



Frames and stresses in Einstein's quest for a generalized theory of relativity

Olivier Darrigol

► To cite this version:

Olivier Darrigol. Frames and stresses in Einstein's quest for a generalized theory of relativity. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 2019, 68, pp.126-157. 10.1016/j.shpsb.2019.06.002 . hal-02950522

HAL Id: hal-02950522

<https://hal.science/hal-02950522>

Submitted on 21 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

FRAMES AND STRESSES
IN
EINSTEIN'S QUEST FOR A GENERALIZED THEORY OF RELATIVITY

Olivier Darrigol, UMR SPHere, CNRS, Paris, France

Summary

We already have extraordinarily detailed and competent accounts of the genesis of general relativity, and a few suggestive summaries of these accounts. This article offers a medium-sized, matter-of-fact account, followed by a critical commentary and a pocket-history for the hurried physicist. It is based on an independent study of Einstein's relevant writings, with special attention to his continual concern with measurement and reference-frames on the one hand and to his central requirement of a stress-based field dynamics on the other hand. It leads to a new evaluation of the relative importance of Einstein's various heuristic principles.

Keywords: General relativity; Albert Einstein; reference-frames; measurement; stress-energy tensor.

FRAMES AND STRESSES
IN
EINSTEIN'S QUEST FOR A GENERALIZED THEORY OF RELATIVITY

Summary

We already have extraordinarily detailed and competent accounts of the genesis of general relativity, and a few suggestive summaries of these accounts. This article offers a medium-sized, matter-of-fact account, followed by a critical commentary and a pocket-history for the hurried physicist. It is based on an independent study of Einstein's relevant writings, with special attention to his continual concern with measurement and reference-frames on the one hand and to his central requirement of a stress-based field dynamics on the other hand. It leads to a new evaluation of the relative importance of Einstein's various heuristic principles.

Keywords: General relativity; Albert Einstein; reference-frames; measurement; stress-energy tensor.

Introduction

By mid-1912 Albert Einstein had the idea of a generalized theory of relativity based on generally covariant equations for the coefficients of a pseudo-Riemannian, 4-dimensional metric. In retrospect, this idea should have quickly led him to general relativity as we

know it. Yet it took him almost four years of outwardly useless detours to arrive at this theory. Whereas historians normally focus on the emergence of new constructive ideas, in the present case they feel compelled to explain why these ideas did not sooner lead to the now expected result. This unusual challenge has attracted some of our best scholars, in the hope of learning something on the nature of theoretical invention in general and on the interpretive difficulties of general relativity in particular.

After a few pioneering studies by John Earman, Clark Glymour, and Abraham Pais, the first editor of Einstein's *Collected papers (ECP)*, John Stachel, gave his own perceptive account and a hint to an important manuscript trace of Einstein's struggles: the "Zürich notebook" written in 1912-1913. In the 1980s, John Norton first exploited this source to offer essentially new historical insights. In addition to his influential "How Einstein found his field equations" of 1984, he produced a commented and annotated version of the notebook for Volume 4 of Einstein's *Collected papers*, published in 1995.. Subsequently, a Berlin-Pittsburgh team including John Stachel, John Norton, Jürgen Renn, Tilman Sauer, Michel Janssen and a few other scholars launched a systematic study of the genesis of general relativity and of related theories and contexts, leading to the monumental *Genesis* published in 2007 in four volumes (*GGR*). Among other accomplishments, this history offers an extremely detailed and competent reconstruction of the calculations in the Zürich notebook, a powerful analysis of Einstein's contemporary heuristics (Renn and Sauer), and a penetrating study of the transition from the so-called *Entwurf* theory to general relativity in 1914-1915 (Renn and Janssen).¹

¹ Earman and Glymour 1978; Pais 1982; Stachel 1989b [1980]; Norton 1984; *ECP* 4; *GGR*. I thank John Norton for clarifying his role in this sequence.

The abundance and the quality of these studies are so high that it may seem preposterous to write still more on their topic. No history, however, is truly final, and new perspectives always remain possible. Three or four years ago, I studied the early reception of general relativity and was struck by the confusing character of Einstein's statements of 1916 regarding the meaning of coordinates and reference frames, and by seeming inconsistencies in the way he derived the three major predictions of his theory (gravitational redshift, gravitational deviation of light, relativistic precession).² I suspected that these obscurities had to do with his persisting reliance on coordination and measurement concepts that had guided his earlier attempts at a generalized theory of relativity. At that point I decided to scrutinize all relevant sources, including the Zürich notebook, before studying the relevant secondary literature and the critical editions. I could thus form an independent opinion about the driving and retarding forces in Einstein's project.³

When I compared my own reading of the sources with the best available histories, I did not find any significant error or misconception in these and I was instead impressed by their depth and accuracy. Yet, there are a few differences. My focus on reference-frames and measurement issues brings more continuity between Einstein's earliest exploitation of the equivalence principle and his later metric-field based approach. Although I agree with Norton, Renn, and Sauer's identification of Einstein's chief heuristic principles, I give them different relative weights. In particular, I believe that Einstein's *stress principle*, according to which there should exist a stress-energy (pseudo-)tensor for the gravitational field, played a much more important role, both critically and

² Darrigol 2015.

³ I deliberately did not rely on Einstein's own retrospective accounts of his past efforts.

constructively, than other considerations for instance about rotating frames and static fields. Also, I find that most of the obstacles that had early diverted Einstein from the Riemann tensor did not persist as long as assumed in earlier histories. Einstein's most resilient prejudices were bound to the stress principle. Lastly, in order to elucidate the mystery of Einstein's three or four years of wandering, I consider not only the conflict between his various heuristic principles but also the absence of a more powerful principle. Had Einstein from the start based his search for a gravitational field equation on the principle of least action, he would probably have reached his aim much faster than he did.⁴

The present article first gives a direct and minimally-commented account of Einstein's writings on generalized relativity and gravitation in the years 1907-1915, so that the reader can form his own impressions on the nature of Einstein's efforts. The second part is a more committed analysis of the dynamics of Einstein's researches, often confirming earlier studies but sometimes shifting the perspective. This analysis is then used to produce a simplified history of the genesis of general relativity, focusing on crucial steps and missteps and eliminating inessential details. The resulting pocket history (section 2.5) can be read independently of the two previous sections. Some readers may want to start with it, for it is technically less demanding and for it directly shows the consistency of the perspective adopted in this study.⁵

Notation

⁴ On principles that forcefully lead to general relativity, cf. Ehlers 1973; Darrigol 2014a, Chaps. 5, 7.

⁵ There already exist a few lucid, short histories of this genesis: see Stachel 1995; Janssen 2014; Sauer 2013 (in part); Janssen and Renn 2015; Norton [2018]. This last essay (which I did not know when I wrote mine) comes closest to my own perspective by arguing the prevalence of energy-momentum considerations over the equivalence principle in Einstein's "two-tiered" heuristics. As John Norton told me, there is a remaining difference: whereas I regard the stress-energy tensor of the gravitational field as a crucial conceptual component of Einstein's developing theory, he tends to see it as an "intermediary."

Now standard notation and conventions (largely inspired from Einstein's *Grundlage* of 1916) are used through this article (except for $\Gamma_{\mu\nu}^\rho$, which Einstein defined differently).

Greek indices run from zero to three, Latin indices from one to three. The Minkowski metric has the signature $(+,-,-,-)$. Summation over repeated indices is understood.

Upper indices are used for contravariant components, lower indices for covariant components (Einstein originally used Latin letters for covariant 2-tensors and Greek letters for contravariant 2-tensors). The product of a tensor by $\sqrt{-g}$ (giving the associated tensor density) is indicated by bold face: for instance, $\mathbf{T}_\mu^\nu = \sqrt{-g}T_\mu^\nu$ (Einstein used gothic letters).

c : velocity of light as defined in special relativity.

$\widehat{c}(\mathbf{r})$: location-dependent velocity of light in Einstein's theory of the static gravitational field.

G : gravitational constant.

x^μ : coordinates of a point of the space-time manifold.

ds^2 : square of the interval between the point-events of coordinates x^μ and $x^\mu + dx^\mu$

$g_{\mu\nu}$: metric tensor such that $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$.

g : determinant of the metric $g_{\mu\nu}$.

∂_μ : partial derivative with respect to x^μ .

D_μ : covariant derivative.

$\{\sigma_{\nu\rho}\}$: Christoffel symbol defined by $g_{\mu\sigma}\{\sigma_{\nu\rho}\} = \frac{1}{2}(\partial_\nu g_{\mu\rho} + \partial_\rho g_{\mu\nu} - \partial_\mu g_{\nu\rho})$.

$R_{\mu\nu\rho\sigma}$: Riemann curvature tensor.

$R_{\mu\nu}$: Ricci tensor, here defined as $R_{\mu\nu}^\rho$.

$t^{\mu\nu}$: stress-energy (pseudo-)tensor of the gravitational field.

$T^{\mu\nu}$: stress-energy tensor of matter (including the electromagnetic field).

$\theta^{\mu\nu}$: stress-energy tensor of a dust.

Frequently used identities are: $g_{\mu\nu} g^{\nu\rho} = \delta_{\mu}^{\rho}$, $g_{\mu\nu} \delta g^{\nu\rho} = -g^{\nu\rho} \delta g_{\mu\nu}$, $\delta g = g g^{\mu\nu} \delta g_{\mu\nu}$,

$D_{\nu} g^{\mu\nu} = 0$, and $g_{\mu\rho} D_{\nu} S^{\mu\nu} = (1 / \sqrt{-g}) \partial_{\mu} (\sqrt{-g} S^{\mu}_{\rho}) - \frac{1}{2} (\partial_{\rho} g_{\mu\nu}) S^{\mu\nu}$ for any symmetric tensor $S^{\mu\nu}$.

1. BASIC ACCOUNT

1.1 Local time and true time (1907)

Einstein first addressed the extension of relativity theory to gravitation at the end of a long review of the present state of this theory for the *Jahrbuch für Radioaktivität und Elektronik*. In the relevant section, he first introduces what he later called the *principle of equivalence*, namely: *the laws of physics should be the same in an accelerated frame without gravity and in an inertial frame with gravity equal and opposed to the acceleration of the former frame*. For the laws of free fall, this equivalence results from Galileo's observation that all bodies fall with the same acceleration in a constant gravity field independently of their mass. Einstein assumes the equivalence to extend to all laws of physics, thus generalizing the principle of relativity to accelerated frames.⁶

There is an evident similarity between the equivalence principle and the reasoning through which Einstein introduced the principle of relativity in 1905. In the received electromagnetic theory, Einstein then remarked, we express the laws of electromagnetism

⁶ Einstein 1907, p. 454. Cf., e.g., Janssen 2012, p. 160.

differently in the ether frame and in another inertial frame. Similarly, in the received theory of gravitation we regard the free motion of a body in a uniform gravity and the motion of a body in an accelerated frame as fundamentally different processes. In the first case, the asymmetry is removed by making the electric and magnetic fields frame-dependent concepts. In the second case, the asymmetry is removed by making the gravitational field a frame-dependent concept. This explains why Einstein regards the principle of equivalence as an extension of the relativity principle to accelerated frames.⁷

Einstein's next step is kinematic: he relates the measurement of space and time in an accelerated frame to their measurement in an inertial frame. For this purpose, he considers a frame Σ moving with the constant acceleration γ along the x axis of the inertial frame S . At every instant of this motion (with respect to S), there is an inertial frame S' whose axes coincide with those of Σ . This frame, now called the *tangent frame*, plays an essential role in the rest of the reasoning.⁸

A priori, Einstein tells us, the lengths of identically built rods and the rate of identically built clocks might differ in the frames Σ and S' owing to the acceleration of Σ . This effect can only be of second order in γ since opposite accelerations obviously produce the same effect. Einstein therefore decides to neglect it. Consequently, within a small enough time lapse, rod-based surveying and optical synchronization of clocks should yield the same space and time measurements in both frames.

Now suppose that at $t = 0$ in S , the origin of Σ coincides with the origin of S and the velocity of Σ with respect to S vanishes. Einstein conceives two ways of defining time in Σ . Identically built clocks being attached to every point of this frame, the first

⁷ Einstein 1905, p. 891.

⁸ Cf. Pais 1982, pp. 180-182.

option is to synchronize these clocks at $t = 0$; call σ the time given by these clocks. The second option is to constantly resynchronize the clocks of Σ with its central clock; call τ the time then given by these clocks. According to these definitions, equal σ times in Σ correspond to equal t times in S ; and equal τ times in Σ correspond to equal t' times in the tangent frame S' . Setting the origin of the time t' so that it coincides with the time t of S at the common origin of S' and Σ , we therefore have

$$(1) \quad \tau - \sigma = t' - t.$$

Calling v the velocity of S' with respect to S at time t' and neglecting second-order terms in v , we have

$$(2) \quad t' - t = -vx'/c^2 = -v\xi/c^2.$$

The σ time and the τ time therefore differ according to

$$(3) \quad \sigma = \tau(1 + v\xi/c^2).$$

As Einstein explains, the σ time is the time we should naturally use when defining physical quantities at a given location in Σ , because these quantities should be measured with standard devices brought to the location. In contrast, the τ time is the one with respect to which the invariance of physical laws is to be stated. For this reason, Einstein calls τ the "time" *tout court*, whereas he calls σ the "local time."⁹

According to the equivalence principle, the accelerated frame Σ should be equivalent to an inertial frame in which the constant gravity $-\gamma$ acts. The gravitational potential in this field being $\Phi = v\xi/c^2$, we may rewrite Eq. (3) as

$$(4) \quad \sigma = \tau(1 + \Phi/c^2).$$

⁹ Einstein 1907, pp. 454-457.

Now suppose we bring two exemplars of the same clock to two different locations. By the former argument, they yield the local time σ . Einstein imagines an observer comparing light signals from the two clocks at a third location. The τ time that a signal takes to travel from one clock to the observer is a constant. Accordingly, the period of the clock immersed in the smaller gravitational potential will appear to be larger to the observer. In other words, this clock runs slower. Extending this law to a non-uniform gravitational field and applying it to spectral emitters, a spectral line from a source situated at the surface of the sun should appear to us red-shifted by about 2×10^{-6} in relative value.¹⁰

Readers familiar with modern general relativity can easily see that Einstein's local time σ corresponds to the proper time given by the invariant metric element ds , whereas the τ time corresponds to the time coordinate with respect to which the metric coefficients are constants in a static gravitational field. Whereas we now regard the proper time as the only fundamental time, Einstein privileges the τ time and consequently talks about the gravitational slowing down of clocks where one should rather speak of the non-conservation of the interval ds between the light signals emitted at two successive ticks of the clock and traveling from the remote source to the observer (the conserved quantity being $dt = (1 + \Phi / c^2)^{-1} ds$).

Imitating special relativity, Einstein proceeds to show that the form of the fundamental equations of electrodynamics is preserved in the accelerated frame Σ , *mutatis mutandis*. In the tangent frame S' , the Maxwell-Lorentz equations for the electric

¹⁰ Einstein 1907, pp. 457-459.

field \mathbf{E}' , the magnetic field \mathbf{H}' , and the current $\rho_e \mathbf{u}$ (for an electric charge of density ρ_e moving at the velocity \mathbf{u}) read

$$(5) \quad \nabla' \times \mathbf{H}' = \frac{1}{c} \left(\rho_e' \mathbf{u}' + \frac{\partial \mathbf{E}'}{\partial t'} \right), \quad \nabla' \times \mathbf{E}' = -\frac{1}{c} \frac{\partial \mathbf{H}'}{\partial t'}.$$

At times infinitesimally close to $t'=0$, the time lapse dt' is identical with the time lapse $d\sigma$ in the accelerated frame, the coordinates (x', y', z') agree with (ξ, η, ζ) , and the fields \mathbf{E} and \mathbf{H} in the accelerated frame are related to the fields in the tangent frame by

$$(6) \quad \mathbf{E}' = \mathbf{E} - \frac{\boldsymbol{\gamma}'}{c} \times \mathbf{H}, \quad \mathbf{H}' = \mathbf{H} + \frac{\boldsymbol{\gamma}'}{c} \times \mathbf{E}.$$

In terms of this new variables, the field equations read

$$(7) \quad \nabla \times \mathbf{H} = \frac{1}{c} \left(\rho_e \mathbf{u} + \frac{\partial \mathbf{E}}{\partial \sigma} \right) - \frac{\boldsymbol{\gamma}}{c^2} \times \mathbf{H}, \quad \nabla \times \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{H}}{\partial \sigma} - \frac{\boldsymbol{\gamma}}{c^2} \times \mathbf{E}.$$

Introducing the fields $\mathbf{E}^* = \mathbf{E}(1 + \gamma^2 / c^2)$, $\mathbf{H}^* = \mathbf{H}(1 + \gamma^2 / c^2)$, and $\rho_e^* = \rho_e(1 + \gamma^2 / c^2)$, this is equivalent to

$$(8) \quad \nabla \times \mathbf{H}^* = \frac{1}{c} \left(\rho_e^* \mathbf{u} + \frac{\partial \mathbf{E}^*}{\partial \sigma} \right), \quad \nabla \times \mathbf{E}^* = -\frac{1}{c} \frac{\partial \mathbf{H}^*}{\partial \sigma}.$$

Einstein thus retrieves the form of the equations in an inertial frame.¹¹

According to Einstein, non-stationary electromagnetic processes must be referred to the true time $\tau = \sigma(1 + \gamma^2 / c^2)^{-1}$. Neglecting the concomitant change in the gradient operator, he gets

$$(9) \quad \nabla \times \mathbf{H}^* = \frac{1}{c_\tau} \left(\rho_e^* \mathbf{u}_\tau + \frac{\partial \mathbf{E}^*}{\partial \tau} \right), \quad \nabla \times \mathbf{E}^* = -\frac{1}{c_\tau} \frac{\partial \mathbf{H}^*}{\partial \tau}, \quad \text{with } c_\tau = c(1 + \gamma^2 / c^2).$$

¹¹ Einstein 1907, pp. 459-460.

Consequently, in a gravitational field of acceleration $-\gamma$, light travels with the location-dependent velocity

$$(10) \quad c_\tau = c(1 + \Phi / c^2) .$$

By analogy with the optics of heterogeneous transparent media, this implies a deviation of the light rays at a rate proportional to γ / c^2 . At that time Einstein considered only the deviation by the gravitational field of the earth, which he judged much too small to be detectable.¹²

Lastly, from Eq. (8) Einstein derives

$$(11) \quad \int \rho_e^* \mathbf{u} \cdot \mathbf{E}^* d\xi d\eta d\zeta = \frac{d}{d\sigma} \int \frac{1}{2} (E^{*2} + H^{*2}) d\xi d\eta d\zeta .$$

In terms of the truly measured field \mathbf{E} and \mathbf{H} and of the true time τ , this gives

$$(12) \quad \int \rho_e \mathbf{u}_\tau \cdot \mathbf{E} (1 + \Phi / c^2) d\xi d\eta d\zeta = \frac{d}{d\tau} \int \frac{1}{2} (E^2 + H^2) (1 + \Phi / c^2) d\xi d\eta d\zeta .$$

This means that the energy $\rho_e \mathbf{u}_\tau \cdot \mathbf{E}$ brought by the currents to the field per time unit and the field energy $(E^2 + H^2) / 2$ per unit volume must both be corrected by a gravitational term corresponding to the potential energy of the associated masses. In other words, an added energy implies not only an added inertial mass but also an added gravitational mass.

To sum up, in these pioneering considerations Einstein investigated the effects of the acceleration of a reference frame on space and time measurement, on electromagnetic processes, and on the energy balance. He thereby introduced two different time variables and used them to derive three promising consequences of general relativity: the gravitational redshift, the gravitational deviation of light rays, and the energy dependence

¹² Einstein 1907, pp. 461-462. I have corrected a number of typos in the published text.

of the gravitational mass. As he later realized, his reasoning based on the transformed Maxwell-Lorentz equations is shaky for it involves uncontrolled approximations in the change of coordinates from an inertial frame to an accelerated frame. As we may retrospectively judge, one consequence of this reasoning, the gravitational deviation of light rays is qualitatively correct; and the other, the dependence of the gravitational mass on energy content, is exact.

1.2 Direct applications of the principle of equivalence (1911)

Four year elapsed before Einstein published again on generalized relativity. He had two motivations to do so: he was no longer satisfied with his considerations of 1907, and he now realized that the gravitational deviation of light rays might be observable for the light seen from stars in a direction close to the surface of the sun (it would then be of the order of a second of arc). His new reasoning was more elementary and rested on direct applications of the equivalence principle, in the strong version of which an accelerated frame simulates uniform gravity in an inertial frame "with respect to all physical processes."¹³

A first consequence of the principle is the equality of the inertial and gravitational masses associated with a given amount of energy. This is immediately seen by noting that the acceleration of a body carrying this energy should be the same in an inertial frame with the constant gravity γ and in an accelerated frame with the acceleration $-\gamma$. In

¹³ Einstein 1911, p. 900. Cf. Pais 1982, pp. 198-200.

addition to this simple reasoning, Einstein considers the following concrete operations.

The source S_2 sends radiation of energy E_2 to the absorber S_1 in the direction of the acceleration of gravity γ (see Fig. 1) in the inertial frame K . According to the principle of equivalence, the process should be the same in a gravity-free frame K' with the acceleration $-\gamma$. In the inertial frame tangent to the K' frame at the time of emission, the radiation reaches S_1 when S_1 has acquired the velocity $-\gamma h / c$ (in a first approximation). The energy of the radiation with respect to S_1 therefore is

$$(13) \quad E_1 = E_2(1 + \gamma h / c^2)$$

(using the transformation laws of special relativity). In the frame K , the product γh is the difference Φ between the gravitational potentials at the locations of S_2 and S_1 . Therefore, the energy of the radiation emitted at S_1 exceeds the energy absorbed at S_2 by the potential energy $(E_2 / c^2)\Phi$ of the mass E_2 / c^2 . In other words, the energy is conserved in this process if any addition of energy to a body contributes to its gravitational mass (from which the potential energy is computed) just as much as it contributes to its inertial energy.¹⁴

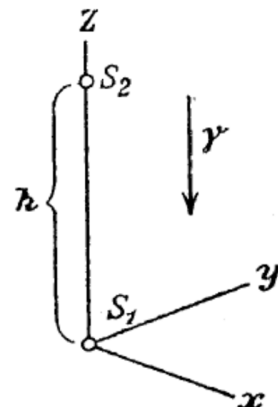


Fig. 1: A source S_2 and an absorber S_1 in the gravitational field γ . From Einstein 1911, p. 901.

¹⁴ Einstein 1911, pp. 900-902. On pp. 902-903, Einstein also gives another reasoning based on a cycle similar to the one used in an earlier derivation of $E = mc^2$.

By similar reasoning, monochromatic radiation sent in K' from S_2 with the frequency ν_2 arrives in S_1 with the Doppler-shifted frequency

$$(14) \quad \nu_1 = \nu_2 (1 + \gamma h / c^2).$$

In the frame K , this means that light from a source placed in a higher gravitational potential appears to be shifted toward the violet, by the relative amount Φ / c^2 if Φ denotes the excess of potential. In his reasoning, Einstein makes clear that the frequency of the same spectral source as judged at the two locations by identically built clocks should be the same. It is only when the clock used to measure the frequency is at a location different from the location of the source that the frequency shift occurs.¹⁵

How could the frequency of the light not be conserved during its travel from the source to the observer? In agreement with his two-time reasoning of 1907, Einstein solves this paradox by arguing that ν_1 and ν_2 are not true frequencies since they are not referred to the "true time" with respect to which the numbers of oscillations of a wave extending between S_2 and S_1 is stationary. It is only with respect to this true time that the laws of physics do not explicitly depend on time. In order to measure this time, clocks must be built so that in a location of potential Φ they run $1 + \Phi / c^2$ times slower than a standard clock brought to this location.¹⁶

The velocity of light, when measured with a standard clock at any location should always have the same constant value c whatever be the gravitational potential. From this

¹⁵ Einstein 1911, pp. 903-905.

¹⁶ Einstein 1911, pp. 905-906.

consequence of the principle of equivalence, Einstein infers that the light velocity \hat{c} measured with respect to the true time depends on the location according to

$$(15) \quad \hat{c}(\mathbf{r}) = c[1 + \Phi(\mathbf{r}) / c^2].$$

The constancy of the velocity of light, Einstein concludes, no longer holds in the new theory.¹⁷

Accordingly, the path of light should be deviated when it travels through an intense gravitational field. Einstein computes the amount of the deviation by Huygens's principle. For the deviation by a celestial body of mass M , he finds $2GM / c^2 \Delta$, wherein G denotes the gravitational constant and Δ the closest distance of the light ray from the center of the body. This gives 4×10^{-6} in the case of stars appearing close to the surface of the sun. Einstein ends by calling astronomers to check for effects of this kind.¹⁸

To sum up, in this brief memoir Einstein retrieved his three predictions of 1907: the gravitational redshift of stellar light, the gravitational deviation of light rays, and the relation between gravitational mass and energy content. Although the reasoning was more direct and more solid, it still involved two effects seemingly at variance with modern general relativity: the slowing down of clocks and the increase of the velocity of light in a gravitational potential. Both oddities derive from Einstein's definition of time as what would now be the time coordinate in a static metric field, and the velocity of light with what is now called the coordinate-velocity. Einstein still saw the equivalence principle as a step toward a fuller relativity in which all frames of reference would be equally acceptable, with a uniform rule for transforming fields from one frame to another. Just as there was no absolute velocity in special relativity, there was no absolute

¹⁷ Einstein 1911, p. 906.

¹⁸ Einstein 1911, pp. 906-908.

acceleration in the generalized relativity. Einstein did not give the transformation formulas, although he hinted he already had them in the case of constant acceleration.

1.3 The static gravitational field (1912)

At the end of 1911, the renowned Göttingen theorist and experimentalist Max Abraham generalized Minkowski's 4-dimensional spacetime to accommodate for Einstein's gravity-dependent velocity of light, without any recourse to the equivalence principle. The formal elegance of this theory, about which more in a moment, challenged Einstein. In a memoir published in February 1912, he rejected Abraham's theory and proposed his own, equivalence-based theory in the more limited case of a static gravitational field.¹⁹

The velocity of light and the static field

Einstein again considers a reference frame K in constant acceleration with respect to the inertial frame Σ (the notation differs from that of 1907). Thanks to Max Born's earlier work on rigid motion, he can now give a precise meaning to the constancy of the acceleration: at every instant the acceleration is constant with respect to the tangent frame to K . The first problem is to find the relation between the coordinates (x, y, z, t) in the accelerated frame and the coordinates (ξ, η, ζ, τ) in the inertial frame. For this purpose, he assumes that the length gauges in K are not affected by its acceleration, and he also assumes that the velocity of light as measured near a given point of K does not depend on

¹⁹ Abraham 1911, 1912a; Einstein 1912a. On the latter memoir and its sequel, cf. Pais 1982, pp. 201-206.

the direction of propagation. In passing, he remarks that the first assumption is far from obvious and that it is likely not to hold in the case of a frame rotating around an axis in an inertial frame, because with respect to the latter frame radial rods are not affected by the rotation while orthoradial rods are subject to the Lorentz contraction. He also notes that the implied rigidity can only be of a kinematic kind.²⁰

From these metric assumptions, Einstein first infers that $y = \eta$ and $z = \zeta$ for the coordinates perpendicular to the acceleration. Then he reasons that for two light-connected, infinitesimally close events one must have $dx^2 + dy^2 + dz^2 - \hat{c}^2 dt^2 = 0$ whenever $d\xi^2 + d\eta^2 + d\zeta^2 - d\tau^2 = 0$. The velocity of light is set to *one* in the inertial frame, and its value \hat{c} in K may depend on the abscissa x . Without further ado, Einstein requires²¹

$$(16) \quad dx^2 - \hat{c}^2(x)dt^2 = d\xi^2 - d\tau^2.$$

Einstein then seeks a non-linear, second-degree transformation for which this relation holds to a sufficient approximation. He finds that such a transformation exists if and only if

$$(17) \quad \hat{c} = \hat{c}_0 + ax.$$

To second order, the transformation is given by

$$(18) \quad \xi = x + a\hat{c}t^2 / 2, \quad \tau = \hat{c}t$$

²⁰ Einstein 1912a, pp. 355-357; Born 1909a, 1909b. In private, Einstein repeatedly discussed the rotating disk in this period: cf. Stachel 1980, 1989a, 2007, who argues that in this context Einstein became aware of the possible lack of direct metric significance of spatial coordinates. In print, however, Einstein was much more concerned with the meaning of the time coordinate.

²¹ The two forms must be proportional, and the proportionality coefficient reduces to *one* if the length unit is the same in both frames.

if the origins of the two frames coincide and if their relative velocity vanishes at $t = 0$.

For small times, the origin of K travels with the constant acceleration a/\hat{c}_0 with respect to Σ .²²

Had Einstein been familiar with the theory of quadratic differential forms or with the theory of Gaussian surfaces, he would have known that the Gaussian curvature of the form $dx^2 - \hat{c}^2 dt^2$, being equal to $-\hat{c}''/\hat{c}$, vanishes if and only if \hat{c} varies linearly in x . This is also the condition for the existence of a transformation that brings this form to the (pseudo-) Euclidean form. By analogy with the polar-coordinate transformation

$X = \rho \cos \theta$, $Y = \rho \sin \theta$ that turns $dX^2 + dY^2$ into $d\rho^2 + \rho^2 d\theta^2$, the exact expression of the transformation sought by Einstein is

$$(19) \quad \xi = \hat{c} \frac{\cosh at}{a} - \frac{\hat{c}_0}{a}, \quad \tau = \hat{c} \frac{\sinh at}{a}.$$

Einstein does not need these formulas because, as we are about to see, he wants \hat{c} to be any function of the spatial coordinates, in which case the curvature does not vanish and the transformation formulas have only local validity.

From his earlier work, Einstein knows that \hat{c} plays the role of the gravitational potential (at least in the case of uniform gravity). He therefore requires the Laplacian $\Delta \hat{c}$ to be proportional to the density ρ of matter. In addition, he requires his field equation to be homogeneous in \hat{c} , because the unit for the time t is arbitrary by the former reasoning (it depends on the location in K of the standard clock with respect to which the other clocks of K are synchronized). As the simplest equation meeting these conditions, he takes

²² Einstein 1912a, pp. 357-359. I use \hat{c} instead of Einstein's c in order to avoid confusion with the usual constant of nature c .

$$(20) \quad \Delta \hat{c} = k \hat{c} \rho ,$$

wherein k is 4π times the usual gravitation constant G .²³

Einstein's next task is to determine the motion of a particle in the static field \hat{c} . For this purpose, he considers the motion for small time t in which the coordinate transformation of Eqs. (18) can locally be used to eliminate the gravitation field. With respect to the transformed coordinates, the motion is rectilinear and uniform. Switching to the (x, y, z, t) coordinates, it is easily seen to satisfy

$$(21) \quad \frac{d}{dt} \left(\frac{\mathbf{v}}{\hat{c}^2} \right) = - \frac{\nabla \hat{c}}{\hat{c}} \quad \text{with } \mathbf{v} = d\mathbf{r} / dt \quad \text{and } \mathbf{r} = (x, y, z) .$$

Although this equation is here established only for small times and for a specific choice of the x axis, Einstein takes it to be valid at any time and for any choice of the function $\hat{c}(\mathbf{r})$, because "the instant $t = 0$ has nothing special compared to other instants."²⁴

In modern terms, Einstein's procedure amounts to introducing local geodesic coordinates at any time t (coordinates for which the metric takes the Minkowskian form and the derivatives of the metric coefficients vanish) for the metric $dx^2 - \hat{c}^2 dt^2$ and assuming rectilinear uniform motion with respect to these coordinates in a sufficient approximation. Equivalently, the trajectory in spacetime should be a geodesic of the metric manifold. Being unaware of this geometric interpretation, Einstein did not immediately realize that his equations of motions derived from the variation

$$(22) \quad \delta \int \sqrt{\hat{c}^2 dt^2 - d\mathbf{r}^2} = \delta \int dt \sqrt{\hat{c}^2 - v^2} = 0 .$$

Einstein arrived at this conclusion a few months later, presumably by generalizing Max Planck's Lagrangian formulation of the equation of motion in special relativity.

²³ Einstein 1912a, p. 360.

²⁴ Einstein 1912a, pp. 361-362.

Indeed this formulation corresponds to the special case of constant \hat{c} in the former variation. The corresponding equation of motion is

$$(23) \quad \frac{d}{dt} \frac{\partial L}{\partial \mathbf{v}} = \frac{\partial L}{\partial \mathbf{r}}, \text{ with } L = \sqrt{\hat{c}^2 - v^2}.$$

This gives

$$(24) \quad \frac{d}{dt} \frac{\mathbf{v}}{\hat{c} \sqrt{1 - v^2 / \hat{c}^2}} = - \frac{\nabla \hat{c}}{\sqrt{1 - v^2 / \hat{c}^2}},$$

which is equivalent to Einstein's Eq. (21) because both equations imply the constancy of $\hat{c}(1 - v^2 / \hat{c}^2)^{-1/2}$ for any given motion.²⁵

Einstein exploits this constancy to define the energy of a moving mass m as

$$(25) \quad E = \frac{m\hat{c}}{\sqrt{1 - v^2 / \hat{c}^2}} \approx m\hat{c} + mv^2 / 2\hat{c}.$$

(This does not imply any dimensional heterogeneity because time and length have the same dimension in his units). In general, he expects forces and energies to be proportional to \hat{c} and thus to depend on the gravitational potential at the location of the system. For instance, he considers the energy of a compressed spring. A particle of mass m projected by this spring acquires a velocity independent of the location if only this velocity is measured with respect to the local time $l = \hat{c}t$. This is a consequence of the equivalence principle, which Einstein here takes to imply that *the relations between quantities measured at a given location with identically built apparatus should give the same result whatever be the value of the gravitational potential at this location*. The associated kinetic energy,

$$(26) \quad mv^2 / 2\hat{c} = m(dx / dt)^2 / 2\hat{c} = m\hat{c}(dx / dl)^2 / 2,$$

²⁵ Einstein 1912b, p. 458; Planck 1906.

is therefore proportional to \hat{c} and so is the force of the compressed spring (since the work of this force is responsible for the velocity of the projectile).²⁶

Just as in his earlier publications, Einstein regards the time t as more fundamental than the local time. He now calls it the "universal time." Since the rate of physical processes at a given location is independent on location when referred to the local time, all processes run faster in a higher gravitational potential. Again, the premise is a consequence of the principle of equivalence: *gravitation does not affect the measuring contraptions when acceleration does not*. Einstein confirms this in two simple cases. The first case is the light-clock obtained by counting the successive bounces of light between two parallel mirrors separated by the unit distance. The number of bounces per unit t -time is evidently proportional to \hat{c} and therefore increases with the gravitational potential. The second is the gravitation-clock obtained by counting the revolutions of a mass m around a mass M at a standard distance. By Eqs. (20) and (21), the acceleration $d^2\mathbf{r} / dt^2$ of the mass m should be proportional to \hat{c}^2 . Kinematically, this acceleration is centripetal and proportional to the square of the angular frequency. This frequency is therefore proportional to \hat{c} , as was to be proved.²⁷

Electromagnetism and gravitation

In a sequel to this memoir, Einstein investigated the electromagnetic field equations in a static gravitational field. This investigation, usually neglected by commentators, is historically important in inaugurating the strategy of covariant field equations and in considering a metrology of the transformed coordinates and fields. In conformity with the

²⁶ Einstein 1912a, pp. 365-367.

²⁷ Einstein 1912a, pp. 325-327.

principle of equivalence, Einstein simulates the gravitational field by acceleration with respect to the tangent frame. In the tangent frame, the Maxwell-Lorentz equations read

$$(27) \quad \nabla' \times \mathbf{H}' = \rho_e' \mathbf{u}' + \frac{\partial \mathbf{E}'}{\partial t'}, \quad \nabla' \times \mathbf{E}' = -\frac{\partial \mathbf{H}'}{\partial t'}, \quad \nabla' \cdot \mathbf{H}' = 0, \quad \nabla' \cdot \mathbf{E}' = \rho_e'.$$

The coordinates and fields in the accelerated frame are related to those in the tangent frame by

$$(28) \quad x' = x + a\hat{c}t^2/2, \quad y' = y, \quad z' = z, \quad t' = \hat{c}t \quad \text{with } \hat{c} = \hat{c}_0 + ax,$$

$$(29) \quad \mathbf{H} = \mathbf{H}' - \gamma t \times \mathbf{E}', \quad \mathbf{E} = \mathbf{E}' + \gamma t \times \mathbf{H}' \quad \text{with } \gamma = \hat{c}^{-1} \nabla \hat{c}.$$

$$(30) \quad \rho_e = \rho_e', \quad \mathbf{u} = \hat{c} \mathbf{u}'.$$

The field equations in the local accelerated frame at $t = t' = 0$ therefore read

$$(31) \quad \nabla \times \hat{c} \mathbf{H} = \rho_e \mathbf{u} + \frac{\partial \mathbf{E}}{\partial t}, \quad \nabla \times \hat{c} \mathbf{E} = -\frac{\partial \mathbf{H}}{\partial t}, \quad \nabla \cdot \mathbf{H} = 0, \quad \nabla \cdot \mathbf{E} = \rho_e.$$

By the equivalence principle, they also are the field equations in the static gravitational field $\hat{c}(\mathbf{r})$. The equations agree with Eqs. (9) of 1907, except that Einstein there used the fields $\mathbf{E}^* = \hat{c} \mathbf{E}$ and $\mathbf{H}^* = \hat{c} \mathbf{H}$ instead of \mathbf{E} and \mathbf{H} .²⁸

The chief novelty, besides the generalization to a heterogeneous gravitational field, is Einstein's discussion of electromagnetic field measurement in the static gravitational field. All measurements are done by what Einstein now calls "pocket" devices, following a suggestion by Paul Ehrenfest. Namely, the measurements are done with identically built devices that can be transported to the location of the measurement. The aforementioned light-clock is a pocket device for time measurement. A spring-scale is a pocket device for the measurement of forces. By the earlier given argument, the force

²⁸ Einstein 1912b, pp. 443-446.

thus measured is \hat{c} times the true force. Consequently, the electric field measured by the force acting on a unit point charge is \hat{c} times the true field \mathbf{E} .²⁹

By reasoning similar to the one used for Poynting's theorem, the field equations (31) imply the identity

$$(32) \quad \int \rho_e \mathbf{u} \cdot \mathbf{E} d^3x = -\frac{d}{dt} \int \frac{1}{2} \hat{c} (E^2 + H^2) d^3x + \int (\hat{c} \mathbf{E} \times \hat{c} \mathbf{H}) \cdot d\mathbf{S},$$

wherein the surface integral is taken over the boundary of the volume integral. The following interpretation suggests itself: the energy furnished to the field by the currents in a given domain equals the variation of the electromagnetic energy in this domain plus the electromagnetic energy flux across the boundary of this domain. The quantity

$(\hat{c}/2)(E^2 + H^2)$ thereby plays the role of the electromagnetic energy density. This means that the quantity $(1/2)(E^2 + H^2)$ plays the role of a gravitational mass, in conformity with the principle of equivalence. As such, Einstein tells us, this quantity should be added to the density of matter ρ in Eq. (20) for the gravitational field. At this point Einstein briefly mentions that the variable \hat{c} in the electromagnetic field Eqs. (31) implies the gravitational deviation of light but he does not repeat the derivation.³⁰

Action and reaction

Lastly, Einstein detects and solves a contradiction in the present theory. The equation of motion of a particle in the gravitational field, as given in Eq. (21), suggests the expression $-\nabla \hat{c}$ for the force acting on a unit mass at rest. Accordingly, the sum of all forces acting

²⁹ Einstein 1912b, pp. 446-448.

³⁰ Einstein 1912a, pp. 448-450. Einstein also notes that in all rigor, including the electromagnetic energy in the source of the gravitational field contradicts its static character. I skip the thermodynamic section of pp. 450-452.

on the static mass distribution ρ is $\int -\rho \nabla \widehat{c} d^3x$. By the principle of equality of action and reaction, this force should vanish (as long as the masses are contained within a finite domain of space). Together with the gravitational field equation (20), this condition gives

$$(33) \quad \int \frac{\Delta \widehat{c}}{\widehat{c}} \nabla \widehat{c} d^3x \equiv \mathbf{0}.$$

Unfortunately, no such identity holds for an arbitrary function $\widehat{c}(\mathbf{r})$. Something must be changed in the theory. After considering several possibilities, Einstein decides that the only reasonable option is to modify the field equation (20). For this purpose, he uses the identity

$$(34) \quad \frac{2\Delta\sqrt{\widehat{c}}}{\sqrt{\widehat{c}}} = \frac{\Delta\widehat{c}}{\widehat{c}} - \frac{1}{2} \left(\frac{\nabla\widehat{c}}{\widehat{c}} \right)^2$$

thanks to which the modified field equation

$$(35) \quad \frac{\Delta\widehat{c}}{\widehat{c}} - \frac{1}{2} \left(\frac{\nabla\widehat{c}}{\widehat{c}} \right)^2 = k\rho$$

is compatible with the equality of action and reaction. Indeed, for $u = \sqrt{\widehat{c}}$ we have

$$(36) \quad \Delta u = ku\rho/2 \quad \text{and} \quad \mathbf{f} = -\rho \nabla \widehat{c} = -2\rho u \nabla u = -4k^{-1} \Delta u \nabla u.$$

By analogy with electrostatics or by $\partial_i u \partial_k u \partial_k u = \partial_j (\partial_i u \partial_j u - \frac{1}{2} \partial_k u \partial_k u)$, the spatial integral of \mathbf{f} vanishes. In other words, there is a stress tensor from which the gravitational force derives, which automatically implies that the principle of action and reaction is globally satisfied. The field equation (35) is still homogeneous in \widehat{c} , and the new term,

being of second order in $\nabla\hat{c}$, does not sensibly affect the earlier physical considerations.³¹

Rewriting the field equation as

$$(37) \quad \Delta\hat{c} = k \left[\hat{c}\rho + \frac{1}{2k} \frac{(\nabla\hat{c})^2}{\hat{c}} \right],$$

Einstein recognizes in $\hat{c}\rho$ the gravitational energy density for the mass distribution ρ and suspects the second term in the square bracket to represent the energy of the gravitational field. He confirms this suspicion by proving that the work δW of the gravitational forces $-\rho\nabla\hat{c}$ during an infinitesimal displacement $\delta\mathbf{r}$ of the masses is equal to the variation of the total energy δU_ρ of the masses plus the variation of the total energy δU_c of the field.

The relevant calculations read:

$$(38) \quad \delta W = \int -\rho\nabla\hat{c} \cdot \delta\mathbf{r} d^3x = \int \hat{c}\nabla \cdot (\rho\delta\mathbf{r}) d^3x = \int \hat{c}\delta\rho d^3x,$$

$$(39) \quad \delta U_\rho = \int (\hat{c}\delta\rho + \rho\delta\hat{c}) d^3x,$$

$$(40) \quad \begin{aligned} \delta U_c &= \frac{1}{2k} \int \delta \frac{(\nabla\hat{c})^2}{\hat{c}} d^3x = \frac{2}{k} \int \delta(\nabla u)^2 d^3x = \frac{4}{k} \int \nabla u \cdot \delta\nabla u d^3x = -\frac{4}{k} \int \Delta u \delta u d^3x \\ &= -2 \int u \rho \delta u d^3x = - \int \rho \delta\hat{c} d^3x. \end{aligned}$$

This reasoning greatly increased Einstein's confidence in the new field equation.³²

From the equivalence principle to the metric field

³¹ Einstein 1912b, pp. 452-457. Cf. Norton 2018, pp. 15-18. Yet, on pp. 455-456, Einstein notes he had difficulty accepting the new term because it contradicts the strict equivalence principle (it does not vanish in a uniform field for which \hat{c} is a linear function of the space coordinates): cf. Norton 2018, pp. 19-20.

³² Einstein 1912b, pp. 457-458.

These twin memoirs of 1912 marked a crucial transition from piecemeal applications of the equivalence principle to a metric approach integrating this principle. The first memoir introduces the form $d\mathbf{r}^2 - \hat{c}^2(\mathbf{r})dt^2$ in order to express the isotropy of light propagation in the static field locally interpreted as an inertial acceleration field. Einstein's derivation of the coordinates (\mathbf{r}', t') in a tangent inertial frame, based on equating the previous form to $d\mathbf{r}'^2 - dt'^2$, imitates a then common derivation of the Lorentz transformation through the invariance of the Minkowski interval. Despite his awareness of Minkowski's 4-dimensional geometric conception, Einstein does not use geometric language or methods in the two memoirs.³³ His derivations betray a lack of familiarity with the methods of differential geometry.³⁴

However, there are two reasons to think that by March 1912 Einstein saw the prospects of a metric interpretation of the differential form $\hat{c}^2(\mathbf{r})dt^2 - d\mathbf{r}^2$. The first reason is found in the aforementioned remark added in the proofs of the second memoir. There he shows that the motion of a particle in the static field corresponds to an extremum of $\int \sqrt{\hat{c}^2(\mathbf{r})dt^2 - d\mathbf{r}^2}$. He could not possibly have missed the analogy with the determination of geodesics on a surface, and he probably recognized the problem of geodesics on a pseudo-Riemannian 4-dimensional manifold with the metric

$$(41) \quad ds = \sqrt{\hat{c}^2(\mathbf{r})dt^2 - d\mathbf{r}^2}.$$

Three months earlier, in December 1911, Max Abraham had tried to integrate Einstein's idea of a variable velocity of light in a Minkowskian framework. According to

³³ In 1908, Einstein had collaborated with Jakob Laub on a 3-vector reformulation of Minkowski's electrodynamics of moving bodies. In texts written after Minkowski's death (1909), he mentioned Minkowski's 4-dimensional approach favorably: see *ECP* 3, pp. 169-170 (1910), 438 (1911).

³⁴ Yet, at the *ETH* in Zürich Einstein had attended Carl Friedrich Geiser's lectures on infinitesimal geometry, which included Gauss's theory of surfaces. Cf. Stachel 2007, pp. 103-104.

Minkowski's representation," Abraham writes, "we regard x, y, z and $u = i\hat{c}t$ as coordinates in a space of four dimensions." The only difference with Minkowski's theory is that the velocity \hat{c} now depends on the gravitational potential Φ . In modern index notation, Abraham's gravitational field equation reads

$$(42) \quad \partial_\mu \partial_\mu \Phi = 4\pi G \rho_0,$$

wherein G denotes the gravitational constant and ρ_0 the density in the frame tangent of the moving matter. The equation of motion of a particle in this potential reads

$$(43) \quad \ddot{x}_\mu = -\partial_\mu \Phi,$$

the dot denoting derivation with respect to Minkowski's proper time τ for which

$$(44) \quad \hat{c} d\tau = \sqrt{-dx_\mu dx_\mu} \quad \text{and} \quad \dot{x}_\mu \dot{x}_\mu = -\hat{c}^2.$$

Combining $\dot{x}_\mu \ddot{x}_\mu = -\hat{c} \dot{\hat{c}}$ with the equation of motion, Abraham gets

$$(45) \quad -\dot{x}_\mu \partial_\mu \Phi = -\hat{c} \dot{\hat{c}}, \quad \frac{d\Phi}{d\tau} = \frac{1}{2} \frac{d\hat{c}^2}{d\tau}, \quad \text{and} \quad \Phi = \frac{\hat{c}^2}{2} - \frac{\hat{c}_0^2}{2}.$$

In a first approximation, this gives $\hat{c} \approx \hat{c}_0(1 + \Phi / \hat{c}_0^2)$, in conformity with the result

Einstein has earlier obtained through the equivalence principle. In addition, Abraham finds $m\hat{c} / \sqrt{1 - v^2 / \hat{c}^2}$ to be a first integral of the equation of motion, just as in Einstein's subsequent theory.³⁵

Einstein was originally impressed by the mathematical elegance of Abraham's theory, although he complained about the lack of proper physical foundation. As he soon communicated to Abraham, the characterization of $u = i\hat{c}t$ as the fourth coordinate is not

³⁵ Abraham 1911, p. 678; 1912a, p. 1.

compatible with \hat{c} being the local value of the velocity of light, because it is

incompatible with $d\tau=0$ for $d\mathbf{r}^2 = \hat{c}^2 dt^2$. Abraham published the following corrective:

Instead of ['we regard x, y, z and $u = i\hat{c}t$ as coordinates in a space of four dimensions'], one should read 'we regard dx, dy, dz and $du = i\hat{c}dt$ as the components of a displacement $d\mathbf{s}$ in four-dimensional space.' –Thus, $ds^2 = dx^2 + dy^2 + dz^2 - \hat{c}^2 dt^2$ is the square of a 4-dimensional line-element, wherein the velocity of light \hat{c} is given by $[\hat{c}^2 - \hat{c}_0^2 = 2\Phi]$.

From this statement, Einstein inferred that Abraham believed the Lorentz transformation to apply infinitesimally with the local velocity \hat{c} . In the first installment of his static theory, he proved this to be mathematically impossible. He recommended to restrict the application of the Lorentz group to regions of space in which \hat{c} had a uniform value, and he predicted that a larger group and more complicated equations would be needed in the general case. In a subsequent publication, Abraham made clear that he meant to apply the Lorentz group (more exactly rotations in Minkowski space) to the 4-vector $(dx, dy, dz, i\hat{c}dt)$ (whereas Einstein believed Abraham was applying a Lorentz boost of velocity \hat{c} to dx, dy, dz, dt). Einstein's chief objection, the incompatibility of Abraham's field equation with the equivalence principle, remained valid. Yet he learned something important in his exchange with Abraham: a Minkowskian integration of the variable velocity of light leads to a Riemannian metric.³⁶

Altogether, in early 1912 Einstein already had in hand the representation of a gravitational field by a quadratic differential form as well as the now usual method for deriving the equations of motion of a particle from this form. He was aware of a possible geometric interpretation of this form and method. Still his static theory of 1912 differed

³⁶ Abraham 1912b; 1912c, p. 433; 1912d, p. 312; Einstein 1912a, pp. 368-369. For a thorough study of Abraham's theory and his polemic with Einstein, cf. Renn 2007a.

from the modern theory in three ways. The first difference is that the quadratic form is of the restricted kind $d\mathbf{r}^2 - \hat{c}^2(\mathbf{r})dt^2$, which Einstein believed to exactly represent a static gravitational field (whereas we now know that the spatial part of a non-trivial static metric is non-Euclidean). The second difference is that Einstein considered only two kinds of coordinates: the coordinates (\mathbf{r}, t) through which the laws of physics in the static gravitational field should be expressed and which Einstein believed to be giving the true space and time relations in some sense (even though the time t differs from the time $\hat{c}t$ given by a standard clock), and the coordinates (\mathbf{r}', t') in a tangent frame in which the gravitational force is eliminated. The third difference concerns the gravitational field equation. With the assumed restriction of the fundamental form, this equation concerns the function $\hat{c}(\mathbf{r})$ only. In order to derive this equation, Einstein relied on three principles: correspondence with Poisson's equation for Newtonian gravitation, homogeneity with respect to \hat{c} (dictated by the kinematic analysis), equality of action and reaction (impossibility of perpetual motion). He verified its compatibility with two additional principles: energy conservation and the equivalence between energy and gravitational mass. He was convinced that the only field equation compatible with these five principles was Eq. (35). The fact that so many principles could be satisfied at the same time comforted him in his choice of the restricted form $d\mathbf{r}^2 - \hat{c}^2(\mathbf{r})dt^2$, although he originally doubted that space measurements would remain Euclidean in a gravitational field.

It is usually believed that Einstein's first realization that non-Euclidean geometry may be needed in general relativity came in his remark, at the beginning of the first memoir of 1912, that the tangent-frame method of measurement leads to a non-Euclidean perimeter/radius ratio in a rotating frame. In reality, the first way in which non-Euclidean

geometry truly entered his theory was in his choice of $d\mathbf{r}^2 - \hat{c}^2(\mathbf{r})dt^2$ for what we would now call the metric in a static field. By somewhat daring generalization from the case of a constantly accelerated frame (for which $\hat{c}(\mathbf{r})$ is an affine function) to an arbitrary static field for which the acceleration varies from place to place, he implicitly introduced a curved 4-manifold for which no change of coordinates can globally turn the metric to its Minkowskian form.

Einstein had great confidence in his static theory of 1912: "I really believe to have found a piece of truth," he then wrote to Ehrenfest. Paradoxically, this work of 1912 both initiated and blocked Einstein's subsequent quest for a theory of gravitation based on a more general metric. On the positive side, the static theory introduced a few basic formal elements of such a theory as well as a few heuristic principles including correspondence with Newton's theory of gravitation, energy-momentum conservation, and existence of a stress-energy tensor for the gravitational field. On the negative side, it blocked further progress by suggesting that the more general field equation should agree with the equation of 1912 in the static case, and, above all, by conflating coordinate systems with physically meaningful reference systems.³⁷

1.4 The Zürich notebook

³⁷ Einstein to Ehrenfest, 10 March 1912, *ECP* 5.

Soon after completing his theory of the static gravitational field, Einstein considered the extension to arbitrary fields. This is already apparent in his remark, in the appendix to his second memoir of 1912, that the variational principle

$$(46) \quad \delta \int \sqrt{\dot{c}^2(\mathbf{r}) dt^2 - d\mathbf{r}^2} = 0$$

is easily extended to dynamical fields.³⁸ Although Einstein did not tell more on this extension, he plausibly reasoned that for a dynamical field more general systems of coordinates and the more general form $g_{\mu\nu} dx^\mu dx^\nu$ would be needed. The resulting principle

$$(47) \quad \delta \int \sqrt{g_{\mu\nu} dx^\mu dx^\nu} = 0$$

and the resulting equations are evidently covariant. This induced Einstein to look for a covariant extension of his static field equation. It will soon be clear that he initially did not expect covariance with respect to any (mathematically acceptable) change of coordinates. In order to have physical, metrological meaning, the choice of coordinates had to be somewhat restricted. All he knew was that the permitted coordinate transformations should at least include transformations corresponding to uniformly accelerated or uniformly rotating frames,³⁹ and also transformations locally turning the form $g_{\mu\nu} dx^\mu dx^\nu$ into the Minkowskian form of special relativity.

First tries: linear covariance

That Einstein reasoned in this or a similar way is confirmed by his first notes on gravitation in a notebook he filled in the years 1912-1913, usually called the Zürich

³⁸ Einstein 1912b, p. 458.

³⁹ See Einstein to Besso, 26 March 1912: I am "still far from being able to conceive rotation as rest."

notebook because most of it was written after his move to Zürich in July 1912. There are no dates in this notebook, and it is not easy to time the various steps taken by Einstein. All we know, from his correspondence, is that in the fall of 1912 he believed to have "the most general equations of gravitation" in hand; that by October at least he had received help from his friend the mathematician Marcel Grossmann and that he had thereby developed a "high regard for mathematics . . . , which [he] heretofore naively considered a luxury in its subtlest parts." At some point, he must have encountered unsuspected difficulties: no sooner than March 1913 could he write to his cousin and future wife that he had "solved the gravitational problem only a few weeks ago after half a year of most strenuous investigation."⁴⁰

In his notebook [p. 39L], Einstein starts from the expression

$$(48) \quad ds^2 = g_{\mu\nu} dx^\mu dx^\nu$$

and writes the equations for the transformation of $g_{\mu\nu}$ and $\partial_\mu \equiv \partial / \partial x^\mu$ under a linear change of coordinates. Although the ds notation indicates analogy with the length element on a (hyper)surface, Einstein proceeds unaware of the general theory of Riemannian manifolds and the attached tensor calculus. This may explain why he confines himself to linear changes of coordinates, which are much easier to handle than general changes.⁴¹

⁴⁰ Einstein to Hopf, 16 Aug. 1912 (cited); Einstein to Freundlich, 27 Oct. 1912; Einstein to Sommerfeld, 29 Oct. 1912 (cited); Einstein to Löwenthal, 23 March 1913 (cited).

⁴¹ The entire contents of the Zürich notebook have been published both in *ECP* 4 with rich annotation by John Norton, and in *GGR* 1 (there also in facsimile). The most detailed and accurate interpretation is given in *GGR* 2, pp. 489-714, under the joint efforts of Janssen, Renn, Sauer, Norton, and Stachel. The page numbering is the one used in these sources (Einstein filled the book from both ends, so that these numbers are not always a growing function of time). I use slightly modernized notation (with the now usual conventions for indices and summation).

His strategy [pp. 39L-40L] is to find linearly covariant generalizations of the equations

$$(49) \quad \frac{\partial \hat{c}}{\partial t} = 0, \quad \frac{\Delta \hat{c}}{\hat{c}} - \frac{1}{2} \left(\frac{\nabla \hat{c}}{\hat{c}} \right)^2 = 0, \quad \text{with } g_{00} = \hat{c}^2$$

for the static field. For the first equation he tries the divergence condition $\partial_\mu g_{\mu\nu} = 0$ and fails since this is not covariant by linear transformations.⁴² He then [pp. 40L-40R] tries to construct generalizations of the second equation (the field equation proper) by taking the most symmetric possible combinations of $g_{\mu\nu}$ and ∂_μ that are of second order with respect to ∂_μ , but he soon gives up. Something will be left from this first unsuccessful attempt: for a while Einstein will keep requiring the equations (49) for the static case to be special cases of the general equations. In particular, in addition to the second-order field equation for the $g_{\mu\nu}$ he will often require a first-order equation.

The stress-energy tensor of a dust

At some point, Einstein introduces energy-momentum considerations as an additional guide in his quest for a field equation [pp. 5R, 43LB].⁴³ Rewriting the geodesic equation (47) as

$$(50) \quad \delta \int L dt = 0 \quad \text{with } L = ds / dt,$$

and interpreting $\partial L / \partial x_i$ (with $i = 1, 2, 3$) as the spatial components of the momentum of the particle, the equations of motion read

⁴² A linearly covariant condition would be $\partial_\mu g^{\mu\nu} = 0$, later used by Einstein. My analysis here differs from *GGR* 2, p. 507.

⁴³ In special relativity, Hermann Minkowski had introduced the energy-momentum tensor of the electromagnetic field in 1908 and Max Laue had generalized this notion in 1911.

$$(51) \quad \frac{dp_\mu}{ds} = \frac{1}{2} \partial_\mu g_{\nu\rho} \frac{dx^\nu}{ds} \frac{dx^\rho}{ds} \quad \text{with} \quad p_\mu = g_{\mu\nu} \frac{dx^\nu}{ds}.$$

For a quasi-continuous dust of spatial density ρ in the small volume V , Einstein

introduces the true density ρ_0 and the volume V_0 in the tangent frame. Combining the

invariance of $\sqrt{-g} d^4x$ (with $g = \det g_{\mu\nu}$) and the invariance of mass, he gets

$$(52) \quad \sqrt{-g} V dt = V_0 ds \quad \text{and} \quad \rho V = \rho_0 V_0 \Rightarrow \rho = \rho_0 \sqrt{-g} dt / ds.$$

The 4-force acting on the unit volume of the dust is

$$(53) \quad \rho \frac{dp_\mu}{dt} = \rho_0 \sqrt{-g} \frac{dp_\mu}{ds} = \frac{1}{2} \sqrt{-g} \partial_\mu g_{\nu\rho} \theta^{\nu\rho} \quad \text{with} \quad \theta^{\nu\rho} = \rho_0 \frac{dx^\nu}{ds} \frac{dx^\rho}{ds}.$$

The 4-momentum flux of the dust is

$$(54) \quad \rho p_\mu \frac{dx^\nu}{dt} = \rho_0 \sqrt{-g} g_{\mu\rho} \frac{dx^\rho}{ds} \frac{dx^\nu}{ds} = \sqrt{-g} g_{\mu\rho} \theta^{\nu\rho}.$$

By energy-momentum conservation, the 4-divergence of this quantity should be equal to the 4-force:

$$(55) \quad \partial_\nu (\sqrt{-g} g_{\mu\rho} \theta^{\nu\rho}) = \frac{1}{2} \sqrt{-g} \partial_\mu g_{\nu\rho} \theta^{\nu\rho}.$$

As Einstein correctly surmises, this equation is equivalent to the equation of motion (55)

because by mass conservation we have [pp. 43LB, 19R]

$$(56) \quad \partial_\mu (\rho dx^\mu / dt) = \partial_\mu (\rho_0 \sqrt{-g} dx^\mu / ds) = 0.$$

Einstein calls $\theta^{\nu\rho}$ the motion tensor, and $g_{\mu\rho} \theta^{\nu\rho}$ the energy-momentum tensor

(although he later favored "stress-energy tensor"). He asserts [p. 5R]⁴⁴ that

$$(57) \quad V_\mu = \partial_\nu (\sqrt{-g} g_{\mu\rho} S^{\nu\rho}) - \frac{1}{2} \sqrt{-g} \partial_\mu g_{\nu\rho} S^{\nu\rho}$$

⁴⁴ Einstein also notes that $V_\mu = 0$ for $S^{\mu\nu} = g^{\mu\nu}$.

is a covariant vector (*zugeordneter Vektor*) for any symmetric tensor $S^{\mu\nu}$. This is almost true since, as he does not yet know, the quotient $V_\mu / \sqrt{-g}$ is the covariant divergence of $S^{\mu\nu}$. At any rate, the 4-divergence equation (55) and its physical derivation will condition Einstein's ulterior considerations in several manners. Firstly, they reinforce his idea that covariance should play a central role in constructing the gravitational field equation. Secondly, the interpretation of $\frac{1}{2}\sqrt{-g}\partial_\mu g_{\nu\rho}\theta^{\nu\rho}$ as the 4-force acting on a dust of stress-energy $\theta^{\mu\nu}$ gives to the derivatives $\partial_\mu g_{\nu\rho}$ prominence in expressing the strength of the gravitational field (thus generalizing the relation between force, mass, and gravitational potential). Thirdly, these considerations suggest that the gravitational field equation might have the form $G^{\mu\nu} = \kappa T^{\mu\nu}$, in which $G^{\mu\nu}$ is a tensor combination of the derivatives of the metric tensor and $T^{\mu\nu}$ is the stress-energy tensor of matter (including the electromagnetic field). Fourthly, they will induce Einstein to demand that the product $\sqrt{-g}\partial_\mu g_{\nu\rho}G^{\nu\rho}$ be expressible as the 4-divergence of a (pseudo-)tensor⁴⁵ representing the stress-energy of the gravitational field (up to a numerical factor). This last condition is a generalization of Einstein's earlier demand that the sum of all gravitational forces should vanish in the static case.

Scalar and tensor generalizations of the d'Alembertian

Einstein does not immediately explore the tensor option $G^{\mu\nu} = \kappa T^{\mu\nu}$, presumably because it is much easier to explore the scalar option in which the field equation in the absence of matter is obtained by setting to zero a scalar combination of the metric tensor

⁴⁵ I will henceforth drop the "pseudo" in conformity with Einstein's usage (as we will see in section 1.8, Einstein was unaware of this distinction until late 1913).

and its first and second derivatives [pp. 6R-11L]. He knows that for any scalar function φ the expression

$$(58) \quad D^2\varphi = (1/\sqrt{-g})\partial_\mu(g^{\mu\nu}\sqrt{-g}\partial_\nu\varphi)$$

is also a scalar.⁴⁶ He also believes (wrongly) that $\sqrt{-g}g^{\mu\nu}\partial_\mu[(1/\sqrt{-g})\partial_\nu\varphi]$ is an invariant.

Inserting $\varphi = \sqrt{-g}$, into the latter expression and using the identity $\partial_\mu g = g g^{\nu\rho}\partial_\mu g_{\nu\rho}$

leads him to

$$(59) \quad g^{\mu\nu}\partial_\mu\partial_\nu\ln(-g) \equiv g^{\mu\nu}\partial_\mu(g^{\rho\sigma}\partial_\nu g_{\rho\sigma}) = 0$$

for the vacuum field equation[6R]. Einstein does not take this candidate seriously, if only because when applied to his static metric $ds^2 = \hat{c}^2(\mathbf{r})dt^2 - d\mathbf{r}^2$ it leads to $\Delta\ln\hat{c} = 0$ instead of Eq. (35).⁴⁷

As Einstein knows, the determinant g and the derived expression $\sqrt{-g}$ are not truly scalars. They are so only with respect a restricted class of transformations, now called unimodular because their determinant is *one*. He is more than willing to adopt this sort of restriction. In his following notes, he tries various would-be scalar combinations of g , $g_{\mu\nu}$, $g^{\mu\nu}$, and ∂_μ , sometimes with additional conditions such as $\partial_\mu(\sqrt{-g}g^{\mu\nu}) = 0$ (harmonic coordinates), $\partial_\mu g^{\mu\nu} = 0$, or $g = -1$, thanks to which the candidates become invariant (for the transformations compatible with the conditions). Ideally, these conditions should be compatible with the static metric and they should be satisfied by the metric of a flat spacetime in a uniformly accelerated or rotating frame (henceforth called "acceleration metric" and "rotation metric" respectively) in order to allow for the frames

⁴⁶ Richard Dedekind had privately discovered this generalized Laplacian, and Beltrami had rediscovered and published it in 1868: cf. Dell'Aglio 1996, pp. 231-235; Darrigol 2014b, p. 68.

⁴⁷ For a more detailed analysis, cf. *GGR* 2, pp. 526-531.

implied in the equivalence principle. As Einstein will realize sooner or later, they are not.⁴⁸

At some point, Einstein abandons the scalar option and turns to the tensor option in which the field equation takes the form $G^{\mu\nu} = \kappa T^{\mu\nu}$.⁴⁹ He hits upon [p. 12L]

$$(60) \quad G^{\mu\nu} = \frac{1}{\sqrt{-g}} \partial_\rho (\sqrt{-g} g^{\rho\sigma} \partial_\sigma g^{\mu\nu}),$$

which is analogous to the scalar invariant expression (58) and reduces to

$G^{\mu\nu} = g^{\rho\sigma} \partial_\rho \partial_\sigma g^{\mu\nu} \equiv \square g^{\mu\nu}$ for harmonic coordinates. For no obvious reason, he then [p.

13R] computes $D_\mu G^{\mu\nu}$ under the simplifying conditions $\sqrt{-g} = 1$ and $\partial_\mu g^{\mu\nu} = 0$. He of

course knows that $D_\mu G^{\mu\nu}$ should vanish as a consequence of the equations $D_\mu T^{\mu\nu} = 0$

and $G^{\mu\nu} = \kappa T^{\mu\nu}$ combined. He is unlikely to have regarded this condition as an identity,

for he did not do so in the rest of the notebook. Instead he probably regarded it as a

further restriction on the metric field $g_{\mu\nu}(x)$. This was the end of his first naive attempts

to derive a conditionally covariant field equation.

Building on the Riemann tensor

Around that time Grossmann told Einstein about the absolute differential calculus of

Elwin Christoffel, Gregorio Ricci-Curbastro, and Tullio Levi-Civita.⁵⁰ Einstein thus

became aware of the Riemann tensor

⁴⁸ Using Eqs. (16-17), the acceleration metric has the form $ds^2 = (c_0 + \mathbf{a} \cdot \mathbf{r})^2 dt^2 - d\mathbf{r}^2$. Injecting

$\xi = x \cos \omega t + y \sin \omega t$, $\eta = -x \sin \omega t + y \cos \omega t$, $\zeta = z$, $\tau = t$ into the Minkowskimetric

$ds^2 = c^2 d\tau^2 - d\xi^2 - d\eta^2 - d\zeta^2$, we get $ds^2 = (c^2 - \omega^2 \mathbf{r}^2) dt^2 - d\mathbf{r}^2 - 2\omega y dx dt + 2\omega x dy dt$ for the rotation metric.

⁴⁹ According to *GGR* 2, pp. 600-602, he rather has in mind $G^{\mu\nu} = \kappa(T^{\mu\nu} + t^{\mu\nu})$. If this were the case, I think he would have considered $\partial_\mu G^{\mu\nu}$ instead of $D_\mu G^{\mu\nu}$ on p. 13R.

$$(61) \quad (\mu\nu, \rho\sigma) = g_{\mu\tau} R^{\tau}_{\nu\rho\sigma},$$

with

$$(62) \quad R^{\mu}_{\nu\rho\sigma} = \partial_{\rho} \{^{\mu}_{\nu\sigma}\} - \partial_{\sigma} \{^{\mu}_{\nu\rho}\} + \{^{\mu}_{\tau\rho}\} \{^{\tau}_{\nu\sigma}\} - \{^{\mu}_{\tau\sigma}\} \{^{\tau}_{\nu\rho}\} \quad \text{and}$$

$$(63) \quad g_{\mu\sigma} \{^{\sigma}_{\nu\rho}\} = \frac{1}{2} (\partial_{\nu} g_{\mu\rho} + \partial_{\rho} g_{\mu\nu} - \partial_{\mu} g_{\nu\rho}).$$

$R_{\mu\nu\rho\sigma}$ is the simplest fully covariant combination of second- and first-order derivatives of the metric tensor (up to a constant coefficient). Einstein forms the Ricci tensor⁵¹

$$R_{\mu\nu} = R^{\rho}_{\mu\nu\rho} \quad \text{and tries [p. 14 L]}$$

$$(64) \quad R_{\mu\nu} = -\kappa T_{\mu\nu}$$

for the field equation in the presence of matter of stress-energy tensor $T_{\mu\nu}$. In continuity

with his previous approach, he tries and fails to reduce $R_{\mu\nu}$ to a simple combination of

$g_{\mu\nu}$, $g^{\mu\nu}$, g , and ∂_{μ} , even under the restriction $g = -1$. He deplores that in the first,

linear approximation (corresponding to the two first terms in $R^{\mu}_{\nu\rho\sigma}$), the Ricci tensor

reduces to

$$(65) \quad R^{(0)}_{\mu\nu} = \frac{1}{2} g^{\rho\sigma} (\partial_{\rho} \partial_{\sigma} g_{\mu\nu} + \partial_{\mu} \partial_{\nu} g_{\rho\sigma} - \partial_{\mu} \partial_{\rho} g_{\nu\sigma} - \partial_{\mu} \partial_{\sigma} g_{\nu\rho}),$$

which contains unwanted terms besides the desired d'Alembertian $\frac{1}{2} g^{\rho\sigma} \partial_{\rho} \partial_{\sigma} g_{\mu\nu}$.

Manipulating the curvature scalar $g^{\mu\nu} R_{\mu\nu}$ under $g = -1$ he tries to extract from it a

gravitation tensor that would not have this defect, to no avail [pp. 14R-18R].⁵²

⁵⁰ Cf. Stachel 2007, pp. 106-107; Sauer 2013.

⁵¹ Einstein's definition of the Ricci tensor differs from the now common $R_{\mu\nu} = R^{\rho}_{\mu\rho\nu}$ by a change of sign.

⁵² Cf. *GGR* 2, pp. 614-622. On the Einstein-Grossmann collaboration, cf. Sauer 2013.

At some point [p. 19L], Einstein realizes that these terms can be eliminated (in the linear approximation) by the harmonic coordinate condition $g^{\mu\nu}\Gamma_{\mu\nu}^{\rho}=0$ (this condition is equivalent to the earlier encountered $\partial_{\mu}(\sqrt{-g}g^{\mu\nu})=0$, thanks to $D_{\mu}g^{\mu\nu}=0$). In the weak-field approximation for which $g_{\mu\nu}$ differs from the Minkowski metric $\eta_{\mu\nu}$ by the small tensor $h_{\mu\nu}$, and for a dust of stress-energy $\theta^{\mu\nu}$, he arrives at the field equation [p. 19R]

$$(66) \quad \square h_{\mu\nu} \equiv \eta^{\rho\sigma}\partial_{\rho}\partial_{\sigma}h_{\mu\nu} = -\kappa\eta_{\mu\rho}\eta_{\nu\sigma}\theta^{\rho\sigma}, \text{ or } \Delta_4 h_{\mu\nu} = \kappa\theta_{\mu\nu}$$

in the imaginary system of coordinates ($x^0 = ict$) for which $\eta_{\mu\nu}$ becomes the identity and the d'Alembertian becomes the 4-dimensional Laplacian Δ_4 .⁵³

In the same approximation, the energy-momentum equation $D_{\mu}\theta^{\mu\nu}=0$ reduces to $\partial_{\mu}\theta^{\mu\nu}=0$ because the product of $\theta^{\mu\nu}$ by a Christoffel symbol is of second order; and the harmonic condition $g^{\mu\nu}\Gamma_{\mu\nu}^{\rho}=0$ reduces to

$$(67) \quad 2\partial_{\mu}h_{\mu\nu} = \partial_{\nu}h_{\mu\mu} \text{ or } \partial_{\mu}(h_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}h_{\rho\rho}) = 0$$

in the imaginary coordinate system. Combining the latter equation with $\Delta_4 h_{\mu\nu} = \kappa\theta_{\mu\nu}$ and $\partial_{\mu}\theta^{\mu\nu}=0$, we get

$$(68) \quad 0 = \partial_{\mu}(\theta_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}\theta_{\rho\rho}) = -\frac{1}{2}\partial_{\nu}\theta_{\rho\rho},$$

which is incompatible with $\theta_{\rho\rho} = \rho_0$. Einstein removes this contradiction by means of the modified field equation [p. 20L]

$$(69) \quad \Delta_4(h_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}h_{\rho\rho}) = \kappa\theta_{\mu\nu}.$$

He also gives the equivalent form⁵⁴

⁵³ Cf. Renn and Sauer 1999, pp. 109-114.

$$(70) \quad \Delta_4 h_{\mu\nu} = \kappa(\theta_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}\theta_{\rho\rho}),$$

which is easily obtained using the contraction $-\Delta_4 h_{\mu\mu} = \kappa\theta_{\mu\mu}$ of the former equation.⁵⁵

As we saw, for Einstein a good field equation of the form $G_{\mu\nu} = -\kappa\theta_{\mu\nu}$ should make the force density $\frac{1}{2}\sqrt{-g}(\partial_\mu g_{\nu\rho})G^{\nu\rho}$ a 4-divergence. Einstein verifies this for the field equation (69) in the linear approximation [pp. 19R, 21L]: $(\partial_\mu h_{\nu\rho})\Delta_4(h_{\nu\rho} - \frac{1}{2}\delta_{\nu\rho}h_{\sigma\sigma})$ is a 4-divergence by analogy with $\nabla\phi\Delta\phi$ in electrostatics. Alas, for the d'Alembertian term

$\frac{1}{2}g^{\rho\sigma}\partial_\rho\partial_\sigma g_{\mu\nu}$ of the exact field equation, Einstein fails to bring the product

$\sqrt{-g}(\partial_\mu g_{\nu\rho})g^{\sigma\tau}\partial_\sigma\partial_\tau g_{\nu\rho}$ into the form of a 4-divergence [p. 21L]. In addition [p. 21R], his older static field theory turns out to be incompatible with the harmonic condition

$g^{\mu\nu}\Gamma_{\mu\nu}^\rho = 0$ (the only non-vanishing coefficients of Einstein's static metric are $g_{00} = \hat{c}^2$,

$g_{11} = g_{22} = g_{33} = -1$, so that $g^{\mu\nu}\Gamma_{\mu\nu}^\rho = \partial_\mu(\sqrt{-g}g^{\mu\rho}) = (0, \nabla\hat{c})$).⁵⁶

At this point, Einstein could have questioned the truth of the static theory. On the contrary, he finds a new argument for the vanishing of the non-diagonal elements of the static metric [p. 21R]. According to the 4-momentum formula in Eqs. (51), the energy E of a particle in the metric field $g_{\mu\nu}$ is given by $g_{0\mu}dx^\mu/ds$. Being certain that

$g_{01} = g_{02} = g_{03} = 0$, Einstein gets $E = g_{00}dx^0/ds$. The metric coefficients being

⁵⁴ I have corrected an algebraic error leading to the wrong sign in front of 1/2.

⁵⁵ As we may retrospectively judge, the new field equation is the linear approximation of the equation $R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = -\kappa\theta_{\mu\nu}$ on which Einstein based his final theory of gravitation.

⁵⁶ Einstein writes only that the static case is "impossible because of the divergence condition." I take this condition to be a first-order differential condition on the metric, the harmonic condition in the present case. An alternative candidate is the weak-field energy-momentum condition $\partial_\mu\theta^{\mu\nu} = 0$, which requires the weak-field equation (69), leading to $\Delta h_{ij} = -(\kappa/2)\delta_{ij}\theta_{00}$ in the static case, in contradiction with Einstein's static metric of 1912 (for which $h_{ij} = 0$). Still another reading is given in *GGR* 2, p. 643.

independent of time, by Eq. (51) the force \mathbf{dp} / dx^0 acting on the particle reduces to $\frac{1}{2} \nabla g_{00} (dx^0 / ds)$ if and only if the gradient of the spatial coefficients of the metric vanish. Therefore, the spatial part of the metric must be uniform, as it is in Einstein's static solution, in order that the energy and the force vary in the same manner with respect to the velocity. Einstein does not justify this condition. Plausibly he wants the kinetic energy to contribute to the gravitational mass of the particle, therefore also to its gravitational potential and to the derived force.

After stumbling over this difficulty, Einstein tries to modify the field equation in such a manner that its second-order terms can be reduced to the d'Alembertian form without the harmonic condition. His strategy is to subtract from the Ricci tensor a unimodular tensor built from g and its derivatives (remember he has already explored such tensors). He thus arrives at the tensor [p. 22R]⁵⁷

$$(71) \quad R_{\mu\nu}^{\times} = R_{\mu\nu} - D_{\nu} \partial_{\mu} \ln \sqrt{-g} = -\partial_{\rho} \{^{\rho}_{\mu\nu}\} - \{^{\sigma}_{\mu\rho}\} \{^{\rho}_{\nu\sigma}\},$$

whose second-order part reduces to the d'Alembertian $\frac{1}{2} g^{\rho\sigma} \partial_{\rho} \partial_{\sigma} g_{\mu\nu}$ under the condition $\partial_{\mu} g^{\mu\nu} = 0$. Unlike the harmonic condition, this condition is compatible with Einstein's static metric. Yet Einstein does not pursue the consequences of the field equation

$$(72) \quad R_{\mu\nu}^{\times} = -\kappa T_{\mu\nu},$$

probably because of his prejudice that the true field equation should only involve a simple combination of $g_{\mu\nu}$ and its derivatives.

In the following pages of his notebook, Einstein restricts the coordinate transformations to those for which the symmetric derivative

⁵⁷ Einstein called this tensor $T_{\mu\nu}^{\times}$. I changed the notation in order to avoid confusion with the energy-momentum tensor.

$$(73) \quad \vartheta_{\mu\nu\rho} = \partial_\rho g_{\mu\nu} + \partial_\mu g_{\nu\rho} + \partial_\nu g_{\rho\mu}$$

behaves like a tensor. With respect to these transformations, the simple expression [p. 23L]

$$(74) \quad R_{\mu\nu}^{\times\times} = \frac{1}{2} \partial_\rho (g^{\rho\sigma} \partial_\sigma g_{\mu\nu}) + g^{\rho\alpha} g^{\sigma\beta} \partial_\alpha g_{\mu\rho} \partial_\beta g_{\nu\sigma}$$

is a tensor. It is easy to understand why the new condition so much simplifies the field equation: it allows replacing the Christoffel symbol $\{\overset{\rho}{\mu\nu}\}$ with the much simpler

$-g^{\rho\sigma} \partial_\sigma g_{\mu\nu}$. Plausibly, Einstein had the idea of this condition because he had earlier

studied the metric for which $\vartheta_{\mu\nu\rho} = 0$ (call it the ϑ -metric) and found it to be simply related to the rotation metric [pp. 42L-42R].⁵⁸

In the notebook, the study of the ϑ -metric comes right after the consideration of geodesics on a surface [p. 41R]. At that time, Einstein plausibly knew that the geodesics for the metric $g_{\mu\nu}$ satisfy Eq. (51), although the relevant calculations appear further in the notebook. This equation,

$$\frac{dp_\mu}{ds} = \frac{1}{2} \partial_\mu g_{\nu\rho} \frac{dx^\nu}{ds} \frac{dx^\rho}{ds} \quad \text{with} \quad p_\mu = g_{\mu\nu} \frac{dx^\nu}{ds},$$

implies

$$(75) \quad \frac{dx^\mu}{ds} \frac{dp_\mu}{ds} = \frac{1}{6} \vartheta_{\mu\nu\rho} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} \frac{dx^\rho}{ds}.$$

⁵⁸ Cf. *GGR* 2, pp. 652-679. The following reconstruction significantly departs from *GGR* by giving a reason why Einstein had earlier explored the ϑ -metric. Whereas *GGR* believe that pp. 42L-42R were filled after Einstein introduced $R_{\mu\nu}^{\times\times}$ (on the other side of the notebook), I tend to believe that they naturally follow the previous pages of the same side of the notebook and that p. 43L is the only page on this side of the notebook that is contemporary to $R_{\mu\nu}^{\times\times}$ (as could be suspected from the fact that this page is written both downwards and upwards).

For a particle at rest in an inertial frame, the motion in a rotating frame is a rotation for which the left side of this equation vanishes (because $dp_0 = 0$ and $d\mathbf{p} \cdot d\mathbf{r} = 0$).

Conversely, one might hope the rotation metric to be a solution of the equation $\vartheta_{\mu\nu\rho} = 0$.

This is a plausible reason why Einstein solves this equation on pp. 42L-42R. The result is amazing: the solutions $g_{\mu\nu}$ of determinant -1 differ from a rotation metric, but the associated contravariant coefficients $g^{\mu\nu}$ are exactly identical with the coefficients of a rotation metric.

Let us return to the context of the $R_{\mu\nu}^{\times\times}$ field tensor. Einstein tries two alternatives to the ϑ -condition, one in which $g_{\mu\nu}$ is replaced with $g^{\mu\nu}$ [p. 43LA], and another in which the permitted transformations are those preserving $\partial_\rho g_{\mu\nu} + \partial_\mu g_{\nu\rho} - 2\partial_\nu g_{\rho\mu}$ [pp. 23L, 25R]. Both fail. Einstein still hopes to conciliate the ϑ -condition with the equivalence principle, by showing that the geodesics in a metric field of vanishing $\vartheta_{\mu\nu\rho}$ resemble the motion of a particle in a centrifugal force field [pp. 42R, 43LA].⁵⁹

In general, Einstein regards his coordinate conditions as restrictions on the choice of physically permitted coordinate systems. This is clear from the fact that he usually tries to determine the transformations satisfying these conditions or the resulting field equation--with little or no success because the relevant differential equations are too complex. For instance, he tries [p. 22L] to determine the transformations for which the condition $\partial_\mu g^{\mu\nu} = 0$ (he knows that the transformations for which the condition

$g^{\mu\nu}\Gamma_{\mu\nu}^\rho = 0$ is preserved are those for which the new coordinates are harmonic functions

⁵⁹ Cf. *GGR* 2, pp. 668-679. On p. 24L Einstein does the same for the motion of a dust obeying Eq. (55).

of the old coordinates, hence the name harmonic condition), and he does the same [p.

23R] for the transformations under which $\vartheta_{\mu\nu\rho}$ is a tensor.⁶⁰

For the sake of the equivalence principle, Einstein wanted these transformations to include (uniform) rotation in Minkowski space.⁶¹ There is no relevant calculation for the two conditions $g^{\mu\nu}\Gamma_{\mu\nu}^{\rho} = 0$ and $\partial_{\mu}g^{\mu\nu} = 0$ in the section devoted to the Riemann tensor. This may be because Einstein remembers his earlier proof [p. 11L] that the rotation metric satisfies the first condition (then the second condition should also be satisfied, for it is the same as the first for unimodular transformations).⁶² Probably for a similar reason, there is almost nothing in the Riemann-tensor-based section of the notebook regarding the compatibility of the ϑ -condition with rotation.⁶³ In contrast, Einstein is repeatedly concerned with the existence of a stress-energy tensor for a metric field obeying the conditioned field equations. This is the stress principle, so far acting as a test for a physically acceptable field equation.

The Entwurf strategy

In the case of the last Riemann-tensor-based candidate, based on the field operator $R_{\mu\nu}^{\times\times}$ of Eq. (74), Einstein fails to re-express the force density $-\frac{1}{2}(\partial_{\rho}g^{\mu\nu})R_{\mu\nu}^{\times\times}$ as the 4-divergence

⁶⁰ Cf. Norton 2005, pp. 88-90 for the importance of coordinate restrictions in the rejection of the Riemann-based tensor. However, I tend to disagree with his opinion, amplified in Norton 2007, that Einstein initially considered coordinate conditions and turned them into restrictions after being misled by a reification of coordinate systems. The distinction between coordinate condition and coordinate restriction belongs to Renn and Sauer 2007, p. 109. It is emphasized in *GGR* 1, p. 11.

⁶¹ One might also want to cover permitted transformations to uniform acceleration in Minkowski space. Einstein's silence in this regard is easily explained by the evident impossibility of satisfying this condition: the acceleration metric is incompatible with the harmonic condition since it is a sub-case of the static metric of 1912; and the unimodular restriction used by Einstein for his other field equations is incompatible with accelerated coordinates.

⁶² Unfortunately, this is not true for finite rotations: cf. Norton 2005, pp. 89-90, and *GGR* 2, pp. 574-577.

⁶³ The exception is p. 24L: cf. *GGR* 2, pp. 674-679.

of a stress-energy tensor. This test involves a repeated application of the identity

$X\partial_\mu Y = \partial_\mu(XY) - Y\partial_\mu X$ to the product $(\partial_\rho g^{\mu\nu})\partial_\sigma(g^{\sigma\tau}\partial_\tau g_{\mu\nu})$ to which the first term of the field operator leads. Einstein thus generates the identity [p. 24R]

$$(76) \quad (\partial_\rho g^{\mu\nu})[\partial_\sigma(g^{\sigma\tau}\partial_\tau g_{\mu\nu}) - \frac{1}{2}\partial_\nu g_{\sigma\tau}\partial_\mu g^{\sigma\tau}] = \partial_\sigma(g^{\sigma\tau}\partial_\rho g^{\mu\nu}\partial_\tau g_{\mu\nu}) - \frac{1}{2}\partial_\rho(g^{\sigma\tau}\partial_\tau g_{\mu\nu}\partial_\sigma g^{\mu\nu}),$$

which suggests to him that the true vacuum field equation might be

$$(77) \quad \partial_\sigma(g^{\sigma\tau}\partial_\tau g_{\mu\nu}) - \frac{1}{2}\partial_\nu g_{\sigma\tau}\partial_\mu g^{\sigma\tau} = 0.$$

Through an erroneous calculation, he finds the rotation metric to be a solution of this equation [p. 25R].⁶⁴

At any rate, identity (76) does not hold. It is easily seen to be equivalent to⁶⁵

$$(77) \quad g^{\sigma\tau}\partial_\tau g^{\mu\nu}\partial_\rho\partial_\sigma g_{\mu\nu} = g^{\sigma\tau}\partial_\tau g_{\mu\nu}\partial_\rho\partial_\sigma g^{\mu\nu}.$$

Einstein believed he could derive the latter identity by differentiating

$$\partial_\sigma g^{\mu\nu}\partial_\tau g_{\mu\nu} = \partial_\sigma g_{\mu\nu}\partial_\tau g^{\mu\nu}, \text{ itself obtained by repeated use of } g_{\mu\nu}\partial_\sigma g^{\nu\rho} = -g^{\nu\rho}\partial_\sigma g_{\mu\nu} \text{ [p.}$$

24R, right column]. In reality, the resulting relation

$$(78) \quad \partial_\sigma\partial_\rho g^{\mu\nu}\partial_\tau g_{\mu\nu} + \partial_\sigma g^{\mu\nu}\partial_\tau\partial_\rho g_{\mu\nu} = \partial_\sigma\partial_\rho g_{\mu\nu}\partial_\tau g^{\mu\nu} + \partial_\sigma g_{\mu\nu}\partial_\tau\partial_\rho g^{\mu\nu}$$

gives only a tautology after contraction with $g^{\sigma\tau}$. Despite this error, Einstein had in hand an efficient strategy to derive a field equation: start with a simple generalization of the d'Alembertian operator, and introduce additional first-order terms so that the associated 4-force density becomes a 4-divergence. The remaining pages of the notebook [pp. 26L-29L] contain fragments of a rigorous implementation of this strategy, soon published

⁶⁴ The error is explained in *GGR* 2, pp. 699-702.

⁶⁵ As noted in *GGR* 2, pp. 683-686, this identity holds to second order in $h_{\mu\nu}$ (the departure of $g_{\mu\nu}$ from the Minkowskian metric) because in this approximation the index-raising and index-lowering tensors needed to convert $g_{\mu\nu}$ into $g^{\mu\nu}$ (and vice versa) can be regarded as constants. I doubt, however, that Einstein reasoned in this approximation.

jointly with Grossmann as "Outline [*Entwurf*] of a generalized theory of relativity and of a theory of gravitation."⁶⁶

To sum up, in his notebook Einstein first tried naive combinations of the metric tensor and its derivatives in order to get conditionally covariant expressions of the gravitational field equation. Conditional covariance here means covariance with respect to transformations compatible with simple, first-order differential conditions on the metric field. At some point, Einstein became aware of the Riemann tensor, and tried to build the field equation from it, using the fully covariant Ricci tensor as well as related, conditionally covariant tensors. None of the resulting candidates passed the four tests of correspondence with the d'Alembertian equation for weak fields, compatibility with the static metric of 1912 (which contains the acceleration metric as a subcase), compatibility with the rotation metric, and existence of a stress-energy pseudo-tensor for the gravitational field. At this point, Einstein realized that this last test could be used to guide the construction of the modified d'Alembertian in the gravitational field equation. This was the final, winning strategy leading to the published *Entwurf*.

All these considerations were of an utterly formalistic nature, at variance with the more intuitive approach of Einstein's earlier memoirs on generalized relativity. The equivalence principle, being now encapsulated in the geodesic principle, no longer played a direct role in the construction (save for the rotation-metric test). There were no thought experiments, and there was much tensor algebra. Not being accustomed with this more mathematical way of thinking, Einstein encountered a number of difficulties. He lacked familiarity with tensor calculus on a differentiable manifold, and he was relying on conflicting heuristic principles. On the one hand, the full covariance of the equation of

⁶⁶ Einstein and Grossmann 1913.

motion of particles and of energy-momentum conservation for a dust suggested to him the full covariance of the gravitational field equation. On the other hand, his earlier theory of the static field, the correspondence with the d'Alembertian operator, and the physical interpretability of the coordinates suggested a conditional covariance.

Earlier commentators of the Zürich notebook did not fail to notice that in the second, Riemann-tensor-based stage of his research, Einstein wrote down the correct weak-field equations and even the equations from which he would later derive the correct value of the relativist precession of the Kepler ellipse. He did not pursue their empirical consequences at that time because they did not pass his tests of correspondence and energy-momentum conservation. Moreover, Einstein was laboring under the prejudice that the field equations should be simple combinations of the metric tensor and its derivatives. The Ricci tensor did not have this kind of simplicity since it was built from the Christoffel symbols, to which Einstein then accorded little significance (he did not even use them in the geodesic equation).

1.5 The *Entwurf* memoir of 1913

Fundamental equations

In the *Entwurf* they published together in the spring of 1913, Einstein and Grossmann start with the generalized d'Alembertian $(1/\sqrt{-g})\partial_\alpha(g^{\alpha\beta}\sqrt{-g}\partial_\beta g^{\mu\nu})$ of Eq. (60),

contract it with $\sqrt{-g}\partial_\sigma g_{\mu\nu}$, and repeatedly apply the identity $X\partial_\mu Y = \partial_\mu(XY) - Y\partial_\mu X$ to the resulting force density until it becomes a 4-divergence. The resulting identity reads⁶⁷

$$(79) \quad \sqrt{-g}(\partial_\sigma g_{\mu\nu})G^{\mu\nu} = \partial_\nu(\sqrt{-g}g_{\sigma\rho}X^{\nu\rho}), \text{ with}$$

$$(80) \quad X^{\mu\nu} = g^{\alpha\mu}g^{\beta\nu}\partial_\alpha g_{\rho\sigma}\partial_\beta g^{\rho\sigma} - \frac{1}{2}g^{\mu\nu}g^{\alpha\beta}\partial_\alpha g_{\rho\sigma}\partial_\beta g^{\rho\sigma}, \quad G^{\mu\nu} = \Delta^{\mu\nu} + \frac{1}{2}X^{\mu\nu},$$

$$(81) \quad \Delta^{\mu\nu} = (1/\sqrt{-g})\partial_\alpha(g^{\alpha\beta}\sqrt{-g}\partial_\beta g^{\mu\nu}) - g^{\alpha\beta}g_{\rho\sigma}\partial_\alpha g^{\mu\rho}\partial_\beta g^{\nu\sigma}.$$

The linear tensor $G^{\mu\nu}$ is the new d'Alembertian for which the gravitational field equation reads⁶⁸

$$(82) \quad G^{\mu\nu} = \kappa T^{\mu\nu}.$$

Einstein interprets the tensor $t^{\mu\nu} = (-1/2\kappa)X^{\mu\nu}$ as the stress-energy tensor of the gravitational field, so that the field equation can be rewritten as

$$(83) \quad \Delta^{\mu\nu} = \kappa(T^{\mu\nu} + t^{\mu\nu}),$$

which means that the energy-momentum of matter and the energy-momentum of the gravitational field both act as sources of the gravitational field. In addition, the equation

$$(84) \quad \frac{1}{2}\sqrt{-g}\partial_\mu g_{\nu\rho}T^{\nu\rho} = \partial_\nu(\sqrt{-g}g_{\mu\rho}T^{\nu\rho})$$

for the 4-force acting on matter and the field equation together lead to

$$(85) \quad \partial_\nu[\sqrt{-g}g_{\mu\rho}(T^{\nu\rho} + t^{\nu\rho})] = 0,$$

which is the local expression of the conservation of the energy and momentum of matter and field together.⁶⁹

⁶⁷ Einstein (1913, p. 15) claims that this identity is unique. It may be true that this is the only identity that can be obtained by repeated partial integration in the expression of the 4-force. However, as John Norton (1984, p. 282) pointed out, this is not the only possible identity of the desired form; it is only the simplest one.

⁶⁸ The plus sign is needed on the right hand side because in the linear approximation $G^{\mu\nu} = \square g^{\mu\nu}$ and $G_{\mu\nu} = -\square g_{\mu\nu} = \kappa T_{\mu\nu}$, which gives the correct Newtonian limit.

Einstein insists that the weak-field approximation of the new theory produces the desired d'Alembertian. He privately believes that the rotation metric is a solution of his field equation. This equation is linearly covariant by construction, and he hopes for a broader covariance in harmony with the equivalence principle. He has not yet been able to specify the relevant class of transformations even though he regards this task as "the most important one." Nor has he solved any concrete problem. The theory truly is nothing but an outline. So much so that one may wonder why Einstein and Grossmann published it. A plausible answer is that Einstein strongly believed in a strategy based on seeking a generalized d'Alembertian compatible with energy-momentum conservation.⁷⁰

Einstein has arguments against the three alternatives that immediately come to mind. First, one cannot obtain a field equation simply by taking the double covariant derivative of $g^{\mu\nu}$ since the first covariant derivative of this tensor vanishes. Second, a scalar theory of gravitation would not comply with the equivalence principle (more on this in a moment). Third, the Ricci tensor cannot serve to write the field equation because this operator does not reduce to the d'Alembertian in the weak-field approximation. Although Einstein does not give details here, we know from his notebook that he has convinced himself that no coordinate condition would solve this difficulty.⁷¹

⁶⁹ Einstein and Grossmann 1913, pp. 15-17, 37-38.

⁷⁰ Einstein and Grossmann 1913, p. 13 (weak field), p. 18 (citation); for the rotating frames, relevant (erroneous) calculations are found on pp. 24R and 25R of the Zürich notebook, and also in the Einstein-Besso MS of 1913, *ECP* 4, pp. 442-445. On the latter, cf. Janssen 1999, 2007. Besso seems to have reached the opposite conclusion by August 1913 (cf. Janssen 2007), and Einstein must have agreed with him for a while since in early 1914 he believed the *Entwurf* theory to be covariant under linear transformations only. In a letter to Lorentz of August 1913 (cited in Janssen 2007, p. 833), Ehrenfest reported that Einstein had already changed his mind five or six times on this issue.

⁷¹ Einstein and Grossmann 1913, pp. 12 ($D_\mu g^{\mu\nu} \equiv 0$), 20-22 (scalar theory), 11 and 36 (Riemann tensor). On p. 11, Einstein writes: "It must be emphasized that it proves to be impossible to find a [second-order] differential expression $G^{\mu\nu}$ that is a tensor generalization of $\Delta\phi$." This is compatible with the statement that the Ricci tensor does not agree with the d'Alembertian of the metric field in the weak-field limit. On p. 36, Grossmann writes: "It turns out that [the Ricci tensor] does not reduce to $\Delta\phi$ in the limit of an infinitely

From the covariance of the geodesic equation $\delta \int ds = 0$ and of the energy-momentum equation of a dust, Einstein judges it plausible that the exact gravitational equation would be fully covariant. However, he also insists that "we lack any clue for a general covariance of the equations of the gravitational equations." He hopes that in a future theory in which the gravitational field equation would include derivatives of order higher than two, general covariance might be reached. But in a theory limited to second-order derivatives, he is convinced that full covariance is impossible.⁷²

Measurement

In the three first paragraphs of his contribution to the *Entwurf*, Einstein lays out the interpretive basis of his theory. He first recalls that the equivalence principle, when applied to the static gravitational field, leads to a value of the velocity of light depending on the gravitational field. These considerations have suggested to him an extension of the relativity principle in which a larger class of coordinate systems would be considered and in which the element ds would take the general form $\sqrt{g_{\mu\nu} dx^\mu dx^\nu}$. Owing to its role in determining the motion of particles in the field, this element has to be an absolute invariant and $g_{\mu\nu}$ is a covariant tensor determining the gravitational field. In a section entitled "Meaning of the fundamental tensor $g_{\mu\nu}$ for the measurement of time and space," Einstein goes on:⁷³

weak static field." This is incorrect if taken literally. Taking the Zürich notebook into account, Grossmann (and Einstein) probably meant that the static weak field case should be the static subcase of a d'Alembertian, which does not agree with the weak-field limit of the Ricci tensor. For a different interpretation, see Stachel 1989b, p. 67.

⁷² Einstein and Grossmann 1913, p. 12 (citation), 18.

⁷³ Einstein and Grossmann 1913, p. 8.

From the previous considerations, we may already infer that between the space-time coordinates x^μ and the results of measurements through rulers and clocks there cannot be relations as simple as in the old theory of relativity. Regarding time, this feature was already apparent in the case of the static gravitational field. Thus we must raise the question of the physical meaning (principal measurability) of the coordinates x^μ .

We remark that ds is to be regarded as the invariant measure of the distance between two infinitesimally close points of spacetime. Therefore, ds must have a physical meaning independent from the selected system of reference. We assume that ds is the "naturally measured" distance of the two points, to be understood in the following manner.

Einstein then introduces a system of coordinates ξ^μ in which ds^2 takes the Minkowskian form at a given point, and he assumes that the usual measuring prescriptions of special relativity, based on a rigid frame, rulers, and clocks, can be used locally to determine the value of these coordinates.⁷⁴ The metric significance of the coordinates x^μ can then be determined by means of the relation between the differentials dx^μ and the differentials $d\xi^\mu$. This relation being a function of the $g_{\mu\nu}$ coefficients,

Einstein concludes:

We see that by given dx^μ , the natural distance belonging to these differences can be specified only if the quantities $g_{\mu\nu}$ that determine the gravitational field are known: the gravitational field influences the measuring bodies and the clocks in a definite manner.

In this citation, we observe the persistence of Einstein's viewpoint, first expressed in his static field theory, that the coordinates in some sense represent the readings of clocks and rulers. These readings being different from the "naturally measured" ds , the relation between ds and dx^μ being dependent on the gravitational field, the rulers and clock are influenced by this field.⁷⁵

⁷⁴ He overlooks the additional requirement that the derivatives of the metric coefficients should vanish at the selected point.

⁷⁵ Einstein and Grossmann 1913, p. 9.

Mercury's perihelion

In June 1913, Einstein collaborated with his close friend Michele Besso to calculate the advance of Mercury's perihelion in the *Entwurf* theory. The calculation involved a center-symmetric, second-order, weak-field solution of the gravitational field equation. Contrary to the static theory of 1912, Einstein and Besso found that the metric had non-diagonal elements in this approximation (although these do not contribute to the equation of motion in the same approximation). Injecting this metric into the equation of motion of the planet, and solving to second order in v/c (v being the velocity of the planet), they arrived at the expression

$$(86) \quad \delta\theta = \frac{5\pi}{4} \frac{\alpha}{a(1-e^2)}$$

for the relativist anomaly in the advance of the planet, wherein $\alpha = 2GM/c^2$, M is the mass of the sun, G the constant of gravitation, a the semi-major axis of the elliptic orbit, and e its eccentricity. This differs from the expression in the final theory only by having $5\pi/4$ instead of 3π . In the case of Mercury, the formula gives 18" instead of the 43" needed to explain the observed anomaly. Einstein does not seem to have taken this failure too seriously, for he did not mention it in later accounts on the *Entwurf* theory. He may have judged that there were too many simultaneous causes of the perihelion advance (ordinary perturbations by other planets, relativistic correction, and the rotation of the sun) for it to be reliably computed. This is an interesting case of asymmetry between confirmation and refutation: whereas a negative result in the *Entwurf* context did not

suffice to reject the theory, the positive result of November 1915 significantly contributed to the credibility of the final theory.⁷⁶

1.6 The scalar theory

Nordström's theory

In 1912, the Finnish theorist Gunnar Nordström published a scalar, Minkowskian theory of gravitation in which the motion of a particle does not depend on its mass. The evident Minkowskian generalization of Poisson's equation for the scalar field ψ reads

$$(87) \quad \square \psi = -\kappa \bar{\rho},$$

wherein $\bar{\rho} = \rho \sqrt{1 - v^2 / c^2}$ is the density of matter in the local rest frame. The naive

Minkowskian generalization of the Newton law of acceleration for a particle of mass m moving in the field ψ reads

$$(88) \quad f_\mu = \frac{dp_\mu}{d\tau} = m \partial_\mu \psi \quad \text{with} \quad p_\mu = m \frac{dx_\mu}{d\tau} \quad \text{and} \quad d\tau = \sqrt{\eta_{\mu\nu} dx^\mu dx^\nu}.$$

This law being incompatible with

$$(89) \quad 2 \frac{dx^\mu}{d\tau} \frac{dp_\mu}{d\tau} = m \frac{d}{d\tau} \left(\frac{dx^\mu}{d\tau} \frac{dx_\mu}{d\tau} \right) = 0,$$

Nordström assumes a variable mass m for which the equation of motion becomes

$$(90) \quad m \frac{d^2 x_\mu}{d\tau^2} + \frac{dm}{d\tau} \frac{dx_\mu}{d\tau} = m \partial_\mu \psi.$$

⁷⁶ *ECP* 4, pp. 344-473, 630-682. Cf. Earman and Janssen 1993. Einstein and Besso originally made a mistake by a factor hundred in the numerical estimate of the anomaly, but later corrected this error.

Contraction with $dx^\mu / d\tau$ then yields $dm = m d\psi$ and $m = m_0 e^{\psi - \psi_0}$. The equation of motion is still independent of the mass m since it may be rewritten as⁷⁷

$$(91) \quad \frac{d^2 x_\mu}{d\tau^2} = \partial_\mu \psi - \frac{d\psi}{d\tau} \frac{dx_\mu}{d\tau}.$$

Einstein's objections

In private Einstein told Nordström he had already considered the scalar theory but did not pursue it for it led to a slower fall of rotating bodies, against the spirit of the equivalence principle.⁷⁸ In the *Entwurf*, he briefly discussed a variant of Nordström's theory in which the equation of motion of a particle in the gravitational field φ reads

$$(92) \quad \varphi \frac{d^2 x_\mu}{d\tau^2} = \partial_\mu \varphi - \frac{d\varphi}{d\tau} \frac{dx_\mu}{d\tau}$$

(as would results from Nordström's equation for $\psi = \ln \varphi$) and the field equation reads

$$(93) \quad \square \varphi = -\kappa T^\mu_\mu$$

if T^μ_μ denotes the trace of the Minkowskian stress-energy tensor. Einstein derives the equation of motion from the variational principle

$$(94) \quad \delta \int \varphi d\tau = 0,$$

and uses the more general "Laue scalar" T^μ_μ instead of the density ρ_0 of a dust (more on this in a moment). He does not repeat his earlier objection based on a free-falling

⁷⁷ Nordström 1912. For a thorough study of this theory and Einstein's contribution to it, cf. Norton 1992; for a broader discussion, cf. Giulini 2008.

⁷⁸ See Nordström 1912, p. 1129 (addendum). Einstein probably obtained this result by (a misleading) analogy with the slower fall of a body with initial horizontal theory in a scalar theory. The latter effect is easily derived by considering the fall from a frame moving uniformly at the initial horizontal theory of the body. By relativistic time dilation, the fall time in the earth-frame is larger than the fall time in the former frame, which is the same as the time of fall of a body initially at rest owing to the invariance of the scalar field.

spinning body, because he must have understood that in such a body the stresses balancing the centrifugal force contribute to the effective mass $\int T_{\mu}^{\mu} d^3x$. Instead he briefly describes a thought experiment in which cavity radiation is used to construct a perpetual motion. Here is a reconstruction of his argument.⁷⁹

Consider a system made of electromagnetic radiation and of a (quasi-) rigid closed vessel whose internal walls are perfect mirrors. As is well known, the trace of the stress-energy tensor for electromagnetic radiation vanishes. The walls of the vessel nonetheless contribute to the total energy-momentum of the system. According to a theorem by Max Laue, for a closed system in equilibrium the volume integrals $\int T_1^1 d^3x$ etc. of the spatial components of the trace vanishes (because for every section of the system the momentum fluxes $\int T_1^1 dx_2 dx_3$ etc. vanish). Therefore, the net source of the gravitational field is the total energy $E = \int T_0^0 d^3x$. From this result Einstein infers that the gravitational mass of the system is E / c^2 .

Einstein then imagines a long rigid cylinder and two rigidly connected pistons that can slide within the cylinder in a frictionless manner. The internal walls of the cylinder and the pistons are perfect mirrors and they contain a fixed amount of electromagnetic radiation. Consider the system made of the radiation and the double piston. Taking the first coordinate axis along the axis of the cylinder, for this system we have

$$(95) \quad \int T_1^1 d^3x = 0, \quad \int T_2^2 d^3x = \int T_3^3 d^3x = -E / 3$$

because by Maxwell's theory the electromagnetic pressure is one third of the energy density. Therefore, the effective source of the gravitational field is $\int T_{\mu}^{\mu} d^3x = E / 3$, and

⁷⁹ Cf. Norton 1993, pp. 18, 20.

the associated gravitational mass is $E / 3c^2$. Imagine a cycle in which 1) the double piston is raised to the height h in the constant gravity g , 2) the cylinder is replaced by a new wall rigidly attached to the piston, 3) the resulting box is brought to the original elevation, and 4) the new wall is replaced with the sliding cylinder. In this cycle the work $(2E / 3c^2)gh$ is produced without compensation. Einstein regards this result as a sufficient reason to reject the scalar theory, although his strongest reason is the lack of relativity with respect to arbitrary frames of reference.⁸⁰

Natural measures

A few months later, Einstein gave Nordström's theory a new hearing in a review of the gravitation problem for the *Naturforscherversammlung* of September 1913 in Vienna. By then he had worked out the scalar option and changed his mind on its value. By analogy between the variational principles of the scalar and tensor theories, which both involve replacing the Minkowskian form $\eta_{\mu\nu}dx^\mu dx^\nu$ with more complex forms ($\phi^2\eta_{\mu\nu}dx^\mu dx^\nu$ and $g_{\mu\nu}dx^\mu dx^\nu$ respectively), he surmised that in the scalar theory just as in the tensor theory the "natural" lengths and times measured by transportable devices differed from the "coordinate-" lengths and times expressed in coordinate differences.⁸¹

Then the cycle in the former thought-experiment no longer is a cycle if the coordinate-section of the cylinder is invariable and if the natural distance between the two pistons is a constant. Under these assumptions, the true volume V of the radiation expands by $2V\delta\alpha$ during step 1, wherein α is the ratio between natural length and

⁸⁰ Einstein and Grossmann 1913, pp. 20-22.

⁸¹ Einstein 1913, p. 1252.

coordinate-length in the field φ . In order to complete the cycle, an additional step is needed in which the box built in step 2 is contracted back to the volume V . The work needed for the contraction is the product of $2V\delta\alpha$ by the radiation pressure $E/3V$. It compensates the work $(2E/3c^2)gh$ produced by the rise and fall of the radiation if and only if $\delta\alpha = gh$. In other words, the impossibility of perpetual motion requires $\alpha = \varphi$ (+ constant) for the ratio between natural length and coordinate-length.⁸²

Since in Nordström's theory the (coordinate-)velocity of light is a constant, the same ratio obtains between natural time and coordinate-time. In his communication, Einstein asserts that a gravitational clock (made of two bodies orbiting around each other under gravitational pull) gives the natural time for $\alpha = \varphi$.⁸³ This is seen as follows. The mass in the scalar theory being proportional to the potential φ , the centrifugal force is proportional to φRT^{-2} if R denotes the distance between the two bodies (for a circular orbit) and T the period of the motion. This force is balanced by the gravitational pull, which is proportional to R^{-2} . This balance is unaffected by the gravitational field if and only if $\varphi\alpha^{-1}\alpha^2 = \alpha^2$, or $\alpha = \varphi$.

The same result obtains directly by assuming that the form $\varphi^2\eta_{\mu\nu}dx^\mu dx^\nu$ in the variational principle of the scalar theory represents the square ds^2 of the naturally measured interval between two events of coordinates x^μ and $x^\mu + dx^\mu$, just as in the

⁸² Einstein 1913, p. 1253. Einstein does not explain how the absurdity is removed. However, Nordström (1913b, pp. 544-545) details a similar thought-experiment in which Einstein imagined a horizontal stressed rod to be lowered, then raised unstressed to its original height, and stressed again. The rod being heavier when stressed than when unstressed, work would be gained in this cycle if it were not for the work associated with the change of length of the rod (this work is done during the descent of the rod if this descent is done at constant natural length). For a lucid discussion of the latter experiment, cf. Norton 1993, pp. 18-24.

⁸³ Einstein 1913, p. 1254.

tensor theory. As we will see in a moment, Einstein will develop this view a few months later in collaboration with the young Dutch theorist Adriaan Fokker.⁸⁴

In his Vienna talk, Einstein judges the scalar theory to be nearly as hopeful as the *Entwurf* theory. Not only his previous objection has turned out to be invalid, but the scalar theory is compatible with the four basic demands of energy-momentum conservation, universality of free fall, local validity of the Minkowskian structure, and observable laws independent of the value of the gravitational potential (in regions of uniform potential). Its only drawbacks are the lack of relativity with arbitrary reference frames and the incompatibility with Mach's principle.⁸⁵

Einstein has earlier remarked that in his theory the inertia of a particle increases when it is brought near large masses, because the gravitational potential energy of the particle contributes to its inertia. This confirms Ernst Mach's idea that the inertia of a body should be traced to its interaction with all other bodies and not to any mysterious property of space. In the scalar theory, the mass does depend on the gravitational potential but it *diminishes* when the particle is brought near large masses. Although this incompatibility with Mach's principle diminishes the probability of the scalar theory in Einstein's eyes, he admits that only experience can decide between the scalar and tensor options, and he urges astronomers to test the gravitational deviation of light during the solar eclipse of 1914.⁸⁶

⁸⁴ Einstein and Fokker 1914.

⁸⁵ Einstein 1913, p. 1250. See also Einstein to Freundlich, mid-Aug. 1913, in which Einstein judges Nordström's theory "very reasonable."

⁸⁶ Einstein and Grossmann, p. 6; Einstein 1913, pp. 1254, 1260-1261, 1262 (eclipse). On Einstein and Mach's principle, cf. Renn 2007b.

1.7 Bridled covariance

The fatal rope?

In the part of his Vienna talk devoted to the *Entwurf*, Einstein dwells on the meaning of coordinates. He recalls that all the information needed to determine the motion of particles and to measure lengths and times is contained in the ds^2 . Somewhat at variance with this view, in his discussion of the Newtonian approximation he still interprets ds / dx^0 as the rate of clocks and he still has the deviation of light hinge on a variable (coordinate-) velocity. Yet he insists that from the observational point of view and in the spirit of the absolute differential calculus "the coordinate by themselves have no physical meaning" and are mere "auxiliary variables." It is therefore natural to require the general covariance of the fundamental equations of the theory:⁸⁷

The space-time coordinates thereby degenerate into intrinsically meaningless, freely selectable auxiliary variables. Then the entire problem of gravitation would receive a satisfactory solution if we succeeded in finding generally covariant equations for the quantities that determine the gravitational field.

Einstein goes on to admit that he has not been able to find fully covariant field equations of second differential order. In the *Entwurf* memoir, he expressed the hope that full covariance would obtain at a higher differential order. In a letter to Lorentz written on 14 August 1913, he deplored his incapacity to reach general covariance or even to find any non-linear transformation that left his equations invariant. Two days later, he again wrote to Lorentz because he had just realized that energy conservation could serve to justify the restriction to linear transformations: "Now that this ugly dark spot [an

⁸⁷ Einstein 1913, pp. 1256-1257 (citations), 1260 (Newtonian approximation).

unjustified restriction of the covariance] has been eliminated, the theory can at last bring me some pleasure."⁸⁸

In September in Vienna, Einstein explains that the equation

$$(85) \quad \partial_\nu [\sqrt{-g} (t_\mu^\nu + T_\mu^\nu)] = 0$$

for energy conservation is covariant only through linear transformations if t_μ^ν has the same transformation properties as T_μ^ν (as Einstein then believes it should). In the discussion following his talk, Gustav Mie denies that a wagon's frame is equivalent to the earth-bound frame because fluctuations in the wagon's motion (owing to irregularities of the rails etc.) cannot be assimilated to a gravitational field. Einstein replies that this lack of equivalence is to be expected in the *Entwurf* theory. In a footnote to the published text of his talk, he mentions that he has just discovered an argument excluding a generally covariant field equation. Almost certainly, this is the hole argument that he will soon publish in reply to other criticism by Mie.⁸⁹

In the latter publication, Einstein reasserts that the "kernel of the equivalence principle," which is the geodesic equation $\delta \int ds = 0$, leads to the demand of general covariance, and he presents the lack of covariance of the *Entwurf* equations as the "fatal rope" with which his colleagues hope to strangle the theory. He distinguishes two ways in which an equation may lack general covariance: 1) There exists a covariant equation from which the given equation can be obtained by specializing the coordinates, 2) There

⁸⁸ Einstein to Lorentz, 14 Aug. 1913, 16 Aug. 1913, *ECP* 5. See also Einstein to Ehrenfest, Nov. 1913, *ECP* 5; Einstein to Mach, Dec. 1913, *ECP* 5: "The reference frame is so-to-say fitted [*angemessen*] to the existing world by means of the energy principle and thus loses its aprioristic nebulous existence. . . . [The coordinates] are otherwise quite arbitrary."

⁸⁹ Einstein 1913, pp. 1258 (linear covariance), 1264 (reply to Mie), 1257n (hole argument). On the argument for linear covariance, cf. Norton 1984, pp. 285-286.

is no such covariant generalization. In the first case, we have a genuine field equation acting as a restriction on possible fields. In the second case, we have only a restriction on the choice of coordinates. An example of the first case would be Eq. (72), which is a restriction of $R_{\mu\nu} = -\kappa T_{\mu\nu}$ for $\sqrt{-g} = 1$. An example of the second case is $\sqrt{-g} = 1$, because any metric (any system of geodesics) can be expressed under this restriction. Einstein firmly believes that the field equation of the *Entwurf* belongs to the first category and that his colleagues' objection therefore reduces to the weaker objection that no sufficient ground has been given for specializing the system of coordinates.⁹⁰

The hole argument

Einstein then gives two such reasons, a "logical" one, and an "empirical" one. We have already encountered the empirical reason, which is the alleged impossibility of expressing energy-momentum conservation in covariant form. The logical reason is what will later be known as the "hole argument." Einstein considers a finite portion of space in which the stress-energy tensor of matter $T^{\mu\nu}$ vanishes. If the field equation is fully covariant, from a given solution $g_{\mu\nu}(x)$ of this equation we may construct another solution $g'_{\mu\nu}(x)$ by the following procedure (x stands for the quadruplet x_0, x_1, x_2, x_3). Perform a (regular) change of coordinates in which the new coordinates agree with the old ones outside the hole. This leads to the field $g'_{\mu\nu}(x')$, which satisfies the transformed field equation by general covariance. In the transformed equation, we have $T'^{\mu\nu}(x') = T^{\mu\nu}(x')$ everywhere because the coordinates are unchanged outside the hole and because $T^{\mu\nu}$ vanishes within

⁹⁰ Einstein 1914a, pp. 177-178.

the hole. Consequently, $g'_{\mu\nu}(x)$ is a solution of the original field equation. As Einstein wants the field $g_{\mu\nu}(x)$ to be uniquely determined by the field equation for a given value of $T^{\mu\nu}(x)$, he feels compelled to specialize the reference frames. The argument is repeated in an appendix to his contemporary republication of the *Entwurf*. Einstein will hold fast to it for two more years.⁹¹

Bridging the scalar and tensor theories

In early 1914, Einstein and Fokker together published a new study of the scalar theory in which the methods of the absolute differential calculus were applied to the form

$$(96) \quad ds^2 = \varphi^2 \eta_{\mu\nu} dx^\mu dx^\nu.$$

From this point of view, the scalar theory differs from the tensor theory only by a specialization of the $g_{\mu\nu}$ coefficients. The characteristic assumption is the existence of a system of coordinates for which the (coordinate-)velocity of light is the constant c . For all such systems of coordinates the above-given form of the metric element applies. As

Einstein knows, the simplest scalar invariant that can be derived from the metric is the

⁹¹ Einstein 1914a, p. 178; 1914b, p. 260. In modern language, Einstein performs an active diffeomorphism instead of a mere change of coordinates. This is made clear in footnote (1) of Einstein 1914a, and also in Einstein 1914c, p. 1067. Cf. Norton 1984, pp. 287-289; Stachel 1989b. The argument almost certainly originated in a private question of Besso to Einstein, as recorded in a memo of 28 August 1913 by Besso: cf. Janssen 2007, pp. 789, 819-830; Renn and Sauer 2007, p. 237-243. There Besso notes that if the field equation was fully covariant, then the metric field outside a domain containing all matter would not be uniquely determined. This looks like the hole argument, inside out. However, Besso argues that observable phenomena (e. g. the motion of a material point) may not be affected by this indetermination (not quite the modern reply, because the motion of the points in a given coordinate system is not an intrinsic notion). In a subsequent insert plausibly recording Einstein's reaction, this escape is excluded by arguing that different solutions of the field equation yield different motions in a given reference system. This response betrays the reification of reference systems implicit in the hole argument. Renn and Sauer suggest that Besso's consideration emerged out of concern for the uniqueness of the solutions of the Mercury problem (in which the matter of the sun is indeed contained within a finite domain). I suggest it may be more directly directed with the rotation problems studied by Einstein and Besso: the metric generated by a rotating sun, the metric of a non-rotating sun in a rotating frame, and the possibility of substituting one for the other. The comparison naturally suggests the intermediate case of coordinates agreeing with those of the rest frame inside the sun, and agreeing with those of the rotating frame outside the sun.

twice contracted Riemann tensor R (the trace of the Ricci tensor). Given the form (96) of the metric, a straightforward calculation yields

$$(97) \quad R = -6\varphi^{-3} \square \varphi.$$

For the field equation, Einstein simply takes

$$(98) \quad R = \kappa' T_{\mu}^{\mu}.$$

In the dust case, $T_{\mu}^{\mu} = \rho_0$, with $\rho_0 = \bar{\rho}\varphi^{-3}$ since the density ρ_0 is defined with respect to the natural volume and the density $\bar{\rho}$ with respect to the coordinate-volume. The field equation is therefore the same as the Einstein-Nordström Eq. (93), with $\kappa = \kappa'/6$.⁹²

These considerations reinforced Einstein's conviction that in any theory compatible with the equivalence principle, the basic field equation could be obtained from a fully covariant field equation by specializing the choice of coordinates:

Since Nature does not provide us with reference systems to which we could relate things [*Dinge*], we refer the four-dimensional manifold to fully arbitrary coordinates (corresponding to the Gaussian coordinates in the theory of surfaces) and we limit the choice of the reference system only when the treated problem induces us to do so.

In the scalar case, the specialization of coordinates rests on the light principle. In the tensor case, it rests on the energy-momentum principle.⁹³

Most important, Einstein now believes that the Riemann tensor could serve to derive the field equation both in the scalar and in the tensor case. As he indicates in a footnote, he no longer holds that a field equation based on this tensor would be incompatible with Poisson's equation in the weak-field static limit. He even hopes to retrieve the *Entwurf* field equation by specializing the coordinate system in a generally covariant, Riemann-tensor-based equation, thus providing a derivation independent of the

⁹² Einstein and Fokker 1914.

⁹³ Einstein and Fokker 1914, p. 321.

more "physical" requirements such as energy-momentum conservation and correspondence with Newton's theory:⁹⁴

Finally, the role played by the Riemann-Christoffel tensor in the present investigation [of the scalar theory] makes it plausible that this tensor may also open the way to deriving the Einstein-Grossmann equations in a manner independent of physical assumptions. A proof of the existence or inexistence of a connection of this sort would be a significant theoretical progress.

1.8 Justified transformations and adapted coordinates

The restricted covariance of the Entwurf theory

A proof of a connection between the Riemann tensor and the *Entwurf* field equation would be conceivable only after determining the transformations that preserve this equation. In the *Entwurf* memoir, Einstein had already declared this question as the most important in his agenda. In particular, he hoped that these transformations would include accelerated and rotating frames. As we saw, in mid-August 1913 he abandoned this hope and came to believe that the equation for momentum-energy conservation restricted the covariance to linear transformations. At some point he realized that this argument was invalid, because it rested on the false assumption that the stress-energy t_{μ}^{ν} "tensor" of the gravitational field was a genuine tensor. In March 1914, Einstein wrote to his friend Heinrich Zangger that, contrary to his earlier view, the *Entwurf* theory allowed for

⁹⁴ Einstein and Fokker 1914, p. 328.

"arbitrarily moving reference frames." The remark made the theory so "harmonious" that he "no longer had the slightest doubt about the correctness of the theory."⁹⁵

The reason for this volte-face was Einstein's renewed collaboration with Grossmann. The two friends had recently proved that the "justified" (*berechtigte*) transformations leaving the *Entwurf* field equation invariant were those preserving the equation

$$(99) \quad B_\sigma \equiv \partial_\nu \partial_\alpha (\sqrt{-g} g^{\alpha\beta} g_{\sigma\mu} \partial_\beta g^{\mu\nu}) = 0.$$

From the form

$$(100) \quad \partial_\alpha (\sqrt{-g} g^{\alpha\beta} g_{\sigma\mu} \partial_\beta g^{\mu\nu}) = \kappa \sqrt{-g} (t_\sigma^\nu + T_\sigma^\nu),$$

of the *Entwurf* field equation (83) and from the conservation equation (85), it is clear that any justified transformation leaves Eq. (99) invariant. The reciprocal statement is less evident. Following an advice from the mathematician Paul Bernays, Einstein and Grossmann based their proof (about which more in a moment) on the invariance properties of the field action

$$(101) \quad S_F = \frac{1}{2} \int g^{\alpha\beta} \partial_\alpha g_{\rho\sigma} \partial_\beta g^{\rho\sigma} \sqrt{-g} d^4x,$$

whose variation leads the *Entwurf* field equation through⁹⁶

$$(102) \quad \delta S_F - \kappa \int T_{\mu\nu} \delta g^{\mu\nu} \sqrt{-g} d^4x = 0.$$

Bernay probably reasoned that it was easier to investigate the covariance properties of a scalar than those of a tensor. Having attended Max Planck's lectures in Berlin and David Hilbert's lectures in Göttingen, he was aware of the power of the principle of least action.

⁹⁵ Einstein and Grossmann 1913, p. 18; Einstein to Zangger, 10 March 1914, *ECP* 5. The argument for linear transformations is given in Einstein 1913, p. 1258; 1914a, p. 178; 1914b, p. 260. It is rejected in Einstein and Grossmann 1914, p. 218n.

⁹⁶ Einstein has 2κ instead of κ in the following equation.

He did a great service to Einstein in directing him to the action of the gravitational field. So far Einstein had used the principle of least action only for the motion of particles and had based his field heuristics on energy-stress considerations.⁹⁷

At first glance, the reduction of the covariance of the *Entwurf* field equation to the covariance of $B_{\sigma} = 0$ does not seem to be a huge progress. In particular, as Einstein admitted to Lorentz, the transformations do not constitute a group, since they depend on the metric $g_{\mu\nu}(x)$ to which they are applied. Einstein accepted this "non-autonomous" character of the transformations and did not worry much about the complexity of the condition $B_{\sigma} = 0$. What mattered to him was that in his opinion this condition allowed for non-linear transformations, in harmony with the equivalence principle.⁹⁸ In private, he also told Besso that the justified transformations included the transformation of the Minkowski metric to a rotating frame.⁹⁹

Formal foundation of the Entwurf theory

In November 1914, Einstein published a bulky memoir entitled "The formal foundation [*formale Grundlage*] of general relativity theory." After a brief introduction of the equivalence *principle* and the fundamental metric field, comes the "theory of covariance"; then the stress-energy tensor of matter, the vanishing of its covariant divergence, and its expressions for a dust and for the electromagnetic field; next, a new "formal" derivation of the *Entwurf* equations based on the invariance properties of the

⁹⁷ Einstein and Grossmann 1914, p. 219n (Bernay).

⁹⁸ Einstein and Grossmann 1914, p. 8. Einstein there asserts the possibility of non-linear transformations without giving a proof. He may have reasoned, as he did in Einstein to Lorentz, 23 Jan. 2015, *ECP* 8, p. 80, that $B_{\sigma} = 0$ was compatible with any choices of the coordinates at the border of a hole.

⁹⁹ Einstein to Besso, 10 March 1914, *ECP* 5. Einstein long hesitated on the validity of this calculation (see note 64 above), until he firmly decided against it : Einstein to Freundlich, 30 Sept. 1915, *ECP* 8.

field action. This derivation capitalizes on Einstein and Grossmann's earlier characterization of the restricted covariance of the *Entwurf* equations.¹⁰⁰

A little more needs to be said on this characterization. It relies on the fact that the condition $B_\sigma = 0$ (Eq. (99)) is satisfied if and only if the coordinates x^μ are chosen so that $\delta_x S_F = 0$ under any infinitesimal transformation $x^\mu \rightarrow x^\mu + \delta x^\mu$. In Einstein and Grossmann's terms, such coordinates are "adapted" (*angepasst*) to the manifold with the metric $g_{\mu\nu}(x)$. The *Entwurf* field equation turns out to be invariant with respect to a transformation preserving the adaptation of coordinates, because the variation $\delta_g S_F$ of the action for $g_{\mu\nu} \rightarrow g_{\mu\nu} + \delta g_{\mu\nu}$ (whose vanishing implies the field equation), is invariant under such transformations. In Einstein's eyes, the adapted coordinates have the additional virtue of solving the hole paradox: the adaptation provides four conditions through which the coordinates in the hole are fully determined when they are known on the border of the hole.¹⁰¹

In his *formale Grundlage* of November 1914, Einstein no longer assumes the expression (101) of the field action and instead tries to derive this expression by "formal" considerations. For the generic action

$$(103) \quad S_F = \int \Lambda(g_{\mu\nu}, \partial_\rho g_{\mu\nu}) \sqrt{-g} d^4x,$$

the variation $x^\mu \rightarrow x^\mu + \delta x^\mu$ leads to¹⁰²

¹⁰⁰ Einstein 1914c.

¹⁰¹ Einstein and Grossmann 1914, pp. 219-224. Remember that for Einstein the existence of two different solutions of the field equation with the same coordinates derives from the possibility of expressing the same solution in two different systems of coordinates.

¹⁰² The correctness of the following formulas is easily judged from

$$\int F[X'^\mu(x')] \sqrt{-g'} d^4x' = \int F[X'^\mu(x)] \sqrt{-g} d^4x, \quad X'^\mu(x) = \partial_\nu x'^\mu X^\nu(x), \text{ with } x' = x + \delta x \text{ and } \delta X'^\mu(x) = X'^\mu(x) - X^\mu(x).$$

$$(104) \quad \delta_x S_F = \int (\delta_x \Lambda) \sqrt{-g} d^4x, \text{ with } \delta_x \Lambda = \frac{\partial \Lambda}{\partial g^{\mu\nu}} \delta_x g^{\mu\nu} + \frac{\partial \Lambda}{\partial (\partial_\rho g^{\mu\nu})} \delta_x (\partial_\rho g^{\mu\nu}),$$

$$(105) \quad \delta_x g^{\mu\nu} = g^{\mu\alpha} \partial_\alpha \delta x^\nu + g^{\nu\alpha} \partial_\alpha \delta x^\mu, \quad \delta_x \partial_\rho g^{\mu\nu} = \partial_\rho (g^{\mu\alpha} \partial_\alpha \delta x^\nu + g^{\nu\alpha} \partial_\alpha \delta x^\mu) - \partial_\alpha g^{\mu\nu} \partial_\rho \delta x^\alpha.$$

A simple calculation gives

$$(106) \quad \delta_x \Lambda = X_\mu^\nu \partial_\nu \delta x^\mu + 2g^{\nu\alpha} \frac{\partial \Lambda}{\partial (\partial_\rho g^{\mu\nu})} \partial_\rho \partial_\alpha \delta x^\mu.$$

At this point, Einstein assumes the Lagrangian Λ to be invariant through linear transformations. Then the coefficients X_μ^ν must vanish, and the variation $\delta_x \Lambda$ reduces to the second term of Eq. (106). A double integration by parts of the resulting expression of leads to

$$(107) \quad \delta_x S_F = 2 \int B_\mu \delta x^\mu d^4x, \quad \text{with} \quad B_\mu = \partial_\rho \partial_\alpha \left[g^{\nu\alpha} \frac{\partial \Lambda}{\partial (\partial_\rho g^{\mu\nu})} \right],$$

granted that the fields $\partial_\rho g^{\mu\nu}$ vanish at the border of the domain in which δx^μ does not

vanish. Hence, the adapted coordinates are those for which $B_\mu = 0$ (I use bold face

instead of Einstein's gothic script for the field density $\mathbf{X} = X \sqrt{-g}$ associated to the field X).¹⁰³

Einstein next proves that $\delta_g S_F$ is invariant for transformations between adapted coordinates. Consequently, the field equation

$$(108) \quad G_{\mu\nu} = \kappa T_{\mu\nu} \quad \text{with} \quad \mathbf{G}_{\mu\nu} = \frac{\partial \Lambda}{\partial g^{\mu\nu}} - \partial_\rho \frac{\partial \Lambda}{\partial (\partial_\rho g^{\mu\nu})},$$

¹⁰³ Einstein 1914c, pp. 1069-1071. The similarity with Noether's theorem should be obvious to the modern reader.

which results from $\delta_g S_F = 0$ is covariant under such transformations.¹⁰⁴ In order to restrict the choice of Λ , besides its invariance under linear transformations Einstein assumes that it is of the first differential order, that it is of second degree with respect to $\partial_\rho g^{\mu\nu}$, and that the field equation $G_{\mu\nu} = \kappa T_{\mu\nu}$ should automatically yield the condition $B_\mu = 0$ for adapted coordinates. This last requirement is needed because the ten $g^{\mu\nu}$ coefficients would otherwise be overdetermined by the fourteen equations $G_{\mu\nu} = \kappa T_{\mu\nu}$ and $B_\mu = 0$.¹⁰⁵

The identity $D^\nu T_{\mu\nu} = 0$ for the stress-energy tensor $T_{\mu\nu}$ of matter leads to

$D^\nu G_{\mu\nu} = 0$, which is equivalent to $B_\rho = \partial_\nu S_\rho^\nu$, with

$$(109) \quad S_\mu^\nu = g^{\nu\rho} \frac{\partial \Lambda}{\partial g^{\mu\rho}} + \partial_\rho g^{\nu\sigma} \frac{\partial \Lambda}{\partial (\partial_\rho g^{\mu\sigma})} + \frac{1}{2} \delta_\mu^\nu \Lambda - \frac{1}{2} \partial_\mu g^{\rho\sigma} \frac{\partial \Lambda}{\partial (\partial_\nu g^{\rho\sigma})}.$$

Therefore, the condition $B_\mu = 0$ will be automatically satisfied if $S_\mu^\nu \equiv 0$. Einstein uses the latter condition as a constraint on the form of the Lagrangian Λ . Unfortunately, he does not see that the quantities S_μ^ν are identical with the coefficients X_μ^ν in the full expression (106) of $\delta_x \Lambda$ (his earlier reasoning on adapted coordinates did not require computing these coefficients). With this knowledge, it is clear that the condition $S_\mu^\nu \equiv 0$ is nothing but the condition for the invariance of Λ with respect to linear transformations. Without this knowledge and with false guesses on the value of S_μ^ν for the various

¹⁰⁴ As Levi-Civita told Einstein, this seemingly obvious implication does not hold for infinitesimal transformations. Einstein believed he could save his proof by appealing to finite transformations. See the correspondence published in *ECP* 8.

¹⁰⁵ Einstein 1914c, pp. 1071-1074.

candidate actions, Einstein wrongly believes that the *Entwurf* action (101) is the only one to meet this condition.¹⁰⁶

The failure of Einstein's reasoning is easily understood from the modern point of view. Einstein's condition for adapted coordinates $\delta_x S_F = 0$ is equivalent to $D^\nu G_{\mu\nu} = 0$ owing to the identities¹⁰⁷

$$(110) \quad \delta_x g^{\mu\nu} = D^\nu \delta x^\mu + D^\mu \delta x^\nu \quad \text{and} \quad \delta_x S_F = \int G_{\mu\nu} \delta_x g^{\mu\nu} \sqrt{-g} d^4x,$$

which together imply

$$(111) \quad \delta_x S_F = -2 \int (D^\nu G_{\mu\nu}) \delta x^\mu \sqrt{-g} d^4x$$

by partial integration. Hence the adapted coordinates are just those for which the field equation $G_{\mu\nu} = \kappa T_{\mu\nu}$ is compatible with the equation $D^\nu T_{\mu\nu} = 0$.

Even though Einstein believes the *Entwurf* field action to be the only possible one, he expresses energy-conservation in the more general case of a linearly covariant Λ . For this purpose, he rewrites Eq. (108) as

$$(112) \quad -\partial_\alpha \left[g^{\nu\rho} \frac{\partial \Lambda}{\partial (\partial_\alpha g^{\mu\rho})} \right] = \kappa T_\mu^\nu - g^{\nu\rho} \frac{\partial \Lambda}{\partial g^{\mu\rho}} - \partial_\alpha g^{\nu\rho} \frac{\partial \Lambda}{\partial (\partial_\alpha g^{\mu\rho})}$$

Applying ∂_ν to this equation and using $B_\mu = 0$, he gets¹⁰⁸

$$(113) \quad \partial_\nu (T_\mu^\nu + \mathbf{t}_\mu^\nu) = 0 \quad \text{with} \quad \mathbf{t}_\mu^\nu = -\frac{1}{\kappa} \left(g^{\nu\rho} \frac{\partial \Lambda}{\partial g^{\mu\rho}} + \partial_\alpha g^{\nu\rho} \frac{\partial \Lambda}{\partial (\partial_\alpha g^{\mu\rho})} \right).$$

Although Einstein does not do so, we may use the equation $S_\mu^\nu = 0$ to retrieve the more familiar form of the stress-energy tensor of the gravitational field:

¹⁰⁶ Einstein 1914c, pp.1074-1076.

¹⁰⁷ See Landau and Lifshitz 1951, §94.

¹⁰⁸ I have corrected sign typos.

$$(114) \quad \mathbf{t}_\mu^\nu = \frac{1}{2\kappa} \left(\delta_\mu^\nu \Lambda - \partial_\mu g^{\rho\sigma} \frac{\partial \Lambda}{\partial (\partial_\nu g^{\rho\sigma})} \right),$$

which remains valid when Λ is not linearly invariant.¹⁰⁹

Toward the end of his memoir, Einstein offers a few remarks about the possibility that the same coordinate, say x_0 , may be time-like in some portion of spacetime and space-like in another; about his inability to exclude closed time-like paths in space-time even though they "badly hurt his physical intuition," and about the necessity of abandoning Euclidean geometry because of its implicit appeal to direct action at a distance (through the notion of rigid body). He also gives the Newtonian approximation in which the metric can be approximated by

$$(115) \quad g_{ij} = \delta_{ij}, \quad g_{i0} = 0, \quad g_{00} = 1 + 2\Phi / c^2.$$

Again, he regards ds as the naturally measured time and dx^0 as the rate of a clock.¹¹⁰

Since his first introduction of the stress-energy tensor of matter in the Zürich notebook, Einstein has sought equations in which the quantity $\partial_\mu g_{\nu\rho}$ occurred in a simple manner, because in Eq. (84),

$$\partial_\nu (\sqrt{-g} g_{\mu\rho} T^{\nu\rho}) = \frac{1}{2} \sqrt{-g} \partial_\mu g_{\nu\rho} T^{\nu\rho},$$

for the stress-energy tensor of matter, the right side represented the 4-force density acting on matter. Granted that $T^{\nu\rho}$ in the field equation for the potentials $g_{\mu\nu}$ is the counterpart of the mass density in Poisson's equation for the potential Φ , then $\frac{1}{2} \sqrt{-g} \partial_\mu g_{\nu\rho}$ should

¹⁰⁹ Einstein 1914c, pp. 1076-1077.

¹¹⁰ Einstein 1914c, pp. 1078-1080 (remarks), 1080-1082 (Newtonian approximation).

be a generalization of the gravitational field $-\nabla\Phi$. Soon after the publication of the *Entwurf*, Einstein favored the simpler form

$$(116) \quad \partial_\nu \mathbf{T}_\mu^\nu = \frac{1}{2} g^{\nu\rho} \partial_\mu g_{\rho\sigma} \mathbf{T}_\nu^\sigma$$

of Eq. (84), which gave a prominent role to the quantities

$$(117) \quad \Gamma_{\mu\sigma}^\nu = \frac{1}{2} g^{\nu\rho} \partial_\mu g_{\rho\sigma}.$$

In his *formale Grundlage* Einstein introduces these Γ coefficients and calls them the "components of the gravitational field." He rewrites the *Entwurf* field equation as¹¹¹

$$(118) \quad \partial_\alpha (\sqrt{-g} g^{\alpha\beta} \Gamma_{\nu\beta}^\sigma) = -\kappa (\mathbf{t}_\nu^\sigma + \mathbf{T}_\nu^\sigma)$$

with¹¹²

$$(119) \quad \mathbf{t}_\nu^\sigma = -\kappa^{-1} \sqrt{-g} (g^{\nu\tau} \Gamma_{\mu\sigma}^\rho \Gamma_{\rho\tau}^\mu - \frac{1}{2} \delta_\nu^\sigma g^{\alpha\beta} \Gamma_{\mu\alpha}^\rho \Gamma_{\rho\beta}^\mu)$$

The field action also is a simple function of the Γ components:

$$(120) \quad S_F = -\int g^{\sigma\tau} \Gamma_{\sigma\beta}^\alpha \Gamma_{\tau\alpha}^\beta \sqrt{-g} d^4x,$$

The relative simplicity of these equations and their analogy with the electromagnetic field equations probably contributed to Einstein's faith in the *Entwurf* theory.¹¹³

Remember that right before determining with Grossmann the transformation properties of the *Entwurf* theory, Einstein had speculated that the field equation of this theory might be obtained by specializing the coordinates in an equation based on the

¹¹¹ The kappa in the following equations is half the kappa of the original *Entwurf* equations.

¹¹² I have added the missing minus sign.

¹¹³ Einstein 1914c, p. 1058. Einstein had earlier introduced the "field strengths" $G_{\mu\sigma}^\nu = \frac{1}{2} \sqrt{-g} g^{\nu\rho} \partial_\mu g_{\rho\sigma}$ (ECP 4, p. 568). The form (118) of the field equations is derived from the original *Entwurf* equation (83) by replacing the equated contravariant tensors with the corresponding mixed tensor densities (as suggested by the simplified form of the divergence equation for the total energy-momentum tensor). I do not believe Einstein got it by analogy with the electromagnetic field equation $\partial_\mu F^{\mu\nu} = j^\nu$. Rather, similar principles produced a similar equation.

Riemann tensor. In early April 1914, Einstein wrote to Ehrenfest: "Grossmann wrote to me that he is now able to derive the gravitational equations from the generally covariant theory. That would be a nice complement to our investigation." Whatever Grossmann did, there is no mention of it in the *formale Grundlage* of November 1914. Plausibly, the concept of adapted coordinates and Einstein's erroneous belief that it could be used to formally determine the field action, had again diverted him from the Riemann tensor.¹¹⁴

1.9 November 1915

Returning to the Riemann tensor

Einstein had long known that the *Entwurf* theory gave a too small value for the relativistic precession of Mercury's perihelion but he did not seem to worry much about it. In September 1915, he firmly decided that the *Entwurf* field equation was incompatible with the rotation metric, after much wavering on this issue. In a letter to his favorite astronomer Erwin Freundlich, he called this result "a logical contradiction of the quantitative kind." Yet he did not yet suspect the foundations of his theory. He opined that "a computational error must be hidden somewhere in the edifice" and that the same

¹¹⁴ Einstein to Ehrenfest, early Apr. 1914, *ECP* 8, #2. It seems impossible to derive the *Entwurf* field equation from the Riemann tensor with coordinate conditions (because there are vacuum solutions of this equation for which $R_{\mu\nu} \neq 0$). It could well be that Grossmann got similar but different equations, as Einstein himself did in November 1915.

error might be responsible for the Mercury perihelion failure. He hesitated between a mistake in the coefficients of the field equation and a misapplication of this equation.¹¹⁵

The most serious blow came a few weeks later, in early October 1915, when Einstein realized that his condition $S^\nu_\mu \equiv 0$ did not further restrict the choice of the field action. More broadly, the condition $B_\mu = 0$ for adapted coordinates follows from the field equation and the vanishing divergence of the energy-stress tensor, for any choice of the Lagrangian.. In the letter in which Einstein announced this failure to Lorentz, he concluded that the correspondence with Newton's theory was the only way to justify the *Entwurf* choice for the field action. However, he soon came to favor a more formal way to restrict this choice, based on general covariance. In the ensuing communication of 4 November 1915, he explains:¹¹⁶

Thus I returned to the requirement of general covariance for the field equations, from which I had departed with a heavy heart three years ago, during my collaboration with my friend Grossmann. In fact at that time we came quite close to the forthcoming solution of the problem.

...

The charm of this theory will not escape anyone who has truly grasped it. It means a true triumph of the methods of the general differential calculus founded by Gauss, Riemann, Christoffel, Ricci-Curbastro, and Levi-Civita.

All the obstacles Einstein had originally imagined against the covariant approach based on the Riemann tensor had successively vanished: since mid-1913 he no longer expected agreement with his scalar theory of 1912; since early 1914, he no longer required the linear part of the field equation to agree with the d'Alembertian equation (at least not without a proper coordinate condition); and he had just ceased to believe that

¹¹⁵ Einstein to Freundlich, 30 Sept. 1915, *ECP* 8. For two space-dimensions, Einstein found $g_{00} = 1 - (3/4)\omega^2(x_1^2 + x_2^2)$ to second order in the angular velocity ω by injecting the first-order rotation metric into the *Entwurf* field equation, whereas for the true rotation metric $g_{00} = 1 - \omega^2(x_1^2 + x_2^2)$.

¹¹⁶ Einstein to Lorentz, *ECP* 8, 12 Oct. 1915. Einstein 1915a, pp. 778-779.

the field equation should involve a simple combination of the field derivatives $\partial_\rho g_{\mu\nu}$. In his first communication of November 1915, , he denounces the "fatal prejudice" (*ein verhängnisvolles Vorurteil*) that the quantities $\frac{1}{2} g^{\nu\rho} \partial_\mu g_{\rho\sigma}$ were the natural candidates for the components $\Gamma_{\mu\sigma}^\nu$ of the gravitational field, as suggested by the form (115) of energy-momentum conservation. He now realizes that the absolute differential calculus and the standard form of the geodesic equation instead favor the Christoffel symbols and he takes $\Gamma_{\mu\nu}^\rho = -\left\{ \begin{smallmatrix} \rho \\ \mu\nu \end{smallmatrix} \right\}$ for the field components.¹¹⁷

The hole argument should not be counted among the obstacles toward general covariance, because Einstein had long ceased to believe that this argument excluded a fully covariant field equation at the most fundamental level. It only meant that the coordinate system had to be adapted to the metric field. It is not clear whether Einstein had yet renounced the hole argument in early November 1915.

For the sake of general covariance, the gravitational field equation should be built from the Ricci tensor. In order to simplify the equation, Einstein subtracts from the Ricci tensor the unimodularly covariant $D_\nu \partial_\mu \ln \sqrt{-g}$, as he has done in the Zürich notebook to get the tensor $R_{\mu\nu}^\times$ of Eq. (71). He thus obtains the unimodularly covariant equation

$$(121) \quad \partial_\rho \Gamma_{\mu\nu}^\rho + \Gamma_{\mu\rho}^\sigma \Gamma_{\nu\sigma}^\rho = -\kappa T_{\mu\nu} ,$$

which is easily seen to derive from

$$(122) \quad \delta S_F - \kappa \int T_{\mu\nu} \delta g^{\mu\nu} d^4x = 0 , \quad \text{with } S_F = \int \Lambda d^4x \text{ and } \Lambda = g^{\sigma\tau} \Gamma_{\sigma\beta}^\alpha \Gamma_{\tau\alpha}^\beta .$$

Einstein next derives the expression

¹¹⁷ Einstein 1915a, pp. 782-783.

$$(123) \quad t_\mu^\nu = \frac{1}{2\kappa} \left(\delta_\mu^\nu \Lambda - \partial_\mu g^{\rho\sigma} \frac{\partial \Lambda}{\partial (\partial_\nu g^{\rho\sigma})} \right)$$

for the stress-energy tensor of the gravitational field such that $\partial_\nu (t_\mu^\nu + T_\mu^\nu) = 0$. For

$\Lambda = g^{\sigma\tau} \Gamma_{\sigma\beta}^\alpha \Gamma_{\tau\alpha}^\beta$, this gives¹¹⁸

$$(124) \quad t_\sigma^\nu = \kappa^{-1} \left(\frac{1}{2} \delta_\sigma^\nu g^{\alpha\beta} \Gamma_{\mu\alpha}^\rho \Gamma_{\rho\beta}^\mu - g^{\nu\tau} \Gamma_{\mu\sigma}^\rho \Gamma_{\rho\tau}^\mu \right).$$

Eqs. (122-124) are the same as in the earlier *formale Grundlage* except that the expression of $\Gamma_{\mu\nu}^\rho$ has changed and the $\sqrt{-g}$ factors have disappeared.¹¹⁹

Combining the field equation, its trace, and the equation for energy-momentum conservation, Einstein gets

$$(125) \quad \partial_\mu (\partial_\alpha \partial_\beta g^{\alpha\beta} - g^{\sigma\tau} \Gamma_{\sigma\beta}^\alpha \Gamma_{\tau\alpha}^\beta) = 0,$$

and integrates it to

$$(126) \quad \partial_\alpha \partial_\beta g^{\alpha\beta} - g^{\sigma\tau} \Gamma_{\sigma\beta}^\alpha \Gamma_{\tau\alpha}^\beta = 0.$$

This condition is covariant because the trace of the covariant field equation leads to¹²⁰

$$(127) \quad \partial_\alpha \partial_\beta g^{\alpha\beta} - g^{\sigma\tau} \Gamma_{\sigma\beta}^\alpha \Gamma_{\tau\alpha}^\beta + \partial_\alpha (g^{\alpha\beta} \partial_\beta \ln \sqrt{-g}) = -\kappa T_\mu^\mu$$

and because T_μ^μ and $\partial_\alpha (g^{\alpha\beta} \partial_\beta \ln \sqrt{-g})$ are unimodular invariants.

Einstein does not see that. On the contrary, he uses the alleged non-covariance of this condition to adapt the coordinates to the manifold.¹²¹ The condition has only one component instead of four in the *Entwurf*. This is one more reason why it cannot be used to compensate the indetermination in the hole paradox. Einstein is silent on this issue.

¹¹⁸ The κ^{-1} factor is missing in Einstein's text.

¹¹⁹ Compare with Eqs. (118-120) above.

¹²⁰ This is equation (21) in Einstein's text.

¹²¹ Einstein 1915a, p. 485.

From a modern point of view, combining energy conservation with the field equation leads to

$$(128) \quad D^\mu R_{\mu\nu}^\times = 0, \text{ with } R_{\mu\nu}^\times = R_{\mu\nu} - D_\nu \partial_\mu \ln \sqrt{-g}.$$

Using the contracted Bianchi identity $D^\mu (R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R) \equiv 0$, this condition is equivalent to

$$(129) \quad D^\mu D_\nu \partial_\mu \ln \sqrt{-g} - \frac{1}{2} \partial_\nu R = 0,$$

which is not an identity. Being (unimodularly) covariant, it restricts the solutions of the field equation without restricting the coordinate system.

In the first, linear approximation, Einstein's condition (126) reduces to

$$(130) \quad \partial_\alpha \partial_\beta g^{\alpha\beta} = 0.$$

This condition, Einstein goes on, "does not yet determine the coordinate system since 4 equations are needed for this purpose. Therefore, in the first approximation we may arbitrarily set $\partial_\beta g^{\alpha\beta} = 0$." This leads to the d'Alembertian equation

$$(131) \quad \square g_{\mu\nu} = 2\kappa T_{\mu\nu},$$

of which Poisson's law is the non-relativistic approximation. Clearly, Einstein no longer regards $\partial_\beta g^{\alpha\beta} = 0$ as a universal condition on physically admissible coordinates. Instead he understands that the linear approximation and the correspondence with Newton's theory in themselves require a specification of the coordinate system. In the end he notes that unimodular transformations include uniform rotation, as he wishes in order to satisfy the equivalence principle. Interestingly, he no longer restricts the transformations to

adapted coordinates (even though he still believes adaptation is needed for the sake of energy conservation) and he considers only the covariance of the initial field equation.¹²²

Addendum

In his article, Einstein remarks that despite the unimodular covariance of the theory, the coordinates cannot be chosen so that $\sqrt{-g} = 1$, because Eqs. (126-127) together imply

$$(132) \quad \partial_{\alpha}(g^{\alpha\beta}\partial_{\beta}\ln\sqrt{-g}) = -\kappa T_{\mu}^{\mu}$$

and because the trace of the stress-energy tensor of matter does not vanish in general (for a dust, the trace is equal to the density of the dust). In an addendum, Einstein speculates that all material energy may be reducible to a combination of electromagnetic and gravitational energy, in which case $T_{\mu}^{\mu} = 0$ for the non-gravitational part of the stress-energy tensor. Then it becomes possible to adopt the fully covariant

$$(133) \quad R_{\mu\nu} = -\kappa T_{\mu\nu}$$

for the gravitational field equation. The equations of the article are retrieved by specializing the coordinates so that $\sqrt{-g} = 1$.¹²³

The gravitational deviation of light and the advance of Mercury's perihelion

Einstein announced his new theory (without the addendum) to the Berlin Academy on 4 November 1915. Two weeks later, on 18 November, he communicated two basic predictions of the theory: a gravitational deviation of light twice as large as in his earlier theory; and an anomaly of about 43" in the advance of Mercury's perihelion, in stunning

¹²² Einstein 1915a, p. 786.

¹²³ Einstein 1915a, pp. 799-801.

agreement with the $45'' \pm 5''$ measured by astronomers. In the first approximation,

Einstein gives

$$(134) \quad g_{ij} = -\delta_{ij} - \alpha x_i x_j / r^3, \quad g_{i0} = g_{0i} = 0, \quad g_{00} = 1 - \alpha / r, \quad \text{with } r^2 = x_1^2 + x_2^2 + x_3^2$$

for the center-symmetric solution of the field equation in quasi-Cartesian coordinates.

The coordinate-velocity of light,

$$(135) \quad c(r) = \sqrt{(\mathrm{d}x^1 / \mathrm{d}x^0)^2 + (\mathrm{d}x^2 / \mathrm{d}x^0)^2 + (\mathrm{d}x^3 / \mathrm{d}x^0)^2} \quad \text{for } g_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu = 0,$$

depends on location. By Huygens's principle (in analogy with familiar reasoning of the

propagation of light in a medium of variable index), this leads to the deviation $2\alpha / \Delta$

instead of the α / Δ given by Einstein's earlier theories. The coefficient α being $2GM / c^2$

for a central mass M , the deviation of rays passing close to the surface of the sun is about

$1.7''$.¹²⁴

In order to compute the motion of a point-mass in a central gravitational field,

Einstein uses the Γ coefficients derived from the expression (134) of the metric, except

for the Γ_{00}^i coefficients, which he gets to the next order of approximation through

$\partial_i \Gamma_{00}^i + \Gamma_{\mu\rho}^\sigma \Gamma_{\nu\sigma}^\rho = 0$. He then injects these coefficients into the geodesic equation

$$(136) \quad \frac{\mathrm{d}^2 x^\mu}{\mathrm{d}s^2} - \Gamma_{\nu\rho}^\mu \frac{\mathrm{d}x^\nu}{\mathrm{d}s} \frac{\mathrm{d}x^\rho}{\mathrm{d}s} = 0.$$

Using the techniques of integration earlier elaborated with Besso in the *Entwurf* context,

he finds

$$(137) \quad \delta\theta = 3\pi \frac{\alpha}{a(1-e^2)},$$

which is 12/5 of the *Entwurf* value given in Eq. (86).

¹²⁴ Einstein 1915b. Cf. Earman and Janssen 1993.

The final equation

In the former communication, Einstein still defends the electromagnetic-gravitational reduction of matter for which the trace T^μ_μ vanishes and the gravitational field equation can consistently be $R_{\mu\nu} = -\kappa T_{\mu\nu}$. Under this hypothesis, he notes, "space and time are deprived of the last remnant of physical reality." However, in a footnote to the printed version of his text, he mentions he will soon show how to do without this hypothesis. The new reasoning is found in a third communication to the Berlin Academy of 25 November.¹²⁵

The basic idea is that a slightly modified version of the field equation,

$$(138) \quad R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = -\kappa T_{\mu\nu} \quad (\text{with } R = R^\mu_\mu = g^{\mu\nu} R_{\mu\nu})$$

no longer requires $T^\mu_\mu = 0$. Einstein thereby uses the kind of calculations with which he related T^μ_μ to $\sqrt{-g}$ in his first November communication. These involve complex, non-covariant manipulations in which the stress-energy tensor of the gravitational field still plays a prominent role. We would now rely on the contracted Bianchi identity

$D^\mu (R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R) \equiv 0$, which is a consequence of the invariance of the field action (using

Eq. (111)). If the field equation is taken to be $R_{\mu\nu} = -\kappa T_{\mu\nu}$, then the former identity

combined with $D^\mu T_{\mu\nu} = 0$ gives $D^\mu (g_{\mu\nu} R) = 0$, which requires $\partial_\mu R = 0$ since

$D^\mu g_{\mu\nu} \equiv 0$. Consequently, R^μ_μ and T^μ_μ must both vanish. This is avoided by adopting Eq.

(137) instead of $R_{\mu\nu} = -\kappa T_{\mu\nu}$.

¹²⁵ Einstein 1915b, p. 831; 1915c. As Einstein mentions in his footnote, the prediction for the relativistic precession is unaltered by the change in the fundamental equation (because this prediction relies on the source-free equation only).

Einstein concludes:¹²⁶

With this, the general theory of relativity is finally made a closed logical construct. The relativity principle in its most general conception, according to which the space and time coordinates become physically meaningless, inevitably leads to a completely determined theory of gravitation explaining the motion of mercury's perihelion.

What happened to the hole argument?

Whereas in the *formale Grundlage* of November 1914 the hole argument played a role in justifying adapted coordinates, no mention of this argument appears in Einstein's later writings. The reason cannot be incompatibility with the general covariance of the fundamental equations, because Einstein believed any physically meaningful equation to be the expression a generally covariant equation in a limited subclass of reference systems. It cannot just be a new awareness that coordinates are mere labels devoid of a physical substratum, because Einstein had earlier expressed this opinion even in the texts in which he presented the hole argument. Yet the lack of any reference to adapted coordinates in the two last communications of November 1915 and in the *Grundlage* of 1916 is a sign that Einstein had abandoned the argument, since the two notions were intimately related. At any rate, the full covariance of the gravitational field equation (already in the addendum to the first communication), signaled a basic error in the argument.

Einstein spelled out this error in a letter to Ehrenfest of 26 December 1915:¹²⁷

From the fact that the two systems $g_{\mu\nu}(x)$ and $g'_{\mu\nu}(x)$ in the same reference frame satisfy the conditions of the gravitational field, one cannot deduce any contradiction to the univocity of evolution. The seeming force of this argument is lost as soon as one considers 1) that the reference system is nothing real, and 2) that the simultaneous realization of two different g -systems (better: two different gravitational

¹²⁶ Einstein 1915c, p. 847.

¹²⁷ Einstein to Ehrenfest, 26 Dec. 1915, *ECP* 8. See also Einstein to Besso, 3 Jan. 1916, *ECP* 8. For a discussion of possible origins of the coincidence argument, cf. Howard and Norton 1991.

fields) in the same domain of the continuum is impossible by the very nature of the theory. §12 [of the *formale Grundlage*, about the hole argument] should be replaced by the following:

The physically real in the world's evolution (in contrast with what depends on the choice of the reference system) consists in coincidences in space and time. For instance, real is the intersection or the non-intersection of two different world-lines. Consequently, the propositions concerning the physically real are unaltered by (univocal) transformations of coordinates. When two $g_{\mu\nu}$ systems (and all variables necessary to the description of the world) are such that one can be deduced from the other by a space-time transformation, then they are completely equivalent. Indeed they share the same point-like coincidences in space and time, that is, they share all observables. At the same time, these considerations show how natural the demand of general covariance is.

In brief, the functions $g_{\mu\nu}(x)$ and $g'_{\mu\nu}(x)$ describe the same metric manifold in different systems of coordinates, even though the notation suggests differently (the argument x is the same for the two functions). When the metric field $g_{\mu\nu}(x)$ is given, we know only that there exists a labeling of the points of the manifold for which the interval between two point labeled by x^μ and $x^\mu + dx^\mu$ is given by $\sqrt{g_{\mu\nu}dx^\mu dx^\nu}$, but we have no independent knowledge of how the labeling is done. Now, the functions $g_{\mu\nu}(x)$ and $g'_{\mu\nu}(x)$ are constructed by means of a diffeomorphism $x \rightarrow x'$ such that $g_{\mu\nu}(x)dx^\mu dx^\nu = g'_{\mu\nu}(x')dx'^\mu dx'^\nu$. The intervals given by the first function with the labeling x of the manifold are therefore equal to the intervals given by the second function with the labeling x' of the manifold.

Einstein did not publicly reject the hole argument in the early years of general relativity. Rather, in his *Grundlage* of 1916 he replaced it with the coincidence argument, just as he told Ehrenfest he would do. Yet there is an interesting formal

connection between Einstein's earlier concept of adapted coordinates and his considerations of 26 October 1916 on the principle of least action in general relativity.¹²⁸

By that time Einstein knew that the field equation (138) could be derived from the invariant action

$$(139) \quad S_F = \int R \sqrt{-g} d^4x .$$

As he also knew, this action differs from the simpler action

$$(140) \quad S_F = \int \Lambda \sqrt{-g} d^4x \quad \text{with} \quad \Lambda = g^{\sigma\tau} (\Gamma_{\sigma\beta}^{\alpha} \Gamma_{\tau\alpha}^{\beta} - \Gamma_{\sigma\tau}^{\alpha} \Gamma_{\alpha\tau}^{\sigma})$$

through a vanishing surface integral. In his memoir he replicates the calculation found in his *formale Grundlage* of November 1914 for the variation $\delta_x S_F$ or the action during the change of coordinates $x^\mu \rightarrow x^\mu + \delta x^\mu$ within a hole, with the result of Eq. (107):

$$\delta_x S_F = 2 \int B_\sigma \delta x^\sigma d^4x, \quad \text{with} \quad B_\sigma = \partial_\rho \partial_\nu \left(g^{\mu\nu} \frac{\partial \Lambda}{\partial (\partial_\rho g^{\mu\sigma})} \right).$$

Whereas in the *formale Grundlage* the action S_F was invariant with respect to linear transformations only, it now is invariant with respect to any smooth transformation.

Therefore, we have $\delta_x S_F = 0$ for any value of x , and the expressions B_μ vanish identically.

Einstein uses this result to prove that the field equation

$$(141) \quad \partial_\rho \left(g^{\mu\nu} \frac{\partial \Lambda}{\partial (\partial_\rho g^{\mu\sigma})} \right) = -\kappa (\mathbf{t}_\sigma^\nu + \mathbf{T}_\sigma^\nu) \quad \text{with} \quad \mathbf{t}_\mu^\nu = \frac{1}{2\kappa} \left(\delta_\mu^\nu \Lambda - \partial_\mu g^{\rho\sigma} \frac{\partial \Lambda}{\partial (\partial_\nu g^{\rho\sigma})} \right)$$

implies the equation $\partial_\nu (\mathbf{t}_\sigma^\nu + \mathbf{T}_\sigma^\nu) = 0$ for the conservation of the total energy-momentum.

Note that the stress-energy tensor of the gravitational field (\mathbf{t}_μ^ν) still plays a role in this reasoning. Remember that in November 1914, Einstein defined the adapted

¹²⁸ Einstein 1916a, pp. 776-777; 1916b. On the latter article, cf. Janssen and Renn 2007, pp. 900-911.

coordinates by $\delta_x S_F = 0$. His reasoning of October 1916 implies the impossibility of this definition when the action is invariant. Using fully covariant reasoning instead of Einstein's recycled calculations, we find

$$(142) \quad \delta_x S_F = -2 \int D^\nu G_{\mu\nu} \delta x^\mu \sqrt{-g} d^4x \quad \text{with} \quad G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R,$$

which yields the contracted Bianchi identity $D^\nu G_{\mu\nu} \equiv 0$. Owing to this identity, the field equations $G_{\mu\nu} = -\kappa T_{\mu\nu}$ are not independent. They determine the 10 functions $g_{\mu\nu}(x)$ only up to 4 arbitrary functions. This freedom exactly corresponds to the indetermination exhibited in the hole argument.

2. RESEARCH DYNAMICS

2.1 Heuristic principles

As recognized by most commentators, Einstein relied on a few heuristic principles and on mathematical techniques and results he learned from Grossmann, Bernay, and others while he was constructing his theory. Although he was a novice in the theory of Riemannian manifolds and although he had to surmount a spontaneous dislike of abstract formalism, he gradually mastered the mathematics he needed for his purpose. His approach was thoroughly algebraic, with rare hints at geometric interpretation. Where we see geodesics, connections, and curvature, he saw trajectories, field strengths, and a covariant 4-tensor, in conformity with the "absolute differential calculus" he inherited from Christoffel, Ricci-Curbastro, and Levi-Civita. A more geometric approach, based on

analogy with the theory of surfaces or on Levi-Civita's later concept of affine connection, would plausibly have oriented Einstein's research differently.¹²⁹

In contrast with the mathematical techniques, Einstein's heuristic principles had physical and intuitive meaning. They guided the construction of the theory, either by suggesting constructive elements, or by testing tentative constructions. In partial conformity with earlier histories, we will consider five principles: the equivalence principle, general covariance, the stress principle, the correspondence principle, and the principle of least action.¹³⁰

The equivalence principle

The equivalence of a gravitational field with the inertial field in an accelerated frame plays a capital role in Einstein's theory. It is based on an extrapolation of the empirically known equality of gravitational and inertial mass. Einstein uses it in a variety of manners: to directly derive observable phenomena including the gravitational redshift, the gravitational curvature of light rays, and the dependence of the gravitational mass on energy content; to determine the space and time coordinates in an accelerated frame; to justify tangent inertial frames and to introduce the invariant ds^2 first in the static case and then in the most general case; and to test possible field equations through their compatibility with the rotation metric.

¹²⁹ On this mathematical background, cf. Reich 1994; Darrigol 2015. According to Lehmkuhl 2014, Einstein never conceived general relativity as a kind of geometrical reduction.

¹³⁰ I do not include electromagnetic analogy among Einstein's heuristic principles, although I agree with Janssen and Renn that there are evident formal analogies between Einstein's gravitational equations and the Minkowskian formulation of electromagnetic theory. The reason is that I do not believe Einstein obtained his equations through this analogy. Rather, the analogy was produced by shared principles, for instance the stress principle, linear covariance, and least action. The stress principle is usually conflated with broader considerations of energy-momentum conservation. I believe it must be singled out and emphasized.

Nowadays, physicists still use the equivalence principle as a springboard to the pseudo-Riemannian manifold of general relativity and to the geodetic principle, but they hurry to say that coordinates are mere labels devoid of physical significance in a generic metric field. In contrast, for Einstein coordinate systems remained tied to reference systems (he indifferently used *Koordinatensystem* and *Bezugssystem*), and he never ceased to believe that accelerated frames were physically meaningful (at least in small domains), in conformity with our common use of earth-bound reference frames despite their not being in free fall. It is with respect to such a frame that the familiar free-falling elevator is falling.¹³¹

General covariance

Einstein arrived at general covariance through the equivalence principle. Firstly, the principle removes the restriction to inertial frames that is characteristic of special relativity. Secondly, the application of the principle by means of local tangent frames leads to the invariant ds^2 and to the equation of motion $\delta \int ds = 0$, which is generally covariant. Thirdly, the motion of a dust is described by a generally covariant equation, $D_\mu \theta^{\mu\nu} = 0$, wherein $\theta^{\mu\nu}$ denotes the stress-energy tensor of the dust. A pure mathematician, at this point, would focus on the intrinsic structure of the Riemannian manifold for space and time and require general covariance for all meaningful field equations on the manifold. Einstein did not do so because he did not divorce general

¹³¹ Cf. Norton 1985 for a lucid discussion of Einstein's version of the equivalence principle and for Einstein's persistent defense of this version. On accelerated frames (local tetrads) in modern general relativity, cf. Synge 1960, pp. 114-118. On the consistency of the equivalence principle with modern general relativity, cf. Darrigol 2015, pp. 186-187. Einstein and Leopold Infeld introduced the free-falling elevator picture in *The evolution of physics* (Cambridge: Cambridge University Press, 1938).

covariance from its physical justification through the equivalence principle. In his view, it could happen that a mathematically acceptable coordinate system would not correspond to a physically acceptable reference system. In practice, he considered local frames in which the metric takes the Minkowski form, and also frames that accelerate or rotate in the former frames. Although it is not clear what he meant a frame to be for the global manifold, one may easily imagine the kind of reference "mollusk" he later popularized: a space-filling array of pre-clocks.¹³²

At any rate, in his earliest writings on generalized relativity Einstein insisted on the frame-dependence of space, time, energy, force, and field measurements. Although this metrological aspect has been generally neglected by commentators, it helps understand why Einstein, in his subsequent quest for a gravitational field equation, could easily imagine a covariance restricted to physically and metrologically significant frames. It also explains the implicit reification of coordinate systems that John Norton sees behind the hole argument.¹³³

The stress principle

In conformity with received field theories, Einstein required the conservation of the total energy and momentum of matter and field. In addition he required the gravitational force density acting on a dust or any matter to derive from stresses in the gravitational field, just as electromagnetic forces acting on a charged matter derive from Maxwell's stress tensor. This is what I called the stress principle. In the static case, this condition

¹³² Einstein 1917, p. 67.

¹³³ Although historians (e. g. Janssen 2014) and philosophers have frequently discussed Einstein's thought experiments based on the equivalence principle, they have usually ignored his operational definitions of physical quantities in accelerated frames (optical synchronization and "pocket" measuring contraptions).

comprehends the vanishing of the total force acting on matter and thereby excludes perpetual motion (by Newton's scholium to his third law). Einstein elevated the four-dimensional generalization of this field-stress assumption to a capital heuristic principle of his new theory of gravitation. Today, we would be content with requiring $D_\mu T^{\mu\nu} = 0$ for the energy-stress tensor of matter $T^{\mu\nu}$, and most of us would be suspicious of any attempt at defining energy, momentum, and stress within the gravitational field. In contrast, Einstein wrote $D_\nu T^{\mu\nu} = 0$ under the form (116)

$$\partial_\nu \mathbf{T}_\mu^\nu = \frac{1}{2} g^{\nu\rho} \partial_\mu g_{\rho\sigma} \mathbf{T}_\nu^\sigma,$$

interpreting the left-hand side as the momentum variation of matter, the right-hand side as the force acting on it. He further required this force to derive from the stress-energy tensor \mathbf{t}_μ^ν for the gravitational field. In the tentative field equation $G_{\mu\nu} = -\kappa T_{\mu\nu}$, the field operator $G_{\mu\nu}$ built from $g_{\mu\nu}$ and its derivatives must be such that $g^{\nu\rho} \partial_\mu g_{\rho\sigma} \mathbf{G}_\nu^\sigma$ can be expressed as the 4-divergence of a symmetric 2-tensor.

This stress principle played an enormous role in Einstein's quest for a field equation, first as a test for tentative equations, then as a means to construct this field equation when it is further assumed to yield a d'Alembertian equation in the weak-field limit (the *Entwurf* strategy). Einstein read the product $\frac{1}{2} g^{\nu\rho} \partial_\mu g_{\rho\sigma} \mathbf{T}_\nu^\sigma$ as a generalization of the Newtonian product of the gravitational vector field by the mass density matter. For this reason, he long regarded the quantities (117)

$$\Gamma_{\mu\sigma}^\nu = \frac{1}{2} g^{\nu\rho} \partial_\mu g_{\rho\sigma}$$

as the "gravitational field components" of which the gravitational field operator would hopefully be a simple combination. This is the "fateful prejudice" he would condemn in November 1915.

The principle of least action

Soon after introducing the invariant element ds^2 , Einstein realized that it could serve to construct the action $\int ds$ for a particle in the $g_{\mu\nu}$ field. From the associated Lagrangian he derived the energy and momentum that guided his construction of the energy-stress tensor of a dust. Yet he did not consider an action for the gravitational field until, in early 1914, the mathematician Paul Bernay advised him to do so in order to study the covariance properties of the *Entwurf* theory. It is not clear why Einstein did not earlier think of basing his theory on an expression for the field action. Had he studied Helmholtz's latest works on electrodynamics or Poincaré's Palermo memoir on the dynamics of the electron, he would have appreciated the the power of the principle of least action in field theory. He would also have seen, as he did with much delay in his *Formale Grundlage*, that this principle offered the simplest and most efficient way to derive field equations that automatically admit a stress-energy tensor. Retrospectively, we can see that some of the $G_{\mu\nu}$ field operators Einstein considered in 1913 derived from a field action and were therefore compatible with the existence of an energy-stress tensor, even though Einstein then believed the contrary.

The correspondence principle

Einstein relied on four different correspondence criteria. Firstly, the new theory had to be compatible with special relativity in a local free falling frame. In the metric approach, this is automatically warranted by the fact that in every small neighborhood of the space-time manifold there is a system of coordinates for which the metric takes the Minkowskian form (Einstein did not specify to which approximation, although Riemann had done so long ago in the Euclidean case by means of geodesic coordinates).

Another of Einstein's correspondence requirements was that in the weak-field limit the gravitational field equation should take the d'Alembertian form $\square g_{\mu\nu} = -\kappa T_{\mu\nu}$, in evident analogy with the electromagnetic field equations. The failure of a given $G_{\mu\nu}$ to meet this condition was not eliminatory, because Einstein felt free to add coordinate conditions that eliminated the unwanted terms in the linearized field operator.

For a while Einstein further required the static solutions of the tentative field equation to agree with his static theory of 1912, in which the metric element has the simple diagonal form $ds^2 = \hat{c}^2(\mathbf{r})dt^2 - d\mathbf{r}^2$. His excessive confidence in this theory could not last long for he knew, by mid-1913, that the *Entwurf* field equation admitted static solutions with non-vanishing off-diagonal elements (of order α^2 / r^2 in the center-symmetric case).¹³⁴ More solidly, Einstein required the Newtonian theory of gravitation to hold in the non-relativistic limit of the static case.

¹³⁴ *ECP* 4, p. 370. Being of second order, the non-diagonal terms do not contribute to the first-correction to the Newtonian motion of planets (whereas they do in the final theory in which they are of first order). They still conflict with the static theory of 1912, because Einstein originally regarded his expression of the static metric as exact (not as an approximation). Some commentators, including Stachel 1989b [1980] and Norton 1884, pp. 299-300 (corrected in Norton 2018, pp. 24-25) have judged differently.

2.2 Early failure with the Riemann tensor

Today we know there are ways of interpreting Einstein's various heuristic principles so that they become mutually compatible and lead to Einstein's field equations of November 1915. We also know that general covariance and the principle of least action are sufficient for this purpose. Unfortunately, Einstein did not rely on a field action before 1914, and he understood his other principles in ways that made them mutually incompatible. This can be seen by examining the reasons why Einstein early rejected gravitational field equations built from the Riemann tensor.¹³⁵

A missed opportunity

In the Zürich notebook, Einstein briefly considered the fully covariant equation

$$R_{\mu\nu} = -\kappa T_{\mu\nu} \text{ and two variants obtained by subtracting from the Ricci tensor terms}$$

invariant through transformations that preserve the determinant g and the symmetric

derivative $\vartheta_{\mu\nu\rho} = \partial_\rho g_{\mu\nu} + \partial_\mu g_{\nu\rho} + \partial_\nu g_{\rho\mu}$ respectively. He rejected the fully covariant

choice because the harmonic coordinate condition that turned $R_{\mu\nu}$ into $\frac{1}{2}\square g_{\mu\nu}$ in the

linear approximation was incompatible with his static metric of early 1912 and because

he could not re-express the product $\sqrt{-g}\partial_\mu g_{\nu\rho}R^{\nu\rho}$ as a 4-divergence. The unimodular and

ϑ -based variants did not have the first defect but they still had the second.

For the sake of the equivalence principle, Einstein also wanted the rotation metric to be a solution of his field equation and coordinate conditions. On the basis of earlier inexact considerations, he probably believed this was the case for the coordinate

¹³⁵ For a review of the reasons evoked in earlier studies, cf. Weinstein 2018.

conditions used in the generally and unimodularly covariant options. He knew the ϑ -covariant option to be incompatible with the rotation metric, but he still hoped there was sufficient analogy between particle motion in a rotating frame and in a metric of vanishing $\vartheta_{\mu\nu\rho}$.

Thus we see that Einstein abandoned the Riemann-tensor-based options essentially for two reasons: incompatibility with the static metric of early 1912 (for the first option), and incompatibility with the existence of a gravitational stress-energy tensor. In the *Entwurf* memoir and in later reminiscences, he suggested that the main reason was incompatibility with the d'Alembertian form in the linear approximation. Before the Zürich notebook was taken into account, this statement was commonly interpreted as Einstein's ignorance of coordinate conditions that bring the desired form. In reality, he was fully aware of such conditions. But in his opinion they failed to solve the d'Alembertian correspondence difficulty because they conflicted with other heuristic requirements: compatibility with the older static metric, existence of an energy-stress tensor.¹³⁶

As noted by Norton, Renn, and Sauer, Einstein thereby understood coordinate conditions as universal restrictions of the class of physically admissible reference frames. In this view, in any physical situation the gravitational field is determined by the original field equation *together with* the coordinate condition. In contrast, the coordinate conditions of general relativity as we know it are adapted to specific problems and they should not be included among the fundamental equations of the theory. Had Einstein used modern coordinate conditions instead of coordinate restrictions, he would still have had

¹³⁶ Cf. Norton 1984.

reasons to reject the Riemann-tensor-based field operators: The equation $R_{\mu\nu} = 0$ allows for non-diagonal static solutions (for instance the Schwarzschild solution) incompatible with his earlier static theory; he would not have seen, without evoking the field action, that the product $\sqrt{-g} \partial_{\mu} g_{\nu\rho} R^{\nu\rho}$ can be turned into a 4-divergence; the field operator is not a simple combination of the expressions $g^{\nu\rho} \partial_{\mu} g_{\rho\sigma}$ that Einstein then regarded as the natural field components.

Evolving obstacles

Einstein's early objections to a Riemann-tensor-based field operator gradually subsided, while new objections emerged. Incompatibility with the static metric of 1912 could no longer be alleged after Einstein and Besso found, in mid-1913, that the center-symmetric solution of the *Entwurf* field equation was non-diagonal. At some point, Einstein may have discovered that none of his coordinate conditions were compatible with the rotation metric. But he had no reason to investigate these conditions in the *Entwurf* context and there is no evidence that he did.

In the late Summer of 1913 Einstein discovered the hole argument following which a generally covariant field equation does not sufficiently determine the metric field. Originally, he believed this argument excluded any generally-covariant field equation. He changed his mind in January 1914, thus appeasing colleagues who deplored his abandonment of general covariance: he now argued that a genuine field equation should be obtainable from a generally covariant equation by specializing the coordinate system. In his subsequent memoir with Fokker on the metric reformulation of the scalar theory, he even expressed the hope that the *Entwurf* field equation would derive from the

fully covariant $R_{\mu\nu} = -\kappa T_{\mu\nu}$ by such specialization. He was therefore willing to resurrect the Riemann tensor, although what truly mattered to him was the effective field equation in the specialized coordinate systems. In his mind this specialization was a necessary precondition for constructing a gravitational stress-energy tensor.

In the same memoir with Fokker, Einstein indicated that he no longer held the non-d'Alembertian character of the linear approximation against a Riemann-tensor-based field operator. The remark evidently resulted from his observation that the Einstein-Nordström field equation could be derived from $R = \kappa T^\mu_\mu$ even though it had the desired d'Alembertian form. Of course, Einstein already knew that proper specialization of the coordinate systems could generate the d'Alembertian from the Ricci tensor. He now believed this could be done without contradicting his other heuristic principles. In particular, the specialized field equation could be compatible with the existence of a gravitational stress tensor as the *Entwurf* field equation was, and it could admit the rotation metric as a solution, as Einstein still hoped.

In the end what blocked Einstein from using a Riemann-tensor-based field-equation was not any prejudice against the heuristic value of this tensor, but the conviction that it could not be used without coordinate restrictions that led to the *Entwurf* field equation. The restrictions seemed necessary in order to avoid the hole paradox. They had to lead to the *Entwurf* field equation because this equation seemed to be the only one compatible with d'Alembertian correspondence and the existence of the stress-energy tensor.

Thus we see that in order to escape from the *Entwurf* charm and move toward a truly general relativity, Einstein had to get rid of two prejudices: the incompatibility of

the Ricci field operator with the existence of the energy-stress tensor, and the necessity of a universal restriction of admissible coordinate systems.

2.3 Einstein's errors

The spotting of errors in works of the past is notoriously risky. Yet there is a clear sense in which Einstein sometimes committed errors he would have himself recognized if anyone had been able to tell him. These are of variable gravity and subtlety.

Trivial errors

The Zürich notebook and other manuscript sources contain numerous errors of calculation, sometimes corrected sometimes not. Most of them are inconsequential. A few are more significant. For instance Einstein repeatedly erred in judging the compatibility of his field equations with the rotation metric, even though the relevant calculations are not especially difficult. Probably, he did not bother to redo a calculation when the result met his expectations. When it did not, he still hoped some minor modification would save the situation. It could also be that, despite the large amount of relevant calculations, the compatibility of his theory with rotating frames was not as crucial as most commentators have assumed. For a short while, in the Fall of 1913, he was willing to limit the covariance of his theory to linear transformations. The rest of the time, he would perhaps have contented himself with non-linear transformations not necessarily including rotation.¹³⁷

¹³⁷ A non-linear transformation $x \rightarrow x'$ can be seen as implying the mutual acceleration $\partial^2 x^i / \partial^2 x^0$ of local frames See Einstein and Grossmann 1914, p. 8: "[The principle of equivalence] is especially convincing when the 'apparent' gravitational field . . . [in an accelerated frame] can be conceived as a 'real' gravitational

Another trivial error occurred in Einstein's argument that energy-momentum conservation, written as $\partial_\nu(\mathbf{t}_\mu^\nu + \mathbf{T}_\mu^\nu) = 0$ was covariant under linear transformations only. There Einstein implicitly assumed the gravitational stress-energy tensor \mathbf{t}_μ^ν to be generally covariant whereas from its expression it is easily seen not to be so.

A more consequential and most embarrassing error occurred in the *formale Grundlage* of November 1914, when Einstein wrongly asserted that the condition $S_\mu^\nu \equiv 0$ was compatible with the *Entwurf* field action only, when it is easily seen to be compatible with any linearly invariant action.

A less trivial error

Although he had no rigorous proof of this, Einstein strongly suspected the Riemann-tensor-based field operators to be incompatible with the stress principle. This is not true for the Ricci tensor and for the $R_{\mu\nu}^\times$ tensor of Eq. (71), but it would be very difficult to derive the associated stress energy-tensor without knowing that the associated field equation derives from an invariant action. This is true for the $R_{\mu\nu}^{\times\times}$ tensor of Eq. (74), but too much covariance is lost in the way. A crucial turning point occurred in Einstein's program when he realized, in the late Fall of 1914, that there existed a stress-energy tensor for any field equation deriving from a field action.

More subtle errors

field, which is the case when accelerative transformations (that is, non-linear transformations) belong to the justified transformations of the theory."

The hole argument was long judged to be based on the trivial error of regarding the expressions of the same metric field in two different coordinate systems, $g_{\mu\nu}(x)$ and $g'_{\mu\nu}(x')$, as two different fields.¹³⁸ In reality, Einstein clearly indicated that he was dealing with two different fields $g_{\mu\nu}(x)$ and $g'_{\mu\nu}(x)$ in the same reference frame, the latter field being the same function of x as $g'_{\mu\nu}(x')$ is a function of x' . Implicitly, Einstein thereby assumed that the reference system to be given before expressing the metric field $g_{\mu\nu}(x)$. Analogously, in Euclidean space an orthonormal basis must be given before expressing a field as a function of the Cartesian coordinates.

The falsity of this analogy is easily seen by considering the simple case of a two-dimensional surface with the metrics $ds^2 = du^2 + dv^2$ and $ds^2 = du^2 + u^2 dv^2$. These metrics differ exactly in the manner of the hole argument, because the latter can be obtained by substituting $u = u' \cos v'$, $v = u' \sin v'$ into the former and then replacing (u', v') by (u, v) . When we write these metric formulas, we of course assume that each doublet (u, v) specifies a point of the surface. But we do not know in which manner until we investigate the consequences of the metric formula. The first formula implies that the lines of constant u and the lines of constant v are geodesics, that they are orthogonal to each other, that the segments of the lines of constant u delimited by two lines of constant v are all of the same length, and reciprocally. We thus discover that we are describing a plane in Cartesian coordinates. The other metric formula implies that the lines of constant v are geodesics converging to a single point at $u=0$, that these lines are orthogonal to the lines of constant u , and that the length of the segments of the latter lines delimited by

¹³⁸ Stachel 1980 first remarked the non-trivial character of the hole argument.

two lines of constant v are proportional to the value of u . That is to say, we are describing the same plane in polar coordinates.

This argument is so simple and so elementary that one may wonder how it eluded Einstein. Even more perplexing is the fact that mathematicians of Hilbert's caliber approved the hole argument. One reason may be that Einstein was not thinking geometrically and that he was instead implementing a highly algebraic "absolute differential calculus" à la Ricci-Curbastro. Another is that in a physical, trans-geometrical context he could imagine a reification of coordinate systems, a concrete reference independent of the choice of the metric field. One could indeed assume such concrete reference; for instance one could, as John Synge suggested in 1921, use local radar coordinates with respect to the world-lines of two clocks.¹³⁹ But then the form of the metric field would be constrained in a complex manner and the task of determining the field equations would become formidable. It is more advantageous to admit any metric field and to regard the metric meaning of the coordinates as a consequence of the metric formula. So we have the choice between an a priori given interpretation of the coordinates with concomitant constraints on the metric field, and an interpretation of the coordinates through an arbitrarily given metric field. Einstein's error was to unconsciously mingle these two incompatible views.¹⁴⁰

Einstein used the hole argument to justify restricted covariance. He had other reasons to restrict covariance even before he introduced this argument. In his view, some general physical requirements such as energy-momentum conservation or correspondence principles could impose a general restriction of the coordinate system, even if general

¹³⁹ Synge 1921.

¹⁴⁰ For a fuller discussion, cf. Darrigol 2015, pp. 165-167.

covariance held at a more fundamental and formal level. For instance, in the Riemann-tensor-based approach he believed the implicit restrictions $\partial_\mu (\sqrt{-g} g^{\mu\nu}) = 0$ or $\partial_\mu g^{\mu\nu} = 0$ were necessary to warrant the d'Alembertian form of the linearized field equation. Nowadays, we would use the harmonic condition for propagation problems but consider it inconvenient for other problems, for instance those involving static gravitational fields. In contrast, Einstein meant his coordinate conditions to be universal, hence the felicitous name "coordinate restrictions" introduced by Renn and Sauer. He tolerated these restrictions, provided they did not conflict with Mach's principle (they could not define a metric structure independent of the distribution of matter) and they did not exclude the accelerated or rotating frames he needed to express the equivalence principle.

Einstein's reliance on coordinate restrictions contributed to his early rejection of a Riemann-tensor-based field operator. One of the restrictions he tried was indeed incompatible with the expected form of the static metric, and he had difficulty judging the compatibility of his other restrictions with the rotation metric. However, recourse to coordinate restrictions was not the chief obstacle and Einstein seems to have quickly shifted to the modern understanding of the harmonic condition once he had overcome the difficulty of conciliating a Riemann-tensor-based field equation with the stress principle.

Should we call Einstein's original reliance on coordinate restrictions an error? It was so from a strategic point of view, since it interfered with the heuristic exploitation of general covariance. From a logical point of view, any problem of general relativity can be solved under any given coordinate restriction since the choice of coordinates is fully arbitrary and untied to the system under consideration. All we can say is that for a given

problem some restrictions are more convenient than others. Well after Einstein arrived at his final equations, there were attempts to reintroduce coordinate restrictions on a physical basis, for instance Vladimir Fock's in the 1950s with harmonic coordinates.¹⁴¹

2.4 Continuities and discontinuities

A superficial reading of Einstein's struggles toward a generalized theory of relativity from 1911 to 1915 could leave the impression that he moved through a chaotic succession of failed attempts until he reached the correct solution. In reality, at every stage of his research Einstein learned something useful for the following stages, both physically and formally. During his first naive guesses at equations for the metric field, he familiarized himself with tensor calculus and a few algebraic properties of the metric tensor, and he already conceived useful coordinate conditions. He soon introduced the energy-stress tensor of matter and the accompanying stress principle, which never ceased to be essential components of his theory. In his aborted attempts at a Riemann-tensor-based field equation, he mastered the relevant calculus of Christoffel, Ricci-Curbastro, and Levi-Civita with Grossmann's help, and he developed the use of coordinate conditions. His failure to conciliate the resulting field equations with the stress principle led him to the *Entwurf* strategy. While working out the relativistic precession of the perihelion in the *Entwurf* theory, he and Besso developed all the techniques he would need for the similar calculation in his final theory. While consolidating the *Entwurf*

¹⁴¹ Fok [Fock] 1959.

theory, he developed the action-based approach that ultimately allowed him to satisfy the stress principle with a Riemann-tensor-based field equation.

In November 1915, Einstein announced that he was giving up the long favored *Entwurf* theory to return to general covariance. This statement has induced most commentators to exaggerate the discontinuity of the transition from the *Entwurf* to the final theory. In reality, there was much continuity, which explains why it took Einstein so little time to complete his theory after returning to full covariance. Janssen and Renn, who properly emphasize this continuity, see it mostly in the "physical" requirements of correspondence, energy-momentum conservation, and analogy with electromagnetic theory, which they oppose to the "formal" requirement of general covariance. Indeed whereas Einstein moved from restricted covariance to general covariance, he preserved the strong electromagnetic analogy Janssen and Renn see in the latest form of the *Entwurf* equations.¹⁴²

In my view, the continuity is best captured by using Einstein's own definition of the physical and formal approaches. In his terms the physical approach is the one he originally used to derive the *Entwurf* field equation by means of the stress principle combined with the correspondence principle: correcting the d'Alembertian equation so that it becomes compatible with the existence of a gravitational stress-energy tensor. The formal approach is based on the covariance properties of the field equation, as investigated by means of the field action. Einstein judged this approach to be superior to the physical approach because it did not rely on the correspondence principle. He originally believed he could re-derive the *Entwurf* field equation in this formal approach. When he discovered a fatal error in this derivation, he saved the formal approach by

¹⁴² Janssen and Renn 2007, p. 840; also *GGR* 2, pp. 500-501.

replacing conditional covariance with general covariance. All he had to do was to change the expression of the field components $\Gamma_{\mu\nu}^{\rho}$ in the field action. In terms of these quantities, most equations of the theory were unchanged, but the field equation was altered to become unimodularly covariant. So with Einstein's definition of "formal" and "physical," the continuity is in the formal approach shared by the *Formale Grundlage* and the November 1915 theory.

Some of the difficulties that had haunted Einstein's theory suddenly disappeared. The hole argument and the dilemmas of restricted covariance went to Einstein's dustbin; the coordinate restrictions became coordinate conditions; the stress principle became compatible with a Riemann-tensor-based field equation. Yet, as Janssen and Renn rightly emphasize, there also was some continuity in the experienced difficulties. The theory remained an evolving, imperfect construct with persisting roots in earlier approaches. Einstein maintained the correspondence between general covariance and the equivalence principle, thus preserving an unexplained reification of coordinate systems; in his discussions of the gravitational redshift or of the gravitational deflection of light, he kept relying on coordinate-based quantities as if they were truer than the invariants built from them; and he maintained the gravitational stress-tensor as an essential component of his theory whereas most physicists now regard it as a physically ill-defined formal intermediate in better founded energy-momentum considerations.¹⁴³

That said, no one would deny that by November 1915 Einstein had the correct general field equations for general relativity (without the cosmological term) and that he

¹⁴³ On Einstein and the stress-energy (pseudo-)tensor for the gravitational physicists, cf. Cattani and De Maria 1993. A few physicists later defended the physical character of this tensor: see Trautmann 1962.

knew how to use them to derive the Newtonian approximation and three crucial departures from it.

2.5 A simplified history

The following pocket history is meant to bring forth two central features of Einstein's quest for a generalized theory of relativity: his concern with the frame-dependence of measurement, and his adherence to a stress-based field ontology.

From the equivalence principle to the metric field

In 1907, Einstein introduced the equivalence principle according to which the effects of gravitation of intensity \mathbf{g} in a small portion of space can be simulated by the acceleration $-\mathbf{g}$ of the reference frame. In a direct application of this principle, as given by Einstein in 1911, the frequency of monochromatic light emitted from a terrestrial source at the moderate elevation h and received at sea-level will appear to be increased by gh/c^2 , because in a frame of upward acceleration g , the receiver acquires the velocity $g(h/c)$ toward the source during the (approximate) traveling time h/c and therefore sees a light Doppler-shifted by the amount $g(h/c)/c$. Yet, for a properly defined time in the accelerated frame, the frequency of light should be conserved during its travel from the emitter to the receiver. Einstein eliminates this discrepancy with the former result by assuming that natural clocks (spectral sources) give a "local time" differing from this truer time. The velocity of light being the constant c with respect to the local time, its true

value must depend on the gravitational potential $\Phi = -m\mathbf{g} \cdot \mathbf{r}$ according to

$\hat{c}(\mathbf{r}) = c(1 + \Phi / c^2)$. Generalizing this law to any form of the potential and using analogy with a transparent medium of variable optical index, Einstein predicts that a heavy spherical body should deviate light by the amount $2GM / c^2\Delta$, wherein G denotes the gravitational constant and Δ the closest distance of the light ray from the center of the body.

In another kind of reasoning developed in 1907 and 1911, Einstein directly compared the time and space measured in an inertial frame with the time and space measured in a constantly accelerated frame. Assuming that the acceleration did not distort the frame and that the velocity of light in the accelerated frame depended on location only, he required

$$(143) \quad dx^2 - \hat{c}^2(x)dt^2 = d\xi^2 - d\tau^2$$

for corresponding differentials of the coordinates in the two frames (Greek letters refer to the inertial frame, Latin letters to the accelerated frame; the x -axis is parallel to the acceleration; and the time unit is chosen so that $c = 1$). For sufficiently small coordinates, this relation is satisfied by the transformation

$$(144) \quad \xi = x + a\hat{c}t^2 / 2, \quad \tau = \hat{c}t$$

if and only is $\hat{c}(x)$ has the affine form

$$(145) \quad \hat{c} = \hat{c}_0 + ax.$$

Together with the equivalence principle, these relations enabled Einstein to predict the effects of a uniform gravitation field on physical phenomena.

In early 1912, Einstein considered an arbitrary static field and mimicked its effect around a given event through an accelerated local frame. The interval between two infinitesimally close events now takes the more general form

$$(146) \quad ds^2 = \hat{c}^2(\mathbf{r})dt^2 - d\mathbf{r}^2.$$

Einstein first obtained the equations of motion in this field by local transformation to a free-falling frame. In the end, he realized that these equations derived from

$$(147) \quad \delta \int ds = 0.$$

They may be written as

$$(148) \quad \mathbf{f} = d\mathbf{p} / dt, \text{ with } \mathbf{p} = \frac{\mathbf{v}}{\hat{c}\sqrt{1 - v^2 / \hat{c}^2}} \text{ and } \mathbf{f} = -\frac{\nabla \hat{c}}{\sqrt{1 - v^2 / \hat{c}^2}}$$

for a unit mass-point. The function $\hat{c}(\mathbf{r})$ thereby plays the role of the gravitational potential.

For the relation between this potential and the density ρ of matter, Einstein generalized Poisson's equation to

$$(149) \quad \frac{\Delta \hat{c}}{\hat{c}} - \frac{1}{2} \left(\frac{\nabla \hat{c}}{\hat{c}} \right)^2 = k\rho,$$

The form of the first-order term is dictated by homogeneity, and the second term is needed so that the force density $-\rho \nabla \hat{c}$ on matter of density ρ derives from the stress system

$$(150) \quad \sigma_{ij} = -4G^{-1}(\partial_i \partial_j \sqrt{\hat{c}} - \frac{1}{2} \partial_k \sqrt{\hat{c}} \partial_k \sqrt{\hat{c}})$$

and therefore satisfies the equality of action and reaction.

With this theory, Einstein believed to have found the unique way of simultaneously satisfying the equivalence principle, Newtonian correspondence, and the

stress principle in the static case. He published it in early 1912 and then tried to generalize it to an arbitrary gravitational field. His successive trials are found in the Zürich notebook written in 1912-1913.

Private attempts

In Einstein's eyes, higher generality of the gravitational field meant higher generality of the permitted reference frames. This is why he looked for a theory in which the motion of particles would be given by the geodesics for the general metric

$$(151) \quad ds^2 = g_{\mu\nu} dx^\mu dx^\nu .$$

The geodesic equation being covariant with respect to any change of coordinates, he looked for a gravitational field equation that would be a covariant generalization of the Poisson equation. He also expected the metric field to satisfy a first-order differential equation that would generalize the equation $\partial_0 g_{00} = 0$ for the static case. That is to say, from the beginning he was willing to restrict the covariance so that the coordinates would be physically meaningful (in the static case, he would exclude coordinates for which the metric coefficients become time-dependent). At the same time, there is ample evidence that he wanted the covariance to be broad enough to include transformations to accelerated or rotating frames, so that the equivalence principle could be concretely implemented.

At some point, Einstein introduced the energy-stress tensor

$$(152) \quad \theta^{\nu\rho} = \rho_0 \frac{dx^\nu}{ds} \frac{dx^\rho}{ds}$$

for a dust of density $\rho_0\sqrt{-g}$, as the natural counterpart of the mass density in a tensor generalization of Poisson's equation. The particles of the dust satisfy the geodesic equation if and only if

$$(153) \quad \partial_\nu(\sqrt{-g} g_{\mu\rho} \theta^{\nu\rho}) = \frac{1}{2} \sqrt{-g} \partial_\mu g_{\nu\rho} \theta^{\nu\rho}.$$

Einstein noticed the general covariance of this equation (which we would now write as $D_\nu \theta^{\mu\nu} = 0$) and took it as an additional reason for seeking a covariant gravitational field equation. In addition, he interpreted $\frac{1}{2} \sqrt{-g} \partial_\mu g_{\nu\rho} \theta^{\nu\rho}$ as the 4-force acting on the dust and required it to derive from a gravitational stress tensor, as he had already done in his static theory. This is what I call the stress principle.

A tensor generalization of the Poisson equation would read

$$(154) \quad G^{\mu\nu} = \kappa \theta^{\mu\nu}.$$

wherein the field operator $G^{\mu\nu}$ is a second-order combination of ∂_μ and $g_{\mu\nu}$. Not being acquainted with the theory of Riemannian manifolds, Einstein tried naive combinations sometimes complemented with coordinated conditions until he learned from Grossmann about the Riemann tensor, which offers a straightforward way to build a fully covariant field operator. Einstein first tried the Ricci tensor $R_{\mu\nu}$, which is the first contraction of the Riemann tensor. In a weak-field approximation, he expected the gravitational field to propagate at the velocity c and he therefore wanted the field equation to agree with the d'Alembertian equation $\square g_{\mu\nu} = -\kappa \theta_{\mu\nu}$. For this purpose, he introduced the harmonic coordinate condition $g^{\mu\nu} \Gamma_{\mu\nu}^\rho = 0$, which reduces the second-order, linear terms of the field operator to the d'Alembertian operator. He gave up after finding this condition to be

incompatible with his earlier static metric. He next tried the unimodularly covariant field operator

$$(155) \quad R_{\mu\nu}^{\times} = R_{\mu\nu} - D_{\nu} \partial_{\mu} \ln \sqrt{-g} = -\partial_{\rho} \{^{\rho}_{\mu\nu}\} - \{^{\sigma}_{\mu\rho}\} \{^{\rho}_{\nu\sigma}\},$$

whose second-order part takes the d'Alembertian form under the condition $\partial_{\mu} g^{\mu\nu} = 0$.

He did not dwell on this option and immediately switched to the much simpler

$$(156) \quad R_{\mu\nu}^{\times\times} = \frac{1}{2} \partial_{\rho} (g^{\rho\sigma} \partial_{\sigma} g_{\mu\nu}) + g^{\rho\alpha} g^{\sigma\beta} \partial_{\alpha} g_{\mu\rho} \partial_{\beta} g_{\nu\sigma},$$

which has the same limited covariance as the expression $\vartheta_{\mu\nu\rho} = \partial_{\rho} g_{\mu\nu} + \partial_{\mu} g_{\nu\rho} + \partial_{\nu} g_{\rho\mu}$.

Einstein hoped that his various restrictions of the covariance still allowed for something like rotating frames, as needed for implementing the equivalence principle. But he stumbled over the incompatibility of the tentative field operators with the stress principle. According to this principle the product $\sqrt{-g} \partial_{\mu} g^{\nu\rho} G_{\nu\rho}$ should be expressible as the 4-divergence of a tensor built from the metric field and its first derivatives. This seemed implausible to Einstein for $G_{\mu\nu} = R_{\mu\nu}$ and for $G_{\mu\nu} = R_{\mu\nu}^{\times}$, and he could easily see it was impossible for $G_{\mu\nu} = R_{\mu\nu}^{\times\times}$. Also, the two first choices of the field operator were incompatible with Einstein's expectation that the gravitational operator should be a simple function of the derivatives $\partial_{\mu} g^{\nu\rho}$. In his opinion, these quantities naturally represented the components of the gravitational field because their contraction with the field-stress tensor of matter yielded the 4-force acting on matter.

The Entwurf theory

While testing whether the field operator $R_{\mu\nu}^{\times\times}$ was compatible with the stress principle, Einstein discovered that he could easily build a compatible operator by correcting the

d'Alembertian operator with non-linear terms generated by partial integration. This led him to the equations he would later rewrite as

$$(157) \quad \partial_\alpha (\sqrt{-g} g^{\alpha\beta} \Gamma_{\nu\beta}^\sigma) = -\kappa (\mathbf{t}_\nu^\sigma + \mathbf{T}_\nu^\sigma), \text{ with}$$

$$(158) \quad \Gamma_{\mu\sigma}^\nu = \frac{1}{2} g^{\nu\rho} \partial_\mu g_{\rho\sigma} \text{ and } \mathbf{t}_\nu^\sigma = -\kappa^{-1} \sqrt{-g} (g^{\nu\tau} \Gamma_{\mu\sigma}^\rho \Gamma_{\rho\tau}^\mu - \frac{1}{2} \delta_\nu^\sigma g^{\alpha\beta} \Gamma_{\mu\alpha}^\rho \Gamma_{\rho\beta}^\mu).$$

The $\Gamma_{\mu\sigma}^\nu$ are the "field components"; the tensor densities \mathbf{T}_ν^σ and \mathbf{t}_ν^σ represent the stress-energy of matter and gravitational field respectively, and they satisfy the equation of conservation

$$(159) \quad \partial_\nu (\mathbf{T}_\mu^\nu + \mathbf{t}_\mu^\nu) = 0.$$

Einstein believed these relatively simple equations to be the only ones compatible with the stress principle and d'Alembertian correspondence. This is why he published them with Grossmann in early 1913 in an outline (*Entwurf*) of a new theory of gravitation, even before testing their compatibility with the equivalence principle.

The *Entwurf* equations are easily seen not to be generally covariant. This did not bother Einstein for he (erroneously) found them to be compatible with the rotation metric while collaborating with Besso on the relativistic precession of a planet's perihelion. During the same collaboration, he arrived at the hole argument following which the metric field within a matter-free hole cannot be uniquely determined by the distribution of matter around the hole if the field equation is fully covariant: if a given field is a solution, any field obtained by diffeomorphic deformation of this field in the hole is also a solution. Originally, Einstein believed the hole argument to exclude full covariance at any level of the theory. In light of his and Fokker's geometric interpretation of the Einstein-Nordström scalar theory of gravitation, he changed his mind and he even came

to hope that the *Entwurf* field equations could be obtained by specializing the coordinate systems in a Riemann-tensor-based field equation.

Resuming his collaboration with Grossmann in early 1914, Einstein succeeded in characterizing the transformations under which the *Entwurf* field equation was covariant. The key, provided by the mathematician Paul Bernay, was to reduce the covariance properties of the field equation to the invariance properties of the field action from which it derives. Einstein and Grossmann found

$$(160) \quad S_F = - \int g^{\sigma\tau} \Gamma_{\sigma\beta}^{\alpha} \Gamma_{\tau\alpha}^{\beta} \sqrt{-g} d^4x$$

for the relevant field action, introduced "adapted coordinates" such that the action is stationary under an infinitesimal variation of the coordinates, and proved that the field equation was invariant under any transformation between adapted coordinates. The four conditions of adaptation,

$$(161) \quad B_\nu \equiv \partial_\alpha \partial_\sigma (\sqrt{-g} g^{\alpha\beta} \Gamma_{\nu\beta}^\sigma) = 0,$$

where just what he needed to avoid the diffeomorphic freedom in the hole. Most important, he believed them to be compatible with non-linear transformations and rotating frames in particular.

Einstein soon exploited these covariance considerations to design a new "formal" derivation of the *Entwurf* theory. For this purpose, he started with an arbitrary, linearly invariant field Lagrangian, and required that the derived field equation $G_{\mu\nu} = \kappa T_{\mu\nu}$, when combined with the energy-momentum conservation $D^\mu T_{\mu\nu} = 0$ should automatically yield the condition for adapted coordinates (in the *Entwurf* case the conditions (161) indeed derive from Eqs. (157) and (159)). Otherwise, there would be 10+4 equations for

determining the 10 unknowns of the metric field. Unfortunately, Einstein erred in developing this condition and falsely concluded that it was compatible with the *Entwurf* action only. From a modern point of view, the impotence of Einstein's condition is easily understood. The variation associated with his condition for adapted coordinates leads to $D^\nu G_{\mu\nu} = 0$, which is precisely the condition for the field equation $G_{\mu\nu} = \kappa T_{\mu\nu}$ to be compatible with the equation $D^\nu T_{\mu\nu} = 0$ for energy-momentum conservation.

The final theory

Einstein became aware of this flagrant error in October 1915. He now understood that any linearly covariant Lagrangian led to a field equation compatible with the stress principle. He also realized that taking $\Gamma_{\mu\nu}^\rho = -\left\{\begin{smallmatrix}\rho \\ \mu\nu\end{smallmatrix}\right\}$ instead of $\Gamma_{\mu\sigma}^\nu = \frac{1}{2} g^{\nu\rho} \partial_\mu g_{\rho\sigma}$ in the field action (160), he could retrieve the unimodularly covariant field operator $R_{\mu\nu}^\times$ he had briefly considered in the Zürich notebook. He reconsidered this option in the following month, now understanding that the condition $\partial_\mu g^{\mu\nu} = 0$ could be used in the weak-field limit to get a d'Alembertian field equation and the static Newtonian sub-case without compromising the unimodular covariance of the theory. He next derived a center-symmetric solution of the field equation $R_{\mu\nu}^\times = 0$ (with $g = -1$) and solved the geodesic equation for a particle moving in this field. The first approximation yielded the Kepler motion, the second a precession compatible with the observed 43" anomaly in Mercury's case.

Instead of the unimodularly covariant field operator $R_{\mu\nu}^\times$, Einstein briefly considered the fully covariant $R_{\mu\nu}$, with the coordinate restriction $g = -1$. As he

understood, this choice is not compatible with the non-vanishing trace of the stress-energy tensor of usual matter. In modern terms, the equation $R_{\mu\nu} = -\kappa T_{\mu\nu}$, the contracted Bianchi identity $D^\mu (R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R) \equiv 0$, and the conservation law $D^\mu T_{\mu\nu} = 0$ together imply the undesired $\partial_\mu T = 0$. Einstein first avoided this difficulty by assuming that all matter was of electromagnetic-gravitational nature (in which case $T_{\mu\nu}$ is purely electromagnetic and has a vanishing trace). In the last week of November he realized that a simple variant of the field equation,

$$(162) \quad R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = -\kappa T_{\mu\nu},$$

was compatible with any value of the trace of the energy-stress tensor. This was his final choice.

Einstein and Hilbert later found out that this equation simply derived from the invariant field action $\int R \sqrt{-g} d^4x$ built from the Riemann scalar R . They also realized, by reasoning similar to Einstein's adaptation of coordinates to the non-invariant *Entwurf* Lagrangian, that the invariance of this action led to the contracted Bianchi identity, as needed for compatibility with energy-momentum conservation. Had Einstein originally based his search for a gravitational equation on an invariant gravitational field Lagrangian, he could have taken this Royal route to general relativity much earlier. This would not have annihilated the difficulties he had conciliating the field operator with the correspondence principle; but he would have thus avoided the chief and most persistent

obstacle in his three years of struggle: seeming incompatibility with the stress principle.¹⁴⁴

Acknowledgments

I thank John Norton and an anonymous reviewer for corrections and helpful suggestions.

Bibliography

ECP: The collected papers of Albert Einstein (Princeton: Princeton University Press).

GGR: J. Renn (general ed.), M. Janssen, J. Norton, T. Sauer, M. Schemmel, J. Stachel (eds.), *The genesis of general relativity*, 4 vols. (Dordrecht: Springer, 2007).

Abraham, Max

1911 Sulla teoria della gravitazione. *Accademia dei Lincei, Rendiconti*, 20: 678-682.

1912a Zur Theorie der Gravitation, *Physikalische Zeitschrift*, 23: 1-4.

1912b Berichtigung. *Physikalische Zeitschrift*, 23: 176.

1912c Sulla conservazione dell' energia e della materia nel campo gravitazionale. *Accademia dei Lincei, Rendiconti*, 21: 432-437.

1912d Die Erhaltung der Energie und der Materie im Schwerkraftfeld. *Physikalische Zeitschrift*, 13: 311-314.

Born, Max

1909a Die Theorie des starren Elektrons in der Kinematik des Relativitätsprinzips. *Annalen der Physik*, 335: 1-56.

1909b Über die Dynamik des Elektrons in der Kinematik des Relativitätsprinzips. *Physikalische Zeitschrift*, 10: 814-817.

Cattani, Carlo, and Michelangelo De Maria

1993 Conservation laws and gravitational waves in general relativity. In J. Earman, M. Janssen, and J. Norton (eds.), *The attraction of gravitation: New studies in history of general relativity* (Boston: Birkhäuser), 63-87.

Darrigol, Olivier

2014a *Physics and necessity: Rationalist pursuits from the Cartesian past to the quantum present*. Oxford: Oxford University Press.

2014b The mystery of Riemann's curvature. *Historia mathematica*, 42: 47-83.

¹⁴⁴ On the Einstein-Hilbert route (and on Einstein's priority in discovering the field equation), cf. Renn and Stachel 2007; Sauer 2005.

2015 Mesh and measure in early general relativity. *Studies in the history and philosophy of modern physics*, 52:163-187.

Dell' Aglio, Luca

1996 On the genesis of the concept of covariant differentiation. *Revue d'histoire des mathématiques*, 2: 215-264.

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: 251-278.

Earman, John, and Michel Janssen

1993 Einstein's explanation of the motion of Mercury's perihelion. In J. Earman, M. Janssen, and J. Norton (eds.), *The attraction of gravitation: New studies in history of general relativity* (Boston: Birkhäuser), 129-172.

Ehlers, Jürgen

1973 Survey of general relativity theory. In Werner Israel (ed.), *Relativity, astrophysics and cosmology* (Dordrecht: Reidel), 1-125.

Einstein, Albert.

1905 Zur Elektrodynamik bewegter Körper. *Annalen der Physik*, 17: 891-921.

1907 Über das Relativitätsprinzip und die aus demselben gezogenen Folgerungen.

Jahrbuch der Radioaktivität und Elektronik, 4: 411-462.

1911 Über den Einfluss der Schwerkraft auf die Ausbreitung des Lichtes. *Annalen der Physik*, 35: 898-908.

1912a Lichtgeschwindigkeit und Statik des Gravitationsfeldes. *Annalen der Physik*, 38: 355- 369.

1912b Zur Theorie des statischen Gravitationsfeldes. *Annalen der Physik*, 38: 443-458.

1913 Zum gegenwärtigen Stande des Gravitationsproblems. *Physikalische Zeitschrift*, 14:1249-1266.

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

1914b Bemerkungen (to a republication of the *Entwurf*). *Zeitschrift für Mathematik und Physik*, 62: 260-261.

1914c Die formale Grundlage der allgemeinen Relativitätstheorie. Königlich Preußische Akademie der Wissenschaften, *Sitzungsberichte* (1914), 1030-1085.

1915a Zur allgemeinen Relativitätstheorie. Königlich Preußische Akademie der Wissenschaften, *Sitzungsberichte* (1915), 778-786, 799-801 (*Nachtrag*).

1915b Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie. Königlich Preußische Akademie der Wissenschaften, *Sitzungsberichte* (1915), 831-839.

1915c Die Feldgleichungen der Gravitation. Königlich Preußische Akademie der Wissenschaften, *Sitzungsberichte* (1915), 844–847.

1916a Die Grundlage der allgemeinen Relativitätstheorie. *Annalen der Physik*, 49: 769–822.

1916b Hamiltonsches Prinzip und allgemeine Relativitätstheorie. Königlich Preußische Akademie der Wissenschaften, *Sitzungsberichte* (1916), 1111–1116.

1917 *Über die spezielle und die allgemeine Relativitätstheorie. (Gemeinverständlich.)* Braunschweig: Vieweg.

Einstein, Albert, and Adriaan Fokker

1914 Die Nordströmsche Gravitationstheorie vom Standpunkt des absoluten Differentialkalküls. *Annalen der Physik*, 44: 321–328.

Einstein, Albert, and Marcel Grossmann

1913 *Entwurf einer verallgemeinerten Relativitätstheorie und einer Theorie der Gravitation*. Leipzig: Teubner.

1914 Kovarianzeigenschaften der Feldgleichungen der auf die verallgemeinerte Relativitätstheorie gegründeten Gravitationstheorie. *Zeitschrift für Mathematik und Physik*, 63: 215–225.

Fok, Vladimir Aleksandrovič

1959 *The theory of space, time, and gravitation*. Oxford: Pergamon.

Giulini, Domenico

2008 What is (not) wrong with scalar gravity? *Studies in history and philosophy of modern physics*, 39: 154–180.

Howard, Don, and John Norton

1991 Out of the labyrinth: Einstein, Hertz and Göttingen answer to the hole argument. In J. Earman, M. Janssen, and J. Norton (eds.), *The attraction of gravitation: New studies in history of general relativity* (Boston: Birkhäuser), 30–62.

Janssen, Michel

1999 Rotation as the nemesis of Einstein's *Entwurf* theory. In H. Goenner, J. Renn, J. Ritter, T. Sauer (eds.), *The expanding worlds of general relativity* (Boston: Birkhäuser), 127–157.

2007 What did Einstein know and when did he know it. *GGR* 2, pp. 785–837.

2012 The twins and the bucket: How Einstein made gravity rather than motion relative in general relativity. *Studies in history and philosophy of modern physics*, 43:159–175.

2014 'No success like failure...': Einstein's quest for general relativity, 1907–1920. In M. Janssen and C. Lehner (eds.), *The Cambridge companion to Einstein* (Cambridge: Cambridge University Press), 167–227.

Janssen, Michel, and Jürgen Renn

2007 Untying the knot: How Einstein found his way back to field equations discarded in the Zürich notebook. *GGR* 2, pp. 839-926.

2015 Arch and scaffold: How Einstein found his field equations. *Physics today*, 68 (11): 30-36.

Landau, Lev Davidovich, and Evgeny Mikhailovich Lifshitz

1951 *The classical theory of fields*. Reading, Mass.: Addison-Westley. French transl. (Paris: Mir, 1970).

Lehmkuhl, Dennis

2014 Why Einstein did not believe that general relativity geometrizes gravity. *Studies in history and philosophy of modern physics*, 46: 316-326.

Nordström, Gunnar

1912 Relativitätstheorie und Gravitation. *Physikalische Zeitschrift*, 13: 1126-1129.

1913 Zur Theorie der Gravitation vom Standpunkt des Relativitätsprinzips. *Annalen der Physik*, 42: 533-554.

Norton, John.

1984 How Einstein found his field equations: 1912-1915. *Historical studies in the physical sciences*, 14: 253-316.

1985 What was Einstein's principle of equivalence? *Studies in history and philosophy of science*, 16: 203-246. 1992 Einstein, Nordström and the early demise of Lorentz-covariant, scalar theories of gravitation. *Archive for history of exact sciences*, 45: 17-94.

1993 Einstein and Nordström: Some lesser known thought experiments in gravitation. In J. Earman, M. Janssen, and J. Norton (eds.) *The attraction of gravitation: New studies in history of general relativity* (Boston: Birkhäuser), 3-29.

2005 A conjecture on Einstein, the independent reality of spacetime coordinate systems and the disaster of 1913. In A. J. Kox and J. Einsenstaedt (eds.), *The universe of general relativity* (Boston: Birkhäuser), 67-102.

2007 What was Einstein's 'fateful prejudice'? *GGR* 2, pp. 715-784.

[2018] Einstein's conflicting heuristics: The discovery of general relativity. Forthcoming in C. Beisbart, T. Sauer, C. Wüthrich (eds.), *Thinking about space and time: 100 years of applying and interpreting general relativity*, Springer. Draft at <http://www.pitt.edu/~jdnorton/homepage/cv.html#forthcoming>

Pais, Abraham

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.

Reich, Karin

1994 *Die Entwicklung des Tensorkalküls. Vom absoluten Differential-kalkül zur Relativitätstheorie*. Basel: Birkhäuser.

Renn, Jürgen

2007a The summit almost scaled: Max Abraham as a pioneer of relativistic theory of gravitation. In *GGR* 3, pp. 305-330.

2007b The third way to general relativity: Einstein and Mach in context. In *GGR* 4, pp. 21-76.

Renn, Jürgen, and Tilman Sauer

1999 Heuristics and mathematical representation in Einstein's search for a gravitational field equation. In H. Goenner, J. Renn, J. Ritter, T. Sauer (eds.), *The expanding worlds of general relativity* (Boston: Birkhäuser), pp. 87-125.

2007 Pathways out of classical physics: Einstein's double strategy in his search for the gravitational field equation. In *GGR* 1, pp. 113-312.

Renn, Jürgen, and John Stachel

2007 Hilbert's foundation of physics. From a theory of everything to a constituent of general relativity. In *GGR* 4, pp. 857-973.

Sauer, Tilman

2005 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 the history of exact sciences*, 59: 577-590.

2013 Marcel Grossmann's contribution to the general theory of relativity.
arXiv:1312.4068.

Stachel, John

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.

1989a The rigidly rotating body as the 'missing link' in the history of general relativity. In R. Howard and J. Stachel (eds.), *Einstein and the history of General relativity* (Boston: Birkhäuser), 48-62.

1989b Einstein's search for general covariance [from a talk given in July 1980]. In R. Howard and J. Stachel (eds.), *Einstein and the history of General relativity* (Boston: Birkhäuser), 63-100.

1995 History of relativity. In L. Brown, A. Pais, B. Pippard (eds.), *Twentieth century physics* (New York: American Institute of Physics Press), 249-356.

2007 The first two acts. In *GGR* 1, pp. 81-111.

Synge, John Lighton

1921 A system of space-time co-ordinates. *Nature*, 108: 275.

1960 *Relativity: The general theory*. Amsterdam: North Holland.

Trautmann, Andrzej

1962 Conservation laws in general relativity. In L. Witten (ed.), *Gravitation: An introduction to current research* (New York: Wiley), 169-198.

Weinstein, Galina

2018 Why did Einstein reject the November tensor in 1912-1913, only to come back to it in November 1915? *Studies in the history and philosophy of modern physics*, 62: 98-122.