering a New One: The Conservation Principle in 1915

The demonstration of energy-momentum conservation for a theory based on the November tensor and the representation of the gravitational field by the Christoffel symbols were, for Einstein, the hallmark of the turnaround in November 1915. His contemporary comments and later recollections not only confirm that his earlier rejection of candidates derived from the Riemann tensor was just as much associated with the difficulty in demonstrating the validity of the conservation principle as with difficulties related to the correspondence principle.[202] They also confirm that it was his revision of the understanding of the components of the gravitational field that was a crucial turning point associated with his return to the mathematical strategy in November 1915.[203]

The decisive progress from Einstein's earlier exploration of a theory based on the November tensor was made possible by the Lagrangian formalism that allowed him to demonstrate that the theory complies with the conservation principle. He thus succeeded in removing an old stumbling block that had earlier forced him to abandon the November tensor, as well as expressions based on it, in the Zurich Notebook. However, the problem of establishing the compatibility between conservation and relativity principles, which had also been a problem with the November tensor in 1912–1913, continued to challenge him even in 1915 when it presented new insights as well as a new obstacle.

The most consequential new insight was related to the fact that the coordinate restriction resulting from the conservation principle turned out to have a remarkably simple structure, being reduced to the requirement that a certain scalar function is constant (cf. eq. (83)). Instead of the usual four equations, Einstein merely obtained a single condition from the requirement of energy-momentum conservation. In spite of this simplification, his physical interpretation of this condition did not change; he still saw it as defining adapted coordinates and admissible transformations (Einstein 1915c, 785).

---

202 Einstein to Michele Besso, 10 December 1915, (CPAE 8, Doc. 162) quoted in the introduction.
203 For extensive discussion, see "Untying the Knot …" (in vol. 2 of this series).

Due to the technical novelties of the November theory, Einstein's perspective on the role of coordinate restrictions changed. Now the requirements arising from the conservation principle and those related to the correspondence principle began to play different roles. The first kind of requirements only led to a minimal but still global constraint on the choice of coordinate systems, the second kind of requirements essentially fixes the coordinate system—but now only in the context of a specific physical situation without global implications. In his paper, Einstein for the first time introduced coordinate conditions in this modern sense albeit without any further explanation. He simply made use of the opportunity that the formalism of the November theory had opened up for him.

Considering his earlier failures, the November theory implemented Einstein's heuristic requirements without requiring much of an adjustment of these requirements. What had changed was, as we have seen, the default setting for the representation of the gravitational field. Furthermore, the way in which the conservation principle had earlier fully determined adapted coordinate systems was, as we have also seen, now changed into a weak constraint that could be harmonized with the requirements of the correspondence principle, thus giving rise to the idea of coordinate conditions in the modern sense. On the other hand, what had not changed was the view of the conservation principle as imposing additional conditions on the choice of coordinates. Finally, the way in which the Newtonian limit is attained in the November theory, that is, via a weak field equation of the form of eq. (33), also remained, by default, the same. That Einstein, at the time, did not regard the November theory as a first step towards a more complete theory—as it must appear to a modern reader. This is evident by a letter he wrote to his son on the day he submitted his first November paper, the 4th of November 1915:

> In the last few days I completed one of the finest papers of my life; when you are older I'll tell you about it.[204]

The formalism of the November theory generated one, apparently minor novelty that could *not* easily be assimilated to Einstein's expectations and that therefore called for a physical interpretation. This new stumbling block was the scalar condition for the choice of adapted coordinates eq. (84). In view of its derivation from the conservation principle, it could not have been surprising to Einstein that this condition determines the choice of adapted coordinates by the properties of the stress-energy tensor of matter.[205] But this general argument does little to make the precise way of this determination plausible, let alone make it understandable that coordinate systems for which $\sqrt{-g} = 1$ are to be excluded, as is implied by Einstein's condition. This was a point where the new formalism of the November theory confronted him with the challenge

204 "Dieser Tage habe ich eine der schönsten Arbeiten meines Lebens fertig gestellt; wenn Du einmal grösser bist, erzähle ich Dir davon." Einstein to Hans Albert Einstein, 4 November 1915, (CPAE 8, Doc. 134).

205 The very existence of an additional coordinate restriction must have been a puzzle in view of Einstein's earlier insights into the relation between covariance and conservation.

to find an adequate physical interpretation. Einstein did not hesitate to take up this challenge and, less than a week after the submission of his first November paper, submitted a short addendum dedicated to the physical interpretation of this condition.

7.16 A Familiar Candidate in a New Context: Einstein's Return to the Ricci Tensor

Einstein's addendum to his first November paper was submitted on 11 November and published on 18 November 1915 (Einstein 1915d). It does not contain a single novel formula with respect to the earlier paper but merely constitutes a reinterpretation of what had been achieved. Yet, it introduced a new, now generally-covariant field equation, replacing that of the November theory, which was covariant only for unimodular transformations. The new field equation is instead based on the Ricci tensor, a candidate that Einstein had also considered earlier while working on the Zurich Notebook (cf. eq. (55)). How did he reinterpret his earlier results?

The point of departure for this reinterpretation was the scalar condition for the choice of adapted coordinates eq. (84). This equation, together with the requirement of unimodularity, were the only obstacles, it appeared, that separated Einstein from the realization of general covariance. Furthermore, there was no general physical interpretation for these two requirements: Why should coordinate transformations be unimodular and why should it nevertheless be impossible to select a coordinate system so that $\sqrt{-g} = 1$? These must have been questions motivating Einstein's further search, beyond what he had achieved in his first November paper. In the new approach presented in the addendum, these two questions were answered in the context of an issue that at first glance appears to be unrelated to gravitation theory; the question of the fundamental constitution of matter.

In the introductory part of his addendum, Einstein discussed a contradiction arising in an electromagnetic theory of matter. He argued that the inclusion of gravitation in the energy-momentum balance could resolve, at least in principle, the following contradiction: The hypothesis that all matter is of electromagnetic origin, and Maxwell's equations imply that the trace of the energy-momentum tensor vanishes:[206]

$$\sum T^{\mu}_{\mu el} = 0. \tag{88}$$

It is also clear that for the default setting of the source-term, i.e. pressureless dust (cf. eq. (4)), the trace of the energy-momentum tensor does not vanish. The conflict between this implication and eq. (88) seems to indicate that matter if conceived of as pressureless dust cannot be constructed on an electromagnetic basis.

However, it is possible to conceive the energy-momentum tensor as being composed of two parts, as is suggested by the parallelism of the energy-momentum of

206 Cf. (Laue 1911, § 13).

matter and of the gravitational field on right-hand side of the field equation (cf. eq. (XLI)):

$$\sum_{\mu} T^{\mu}_{\mu} = \sum_{\mu} (T^{\mu}_{\mu el} + t^{\mu}_{\mu grav}) = \sum_{\mu} (0 + t^{\mu}_{\mu grav}), \tag{89}$$

where $T^{\mu}_{\mu el}$ is due to the electromagnetic origin of matter and $t^{\mu}_{\mu grav}$ to gravitational fields, which are now assumed to play a role in the constitution of matter as well. It follows that the non-vanishing trace of the energy-momentum tensor for matter no longer necessarily contradicts with eq. (88) since it seems possible that the vanishing of $T^{\mu}_{\mu el}$ is compensated by $t^{\mu}_{\mu grav}$. In other words, the additional assumption that gravitational fields play a role in the constitution of matter might be considered as hinting at the solution of a puzzle in a purely electromagnetic theory of matter (Einstein 1915d, 800).[207]

The discussion of an electromagnetic theory of matter in the introductory part of Einstein's addendum raises the obvious question of its function in his theory of gravitation. In his introductory paragraph he stakes the following claim:

> In a recent investigation I have shown how Riemann's theory of covariants in multidimensional manifolds can be utilized as a basis for a theory of the gravitational field. I now want to show here that an even more concise and logical structure of the theory can be achieved by introducing an admittedly bold additional hypothesis on the structure of matter.[208]

The hypothesis of an electromagnetic constitution of matter on the basis of "a theory more complete than Maxwell's theory" allowed Einstein to invalidate an important implication regarding the source term of the gravitational field equation—its non-vanishing trace. In his first November paper, this default assumption had forced him to introduce the coordinate condition that excluded coordinate systems with $\sqrt{-g} = 1$. Even earlier, in the Zurich Notebook, he had discarded the Ricci tensor as a candidate for the left-hand side of a gravitational field equation (if only on the level of linear approximation) because this candidate implies the vanishing of the trace of the stress-energy tensor in contrast to the default properties of Einstein's standard model of matter (cf. eq. (LXXV)). The hypothesis of an electromagnetic origin of

---

207 Apart from the fact that Einstein does not elaborate on the question as to how matter might be conceived of on the basis of "einer gegenueber Maxwells Theorie vervollstaendigten Elektrodynamik" (p. 800), Einstein's suggestion suffers, however, from a rather conspicuous difficulty: Since t^{μ}_{ν} is no tensor but a coordinate-dependent expression, it can in fact not replace the stress-energy tensor of matter. In particular, the claim that the coordinate-dependent expression t^{μ}_{μ} is positive remains unproven in Einstein's paper and can be refuted in a rather simple way by considering a coordinate system in which this quantity vanishes as well. See (Earman and Glymour 1978, 298).

208 "In einer neulich erschienenen Untersuchung habe ich gezeigt, wie auf Riemanns Kovariantentheorie mehrdimensionaler Mannigfaltigkeiten eine Theorie des Gravitationsfeldes gegründet werden kann. Hier soll nun dargetan werden, daß durch Einführung einer allerdings kühnen zusätzlichen Hypothese über die Struktur der Materie ein noch strafferer logischer Aufbau der Theorie erzielt werden kann." (Einstein 1915d, 799)

matter made it possible to resolve both these problems at the same time. In fact, under the condition $\sqrt{-g} = 1$, admissible under this hypothesis, the November tensor coincides with the generally-covariant Ricci tensor. Together with the generalized principle of relativity, this mathematical feature led Einstein to choose the Ricci tensor rather than the November tensor as the more appropriate candidate for the left-hand side of a gravitational field equation. With the introduction of the generally-covariant Ricci tensor, the other problem of the November theory—its restriction to unimodular coordinate systems—disappeared.

All aspects of the new Ricci theory are simply straightforward consequences of the November field equation plus the condition $\sqrt{-g} = 1$ which can now be conceived of as a coordinate condition in the sense that its stipulation does not affect the physical validity of the equations. In particular, Einstein did not present a new derivation of the new field equation from a Hamiltonian variation principle, now to be formulated for the Ricci tensor. He did not write down the free field Lagrangian that would produce the Ricci tensor in the field equation. Instead, Einstein still used the technique of reducing his gravitational field equation, using the condition $\sqrt{-g} = 1$, to the November field equation in his conclusive 1915 paper, and even in the 1916 review paper on general relativity. Similarly, in his addendum, he neither provided an independent discussion of energy-momentum conservation nor of the Newtonian limit, but just assumed that everything would carry over unchanged from the November theory. In his paper, he explicitly claimed that the physically relevant relations remain unchanged by the transition from the November to the Ricci theory:

> Based upon this system one can—by retroactive choice of coordinates—return to those laws which I established in my recent paper, and without any actual change in these laws, ...[209]

He emphasized that the only difference was the increased freedom in choosing a coordinate system:

> The only difference in content between the field equations derived from general covariance and those of the recent paper is that the value of $\sqrt{-g} = 1$ could not be prescribed in the latter.[210]

Even more radically, Einstein claimed in his letter to Hilbert of 12 November that his latest modification implied that Riemann's tensor would now directly produce the gravitational equations but would not change the equations of the theory.[211] In short, his new generally-covariant field equation based on the Ricci tensor represented for

209 "Von diesem System aus kann man durch nachträgliche Koordinatenwahl leicht zu dem System von Gesetzmäßigkeiten zurückgelangen, welches ich in meiner letzten Mitteilung aufgestellt habe, und zwar ohne an den Gesetzen tatsächlich etwas zu ändern." (Einstein 1915d, 801)

210 "Der Unterschied zwischen dem Inhalte unserer aus den allgemein kovarianten gewonnenen Feldgleichungen und dem Inhalte der Feldgleichungen unserer letzten Mitteilung liegt nur darin, daß in der letzten Mitteilung der Wert für $\sqrt{-g} = 1$ nicht vorgeschrieben werden konnte." (Einstein 1915d, 801)

211 Einstein to David Hilbert, 12 November 1915, (CPAE 8, Doc. 139).

Einstein largely a reinterpretation of his earlier results from the November theory.[212] Only the condition $\sqrt{-g} = 1$ had changed its status from being an excluded special case to a key relation for translating results from the older theory into the new one. What had also changed was the physical interpretation of the theory, in particular with regard to its implications for physics outside of gravitation theory.

Concerning Einstein's new gravitation theory, the only significant property of his new model of matter that replaces his standard instantiation of the slot for the source term of pressureless dust is the vanishing of the trace of the stress-energy tensor. All other aspects of such a theory of matter were irrelevant. As we have seen, he had merely two tenuous arguments to support his audacious new approach. According to the first argument, the inclusion of gravitation in an electromagnetic theory of matter could help to avoid the conflict between the vanishing of the trace of the stress-energy tensor for electromagnetic fields and its non-vanishing for matter. This argument was, however, only an unelaborated idea and quite problematic. According to the second argument, an electromagnetic theory of matter was rendered plausible by the greater consistency of the theory of gravitation that it made possible.

Since all essential equations, according to Einstein's assertion, remain the same in the Ricci and November theories, the question of which theory was to be given preference was thus a matter of choice between the following two options:

1. to rely on a standard model of matter and to accept the physically unmotivated restriction of the theory to unimodular transformations and an inexplicable exclusion of certain coordinate systems (November theory);
2. to achieve a generally-covariant theory without special requirements on coordinate systems and with a logically simple structure, but to accept the introduction of non-trivial consequences for a highly problematic fundamental theory of matter (Ricci theory).

Einstein's preference for the second option was affected by the context in which he formulated his new approach, in particular, by the contemporary discussion about an electrodynamic worldview and the parallel work of David Hilbert, which constituted serious competition for Einstein.[213] The context of this discussion lent some credibility to the introduction of speculative assumptions about a fundamental theory of matter.

---

212 In his paper, Einstein did not address the conflict between the Ricci tensor and the correspondence principle, cf. (Stachel 1989; Norton 1984). This conflict was somewhat hidden by the fact that the physical consequences of the Ricci theory were elaborated in terms of the November theory in which the Newtonian limit can be attained via the Hertz condition which is not in conflict with Einstein's default assumption about the static metric. The conflict here arises from to the condition $\sqrt{-g} = 1$, mediating between the two theories and implying, together with the Hertz condition, the harmonic condition. It seems that Einstein was, either at that time or when working on the Zurich Notebook, not aware of this conflict.

213 See (Corry, Renn and Stachel 1997; Sauer 1999).

7.17 The Mercury Problem as a Theoretical Laboratory for the Ricci Tensor

7.17.1 Einstein's Motivation

Only seven days after his last note, on 18 November 1915, Einstein presented an application of his newly found field equation, his "Explanation of the Perihelion Motion of Mercury from the General Theory of Relativity" (*Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie*) to the Academy. It was the only one of his November papers which he submitted as a manuscript to the assembly of the Academy accompanied by oral comment, as is documented by the protocols.[214] He saw the excellent agreement between his calculated value of the perihelion shift of Mercury (43") and astronomical observations (45" +/– 5") as a breakthrough for his new theory. Einstein may also have commented publicly on his note because he hoped that his achievement would attract the attention of the astronomers attending his presentation to the Academy, such as Karl Schwarzschild.[215]

Einstein achieved his result in only seven days, a very short time for an involved calculation. David Hilbert showed himself impressed by Einstein's rapid success:

> Many thanks for your postcard and cordial congratulations on conquering perihelion motion. If I could calculate as rapidly as you, in my equations the electron would correspondingly have to capitulate, and simultaneously the hydrogen atom would have to produce its note of apology about why it does not radiate.[216]

Einstein must have been eager to quickly find convincing physical consequences for his new theory for three main reasons:[217]

1. he had produced several candidate theories among which no definite decision had yet been possible; the *Entwurf* theory, the November theory, and the Ricci theory,
2. he was in close competition with Hilbert who had just sent him a manuscript about his own gravitational field theory and had to make an effort in order to secure his priority, and
3. he may have been looking for further confirmation for his bold hypothesis of a combined electromagnetic and gravitational origin of matter, a hypothesis which so far had been based mainly on reasons of internal consistency or on general philosophical arguments.

All three motivations for Einstein's concern with the perihelion problem are well documented by his contemporary correspondence with Hilbert.[218] Einstein consid-

---

214 See (Archiv der Berlin-Brandenburgischen Akademie der Wissenschaften, II–V, vol. 91, sheet 64).

215 See previous note, and for historical discussion, see (Renn, Castagnetti and Damerow 1999).

216 "Vielen Dank fuer Ihre Karte und herzliche Gratulation zu der Ueberwältigung der Perihelbewegung. Wenn ich so rasch rechnen könnte wie Sie, muesste bei meinen Gleichg entsprechend das Elektron kapituliren und zugleich das Wasserstoffatom seinen Entschuldigungszettel aufzeigen, warum es nicht strahlt." David Hilbert to Einstein, 19 November 1915, (CPAE 8, 149).

217 See also the discussion in (Earman and Janssen 1993), on which the following relies.

218 See Einstein to David Hilbert, 12 November 1915, Einstein to David Hilbert, 18 November 1915, and David Hilbert to Einstein, 19 November 1915, (CPAE 8, Docs. 139, 148, 149).

ered the Mercury calculation as a piece of evidence in favor of an electromagnetic theory of matter is also confirmed by the abstract of his paper in the Academy proceedings:

> It is shown that the general theory of relativity explains qualitatively and quantitatively the perihelion motion of Mercury, which was discovered by Leverrier. The hypothesis of the vanishing of the stress-energy tensor of matter is thus confirmed. Furthermore, it is shown that the examination of the bending of light rays in the gravitational field makes it also possible to verify this important hypothesis.[219]

7.17.2 The Advantages of a Second Attempt

What enabled Einstein to check this physical consequence of the anomalous perihelion advance of Mercury on the basis of the new field equation so rapidly was his earlier attempt in 1913, undertaken jointly with Michele Besso, to calculate the perihelion shift for the *Entwurf* theory.[220] This earlier attempt had given him a quantitative result that is too small (18") if compared to the empirical value. But this attempt had given Einstein the tools that could now be applied without any essential modification to the new field equation based on the Ricci tensor. The additional resource which the earlier work had laid in Einstein's hands not only allowed him to achieve quick success by applying his new theory to a challenging problem. Even more remarkably, this application also had a profound repercussion on the applied theory itself. The employment of the Mercury calculation scheme in the context of the Ricci theory effectively changed the heuristic criteria of Einstein's search for the field equation and resulted in a more sophisticated understanding of the correspondence principle.

This far-reaching consequence emerged only after Einstein's theory was explored in greater depth with his new calculation of the Mercury problem. His first two November papers were short and contain hardly any discussion of the physical consequences of the postulated field equations. The addendum of November 11 refers entirely to the considerations, including the Newtonian limit, that are presented in the main paper of November 4. The study of the Mercury problem hence constitutes the first elaboration of the Ricci theory, giving it a justification beyond the field equation and its immediate consequences. This holds even if one takes into account Einstein's earlier consideration of this theory in the Zurich Notebook. But now, in mid-November 1915, a fully fledged calculation scheme permitted the determination of approximate solutions to the gravitational field equation. This calculation scheme was inherited form the earlier calculation of the Mercury problem in the context of the

---

219 "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." (Einstein 1915e)

220 See (CPAE 4, Doc. 14; Earman and Janssen 1993).

Entwurf theory; it had been developed after Einstein's struggle with various candidate field equations in the Zurich Notebook at a time when he believed in the validity of the *Entwurf* theory. This scheme had therefore not yet been applied to different candidate field equations and had thus not have had effect on the balance between Einstein's heuristic criteria. The lasting impact of the Mercury problem on the development of the field equations of general relativity in 1915 was to provide the judgement about candidate field equations with knowledge about the formalism of a gravitational field theory that was essentially independent of the field equations and that had been acquired as early as 1913. The accumulation of this knowledge triggered a process of reflection which guided Einstein to the definite field equation of general relativity.

Einstein's calculation of Mercury's perihelion shift was based on finding an approximate solution to the gravitational field equation by an iterative procedure. To find the solution of first order, Einstein and Besso, in 1913, turned directly to the first-order field equation with which Einstein was familiar from his consideration of the Newtonian limit (cf. eq. (33)), (CPAE 4, 360). The solution to this equation was hence given in terms of the canonical metric for a static field (25) which Einstein used to obtain the Newtonian limit. To obtain the second approximation, Einstein and Besso wrote down the general form of a spherically symmetric metric in Cartesian coordinates in terms of three unknown functions so as to immediately satisfy one of the constraints of the problem with an appropriate ansatz (CPAE 4, 364). These functions were then determined by the iterative procedure, starting from the first approximation.

In his 1915 paper, Einstein no longer proceeded in two separate steps but immediately started from the generic ansatz for a spherically symmetric metric (Einstein 1915b, 833). In 1915, this approach was not only natural but also necessary. It was natural because the procedure Einstein and Besso had constructed in 1913 worked just as well for the first as for the second approximation so that there was really no reason for proceeding in two steps as they had done when first developing their method for calculating the Mercury problem. In 1915 it was necessary to begin right away with the second step since the first step of 1913 no longer worked for the field equation of the Ricci theory. While Einstein's default assumption about the metric for a static field presented no manifest problem in the November theory, as it was compatible with the Hertz condition that served to obtain its Newtonian limit, this default assumption was no longer acceptable in the Ricci theory due to the additional condition $\sqrt{-g} = 1$. The conflict between this condition and Einstein's earlier understanding of the Newtonian limit is also addressed in a letter Einstein wrote to Schwarzschild in early 1916, probably referring to a problem analogous to the conflict represented by eq. (LXXV):

> My comment in this regard in the paper of November 4 no longer applies according to the new determination of $\sqrt{-g} = 1$, as I was already aware. [At this point he added in footnote: The choice of coordinate system according to the condition $\sum \partial g^{\mu\nu}/\partial x_\nu = 0$ is not consistent with $\sqrt{-g} = 1$.] Since then, I have handled Newton's case differently, of course, according to the final theory.[221]

In the context of the Ricci theory, Einstein's generic ansatz for a spherically symmetric metric pointed almost without any further calculation to the existence of non-trivial values for $g_{11}\dots g_{33}$ in contrast to his default assumption about the metric for static gravitational fields. Einstein considered the difference to his earlier assumption about such a metric a remarkable consequence of the application of his methods for solving the perihelion problem to the new field equation. This is evident from his contemporary correspondence. After the completion of the final version of general relativity, he repeatedly mentioned this fact in letters to Michele Besso. In a letter from 10 December, he remarked:

> You will be surprised by the appearance of the $g_{11}\dots g_{33}$.[222]

A little more than a week later, Einstein returned to this point, again emphasizing the remarkable nature of the deviation from what he expected to be the metric for weak static fields. He was now able to point out how the conflict between this deviation and the Newtonian limit could be avoided:

> Most gratifying is the agreement with perihelion motion and the general covariance; strangest, however, is the circumstance that Newton's theory of the field is incorrect already in the 1st order eq. (appearance of the $g_{11}\dots g_{33}$). It is just the circumstance that the $g_{11}\dots g_{33}$ do not appear in first-order approximations of the motion eqs. which determines the simplicity of Newton's theory.[223]

The scheme for calculating a spherically symmetric static metric did not in itself lead to a way in which the deviation from Einstein's standard metric could be reconciled with the correspondence principle. However, it was clear that gravitational fields cannot be observed directly but only via the motion of bodies within these fields—a point stressed in Besso's 1913 memo[224]—so that the equation of motion, at second glance, suggested a natural way out of this dilemma. This second glance showed that in first-order approximation only the g_{44}-component of the metric tensor determines the motion of a material point and that, accordingly, non-trivial values for $g_{11}\dots g_{33}$

---

221 "Meine diesbezügliche Bemerkung in der Arbeit von 4. November gilt gemäss der neuen Festsetzung $\sqrt{-g} = 1$ nicht mehr, wie mir schon bekannt war. [At this point he added in footnote: Die Wahl des Koordinatensystems gemäß der Bedingung $\sum \partial g^{\mu\nu}/\partial x_\nu = 0$ ist nicht vereinbar mit $\sqrt{-g} = 1$.] Seitdem habe ich ja den Newton'schen Fall nach der endgültigen Theorie ja anders behandelt." Einstein to Karl Schwarzschild, 19 February 1916 (CPAE 8, Doc. 194). For a discussion of the role of coordinate conditions in general relativity, see also Einstein's paper on gravitational waves (Einstein 1916c).

222 "Du wirst über das Auftreten der $g_{11}\dots g_{33}$ überrascht sein." Einstein to Michele Besso, 10 December 1915, (CPAE 8, Doc. 162).

223 "Das Erfreulichste ist das Stimmen der Perihelbewegung und die allgemeine Kovarianz, das Merkwürdigste aber der Umstand, dass Newtons Theorie des Feldes schon in Gl. 1. Ordnung unrichtig ist (auftreten der $g_{11}\dots g_{33}$). Nur der Umstand, dass die $g_{11}\dots g_{33}$ nicht in den ersten Näherungen der Bewegungsgleichungen des Punktes auftreten, bedingt die Einfachheit von Newtons Theorie." Einstein to Michele Besso, 21 December 1915, (CPAE 8, Doc. 168).

224 For a facsimile of the relevant passage, see Fig. 2 on p. 300 of "What Did Einstein Know..." (in vol. 2 of this series) and (Renn 2005a, 128).

do not affect the first-order equation of motion and hence the Newtonian limit of Einstein's theory.

In a third letter to his friend Besso, Einstein once more returned to this point, now in order to explain that the new way of obtaining the Newtonian limit is closely related to the perihelion shift of Mercury and hence to the excellent agreement between theory and experiment.

> The great magnification of the effect against our calculation stems from that, according to the new theory, the $g_{11} \cdots g_{33}$'s also appear in the first order and hence contribute to the perihelion motion.[225]

This close connection between the empirical success of the theory and the deviation from the correspondence principle, as originally conceived by Einstein, stabilized the modified understanding of this principle and freed it from the aura of a dubious technical trick.

7.17.3 A New Problem Meets an Old Solution

Einstein repeatedly stressed the fact that only g_{44} matters for the equation of motion, a circumstance that must have seemed a strange but lucky coincidence to him. This solution to the dilemma created by the occurrence of non-trivial diagonal components in the first-order static metric was in itself as little new in 1915 as the dilemma itself. We have seen that in 1912–1913 the harmonically reduced and linearized Einstein tensor had been discarded because it led to a metric for weak static fields with non-trivial diagonal components (cf. eq. (74)). Furthermore, the very ansatz for a spherically symmetric static metric used to treat the Mercury problem pointed to the possibility of such non-trivial components. When Einstein first developed this ansatz in 1913, he did not give this possibility serious consideration because he was convinced of the validity of the *Entwurf* equation which does not give rise to such components.

It is remarkable is that, even though the dilemma of a non-spatially flat static metric was in mid-1913 no longer (and not yet) a real one for Einstein, it was nevertheless at that time already considered and resolved by Besso, and probably also by Einstein. This is documented by a page in the Einstein-Besso manuscript, written by Michele Besso on the back of a letter to Einstein. The page can be dated to June 1913 when both worked together in Zurich (CPAE 4, 392). It is one of a couple of pages on which Besso recapitulated the procedure he and Einstein had applied in order to determine the perihelion shift of Mercury. The purpose of this recapitulation was evidently not only Besso's wish to understand more thoroughly a method that in essence had probably been developed by Einstein, but also his intention to apply this method to more complex cases such as the field of a rotating sun or the inclusion of the sun's pressure in the calculation. Given the reflective character of Besso's notes, we also

225 "Die starke Vergrösserung des Effektes gegenüber unserer Rechnung führt daher, dass gemäss der neuen Theorie auch die $g_{11} \ldots g_{33}$). in Grössen erster Ordnung auftreten und so zur Perihelbewegung beitragen." Einstein to Michele Besso, 3 January 1916, (CPAE 8, Doc. 178).

find, along with the recapitulation of essentials, general remarks concerning the nature and plausibility of assumptions which Einstein and Besso had made in the course of their application of the method. It is among such remarks that one finds the brief reflection on the assumption of the spatially flat static metric quoted above, which shows that the assumption of this form of the metric was not an unquestioned prejudice.[226] Besso also considered the more general possibility of a weak-field metric in which components other than the 4–4 one deviated from the Minkowski metric. He came to the same conclusion as Einstein in his perihelion paper of November 1915, namely that, to first order, only the g_{44}-component is relevant for the equation of motion:

> The values so derived and inserted in the equations for the motion of the material point lead to the result that in the latter, [as far as] deviations of the magnitudes g from the relativity scheme [i.e. the Minkowski metric] [are concerned], only the elements g_{44} have any influence.[227]

This observation from June 1913 shows that the possibility of attaining the Newtonian limit also for spatially non-flat static metrics was not new. What was new was the necessity to bring this knowledge to bear on a field equation which seemed to forclude any other way of satisfying the correspondence principle. The novelty on 18 November 1915 was thus the combination of two chunks of knowledge that had been available independently for years, i.e., to base field equations on the Ricci tensor and to attain the Newtonian limit also for spatially non-flat static metrics.[228] In the following, we shall see that this combination triggered a new development that would very soon lead Einstein beyond the Ricci tensor.

7.18 Completing the Circle: Einstein's Return to the Einstein Tensor

7.18.1 Finding the Capstone of General Relativity by Double-Checking a New Theory of Matter

Einstein's completion of general relativity in November 1915 was essentially a solitary phase during which he had little correspondence and no collaboration on this subject, except for the mathematician David Hilbert, with whom Einstein's corresponded on the progress of their respective efforts. Hilbert had a long-standing interest in physics and was especially interested in foundational issues within his program of an axiomatization of the natural sciences.[229] When Gustav Mie published a special

226 See (CPAE 4, Doc. 14 [p. 16]). The mention of § 1 is probably a reference to (Einstein and Grossmann 1913).

227 "Die so ermittelten Werte in die Gleichg[ung]en für die Beweg[ung] des Materiellen Punktes eingesetzt, ergeben dass in denselben Abweichngen der Grössen g vom Relativitätsschema nur die Glieder g_{44} von Einfluss sind." (CPAE 4, Doc. 14 [p. 16])

228 For Einstein's continued concern with the problem of the appearance of other components of the metric tensor than g_{44}, see Albert Einstein to Erwin Freundlich, 19 March 1915, (CPAE 8, Doc. 63).

229 See (Corry 2004).

relativistic, electromagnetic theory of matter in 1912, he was particularly intrigued by it, and after Einstein's visit to Göttingen, at Hilbert's invitation, in the summer of 1915, Hilbert engaged in an attempt to find a synthesis between Mie's theory and Einstein's approach to gravitation. In November 1915, he was close to finishing his work and became Einstein's competitor for priority of the field equation of general relativity. The two scientists exchanged criticism and preliminary results, directly and possibly also indirectly via others, so that the question arises of the extent to which their results can be considered independent achievements. A set of proofs of Hilbert's "First Communication on the Foundations of Physics" (Hilbert 1915) rules out the possibility that Einstein took the last and crucial step in completing general relativity from the work of David Hilbert. Since this issue is discussed elsewhere in detail,[230] we limit ourselves here to the analysis of how Einstein completed this last step along the pathways of his own prior research.

Einstein considered the calculation of the perihelion shift of Mercury as the success of a generally-covariant theory of gravitation based on the Ricci tensor, but also as confirming the possibility of a new theory of matter. This is clear from the abstract of his paper quoted above and also from a letter he wrote to his friend Besso:

> In these last months I had great success in my work. *Generally covariant* gravitation equations. *Perihelion motions explained quantitatively.* The role of gravitation in the structure of matter. You will be astonished. I worked horrendously intensely; it is strange that it is sustainable.[231]

But the agreement between theoretical and empirical values for the perihelion shift of Mercury supported the new theory of matter only through the condition $\sqrt{-g} = 1$. The precarious role of this condition for further considerations by Einstein is made evident by a footnote appended to the perihelion paper:

> In a forthcoming communication it will be shown that this hypothesis is unnecessary. It is because such a choice of reference frame is possible that the determinant $|g_{\mu\nu}|$ takes on the value –1. The following investigation is independent of this choice.[232]

Einstein reexamined the connection between this determinant condition and his new theory of matter, which found its essential expression in the vanishing of the trace of the stress-energy tensor of matter. The requirement of the vanishing trace resulted

230 See (Corry, Renn, and Stachel 1997; Sauer 1999, 2002, 2005; Corry 2004) and further references cited therein. For a facsimile reproduction of both the proofs and the published version of (Hilbert 1915), see (Renn 2005, 146–173). In this series, the relation between Einstein's and Hilbert's work is further discussed in the section "Including Gravitation in a Unified Theory of Physics" (vol. 4 of this series).

231 "Ich habe mit grossem Erfolg gearbeitet in diesen Monaten. *Allgemein kovariante* Gravitationsgleichungen. *Perihelbewegungen quantitativ erklärt*. Rolle der Gravitation im Bau der Materie. Du wirst staunen. Gearbeitet habe ich schauderhaft angestrengt; sonderbar, dass man es aushält." Einstein to Michele Besso, 17 November 1915, (CPAE 8, Doc. 147).

232 "In einer bald folgenden Mitteilung wird gezeigt werden, daß jene Hypothese entbehrlich ist. Wesentlich ist nur, daß eine solche Wahl des Bezugssystems möglich ist, daß die Determinante $|g_{\mu\nu}|$ den Wert –1 annimmt. Die nachfolgende Untersuchung ist hiervon unabhängig." (Einstein 1915b, 831)

from the comparison of two equations Einstein had derived in his paper of 4 November 1915, one with the help of the energy-momentum balance, the other directly from the field equation. With the help of the trace t of the energy-momentum expression of the gravitational field, these equations can be rewritten as:[233]

$$\sum_{\alpha\beta} \frac{\partial^2 g^{\alpha\beta}}{\partial x_\alpha \partial x_\beta} - \kappa t = 0, \text{ and} \tag{90}$$

$$\sum_{\alpha\beta} \frac{\partial^2 g^{\alpha\beta}}{\partial x_\alpha \partial x_\beta} - \kappa t + \frac{\partial}{\partial x^\alpha}\left(g^{\alpha\beta} \frac{\partial}{\partial x^\beta} \log \sqrt{-g}\right) = \kappa t. \tag{91}$$

Comparing these two requirements Einstein had derived the scalar coordinate restriction of his first November paper, eq. (84). Both the empirical success of his perihelion calculation and the support for his new theory of matter were hinging on this condition. But there was another way of bringing the trace of the full field equation into agreement with eq. (83). Possibly feeling uneasy about the far-reaching consequences that this delicate compatibility argument had to support, Einstein reexamined his earlier reasoning.

From this perspective, the system of equations (90) and (91) provided a representation in which to explore the optimal way of putting together the pieces of his puzzle. This exploration led to yet another modification of the field equation. A reflection on how conditions (90) and (91) had been derived from the November field equation may have sufficed for the identification of an appropriate modification of this field equation by adding a multiple of the trace of the stress-energy tensor of matter to its right-hand side so as to yield instead:[234]

$$\sum_{\alpha\beta} \frac{\partial^2 g^{\alpha\beta}}{\partial x_\alpha \partial x_\beta} - \kappa(t + T) = 0, \text{ and} \tag{92}$$

$$\sum_{\alpha\beta} \frac{\partial^2 g^{\alpha\beta}}{\partial x_\alpha \partial x_\beta} - \kappa(t + T) + \frac{\partial}{\partial x^\alpha}\left(g^{\alpha\beta} \frac{\partial}{\partial x^\beta} \log \sqrt{-g}\right) = 0. \tag{93}$$

This set of equations no longer gives rise to problematic additional conditions.

The compatibility of Einstein's two conditions could thus be achieved without requiring $\sqrt{-g} = 1$ to imply that the trace of the stress-energy tensor must vanish, i.e. without eq. (61). This was the final step by which Einstein arrived at the definitive field equations of general relativity, which were presented in his paper of 25 November 1915 (cf. eqs. (69), (70)).

The modified source term in the new field equation violated the default assumption eq. (XLI) about the right-hand side of Einstein's mental model of a gravitational field equation. But Einstein could accept this violation since the energy-momentum

---

233 Cf. (Einstein 1915a, eqs. 9 and 10).

234 See "Untying the Knot …" (in vol. 2 of this series), sec. 7, eqs. 85–91.

tensor of matter and the energy-momentum expression of the gravitational field now entered the right-hand side of the field equation in a completely analogous way, (cf. eq. (68) to eq. (81)). It may also have played a role that the step from the Ricci to the Einstein tensor was, after all, not unfamiliar given his earlier experience in the Zurich Notebook. In his paper, Einstein lapidarily noted:

> I now quite recently found that one can get away without this hypothesis about the energy tensor of matter merely by inserting it into the field equations in a slightly different way than is done in my earlier papers.[235]

What had earlier prevented Einstein from accepting the (harmonically reduced and linearized) Ricci and Einstein tensors—his understanding of the correspondence principle—had meanwhile been transformed in the context of the perihelion calculation. The success of his solution to the Mercury problem included a solution to the problem of the Newtonian limit, and this solution now effectively replaced the correspondence principle as a criterion for an acceptable field equation.

In summary, the final phase of Einstein's work in November 1915 was not so much a phase in which new results challenged old prejudices, but rather one of reflection on the knowledge that was already available to him and in which different options were weighed against each other. One of the results of this process of reflection was that there was no support for a new theory of matter as Einstein had believed, possibly following Hilbert, in his addendum of 11 November. In the conclusion of his last November paper Einstein explicitly revoked his earlier claim:

> With this, we have finally completed the general theory of relativity as a logical structure. The postulate of relativity in its most general formulation (which makes spacetime coordinates into physically meaningless parameters) leads with compelling necessity to a very specific theory of gravitation that also explains the movement of the perihelion of Mercury. However, the postulate of general relativity cannot reveal to us anything new and different about the essence of the various processes in nature than what the special theory of relativity taught us already. The opinions I recently voiced here in this regard have been in error. Every physical theory that complies with the special theory of relativity can, by means of the absolute differential calculus, be integrated into the system of general relativity theory—without the latter providing any criteria about the admissibility of such physical theory.[236]

---

235 "Neuerdings finde ich nun, daß man ohne Hypothese über den Energietensor der Materie auskommen kann, wenn man den Energietensor der Materie in etwas anderer Weise in die Feldgleichungen einsetzt, als dies in meinen beiden früheren Mitteilungen geschehen ist." (Einstein 1915a, 844)

236 "Damit, ist endlich die allgemeine Relativitätstheorie als logisches Gebäude abgeschlossen. Das Relativitätspostulat in seiner allgemeinsten Fassung, welches die Raumzeitkoordinaten zu physikalisch bedeutungslosen Parametern macht, führt mit zwingender Notwendigkeit zu einer ganz bestimmten Theorie der Gravitation, welche die Perihelbewegung des Merkur erklärt. Dagegen vermag das allgemeine Relativitätspostulat uns nichts über das Wesen der übrigen Naturvorgänge zu offenbaren, was nicht schon die spezielle Relativitätstheorie gelehrt hätte. Meine in dieser Hinsicht neulich an dieser Stelle geäußerte Meinung war irrtümlich. Jede der speziellen Relativitätstheorie gemäße physikalische Theorie kann vermittels des absoluten Differentialkalküls in das System der allgemeinen Relativitätstheorie eingereiht werden, ohne daß letztere irgendein Kriterium für die Zulässigkeit jener Theorie lieferte." (Einstein 1915a, 847)

7.18.2 Reorganizing the Structure of General Relativity

The further history of general relativity shows that this theory could not yet be considered "logically complete," as Einstein formulated in the last paragraph of his conclusive paper. Even if one disregards later developments such as his modification of the field equations with a cosmological term, fundamental issues such as the status of energy-momentum conservation as an independent postulate of the theory still remained to be clarified. Without this clarification, the theory was initially unconvincing even to those physicists, such as Ehrenfest and Lorentz, who supported Einstein and closely followed his work.

Ehrenfest argued that one can eliminate the stress-energy tensor of matter from the two postulates of the theory, the conservation equation and the field equation, and thus arrive at a new differential equation, which the metric tensor has to satisfy in addition to the field equations. He therefore doubted, apparently following Einstein's earlier line of argumentation, that the new field equation was actually generally covariant. On 1 January 1916 Einstein wrote to Lorentz:

> I am conducting a discussion with Ehrenfest at present essentially on whether the theory really does fulfill the general covariance requirement. He also indicated to me that you had encountered problems or objections to it as well; you would do me a great favor if you were to inform me of them briefly. I have broken in my hobbyhorse so thoroughly that with a short hint I certainly also would notice where the crux of the problem lies.[237]

It was in the exchange with Ehrenfest that Einstein arrived at the conclusion that energy-momentum conservation was not an independent postulate but a consequence of the field equation.[238] The substantial clarification of the conservation principle that Einstein achieved in this debate became a starting point for a rearrangement of the foundational elements of his theory. The first step was taken in a lengthy letter that Einstein wrote to Ehrenfest.[239] In this letter he presented a derivation of the field equation from scratch and showed how energy-momentum conservation can be derived from it. Einstein proceeded in four steps:

1. He first derived the Lagrangian form of the field equation.
2. He next turned to the conservation principle. However, he did not yet derive the conservation of energy and momentum from the field equation. Rather, he assumed an equation that includes an unspecified function that has the form of energy-momentum conservation of matter, as he had postulated it in the earlier

237 "Mit Ehrenfest stehe ich in einer Diskussion im Wesentlichen darüber, ob die Theorie die Forderung der allgemeinen Kovarianz wirklich erfülle. Er deutete mir auch an, dass Sie Schwierigkeiten bezw. Einwendungen gefunden hätten; Sie würden mir große Freude machen, wenn Sie mir dieselben kurz mitteilten. Mein Steckenpferd habe ich so gründlich eingeritten, dass ich gewiss auch nach kurzer Andeutung merke, wo das Wesen der Schwierigkeit liegt." Einstein to H. A. Lorentz, 1 January 1916 (CPAE 8, Doc. 177).

238 Einstein to Paul Ehrenfest, 29 December 1915, Einstein to Paul Ehrenfest, 3 January 1916, Einstein to Paul Ehrenfest, 24 January 1916 (CPAE 8, Docs. 174, 179,185).

239 Einstein to Paul Ehrenfest, 24 January 1916 or later (CPAE 8, Doc. 185).

versions of his theory, and then derived from this equation another equation that has the form of energy-momentum conservation for matter *and* gravitation.

3. In the next step, Einstein wrote the gravitational field equation in terms of mixed tensor densities. He had apparently two reasons for doing so, the first being the possibility of an immediate physical interpretation of the equation in this form. The second reason was the preparation of the fourth and final step of his argument in which the conservation principle is demonstrated.
4. In his last step, Einstein derived energy-momentum conservation with the help of an indirect proof. He showed that one obtains a contradiction with the field equation in the mixed form if one does not assume that the unspecified function in the hypothetical equation for energy-momentum conservation (step 2) vanishes.

Einstein considered this line of argument as a new achievement clarifying the foundations of the theory, as becomes evident from the final passage of his letter:

> You will certainly not encounter any more problems now. Show this thing to Lorentz as well, who also does not yet perceive the need for the structure on the right-hand side of the field equations. I would appreciate it if you would then give these pages back to me, because nowhere else do I have these things so nicely in one place.[240]

Einstein made this new derivation the basis for his exposition in the 1916 review paper (Einstein 1916a), submitted on 20 March 1916.[241]

In the 1916 review, however, Einstein introduced a further rearrangement of the foundational elements of his theory. His main new results were a transformation of the indirect proof of the letter to Ehrenfest into a direct proof of energy-momentum conservation and the establishment of a connection between this derivation and a mathematical theorem by Hilbert, which was later generalized by Emmy Noether. The latter result is particularly important as it amounted, in effect, to a recognition of the contracted Bianchi identities and their role as integrability conditions for the sources of the field equation of general relativity.

In his review Einstein proceeded in six steps. We will briefly review these steps and show how a new deductive structure of general relativity emerged from Einstein's reflection on his discovery process and from the insights obtained in the controversy with Ehrenfest:

1. Einstein first introduced the field equations for the source-free case. In this step he transformed his own pathway from the Ricci to the Einstein tensor into a strategy for justifying the foundations of his theory. He introduced the Ricci equation as the appropriate gravitational field equation for empty space conceiving it as a weakening of an equation based on the Riemann tensor (Einstein 1916a, 803).

---

240 "Du wirst nun wohl keine Schwierigkeit mehr finden. Zeige die Sache auch Lorentz, der die Notwendigkeit der Struktur der rechten Seite der Feldgleichungen auch noch nicht empfindet. Es wäre mir lieb, wenn Du mir diese Blätter dann wieder zurückgäbest, weil ich die Sachen sonst nirgends so hübsch beisammen habe." (CPAE 8, Doc. 185)

241 For historical discussions of this paper, see (Janssen 2005, Sauer 2005).

2. He then developed the Lagrangian formalism and derived an equation for energy-momentum conservation of the gravitational field alone using the pseudo-tensor for the stress-energy of the gravitational field.
3. Einstein next reformulated the field equation in “mixed” form, including the trace term of the pseudo-tensor that suggested the new default setting eq. (81). The peculiar way in which the matter tensor has to be introduced as the source term of the Einstein field equation was thus prepared.
4. Einstein then supplemented the source-free field equation with this matter term, which was introduced in analogy to the pseudo-tensor for the stress-energy of the gravitational field, and thus arrived at the complete Einstein field equation. With respect to the paper of 25 November 1915, the context for Einstein’s justification of his new field equations had changed: Equations corresponding to eq. (92) and eq. (93) no longer appear in the 1916 review paper since energy-momentum conservation is not introduced as an independent postulate. As a consequence, these equations were no longer available as a justification for the new field equation. Instead, Einstein introduced the requirement that the energy of matter and the energy of gravitation enter the field equation on the same footing as the primary motivation for postulating the particular form of the Einstein field equation (Einstein 1916a, 808). He made it additionally clear that the main justification for his postulated field equation were the deductive consequences following from it.
5. Again in analogy to the source-free case, Einstein next showed that an energy-momentum equation holds for matter and the gravitational field. Previously, the equivalents of this equation in the earlier versions of the theory, going back to and including the *Entwurf* theory, were derived from the field equation, together with the independent postulate of energy-momentum conservation. Einstein had now succeeded in deriving this equation from the field equation alone.
6. In his final step, Einstein shows how his usual equation for the energy-momentum conservation of matter in the presence of a gravitational field, which was represented by the vanishing covariant divergence of the stress-energy tensor of matter, actually follows from his field equation. In other words, what had been a heuristic principle useful for selecting appropriate field equations now became a consequence of the field equation that was useful for selecting an appropriate stress-energy tensor of matter suitable to act as a source of the field equation (Einstein 1916a, 809–810).

In the last step of his deductive construction, Einstein also established a bridge to Hilbert’s contemporary work integrating one of its mathematical corner stones into his own newly established framework of general relativity. As we have just seen, within this framework the stress-energy tensor of matter is no longer conceived as an independent ingredient of the theory with properties that affect its physical interpretation (such as the selection of preferred coordinates) but this tensor has itself to satisfy certain constraints imposed by the theory. In a short remark, Einstein characterized this partly dependent and partly independent status of the material process in his theory of gravitation:

> Thus the field equations of gravitation contain four conditions which govern the course of material phenomena. They give the equations of material phenomena completely, if the latter is capable of being characterized by four differential equations independent of one another.[242]

At this point Einstein appended a footnote in which he referred to Hilbert.[243] Einstein provides here a reinterpretation of the mathematical claim central to Hilbert's theory, which constitutes the core of what later became Noether's theorem.[244] In fact on the page referred to by Einstein we find the following passage:

> ... then in this invariant system of n differential equations for the n quantities there are always four that are a consequence of the remaining $n—4$ in this sense, that among the n differential equations and their total derivatives there are always four linear and linearly independent combinations that are satisfied identically.[245]

By referring this general theorem to the relation between his gravitational field equation and the four differential equations corresponding to the vanishing of the covariant divergence of the stress-energy tensor, Einstein gave a physical interpretation of this theorem that was quite different from Hilbert's. Combining his own results with those of Hilbert, he was able to understand that energy-momentum conservation follows from the field equations. He had thus finally realized the structural role which the four differential equations, expressing energy-momentum conservation and mathematically corresponding to the contracted Bianchi identities, play for the conservation principle in the general theory of relativity as we understand it today.

8. THE TRANSITION FROM CLASSICAL PHYSICS TO GENERAL RELATIVITY AS A SCIENTIFIC REVOLUTION

In the preceding sections, we have reconstructed the complex process by which Einstein's heuristics led to the formulation of the general theory of relativity. We have shown that a key role was played by the interaction between the heuristics guiding the search for the new theory and the concrete representations of intermediate results in terms of physically interpreted mathematical formalisms. These representations opened up new possibilities for further development and often required adjustments

---

242 "Die Feldgleichungen der Gravitation enthalten also gleichzeitig vier Bedingungen, welchen der materielle Vorgang zu genügen hat. Sie liefern die Gleichungen des materiellen Vorganges vollständig, wenn letzterer durch vier voneinander unabhängige Differentialgleichungen charakterisierbar ist." (Einstein 1916a, 810)

243 Cf. (Hilbert 1915, 3). Einstein's page number actually refers to an offprint of Hilbert's paper, not to the published version. Offprints were available to Hilbert already by mid-February 1916, the published paper itself appeared on 31 March 1916, see (Sauer 1999, note 74).

244 See (Sauer 1999). For the roots of this theorem in Einstein's own work, see sec. 3 of "Untying the Knot ..." (in vol. 2 of this series).

245 "... 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." (Hilbert 1915, 3). See vol. 4 of this series.)

of physical concepts and heuristic principles. Einstein's heuristics, together with such concrete intermediate results, was evidently capable of generating enough of those arguments on which the justification of general relativity as an essential part of modern physics, is still based today.

This heuristics itself and some of Einstein's conceptual starting points in classical physics underwent changes that justify the designation of this process as a scientific revolution. In this final section, we shall first review the beginning and the end of the development of Einstein's heuristics in order to highlight the conceptual innovations brought about by this development with respect to classical physics. We shall then summarize our answers to the three epistemic paradoxes raised by this scientific revolution. These answers make use of the key elements for an understanding of a scientific revolution that is suggested by historical epistemology: the long-term character of knowledge development, the architecture of knowledge, and the mechanisms of knowledge dynamics.

8.1 The Lorentz Model Remodelled

Our analysis has shown that for Einstein's search, the Lorentz model was structurally the most significant heuristic element inherited from classical physics. At each stage of its development, the structure of this mental model and its default settings determined the way in which the specific problems of finding the field equations could be addressed. As long as it remained unquestioned, the model thus opened (or closed) the viable paths of further exploration and determined the possibilities of conceptual unfolding. In classical physics, the two basic structures of the Lorentz model, the field equation and equation of motion, are related to each other as independent components of which the first determines the creation of a global field by a local source, while the second determines the effect of the global field on a local probe. Within the classical framework, source and probe are essentially independent entities entering into this model.

In general relativity, this basic structure has changed. First, the source can no longer be independently prescribed from the field. The distribution of matter and energy acting as a source of the gravitational field can only be described in a given geometry of spacetime, which in turn is only another aspect of the gravitational field determined by the field equation. Second, the equation of motion is no longer an independent aspect of the problem, linked to the description of the gravitational field by an overarching force concept, but is constrained and in special cases even completely determined by the field equation.[246] These features of general relativity, which mark its conceptual distinction from classical physics were not yet evident in 1915 when Einstein formulated his field equations. In other words, the corresponding conceptual innovation was not the presupposition but the result of his research. Thus Einstein's heuristics, which was structured by the Lorentz model, led to the develop-

246 For historical discussion of this point, see (Havas 1989, Kennefick 2005).

ment of a theory whose cognitive content can no longer be adequately captured by this mental model.

The discovery of general relativity would, however, have been impossible if the Lorentz model had not at least been adequate for capturing just those partial aspects of the final theory that made its discovery possible. As we have seen in the previous sections, it was even possible to construct and interpret the definitive field equation of general relativity according to this model. Furthermore, the kind of solutions that Einstein had in mind when he searched for the field equation obscured the new relation between matter distribution and geometry mentioned above. The solutions that he seriously considered were given either by Minkowski spacetime (a vacuum solution) described in various coordinate systems, or weak field solutions that could be obtained from it by an iterative procedure. The problem of having to first specify the geometry and then the distribution of matter and energy in order to solve the field equation turns into an approximation procedure. The further elaboration of the consequences of Einstein's field equation revealed the changes with respect to the Lorentz model. That a revision of this mental model was implied by the field equation of general relativity was clear to Einstein as soon as he noticed that the field equation of the new theory would have to be non-linear. As early as 1912, he interpreted this technical feature as representing the conceptual conclusion that the gravitational field possessing energy must also act as its own source. However, at that time, this modification of the model did not appear to be a radical break, since a modification of only a default setting of the mental model ("linearity of the field equation") was sufficient to account for the insight that gravitation can act as its own source.

8.2 The Ill-Conserved Conservation Principle

In classical physics, the conservation of energy and momentum is a consequence of the fundamental laws governing gravitational and electrodynamic interaction. In special relativity, the conservation principle has found an elegant formulation as a tensorial equation that unifies the conservation of momentum and energy. In both classical and special-relativistic physics, momentum and energy are conceived as localizable physical quantities whose conservation can be described by a partial differential equation which describes a local balance between the various contributions to the energy and momentum of a physical system. Einstein's consideration of a particular example (the behavior of a pressureless dust of particles in a gravitational field) formed the basis, as we have seen, for a tentative generalization of the equation expressing the conservation principle in special relativity, which now also included the effect of gravitation, interpreted as the effect of an external force. Two distinct perspectives on this equation exist, one from classical physics and special relativity, the other from general relativity. The possibility of having these two perspectives on the same mathematical expression turned out to be crucial for the emergence of general relativity.

From the point of view of classical physics and of special relativity, Einstein's postulated equation represented a twofold constraint for the gravitational field equation to be found: the resulting field theory of gravitation had to be compatible with this equation, even at the price of restricting its range of applicability, and, furthermore, the field equation should allow this equation to be rewritten as a local, frame-independent balance between the energy-momentum of matter and that of gravitation.

Further elaboration of the consequences of this equation, however, made this latter request questionable. In the course of his search, Einstein was forced to realize that the expression for energy-momentum conservation which he had postulated turned out to be incompatible with the assumption of a frame-independent stress-energy tensor of gravitation. If this postulate is accepted, then energy and momentum of a gravitational field cannot, in contrast to classical physics, be localizable physical quantities. In this way, a feature of general relativity that is incompatible with classical physics was suggested by a framework still anchored in its fundamental concepts. Einstein's insight into the character of the expression representing the stress-energy of the gravitational field might have given him good reason to abandon this entire approach since its results conflicted with his well-founded expectation that the gravitational field has localizable energetic properties just like all the other known physical fields. Why did he hold on to this equation in spite of its, from the point of view of classical physics, problematic implications? His reasons were in any case not an anticipation of those of the later theory of general relativity.

The equation expressing the energy-momentum balance in a gravitational field that Einstein had postulated at the beginning of his search, and from which the problematic conceptual consequences summarized above can be inferred is obtained in general relativity as an integrability condition of the field equation. Technically speaking, it is a condition to be imposed on an admissible energy momentum tensor, representing the right-hand side of the field equation, required in order to be compatible with a mathematical identity—the contracted Bianchi identity—valid for the left-hand side of the field equation. The Bianchi identity ensures that the gravitational field equation determines the dynamics of the geometry of spacetime without determining also the coordinate system. It reduces the 10 components of the field equation for the 10 components of the metric tensor to only 6 component-equations, thus leaving open the choice of four arbitrary functions corresponding to the choice of a coordinate system. The Bianchi identity together with the gravitational field equation then also determines the evolution of energy and momentum in space and time by way of the equation which Einstein interpreted as the expression for the conservation of energy and momentum in the presence of a gravitational field.

Clearly this argument could not have played a role for Einstein when he was still searching for the correct field equation. He was not even familiar with the Bianchi identity at the time when he concluded his search with the publication of the field equation of general relativity in 1915. Instead he only had two comparatively weak arguments to hold on to this equation even when he recognized that it did not lead him to an invariant local expression for the energy-momentum of the gravitational field. The first argument was that, for the special case of a dust-like cloud of particles, it was possible to obtain this equation from the equation of motion of a single point-particle in a gravitational field described by the metric tensor. The second argument was related to the mathematical form of energy-momentum conservation. The corresponding equation has the form of a generally-covariant divergence equation which is not only the precise analogon for the corresponding special relativistic equation but which also reduces to the latter in the absence of a gravitational field. These two arguments reinforced each other and are in turn supported by other aspects of Einstein's heuristics, in particular by the generalized relativity principle and all those aspects which underlay his understanding of motion in a gravitational field and the requirement of a close correspondence between special relativistic insights and their generalizations in the new theory to be constructed.

But in whatever way Einstein could support his understanding of energy-momentum conservation by drawing on special cases and analogies, it was, from the point of view of the deductive structure of the later theory, support for the wrong side of his argumentative construction, in so far as it stabilized the role of energy-momentum conservation as an independent first principle rooted in the conviction of the fundamental status of energy and momentum conservation for any physical theory. This understanding motivated its use both as a compatibility requirement and as an additional constraint on trial field equations. From the perspective of general relativity, Einstein had thus developed an improper argumentative structure around a proper equation, whereas from his own perspective at the time, he had attained a partial insight into the deductive structure of the theory which he attempted to construct. Only after his achievement of 1915 he was able to reverse this deductive structure and obtain the vanishing of the covariant divergence of the energy-momentum tensor as a consequence of the gravitational field equation in the sense explained above. As was the case for the development of the mental model in Einstein's research, the structural and conceptual insights associated with understanding the role of energy-momentum conservation in general relativity were thus the result and not the presupposition of finding the correct equations.

8.3 The Lack of Correspondence between the Correspondence Principle as seen from Classical Physics and from General Relativity

The way in which the classical theory of gravitation is contained in the theory of general relativity could, of course, not be anticipated on the basis of classical physics before that theory was actually formulated. Nevertheless, the same heuristics which led to the introduction of the principal building block of the new theory, the metric tensor, also determined, to a large extent, Einstein's understanding of the relation between the theory of gravitation which he was looking for and Newton's theory. All in all, he developed, as we have seen, in the course of his research three different arguments in favor of the representation of static gravitational fields by a spatially flat metric tensor in which, for an appropriate coordinate representation, only one component is variable and a function of the three space coordinates; this function then, so Einstein's conclusion, corresponded, under certain limiting conditions, to the Newtonian gravitational potential in his new theory, whatever the precise field equation would be.

The first argument was directly related to the introduction of the metric tensor as representing a gravito-inertial field, a step that was, as we have seen, motivated by the equivalence principle. Einstein conceived Newtonian gravitation and inertia as special cases of a more general interaction. For the case of uniform acceleration he was able to directly identify inertial effects with a scalar Newtonian gravitational field and he expected that he would be able to do the same for more general cases by generalizing the notion of the gravitational field. A model for that generalization was delivered by electrodynamics. In spite of the obvious differences between gravitation theory and electrodynamics, the analogy between them was in fact the only available one and hence determined Einstein's view of the general pattern according to which a theory of the static field should be contained as a special case in a general field theory. According to this pattern, the general potential was represented by a many-component object such as a vector or a tensor which, in the special case of a static field, reduces to a single-component object. In the case of gravitation it should naturally be possible to identify this single-component object with Newton's gravitational potential. This expectation was reinforced by the fact that Einstein had developed, even before introducing a metric formalism, a theory of static gravitational fields in which these are represented by a single function. When he began to employ a metric formalism, it was hence natural to describe static fields by a metric with one variable component and to identify this component with the gravitational potential of his theory of static fields.

Einstein's "classical" understanding of the transition from his general theory to Newton's theory was stabilized by further arguments developed in the course of his research. The second argument was based on the role of special relativity as an intermediate step in this transition. In order to describe the gravitational effects known from classical physics as aspects of a more general gravitational field it is necessary to specify also the conditions under which such an identification is possible. These physical circumstances require, in particular, the general field to be weak and static. These conditions are, however, not sufficient for restricting the realm of gravitational effects to that covered by Newton's theory. The case in which the masses involved perform motions of high velocities requires a treatment by the special theory of relativity. According to this line of reasoning, weak fields, and in particular weak static fields, should hence play the role of an intermediate case in the transition to Newtonian gravitation, an intermediate case to which the special theory of relativity should be applicable. It should hence be possible to formulate a special relativistic gravitational field equation which holds under these circumstances. As it turned out, the solutions of such a weak-field equation, as suggested by the appropriate default-settings of the Lorentz model for this case, exactly correspond to Einstein's classical expectations.

Nevertheless, these expectations were, as we have seen, challenged in the course of Einstein's research which pointed on several occasions towards a representation of static fields by a metric tensor whose form does not correspond to the one which he expected. He therefore felt, at some point, the necessity of developing yet another argument in favor of this expectation. His third argument, which we have also discussed above, was completely independent of a particular gravitational field equation. In essence it consisted in a problematic attempt to deriving the form of the metric tensor for static gravitational fields from the postulate underlying the equivalence principle that all bodies—no matter what their energy content—fall with the same acceleration in a gravitational field.

The assertion that the metric for static fields is of the canonical form expected by Einstein does not belong to the realm of classical physics. It rather appears to be a specific technical assumption which entered his preliminary gravitational theories as an inconspicuous and perhaps precisely for this reason fateful prejudice delaying his progress towards the correct field equation. However, the preceding synopsis of the reasoning by which this assumption was actually anchored in Einstein's thinking shows that, once the metric tensor was introduced as a representation of gravitational fields, the association of static fields with a metric tensor of the canonical form was a necessary consequence of Einstein's understanding of classical physics applied to this representation.

This entire network of reasoning, and in particular, the procedure for attaining the Newtonian limit which forms its core, is not compatible with the final theory of general relativity. According to this theory, static fields are, in general, not represented by a metric tensor of the canonical form. A consistent treatment of the problem of the Newtonian limit in general relativity is an intricate problem[247] and indeed requires a mathematical formalism which did not even exist when Einstein first formulated the theory in 1915; it was only introduced much later by Cartan and others (Cartan 1923, 1924). It is only in this formalism, by using the concept of an affine connection, that it is possible to formulate both general relativity and Newton's theory of gravitation in a way that makes them mathematically comparable.[248] In fact, whereas in general relativity, the geometry of spacetime is described by a metric structure, in Newtonian theory of gravitation, the four-dimensional metric structure is degenerate and only an affine structure can be introduced for spacetime. But since a metric determines also an affine structure, both theories can, with the help of this mathematical concept, actually be expressed in the same mathematical terms. Vice versa, the fact that the spacetime of general relativity carries not only an affine but also a metric structure represents a conceptual leap with respect to Newtonian physics that cannot be bridged by considering the special theory of relativity, which also comprises a metric structure of spacetime, as an intermediate case. It is simply impossible to describe Newtonian gravitational fields by a non-degenerate four-dimensional metric tensor.

---

247 See note 44 above.
248 See "The Story of Newstein …" (in vol. 4 of this series).

This conceptual leap between general relativity and Newtonian physics, together with the strong arguments which Einstein had in favor of his classical conception of the correspondence principle, raise the question as to how he could have ever overcome the crucial hurdle of dealing with the Newtonian limit of his new theory. The surprising solution is that it was in fact not the removal of this major stumbling block which freed his way, but rather its circumvention for the specific problems in the focus of his attention at the time, in particular for the treatment of the motion of a point mass in a gravitational field. By showing that, although the spatial curvature is present even under Newtonian conditions, it remains unobservable if only slowly moving particles are considered, Einstein found a technical loop-hole through which he could escape from his dense network of reasons supporting the canonical form of the metric.

8.4 The Ambiguity of the Equivalence Principle

We have identified the beginning and end points of the development of those aspects of Einstein's heuristics which were obviously rooted in classical physics. We have concluded that this classical heuristics was just sufficient to allow for the formulation of the key equations of general relativity, whose exploration then, however, led to conceptual insights with which the original expectations were no longer compatible. We now turn to those elements of Einstein's heuristics which were peculiar to his specific research strategy, the equivalence principle and the generalized relativity principle. No such principles belonged to the accepted core of classical physics at the time when he took up his research.

If one separates, however, the mathematical development from that of the physical theories, then Einstein's introduction of the principle of equivalence appears to be much less of an idiosyncrasy than it may seem at first sight. To make this clear, consider the reformulation of the classical Newtonian theory of gravitation as a spacetime theory with a non-trivial geometry. This geometry can be described in terms of an affine connection determining the notion of parallel transport of vectors and hence also the geodesic lines in that spacetime. It is a fundamental statement of this reformulation of Newton's theory that the geodesic lines represent the motions of freely falling particles according to the law of gravitation. Remarkably, the equality of gravitational and inertial mass has become, in this modern formulation, an in-built feature rather than a contingent fact as in the traditional formulation. In the formulation of the law of motion as being given by the geodesic lines, the notion of mass appears not at all, while only the notion of gravitational mass enters the field equation of the theory which determines the geometry of spacetime.

From the point of view of this mathematically advanced formulation, Einstein's adoption of the principle of equivalence can hence be recognized as expressing a fundamental feature of Newton's theory of gravitation, shaped, however, by the particular mathematical formulation of the classical theory which formed his starting point and which suggested a sharp conceptual and technical distinction between gravitational and inertial forces. This conclusion indeed frees Einstein's adherence to the principle of equivalence from its idiosyncratic appearance. One might even conjecture that, had the development of the appropriate mathematical tools come a little earlier, others as well might have found it attractive to employ them for a new formalization of the classical knowledge on gravitation, thus arriving at the considerations outlined above even before the advent of general relativity. But it now emerges, on the other hand, even more as a riddle how a principle expressing the knowledge of classical mechanics could have served as a crucial heuristic guidance for overcoming this theory in favor of a theory incompatible with it.

The key to resolving this riddle comes from considering the fact that Einstein used the principle of equivalence not in order to reorganize the knowledge of classical mechanics but the knowledge embodied in both, classical mechanics and the special theory of relativity. His theory of the static gravitational field as well as his early attempts to generalize that theory were nothing but a reinterpretation of the special theory of relativity with the help of the introduction of accelerated frames of reference. His systematic consideration of such accelerated frames induced him to make use of generalized Gaussian coordinates in order to describe the coordinate systems adapted to these frames. It was then a short step for him to consider the metric tensor, coming with the introduction of such coordinates, also as the representation of gravitational effects when these could not be generated by acceleration. In other words, with the introduction of the metric tensor Einstein had found an object that was capable of representing gravitational and inertial effects on the same footing, just as is the affine connection within the modern reformulation of Newton's theory.

It was, however, not a mere coincidence governed by the availability of mathematical methods that Einstein directly attempted to implement the principle of equivalence in a theory that was to generalize special relativity rather than concentrating on a reformulation of classical mechanics. He was aiming from the beginning at a new theory of gravitation which was to comprise both the knowledge on gravitation and inertia represented by classical mechanics and the knowledge on the structure of space and time embodied by special relativity. Effectively, the principle of equivalence acted, according to this reconstruction, as a demand for integrating the knowledge on gravitation and inertia from classical mechanics, which in a modern formulation can be expressed by means of an affine connection, with the knowledge on the metric structure of spacetime from special relativity. It thus acted as a particular instance of Einstein's general strategy to exploit the entire range of classical and special-relativistic physics for constructing his new theory of gravitation. The analysis given here does, however, not square with Einstein's own interpretation of the principle of equivalence as guiding the development of classical and special-relativistic physics with its privileged systems of reference towards a theory of gravitation which would have to encompass also a generalized principle of relativity.

8.5 The Relativity Principle Relativized

Einstein's view that it made sense to search for a generalization of the relativity principle of classical mechanics and special relativity was, as we have seen, based on his acceptance of a philosophical critique of classical mechanics raised by Mach and others. According to this critique, the justification of the privileged role of inertial frames of reference by the notion of absolute space was problematic, while the inertial forces experienced in accelerated frames of reference require an explanation in terms of the interaction between physical masses. Such an explanation would then eliminate any need for absolute space as a causal agent in the analysis of motion. The generalized relativity principle would go, so at least was Einstein's expectation, a long way, and might actually go all the way, towards an implementation of Mach's critique of classical mechanics in the new theory of gravitation.

The implementation of Mach's critique of classical mechanics by way of the generalized relativity principle in Einstein's new theory of gravitation was, however, rather indirect. Rather than explaining inertial properties directly by a physical interaction of masses, they were described by a gravito-inertial field represented by the metric tensor in a way that in fact depends on the frame of reference. But the gravito-inertial field itself would be determined only by the distribution of masses in the universe via a generally-covariant field equation. It follows that the question of whether or not this approach would lead to an exhaustive explanation of inertial properties by the relative distribution of masses depends on the precise nature of the field equation and its solutions. While Einstein was initially convinced that his theory would fully do justice to the Machian roots of the generalized relativity principle, he felt eventually forced to introduce what he called Mach's principle as a separate and additional criterion to be satisfied by the field equation and its solutions. With the establishment of the General Theory of Relativity in 1915, Einstein succeeded in formulating a theory which implemented the generalized relativity principle in its utmost form, the theory being generally covariant; whether it also satisfied Mach's principle, demanding a complete determination of the gravito-inertial field by the distribution of matter in the universe, remained, on the other hand, a much debated issue for a long time to come.[249]

On closer inspection, however, even Einstein's realization of the generalized relativity principle by his formulation of a generally-covariant theory of gravitation represented a questionable success of this heuristic principle. In fact, not only the general theory of relativity of 1915, but also several other theories of gravitation and in particular also the classical Newtonian theory can be given a generally-covariant formulation. The demand for general covariance has to be considered as nothing but a minimal requirement to be imposed on any sensible physical theory, namely to make assertions about physical processes which do not depend on the specific coordinates used for describing them. But Einstein's generalized relativity principle—together with its broader Machian understanding—effectively corresponded, as we have seen,

249 For extensive discussion, see "The Third Way to General Relativity …" (in vol. 3 of this series).

to further requirements beyond the demand for a generally-covariant formulation of the theory of gravitation. It also comprised the demand for treating inertia and gravitation as aspects of a more general interaction as well as the demand for the absence of any prior geometry of spacetime. The latter requirement excludes, for instance, Nordström's theory as being not compatible with Einstein's heuristics since it assumes the geometry of spacetime a priori to be Minkowskian, up to a conformal factor representing the gravitational potential.

Nevertheless, Einstein's own interpretation of this heuristics can hardly be vindicated by the modern understanding of general relativity. First, the demand for treating inertia and gravitation as aspects of a more general interaction can, as we have seen in our discussion of the equivalence principle, already be fulfilled by classical mechanics in an appropriate reformulation. Second, the general covariance of Einstein's theory does not embody a generalization of the relativity principle from classical mechanics and the special theory of relativity, since, in the modern understanding, relativity principles are represented by the symmetry properties of a theory and not by their behavior under coordinate transformations. Third, "Mach's principle" in the sense of Einstein's demand that the metric structure of space be completely determined by the material masses makes little sense according to the modern understanding of general relativity, since the very notion of material bodies acting as a "source" of the gravitational field that can be prescribed independently from the field has turned out to be problematic.

8.6 The Long-Term Development of Knowledge

In the preceding discussion we have emphasized the differences between Einstein's heuristics and the conceptual consequences of the theory whose development was guided by this heuristics. These differences were the result of a process covering two eras stretching from the beginning of the relativity revolution in 1905 to the present: The first era comprised the elaboration of the foundational equations of the new theory guided by the original heuristics, a process that was essentially complete with Einstein's formulation of general relativity in 1915 and that also included, as we have seen, adjustments of the original heuristics. The second era consists in the exploration of the conceptual consequences of the new theory on the basis of an interpretation of the results achieved in the first era as well as in the course of its further elaboration, a process that has still not come to a hold today. In view of the often striking differences between the modern interpretation of general relativity and Einstein's original motivations for searching for such a theory, it represents a remarkable challenge for the historical reconstruction to explain how these original motivations could have led him to such a definitive formulation of the new theory of gravitation. In the beginning we have formulated this challenge in terms of the three epistemic paradoxes of the emergence of general relativity, the paradox of missing knowledge, the paradox of deceitful heuristics, and the paradox of discontinuous progress.

As our reconstruction has shown, an adequate response to the missing-knowledge paradox can only be found when the long-term development of scientific knowledge is taken into account. This development led, after all, to the emergence of a theory whose understanding of how gravity affects motion in terms of spacetime structure is closer to Aristotle's concept of natural motion than to Newton's explanation in terms of an anthropomorphic force. The knowledge on which the astonishing stability of general relativity is founded was, as we have seen, accumulated long before its creation by centuries of physics, astronomy, and mathematics. Our modern acceptance of general relativity is not only based on experiments or observations related to some of its special predictions but also on the fact that it incorporates the entire Newtonian knowledge on gravitation, including its relation to other physical interactions, that has been accumulated over a long period of time in classical physics and in the special theory of relativity. This knowledge embraces, among other aspects, Newton's law of gravitation including its implications for the conservation of energy and momentum, the relation between gravitation and inertia, the understanding that no physical action can propagate with a speed greater than that of light, which was first achieved by the field theoretic tradition of classical physics and then finally established with the formulation of special relativity, and, more generally, the local properties of space and time, also formulated in special relativity.

After special relativity had elevated the causality requirements implicit in field theory to a universal status, gravitation, traditionally a subject at the core of mechanics, had effectively turned into a borderline problem between mechanics and field theory. As was the case for other borderline problems, its successful solution depended on the shared knowledge resources taken into account. In the case of the creation of special relativity, Einstein's success depended on his combining the heritage of mechanics, embodied in the relativity principle, with the heritage of electrodynamics, embodied in the principle of the constancy of the speed of light. In the case of a relativistic theory of the gravitational field, the combination of the heritage of mechanics represented by the Newtonian theory of the static gravitational field with what was known about dynamic fields from electrodynamics was, however, insufficient to create a new and satisfactory theory—as Einstein's competitors experienced to their chagrin. There was, in particular, no clue to the properties of dynamic gravitational fields so that the challenge to build a relativistic field theory of gravitation was comparable to the development of the entire theory of electromagnetism knowing only Coulomb's law.

It was at this point that Einstein's broad perspective, including the philosophical critique of classical mechanics by Mach, allowed him to muster additional resources from classical physics. Einstein exploited the Machian interpretation of the inertial forces in an accelerated reference frame as being due to the interaction of moving masses in order to fill the above-described gap in a field theory of gravitation. By conceiving the inertial forces in accelerated reference frames, such as Newton's rotating bucket, as embodying dynamic gravitational fields he managed in fact to anticipate essential properties of the relativistic theory of gravitation he was about to construct, in particular the necessity to generalize the spatio-temporal framework of special relativity, which led to the notion of a curved spacetime.

8.7 The Architecture of Knowledge

The answer to the second paradox of how Einstein could have formulated the criteria for a gravitational field equation years before finding the solution comes, as we have seen, from considering the architecture of the shared knowledge resources available to him. These resources were in fact part of a system of knowledge with active components capable of providing heuristic guidance to his research.

The characteristics of Einstein's search have become comprehensible by realizing that it was guided by a qualitative knowledge representation structure inherited from classical physics: the mental model of a field theory as embodied in an exemplary way by Lorentz's electron theory. Einstein's preliminary research on a relativistic theory of gravitation in the years between 1907 and 1912 had established default-settings for two of its terminals; the field-slot (filled by assuming that the gravitational potential is represented by the metric tensor), and the source-slot (filled by the stress-energy tensor of matter as suggested by relativistic continuum mechanics). In the context of his research the differential operator describing how the source generates the field represented an open slot for which Einstein was unable to identify a satisfactory instantiation.

As we have discussed, Einstein's difficulty did not result from the fact that too little was known but rather from the fact that too much knowledge had to be taken into account to formulate a field equation that responded to the understanding of gravitation as a borderline problem of mechanics and field theory. On the one hand, a physically plausible instantiation for the differential operator was suggested by knowledge of the Newtonian static gravitational field as well as of the relation between static and dynamic fields in electrodynamic field theory. Constructed in this way, the new theory would automatically be compatible with Newton's theory, thus fulfilling the correspondence principle. On the other hand, a mathematically plausible way to obtain an instantiation of the differential operator was offered by the knowledge about dynamic fields incorporated in Einstein's equivalence principle, which suggested taking generally-covariant objects such as the Riemann tensor as the starting point. Constructed in this way, the new theory would automatically fulfill the generalized relativity principle. The equivalence principle and the generalized relativity principle had helped, in addition, to reveal just those elements of the traditional knowledge on whose integration the new theory could be based. In the modern formulation, they posed the problem of the compatibility between chronogeometry and gravito-inertial structure. Within the knowledge system of classical physics, the Lorentz model was, furthermore, embedded in a network of relations to other frames and mental models; this network served as a control structure for any acceptable implementation of the model. In particular, the new theory had to satisfy the conservation principle, generalizing similar principles from classical and special-relativistic physics.

In short, Einstein's heuristics was overdetermined by the knowledge available to him, explaining why it was so powerful and yet so fortuitous at the same time. The compatibility of the various requirements it imposed could not be established *a priori* but had to be checked by elaborating a mathematical representation of the Lorentz model, starting from one or the other default setting and shaping it according to the remaining heuristic criteria. Einstein's oscillation between a physical strategy starting from an implementation of the correspondence principle, and a mathematical strategy starting from an implementation of the generalized relativity principle could thus be interpreted as realizing alternative and ultimately converging pathways with which to integrate the knowledge of classical physics.

8.8 Knowledge Dynamics

The third paradox, of discontinuous progress, could only be resolved by taking into account that the development of knowledge does not only consist of enriching a given architecture but also comprises processes of reflection by which this architecture is being transformed. Einstein's learning experience was, in fact, characterized by a bottom-up process that accommodated the higher-order structures at the core of his heuristic principles to the outcome of the experiences he made implementing these principles. The interplay between assimilation and accommodation mediated by the mathematical representation has turned out to be the crucial process determining the knowledge dynamics leading to the creation of general relativity as a non-classical theory. Against this background four stages of Einstein's search for the gravitational field equation could be distinguished.

The *tinkering phase* of fall 1912 is documented in the early pages of Einstein's Zurich Notebook. It is characterized by his unfamiliarity with the mathematical operations suitable for constructing a field equation for the metric tensor. Nevertheless, reflecting on his first attempts to formulate a field equation that satisfied his heuristic principles, Einstein built up higher-order structures operating on a strategic level that would later guide his systematic implementation of these principles, in particular, the physical and the mathematical strategy.

The *systematic searching phase* from late 1912 to early 1913 is also extensively documented by the Zurich Notebook. In this phase Einstein systematically examined candidates according to his heuristic principles alternating between physical and mathematical strategies. Meanwhile, the relative weight of the heuristic principles kept changing with the conservation principle emerging as the principal challenge. Paradoxically, the main result of the pursuit of the mathematical strategy was the derivation of an erroneous theory—the *Entwurf* theory—along the physical strategy.

The *consolidation phase* is documented by Einstein's publications and correspondence between 1913 and mid-1915. During this phase he elaborated the *Entwurf* theory, essentially following his earlier heuristics but now under the perspective of consolidation rather than exploration. Paradoxically, however, the main result of the consolidation period was the creation of the presuppositions for a renewed exploration of candidate field equations. Adapting the mathematical strategy to legitimize the *Entwurf* theory, Einstein found that the resulting mathematical formalism did not single out this theory but reopened the perspective of examining other candidates, removing, in particular, the difficulty of implementing the conservation principle. Because of the extended network of results meanwhile assembled, this reexamination could now take the form of a reflective reorganization of Einstein's earlier achievements.

The *reflection phase*, decisive in resolving the paradox of discontinuous progress, is documented by the dramatic series of four communications Einstein submitted to the Prussian Academy in November 1915. The essence of Einstein's return in the first of these communications to a field equation related to the Riemann tensor consists in reinterpreting results achieved in the context of the *Entwurf* theory. As a consequence, also Einstein's original heuristic principles received a revised physical interpretation. The crucial step of the transition from the *Entwurf* theory, still rooted in classical physics, to the non-classical theory of general relativity was, however, the shift in the physical interpretation of the representation he had unfolded in the preceding years.[250] This transition was a Copernicus process resembling Einstein's reinterpretation of Lorentz's auxiliary variable for local time as the time measured in a moving reference frame. But in passing from the *Entwurf* theory to general relativity, however, Einstein was, in a sense, his "own Lorentz"—hence the more isolated character of the second phase of the relativity revolution. In the case of the transition to general relativity, it was, in particular, the Christoffel symbol, initially only an auxiliary quantity, that assumed a new physical meaning, now representing the gravitational field.

The synthesis represented by general relativity was not without alternatives at the time of its establishment—nor is it today. Some of these alternatives were even distinguished by consequences which could be tested empirically. The observational consequences which distinguish general relativity from its main competitor at the time, Newton's and Nordström's theory of gravitation, were, however, by no means momentous and could have easily gone unnoticed for a long time, or might have remained irrelevant for a decision between alternative theories of gravitation had Einstein's research not drawn attention to them. The contemporary discussion about these alternatives and their elaboration document a process of equilibration between individual perspectives and shared knowledge resources.[251] Even the most ingenious phase of the relativity revolution—the phase of reflection—was, from the point of view of historical epistemology, not the privilege of an outstanding individual, but just one aspect of the transformation of a system of knowledge.

---

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

251 See vols. 3 and 4 of this series.

ACKNOWLEDGEMENTS

First drafts of this chapter resulted from a collaboration in the context of the Arbeitsstelle Albert Einstein, funded in the years 1990–1995 by the Berlin Senate, see (Castagnetti et al. 1994). First and above all, we wish to thank Peter Damerow for his inspiration and support. It is through him that we learned about the potential use of concepts from cognitive science for the history of science. We are also grateful to our colleagues John Norton, John Stachel, and especially to Michel Janssen for many discussions and all their unrelenting criticism. The final version owes much to recent joint work with Michel Janssen (see "Untying the Knot ..." and "Commentary ..." (in this volume)), who also co-edited the text. We also wish to thank Diana Buchwald for her support and understanding. Last not least, we wish to express our sincere gratitude to Lindy Divarci for her reliable help with copy-editing this paper.

REFERENCES

Abraham, Max. 1914. "Die neue Mechanik." *Scientia* 15: 8–27.

Bianchi, L. 1910. *Vorlesungen über Differentialgeometrie*. Translated by Max Lukat. Leipzig: B.G.Teubner.

Buchwald, Jed. 1985. *From Maxwell to Microphysics. Aspects of Electromagnetic Theory in the Last Quarter of the Nineteenth Century.* Chicago: University of Chicago Press.

Capria, Marco Mamone. 2005. "General Relativity: Gravitation as Geometry and the Machian Programme." In M.M. Capria (ed.), *Physics Before and After Einstein*. Amsterdam: IOS, 93–128.

Cartan, Elie. 1923. "Sur les variétés à connection affine et la théorie de la relativité généralisée (première partie)." *Ecole Normale Supérieure* (Paris). *Annales* 40: 325–412.

———. 1924. "Sur les variétés à connexion affine et la théorie de la relativité généralisée (première partie, suite)". *Ecole Normale Supérieure* (Paris). *Annales* 41: 1–25.

Castagnetti, Giuseppe, Peter Damerow, Werner Heinrich, Jürgen Renn and Tilman Sauer. 1994. *Wissenschaft zwischen Grundlagenkrise und Politik: Einstein in Berlin*. Arbeitsbericht der Arbeitsstelle Albert Einstein 1991–1993. Max-Planck-Institut für Bildungsforschung.

Cattani, Carlo and Michelangelo De Maria. 1989a. "Max Abraham and the Reception of Relativity in Italy: His 1912 and 1914 Controversies with Einstein." In (Howard and Stachel 1989, 160–174).

———. 1989b. "The 1915 Epistolary Controversy between Einstein and Tullio Levi-Civita." In (Howard and Stachel 1989, 175–200).

Corry, Leo. 2004. *David Hilbert and the Axiomatization of Physics, 1898–1918: From "Grundlagen der Geometrie" to "Grundlagen der Physik."* Dordrecht: Kluwer.

Corry, Leo, Jürgen Renn and John Stachel. 1997. "Belated Decision in the Hilbert-Einstein Priority Dispute." *Science* 278: 1270–1273.

CPAE 2. 1989. John Stachel, David C. Cassidy, Jürgen Renn and Robert Schulmann (eds.), *The Collected Papers of Albert Einstein*. Vol. 2. *The Swiss Years: Writings, 1900–1909*. Princeton: Princeton University Press.

CPAE 3. 1993. Martin J. Klein, A. J. Kox, Jürgen Renn, and Robert Schulmann (eds.), *The Collected Papers of Albert Einstein*. Vol. 3. *The Swiss Years: Writings, 1909–1911*. Princeton: Princeton University Press.

CPAE 4. 1995. Martin J. Klein, A. J. Kox, Jürgen Renn, and Robert Schulmann (eds.), *The Collected Papers of Albert Einstein*. Vol. 4. *The Swiss Years: Writings, 1912–1914*. Princeton: Princeton University Press.

CPAE 5. 1993. Martin J. Klein, A. J. Kox, and Robert Schulmann (eds.), *The Collected Papers of Albert Einstein*. Vol. 5. *The Swiss Years: Correspondence, 1902–1914*. Princeton: Princeton University Press.

CPAE 6. 1996. A. J. Kox, Martin J. Klein, and Robert Schulmann (eds.), *The Collected Papers of Albert Einstein*. Vol. 6. *The Berlin Years: Writings, 1914–1917*. Princeton: Princeton University Press.

CPAE 8. 1998. Robert Schulmann, A. J. Kox, Michel Janssen, and József Illy (eds.), *The Collected Papers of Albert Einstein*. Vol. 8. *The Berlin Years: Correspondence, 1914–1918*. Princeton: Princeton University Press. [Part A: 1914-1917, pp. 1–590. Part B: 1918, pp. 591–1118.]

CPAE 9. 2004. Diana Kormos Buchwald, Robert Schulmann, József Illy, Daniel Kennefick, and Tilman Sauer (eds.), *The Collected Papers of Albert Einstein*. Vol. 9. *The Berlin Years: Correspondence, January 1919–April 1920*. Princeton: Princeton University Press.

Damerow, Peter. 1996. *Abstraction and Representation: Essays on the Cultural Revolution of Thinking*. Dordrecht: Kluwer.

Damerow, Peter, Gideon. Freudenthal, Peter McLaughlin and Jürgen Renn. 2004. *Exploring the Limits of Preclassical Mechanics,* 2nd edition. New York: Springer.

Darrigol, Oliver. 2000. *Electrodynamics from Ampère to Einstein*. Oxford: Oxford University Press.

Davis, R. 1984. *Learning Mathematics: The Cognitive Approach to Mathematics Education*. New Jersey: Ablex Publishing Corporation.

Dongen, Jeroen van. 2002. *Einstein's Unification: General Relativity and the Quest for Mathematical Naturalness*. Dissertation, University of Amsterdam.

———. 2004. "Einstein's methodology, semivectors and the unification of electrons and protons." *Archive for History of Exact Sciences* 58 (2004) 219–254.

Earman, John and Clark Glymour. 1978. "Lost in the Tensors: Einstein's Struggles with Covariance Principles 1912–1916." *Studies in History and Philosophy of Science* 9(4):251–278.

Earman, John and Michel Janssen. 1993. "Einstein's Explanation of the Motion of Mercury's Perihelion." In (Earman et al. 1993, 129–172).

Earman, John, Michel Janssen and John Norton (eds.). 1993. *The Attraction of Gravitation. New Studies in the History of General Relativity.* (*Einstein Studies,* Vol. 5), Boston/Basel/Berlin: Birkhäuser.

Earman, John and John Norton. 1987. "What Price Space-Time Substantivalism? The Hole Argument." *British Journal for the Philosophy of Science* 38: 515–525.

Ehlers, Jürgen 1981. "Über den Newtonschen Grenzwert der Einsteinschen Gravitationstheorie." In J. Nitsch et al. (eds.), *Grundprobleme der modernen Physik*. Mannheim: Bibliographisches Institut, 65–84.

———. 1986. "On Limit Relations Between, and Approximative Explanations of, Physical Theories." In R. Marcus et al. (eds.), *Logic, Methodology, and Philosophy of Science VII*. New York: Elsevier Science Publishers, 387–403.

Einstein, Albert. 1907. "Über das Relativitätsprinzip und die aus demselben gezogenen Folgerungen." *Jahrbuch der Radioaktivität und Elektronik* 4: 411–462 (CPAE 2, Doc. 47).

———. 1911. "Über den Einfluss der Schwerkraft auf die Ausbreitung des Lichtes." *Annalen der Physik* 35: 898–908 (CPAE 3, Doc. 23).

———. 1912a. "Gibt es eine Gravitationswirkung, die der elektrodynamischen Induktionswirkung analog Ist?" *Vierteljahrsschrift für gerichtliche Medizin und öffentliches Sanitätswesen* 44: 37–40 (CPAE 4, Doc. 7).

———. 1912b. "Lichtgeschwindigkeit und Statik des Gravitationsfeldes." *Annalen der Physik* 38: 355–369 (CPAE 4, Doc. 3).

———. 1912c. "Zur Theorie des statischen Gravitationsfeldes." *Annalen der Physik* 38, 443–458 (CPAE 4, Doc. 4).

———. 1913. "Zum gegenwärtigen Stande des Gravitationsproblems." *Physikalische Zeitschrift* 14:1249–1266 (CPAE 4, Doc. 17). (English translation in vol. 3 of this series.)

———. 1914a. "Die formale Grundlage der allgemeinen Relativitätstheorie." *Sitzungsberichte der Preussischen Akademie der Wissenschaften*. 2. Halbband, XLI: 1030–1085 (CPAE 6, Doc. 9).

———. 1914b. "Prinzipielles zur verallgemeinerten Relativitätstheorie und Gravitationstheorie." *Physikalische Zeitschrift* 15: 176–180 (CPAE 4, Doc. 25).

———. 1914c. "Zum Relativitäts-Problem." *Scientia* 15: 337–348 (CPAE 4, Doc. 31).

———. 1915a. "Die Feldgleichungen der Gravitation." *Sitzungsberichte der Preussischen Akademie der Wissenschaften*. 2. Halbband, XLVIII-XLIX: 844–847 (CPAE 6, Doc. 25).

———. 1915b. "Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie." *Sitzungsberichte der Preussischen Akademie der Wissenschaften*. 2. Halbband, XLVII: 831–839 (CPAE 6, Doc. 24).

———. 1915c. "Zur allgemeinen Relativitätstheorie." *Sitzungsberichte der Preussischen Akademie der Wissenschaften*. 2. Halbband, XLIV: 778–786 (CPAE 6, Doc. 21).

———. 1915d. "Zur allgemeinen Relativitätstheorie (Nachtrag)". *Sitzungsberichte der Preussischen Akademie der Wissenschaften*. 2. Halbband, XLVI: 799–801 (CPAE 6, Doc. 22).

———. 1915e. "Zusammenfassung der Mitteilung 'Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie.'" *Königlich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (XLVII):803, (CPAE 6, Doc. 24).

———. 1916a. "Die Grundlage der allgemeinen Relativitätstheorie." *Annalen der Physik* 49, 7: 769–822 (CPAE 6, Doc. 30).

———. 1916b. "Eine neue formal Deutung der Maxwellschen Feldgleichungen der Elektrodynamik." *Königlich Preußische Akademie der Wissenschaften* (Berlin). *Sitzungsberichte*, 184–188 (CPAE 6, Doc. 27).

———. 1916c. "Näherungsweise Integration der Feldgleichungen der Gravitation." *Königlich Preußische Akademie der Wissenschaften* (Berlin). *Sitzungsberichte*: 668–696, (CPAE 6, Doc. 32).

———. 1933. *About the Origins of General Theory of Relativity.* [George A. Gibson Foundation Lecture, delivered at Glasgow, 20 June 1933.] University of Glasgow: MS general 1314, (a typescript amended in Einstein's handwriting).

Einstein, Albert and A. D. Fokker. 1914. "Die Nordströmsche Gravitationstheorie vom Standpunkt des absoluten Differentialkalküls." *Annalen der Physik* 44: 321–328 (CPAE 4, Doc. 28).

Einstein, Albert and Marcel Grossmann. 1913. *Entwurf einer verallgemeinerten Relativitätstheorie und einer Theorie der Gravitation*. Leipzig: Teubner (CPAE 4, Doc. 13).

———. 1914. "Kovarianzeigenschaften der Feldgleichungen der auf die verallgemeinerte Relativitätstheorie gegründeten Gravitationstheorie." *Zeitschrift für Mathematik und Physik* 63: 215–225 (CPAE 6, Doc. 2).

Einstein, A. and L. Infeld. 1938. *The Evolution of Physics: The Growth of Ideas from Early Concepts to Relativity and Quanta*. New York: Simon & Schuster.

Eisenstaedt, Jean and A.J.Kox (eds.). 1992. *Studies in the History of General Relativity.* (*Einstein Studies,* Vol. 3). Boston/Basel/Berlin: Birkhäuser.

Galison, Peter. 2003. *Einstein's Clocks and Poincaré's Maps: Empires of Time*. New York: Norton.

Gentner, Dedre. and A. L. Stevens. 1983. *Mental Models*. Hillsdale, NJ: Erlbaum.

Gray, Jeremy (ed.). 1999. *The Symbolic Universe: Geometry and Physics, 1890–1930*. Oxford: Oxford University Press.

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

Havas, Peter. 1989. "The Early History of the "Problem of Motion" in General Relativity." In (Howard and Stachel 1989, 234–276).

Hilbert, David. 1915. "Die Grundlagen der Physik. (Erste Mitteilung)." *Nachrichten der Königlichen Gesellschaft zu Göttingen*, 395–407.

Hoffmann, Banesh. 1972. "Einstein and Tensors." *Tensor* 26: 439–162.

Holmes, Frederick L., Jürgen Renn and Hans-Jörg Rheinberger (eds.). 2003. *Reworking the Bench: Research Notebooks in the History of Science*. Dordrecht: Kluwer.

Howard, Don and John D. Norton. 1993. "Out of the Labyrinth? Einstein, Hertz, and the Göttingen Answer to the Hole Argument." In (Earman et al. 1993, 30–62).

Howard, Don, and John Stachel (eds.). 1989. *Einstein and the History of General Relativity* (*Einstein Studies,* Vol. 1). Boston: Birkhäuser.

———. 1989–2005. *Einstein Studies,* 11 vols. Boston/Basel/Berlin: Birkhäuser.

Janssen, Michel. 1995. *A Comparison between Lorentz's Ether Theory and Special Relativity in the Light of the Experiments of Trouton and Noble*. Dissertation, University of Pittsburgh.

———. 1999. "Rotation as the Nemesis of Einstein's *Entwurf* Theory." In (Goenner et al. 1999, 127–157).

———. 2005. "Of Pots and Holes: Einstein's Bumpy Road to General Relativity." *Annalen der Physik* 14: *Supplement* 58–85. Reprinted in J. Renn (ed.) *Einstein's Annalen Papers. The Complete Collection 1901–1922*. Weinheim: Wiley-VCH, 2005.

Janssen, Michel and John Stachel 2004. *The Optics and Electrodynamics of Moving Bodies*. Preprint 265: Max Planck Institute for the History of Science.

Katzir, Shaul. 2005. "Poincaré's Relativistic Theory of Gravitation." In (Kox and Eisenstaedt 2005, 15–38).

Kennefick, Daniel. 2005. "Einstein and the Problem of Motion: A Small Clue." In (Kox and Eisenstaedt 2005, 109–124).

Kox, A.J. and Eisenstaedt, Jean (eds.). 2005. *The Universe of General Relativity.* (*Einstein Studies,* Vol. 11) Boston/Basel/Berlin: Birkhäuser.

Künzle, H.P. 1976. "Covariant Newtonian Limit of Lorentz space-times." *General Relativity and Gravitation* 7(5): 445–457.

Lanczos, C. 1972. "Einstein's Path From Special to General Relativity." In L. O'Raifertaigh (ed.), *General Relativity: Papers in Honour of J. L. Synge*. Oxford: Clarendon Press, 5–19.

Laue, Max von. 1911. *Das Relativitätsprinzip*. Braunschweig: Friedr. Vieweg & Sohn.

Lorentz, Hendrik A. 1915. "Het beginsel van Hamilton in Einstein's theorie der zwaartekracht." *Koninklijke Akademie van Wetenschappen te Amsterdam. Wis- en Natuurkundige Afdeeling. Verslagen van de Gewone Vergaderingen* 23: 1073–1089. Reprinted in translation as "On Hamilton's Principle in Einstein's Theory of Gravitation." *Koninlijke Akademie van Wetenschappen te Amsterdam. Section of Sciences. Proceedings* 19 (1916–1917): 1341–1354.

———. 1895. *Versuch einer Theorie der electrischen [elektrischen] und optischen Erscheinungen in bewegten Körpern*. Leiden: Brill.
Maltese, Giulio. 1991. "The Rejection of the Ricci Tensor in Einstein's First Tensorial Theory of Gravitation." *Archive for History of Exact Sciences* 41(4): 363–-381.
Maltese, Giulio, and Lucia Orlando. 1995. "The Definition of Rigidity in the Special Theory of Relativity and the Genesis of the General Theory of Relativity." *Studies in History and Philosophy of Modern Physics* 26(3): 263–306.
Mehra, Jagdesh. *Einstein, Hilbert, and the Theory of Gravitation. Historical Origins of General Relativity Theory.* Dordrecht: Reidel.
Miller, Arthur I. 1992. "Albert Einstein's 1907 Jahrbuch Paper: The First Step from SRT to GRT." In (Eisenstaedt and Kox 1992, 319–335).
Minkowski, Hermann. 1908. "Die Grundgleichungen für die elektromagnetischen Vorgänge in bewegten Körpern." *Königliche Gesellschaft der Wissenschaften Zu Göttingen, Nachrichten. Mathematisch-physikalische Klasse*, 53–111.
Minsky, Marvin. 1975. "A Framework for Representing Knowledge." In P. H. Winston (ed.), *The Psychology of Computer Vision*. New York: McGraw-Hill, 211–281.
———. 1987. *The Society of Mind*. London: Heinemann.
Norton, John D. 1984. "How Einstein Found his Field Equations, 1912–1915." *Historical Studies in the Physical Sciences* 14: 253–316. Reprinted in (Howard and Stachel 1989, 101–159). Page references are to this reprint.
———. 1985. "What Was Einstein's Principle of Equivalence?" *Studies in History and Philosophy of Science* 16. Reprinted in (Howard and Stachel 1989, 5–47.)
———. 1992a. "Einstein, Nordström and the Early Demise of Scalar, Lorentz Covariant Theories of Gravitation." *Archive for the History of Exact Sciences* (45) 1: 17–94.
———. 1992b. "The Physical Content of General Covariance." In (Eisenstaedt and Kox 1992, 281–315).
———. 1994. "General Covariance and the Foundations of General Relativity: Eight Decades of Dispute." *Reports on Progress in Physics* 56: 791–858.
———. 1999. "Geometries in Collision: Einstein, Klein, and Riemann." In (Gray 1999, 128–144).
———. 2000. "'Nature is the Realisation of the Simplest Conceivable Mathematical Ideas': Einstein and the Canon of Mathematical Simplicity." *Studies in History and Philosophy of Modern Physics* 31: 135–170.
O'Neill, Barrett. 1983. *Semi-Riemannian Geometry with Applications to Relativity.* New York: Academic Press.
Pais, A. 1982. *'Subtle is the Lord ...' The Science and the Life of Albert Einstein*. Oxford: Oxford University Press.
Planck, Max. 1906. "Das Prinzip der Relativität und die Grundgleichungen der Mechanik." *Deutsche Physikalische Gesellschaft. Verhandlungen* 8: 136–141.
———. 1907. "Zur Dynamik bewegter Systeme." *Annalen der Physik* 26: 1–34.
Reich, Karin. 1994. *Die Entwicklung des Tensorkalküls. Vom absoluten Differentialkalkül zur Relativitätstheorie*. Basel: Birkhäuser.
Renn, Jürgen. 1993. "Einstein as a Disciple of Galileo: A Comparative Study of Conceptual Development in Physics." *Science in Context* (special issue) 6: 311–341.
———. 1997. "Einstein's Controversy with Drude and the Origin of Statistical Mechanics in his Atomism: A New Glimpse from the 'Love Letters'." *Archive for History of Exact Sciences* 4: 315–354.
———. 2000. "Mentale Modelle in der Geschichte des Wissens: Auf dem Wege zu einer Paläontologie des mechanischen Denkens." In E. Henning (ed.), *Dahlemer Archivgespräche*. Berlin: Archiv zur Geschichte der Max-Planck-Gesellschaft, 83–100.
———. 2004. "Die klassische Physik vom Kopf auf die Füße gestellt: Wie Einstein die Spezielle Relativitätstheorie fand." *Physik Journal* 3: 49–55.
——— (ed.). 2005a. *Documents of a Life's Pathway.* Weinheim: Wiley-VCH.
———. 2005b. "Before the Riemann Tensor: The Emergence of Einstein's Double Strategy." In (Kox and Eisenstaedt 2005, 53–66).
———. 2005c. "Standing on the Shoulders of a Dwarf: General Relativity—a Triumph of Einstein and Grossmann's Erroneous Entwurf Theory." In (Kox and Eisenstaedt 2005, 39–52).
Renn, Jürgen, Giuseppe Castagnetti and Peter Damerow. 1999. "Albert Einstein: alte und neue Kontexte in Berlin." In J. Kocka (ed.), *Die Königlich Preußische Akademie der Wissenschaften zu Berlin im Kaiserreich*. Berlin: Akademie Verlag, 333–354.
Renn, Jürgen, Peter Damerow and S. Rieger. 2001. "Hunting the White Elephant: When and How did Galileo Discover the Law of Fall? (with an Appendix by Domenico Giulini)." In J. Renn (ed.), *Galileo in Context*. Cambridge: Cambridge University Press, 29–149.

Renn, Jürgen and Tilman Sauer. 1996. “Einstein’s Züricher Notizbuch. Die Entdeckung der Feldgleichungen der Gravitation im Jahre 1912.” *Physikalische Blätter* 53 (1996) 856–872.

———. 1999. “Heuristics and Mathematical Representation in Einstein’s Search for a Gravitational Field Equation.” In (Goenner et al. 1999, 87–125).

———. 2003a. “Errors and Insights: Reconstructing the Genesis of General Relativity from Einstein’s Zurich Notebook.” In (Holmes et al. 2003, 253–268).

———. 2003b. “Eclipses of the Stars - Mandl, Einstein, and the Early History of Gravitational Lensing.” In A. Ashtekar et al. (eds.), *Revisting the Foundations of Relativistic Physics - Festschrift in Honour of John Stachel*. Dordrecht: Kluwer Academic Publishers.

Renn, Jürgen, Tilman Sauer and John Stachel. 1997. “The Origin of Gravitational Lensing. A Postscript to Einstein’s 1936 Science Paper.” *Science* (275) 5297: 184–186.

Ricci, G. and T. Levi-Civita. 1901. “Méthodes de calcul différentiel absolu et leurs applications.” *Mathematische Annalen* 54: 125–201.

Salmon, Merrilee H. et al. 1999. *Introduction to the Philosophy of Science: A Text by Members of the Department of the History and Philosophy of Science of the University of Pittsburgh*. Indianapolis: Hackett.

Sauer, Tilman. 1999. “The Relativity of Discovery: Hilbert’s First Note on the Foundations of Physics.” *Archive for History of Exact Sciences* (53) 529–575.

———. 2002. “Hopes and Disappointments in Hilbert’s Axiomatic ‘Foundations of Physics.” In M. Heidelberger and F. Stadler (eds.), *History of Philosophy and Science: new trends and perspectives*. Dordrecht: Kluwer, 225–237.

———. 2005a. “Albert Einstein, Review Paper on General Relativity Theory.” In I. Grattan-Guiness (ed.), *Landmark Writings in Western Mathematics* 1640–1940. Amsterdam: Elsevier, 802–822.

———. 2005b. “Einstein Equations and Hilbert Action. What is Missing on Page 8 of the Proofs for Hilbert’s First Communication on the Foundations of Physics?” *Archive for History of Exact Sciences* 59 (2005) 577–590.

———. 2006. “Field Equations in Teleparallel Spacetime. Einstein’s *Fernparallelismus* Approach Toward Unified Field Theory.” *Historia Mathematica*, in press.

Scheibe, Erhard. 1997. *Die Reduktion physikalischer Theorien. Ein Beitrag zur Einheit der Physik. Teil I: Grundlagen und elementare Theorie*. Berlin: Springer.

———. 1999. *Die Reduktion physikalischer Theorien. Ein Beitrag zur Einheit der Physik. Teil II: Inkommensurabilität und Grenzfalreduktion*. Berlin: Springer.

Sommerfeld, A. 1910a. “Zur Relativitätstheorie. I. Vierdimensionale Vektoralgebra.” *Annalen der Physik* 32(9): 749–776.

———. 1910b. “Zur Relativitätstheorie. II. Vierdimensionale Vektoranalysis.” *Annalen der Physik* 33(14): 649–689.

Stachel, J. 1980. “Einstein and the Rigidly Rotating Disk.” In A. Held (ed.), *General Relativity and Gravitation: A Hundred Years after the Birth of Einstein*. New York: Plenum, 1–15.

———. 1982a. “The Genesis of General Relativity.” In H. Nelkowski, H. Poser, R. Schrader and R. Seiler (eds.), *Einstein Symposion Berlin aus Anlass der 100. Wiederkehr seines Geburtstages*. Berlin: Springer, 428–442.

———. 1987b. “How Einstein Discovered General Relativity.” In M. A. H. MacCallum (ed.), *General Relativity and Gravitation: Proceedings of the 11th International Conference on General Relativity and Gravitation*. Cambridge: Cambridge University Press, 200–208.

———. 1989a. “The Rigidly Rotating Disk as the ‘Missing Link” in the History of General Relativity.” In (Howard and Stachel 1989, 48–62).

———. 1989b. “Einstein’s Search for General Covariance, 1912–1915.” In (Howard and Stachel, 63–100).

———. 1994. “Changes in the Concepts of Space and Time Brought About by Relativity.” In C. C. Gould and R. S. Cohen (eds.), *Artifacts, Representations, and Social Practice*. Dordrecht: Kluwer Academic Publishers, 141–162.

———. 1995. “History of Relativity.” In L. M. Brown, A. Pais and B. Pippard (eds.), *Twentieth Century Physics*, Vol. 1. Bristol: Institute of Physics Pub., Philadelphia: American Institute of Physics Press, 249–356.

———. 2002. *Einstein from ‘B’ to ‘Z’*. (*Einstein Studies,* Vol. 9) Boston/Basel/Berlin: Birkhäuser.

Vizgin, Vladimir P. 2001. “On the discovery of the gravitational field equations by Einstein and Hilbert: new materials.” *Physics-Uspekhi* 44(12): 1283–1298.

Vizgin, Vladimir P. and Y. A. SmorodinskiÎ. 1979. “From the Equivalence Principle to the Equations of Gravitation.” *Soviet Physics-USPEKHI* 22 (1979) 489–513.

Walter 1999. “Minkowski, Mathematicians, and the Mathematical Theory of Relativity.” In (Goenner et al. 1999, 45–86).

Whittaker, Edmund. 1951. *A History of the Theories of Aether and Electricity. Vol. 1: The Classical Theories*. London: Nelson.

———.1953. *A History of the Theories of Aether and Electricity. Vol. 1: The Modern Theories 1900–1926*. London: Nelson (both volumes reprinted by Dover in 1989).

Wright, J. E. 1908. *Invariants of Quadratic Differential Forms*. Cambridge: Cambridge University Press.