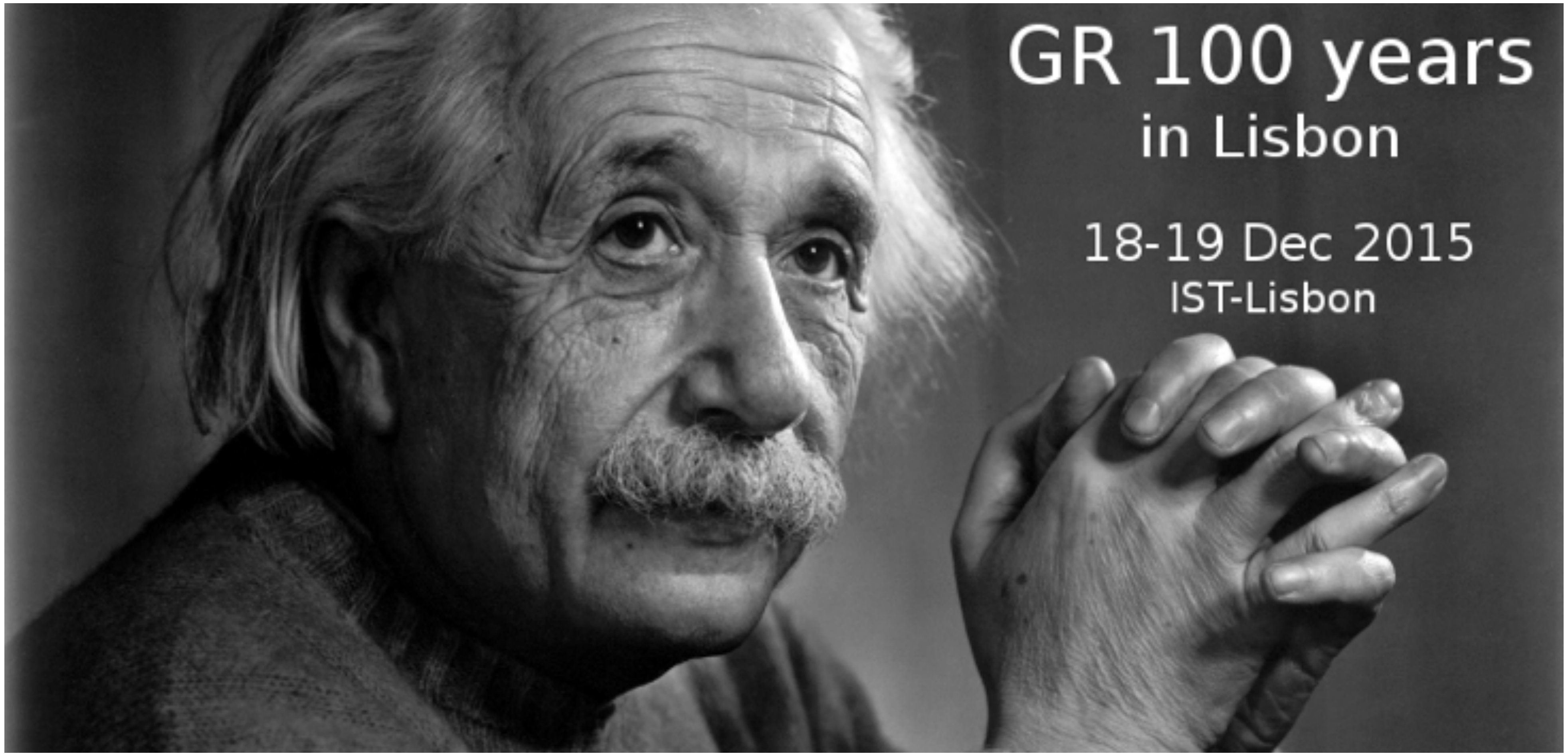


The ‘hole argument’ and the genesis of general relativity

Paulo Crawford
Lisboa, 18 December 2015
Instituto de Astrofísica e Ciências do Espaço UL

Coordinate conditions or restrictions

- In “Autobiographical Notes”, Einstein points out that the importance of the EP in requiring a generalisation of SR was clear to him in 1908 (actually it was in 1907).

  And he adds: “Why were another seven (eight) years required for the construction of the general theory of relativity? The main reason lies in the fact that it is not so easy to free oneself from the idea that co-ordinates must have an immediate physical meaning.”
  (Einstein 1949, p.67).

- It was exactly the resolution of this puzzle that separated Einstein from the final theory particularly from 1913 to 1915.

From the special to the general theory

- Einstein saw his work on general relativity as something quite unique in his life.

- He felt that if he had not created the special theory of relativity, someone else would have done so (P. Langevin?).

- His approach to a new theory of gravitation was entirely his own, carried through with considerable hard work and facing scepticism, if not active opposition, from physicists he respected (Max Planck or Max Abraham).

- He characterised his efforts on special relativity as mere child’s play compared to what was needed to complete general relativity.

Einstein comes to Zurich

- It took some time for Einstein to embrace Minkowski's reformulation of special relativity in terms of a 4-dimensional space-time manifold, a crucial instrument for the further development of a relativistic theory of gravitation. As late as July 1912, Einstein had not adopted the 4-dimensional geometrical approach of Minkowski.

- Apparently, Einstein became acquainted with Minkowski's formalism through M. Laue's book (Laue 1911). All this changed with **Einstein's move to Zurich in August 1912 where he began collaborating with his old mate Marcel Grossmann.** Einstein was then introduced by Grossmann to the 'absolute differential calculus of Ricci and Levi-Civita.

- There we find Einstein recounting the elements of the four-dimensional approach to relativity and Minkowski's electrodynamics, starting with four space-time coordinates $(x, y, z, ict) = (x^1, x^2, x^3, x^4)$ and going on through scalars, four-vectors and six-vectors and their operations.

The Zurich Notebook

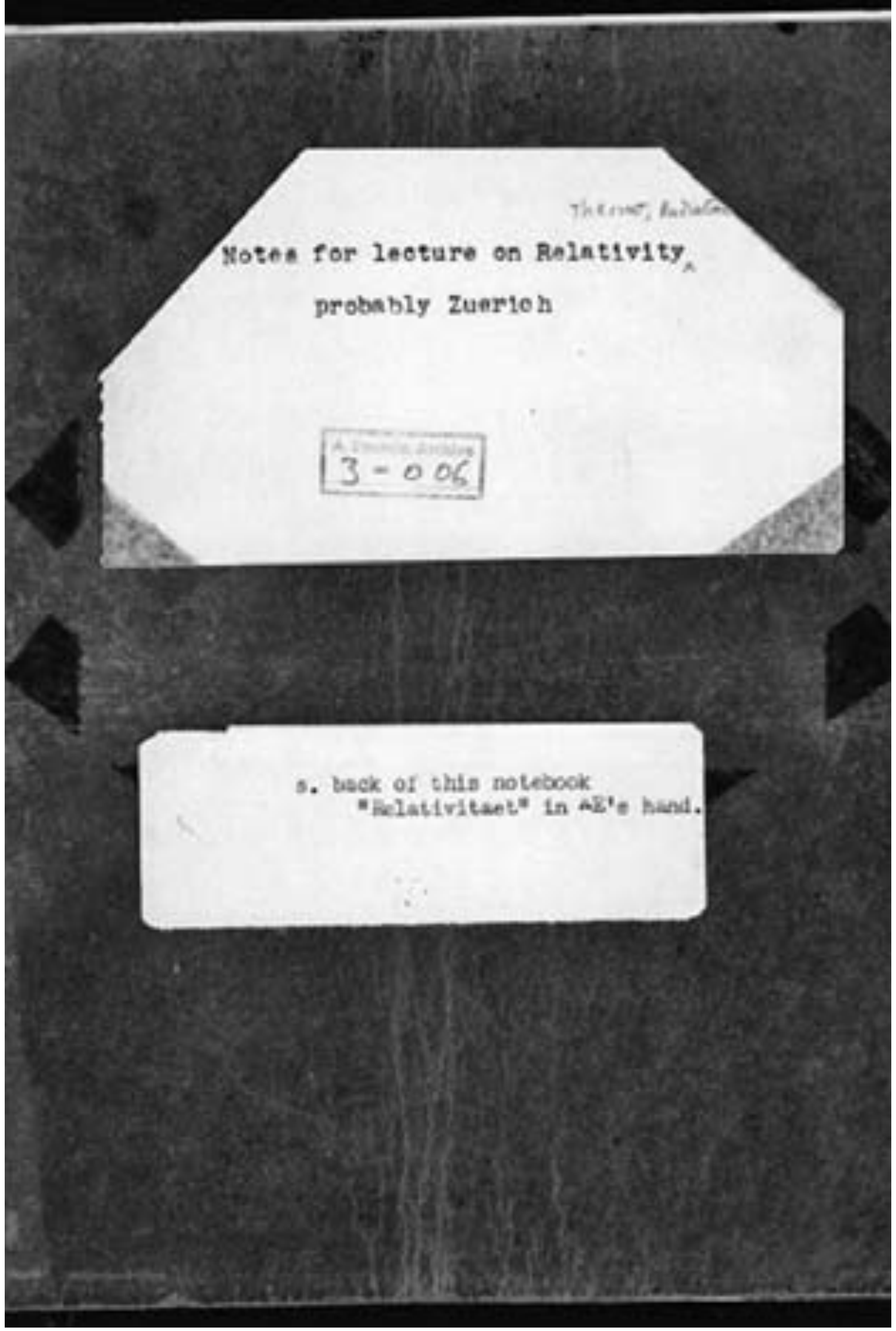
- The task of a reconstruction of Einstein's building of the theory of GR has challenged several historians of science for a long time.
- A major step forward in this venture is due to John Stachel's and John Norton's groundbreaking investigations.
- A very important interpretative tool for understanding Einstein's search for the gravitational field equations is the so-called **Einstein's Zurich Notebook, a document written between Summer 1912 and Spring 1913**, during his time in Zurich (Stachel 1980).
- A little later, John Norton also published a comprehensive reconstruction of Einstein's discovery process (Norton 1984).

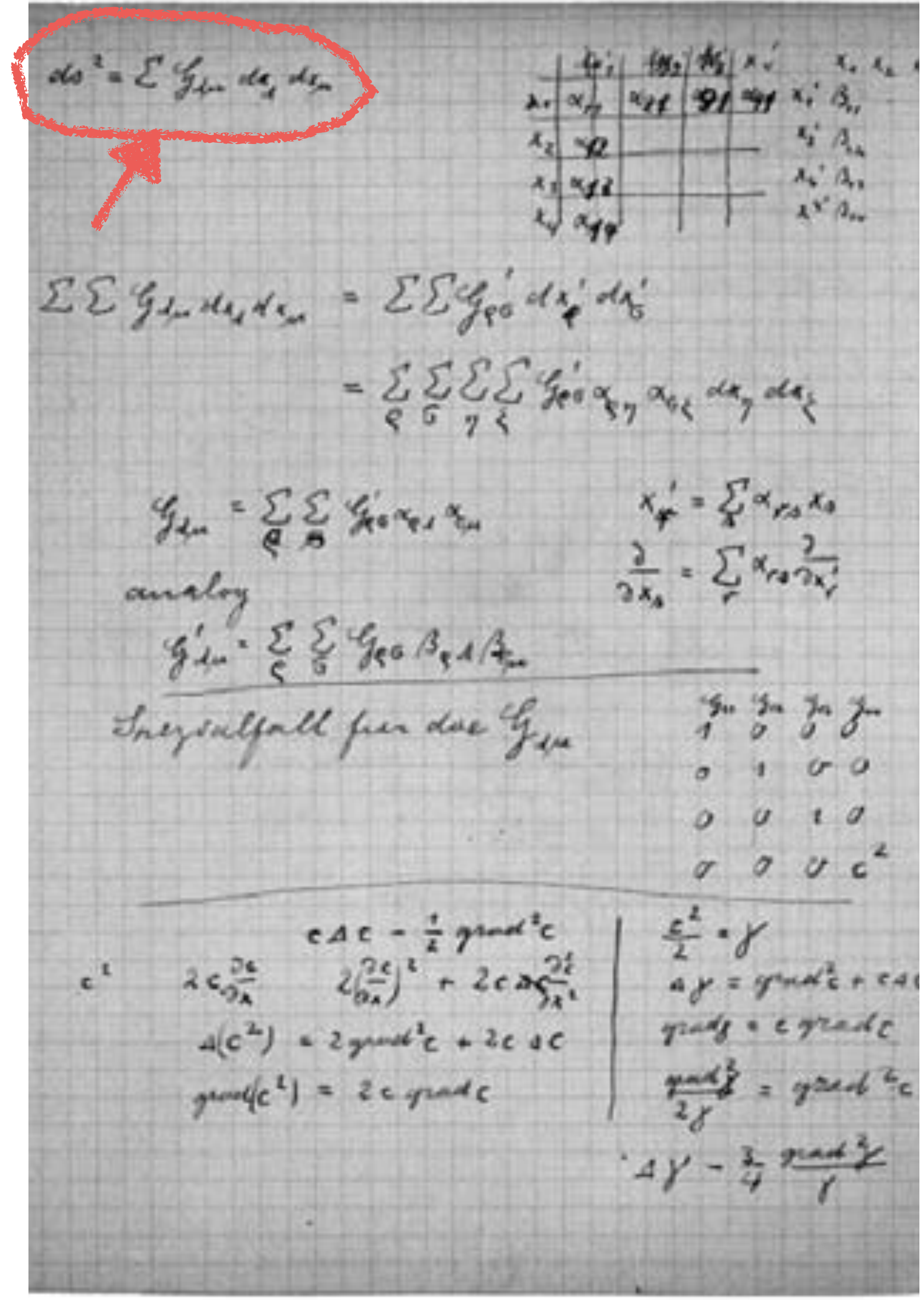
A systematic analysis of the Zurich notebook

- A group of scholars, including John Stachel, John D. Norton and Jürgen Renn and his group at the MPI, undertook a systematic analysis of this notebook and revealed an unexpected result: Einstein had written down, in 1912, an approximation to his final field equations of gravitation, which were derived by him in 1915.

- In 1997, Jürgen Renn and Tilman Sauer have shown that the clarification reached by deciphering Einstein's research notes would have serious consequences for our understanding of the genesis of GR: **the Zurich Notebook shows that in1912-1913 "Einstein had already come within a hair's breadth of the final GTR"**.

- However, he failed to recognise the physical meaning of his mathematical results. In any case, the period between 1913 and November 1915 should not be considered as a period of stagnation. It was, rather, a period during which Einstein arrived at a number of insights that created the prerequisites for his final triumph.





ZN originally comprised 96 pages. If we flip it over, we find a second cover with the word “Relativität” in Einstein’s handwriting.

The turning point in the history of Einstein’s discovery of the gravitational field equations was when he realised the significance of the metric tensor…

The page 39L of ZN

- But let's go back to the ZN pages where Einstein was starting to deal with Minkowski's approach. A central element of Minkowski's geometrical representation of SR was the manifest invariance under linear, orthogonal transformations of the quantity

$$ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2$$

- The development continues for 13 pages, recounting notions in electrodynamics and thermal physics. All of a sudden we stumble on the basic notion of GR, the "line element", written at the top of page 39L, the first exploration of a metric theory. This was possibly the first time Einstein had written down this expression. The coefficients $G_{\mu\nu}$ of what we now know as the "metric tensor" are written with an uppercase *G*.

- Einstein changed within a few pages to a lowercase *g*, which remained his standard notation from then on,

$$ds^2 = \sum_{\alpha,\beta=1}^{4} g_{\alpha\beta} dx^\alpha dx^\beta$$

The Equivalence Principle

- For Einstein, the big project was to find how $g_{\mu\nu}$, the metric tensor, is generated by sources (masses or fields). This would lead to the new gravitational field equation, that is, Einstein's analogue of Newton's inverse square law of gravity. The lower half of the ZN's page is clearly making rudimentary efforts in that direction.

- Let's recall the principal steps taken by Einstein in his path towards a new theory of gravitation. In 1907, Einstein, still at the patent office in Bern, discovered a practical way to deal with gravity and with accelerated observers. He realised then that the effects of acceleration were indistinguishable from the effects of gravity.

- Somehow, Einstein succeeded in unifying all kinds of motion. Uniform motion is indistinguishable from rest and acceleration is no different from being at rest in a gravitational field, at least locally. Einstein saw this "Generalised Principle of Relativity" as guaranteeing the satisfaction of the Equivalence Principle (EP).

- As early as 1907, he had come to consider two possible physical consequences of the EP: the bending of light in a gravitational field and the gravitational red-shift.

Generalised PR versus GC

- That is, the equivalence principle extends the covariance of special relativity beyond Lorentz covariance but not as far as general covariance. Only later did Einstein formulate a “Generalized Principle of Relativity” which would be satisfied if the field equation of the new theory could be shown to possess general covariance. But Einstein’s story, appealing to this mathematical property, is full of ups and downs.

- The turning point in the history of Einstein’s discovery of the gravitational field equations was in the early summer of 1912, when he realised the significance of the metric tensor and the general line element for a generalized theory of gravitation (Pais 1982, section 12b, and Stachel 1980).

- Then Einstein started to study the mathematics of Gaussian surface theory, in collaboration with Grossmann, who discovered for Einstein the existence of the “absolute differential calculus” of Ricci and Levi-Civita (1901) that would enable Einstein to construct a generally covariant theory of gravitation.

Einstein, The Meaning of Relativity, Appendix to 5th ed. (1955)

- The development … of the mathematical theories, essential for the setting up of general relativity, had the result that at first the Riemannian metric [chrono-geometry] was considered the fundamental concept on which the general theory of relativity and thus the avoidance of the inertial system were based.

- Later, however, Levi-Civita rightly pointed out that the element of the theory that makes it possible to avoid the inertial system is rather the infinitesimal displacement field Γ^m_{ik} [the inertio-gravitational field]. The metric or the symmetric tensor field g_{ik} which defines it is only indirectly connected with the avoidance of the inertial system in so far as it determines a displacement field.

The "Entwurf" theory

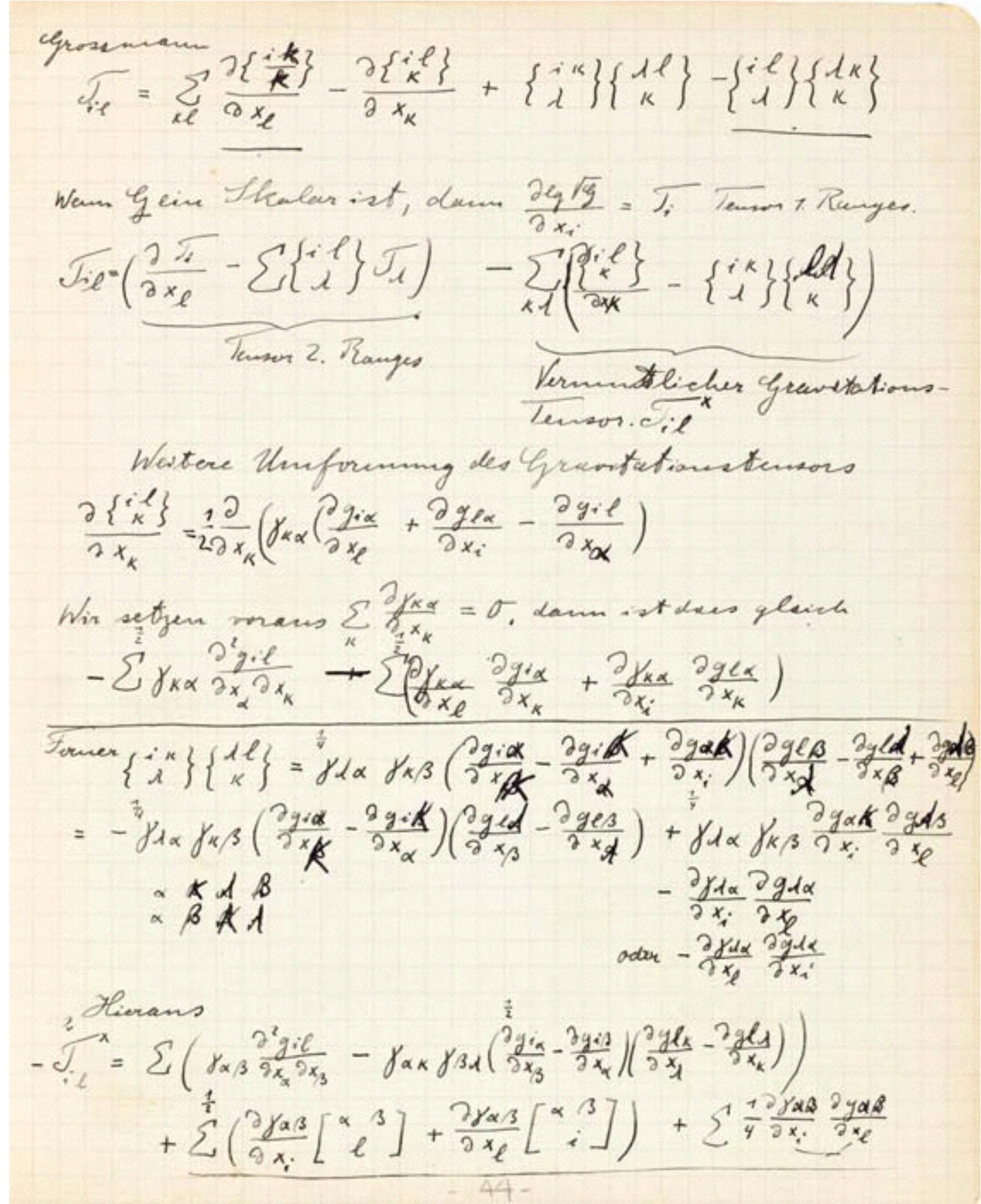
- However, when Einstein & Grossmann published the results of their own research (early 1913), the theory of the resulting paper, known as the "Entwurf" (outline or draft) theory from the title of the paper, failed to comply with the generalized principle of relativity, since **this theory offered a set of gravitational field equations that was not generally covariant**.

- Until the Autumn of 1915, Einstein continued to elaborate on and improve the "Entwurf" theory and explored many of its consequences. Already in 1913, Einstein and his friend Michele Besso had found that **"Entwurf" equations did not account for the anomalous advance of the perihelion of Mercury**, something that Einstein hoped to explain.

- Although Einstein knew the failure of its "Entwurf" theory to resolve the Mercury anomaly, he continued to hold on to this theory in spite of everything.

At the top of this particular page, 22R, Einstein wrote down the generally covariant Ricci tensor T_{il} under the heading "Grossmann," "If G is a scalar" (unimodular transformations) Einstein noted, the first half of T_{il} transforms as a tensor. If the first half transforms as a tensor under unimodular transformations, the second half must too, since their sum transforms as a tensor under arbitrary transformations.

Underneath the second half of the second equation, Einstein wrote, "probable gravitation tensor"



Grossmann

$$T_{il} = \sum_{\kappa l} \frac{\partial \left\{ {i\kappa \atop \kappa} \right\}}{\partial x_l} - \frac{\partial \left\{ {il \atop \kappa} \right\}}{\partial x_\kappa} + \left\{ {i\kappa \atop \lambda} \right\}\left\{ {\lambda l \atop \kappa} \right\} - \left\{ {il \atop \lambda} \right\}\left\{ {\lambda\kappa \atop \kappa} \right\}$$

Wenn G ein Skalar ist, dann $\frac{\partial \lg \sqrt{G}}{\partial x_i} = T_i$ Tensor 1. Ranges.

$$T_{il} = \left(\frac{\partial T_i}{\partial x_l} - \sum \left\{ {il \atop \lambda} \right\} T_\lambda \right) - \sum_{\kappa\lambda} \left(\frac{\partial \left\{ {il \atop \kappa} \right\}}{\partial x_\kappa} - \left\{ {i\kappa \atop \lambda} \right\}\left\{ {\lambda l \atop \kappa} \right\} \right)$$

Tensor 2. Ranges

Vermutlicher Gravitationstensor T_{il}^{x}

Weitere Umformung des Gravitationstensors

Wir setzen voraus $\sum_\kappa \frac{\partial \gamma_{\kappa\alpha}}{\partial x_\kappa} = 0$, dann ist dies gleich

Ferner

oder

Hieraus

-44-

$$G_{im} = \{il, lm\} = R_{im} + S_{im}$$

$$R_{im} = -\frac{\partial \left\{ \begin{matrix} im \\ l \end{matrix} \right\}}{\partial x_l} + \sum_{\varrho} \left\{ \begin{matrix} il \\ \rho \end{matrix} \right\} \left\{ \begin{matrix} \rho m \\ l \end{matrix} \right\}$$

$$S_{im} = \frac{\partial \left\{ \begin{matrix} il \\ l \end{matrix} \right\}}{\partial x_m} - \left\{ \begin{matrix} im \\ \rho \end{matrix} \right\} \left\{ \begin{matrix} \rho l \\ l \end{matrix} \right\}.$$

- In his paper of 4 November 1915, Einstein split up the Ricci tensor G_{im}, which encodes spacetime curvature. (The indices *i, l, m,* and ρ take on the values 1 through 4.) As his new field equations, he proposed $R_{im} = -\kappa T_{im}$, where T_{im} is the energy–momentum tensor for matter and κ is proportional to Newton's gravitational constant. These are equations of broad but not yet general covariance.

$$G_{im} = -\varkappa\left(T_{im} - \frac{1}{2} g_{im} T\right),$$

$$\sum_{\varrho\sigma} g^{\varrho\sigma} T_{\varrho\sigma} = \sum_{\sigma} T_{\sigma}^{\sigma} = T$$

- **The Einstein field equations** first appear in Einstein's 25 November 1915 paper. Here, G_{im} is the Ricci tensor; g_{im}, the metric tensor; and T_{im}, the energy–momentum tensor for matter. Three weeks earlier Einstein had proposed the field equations $R_{im} = -\kappa T_{im}$ which retain their form under unimodular transfs.

The “hole” argument…

- By November 1913, Einstein had developed the “hole” argument against general covariance. He wrote to Ludwig Hopf on 2 November:
    - “I am now very content with the gravitation theory. The fact that the gravitational equations are not generally covariant, which a short time ago still disturbed me so much, has proved to be unavoidable; **it is easily proved that a theory with generally covariant equations cannot exist if one demands that the field be mathematically completely determined by matter**.”
- The proof alluded to in the letter, is the infamous ‘hole’ argument first published in the addendum to the “Entwurf” paper (Einstein and Grossmann 1913), signed by Einstein alone and not published in the original printing of the paper.
- Einstein repeats just about the same argument in two subsequent papers in 1914 and in several letters to friends and colleagues, and the core of his reasoning was complete by November 1913.

… in action

- Let there be a region of space-time *H* (the "hole"), an open subspace of a manifold *M* devoid of matter and energy, and a set of generally covariant field equations valid for the entire space-time manifold *M*, both inside and outside *H*. Given a coordinate system of the manifold, *K*, what happens physically in *H* is then completely determined by the solutions of the field equations, $g_{\mu\nu}$. The totality of these functions will be represented by *G(x)*.

- Given a 2nd coordinate system *K'* that coincides with *K* everywhere outside and on the boundary of *H*, and diverges from *K* within *H* but in such a way that the metric components $g'_{\mu\nu}$ referred to *K'*, like $g_{\mu\nu}$ and their derivatives, are everywhere continuous. The totality of $g'_{\mu\nu}$ expressed in terms of the new coordinates x'_ν will also be represented by *G'(x')*. **Note that *G'(x')* and *G(x)* describe the same gravitational field. That is, they are two different mathematical representations of the same physical field.**

- However, if we replace the coordinates x'_ν by the coordinates x_ν in the functions $g'_{\mu\nu}$ and represent them by *G'(x)* then *G'(x)* also describes a gravitational field with respect to *K*, **which is different from the original gravitational field within the "hole" *H*.**

- However, the two different solutions *G'(x)* and *G(x)*, written in the same coordinate system, correspond to the same "reality" (e.g. the same sources and same boundary conditions).

An equivalent class of solutions

- **In particular, any set of generally covariant field equations that has *G(x)* as a solution in some empty region of space-time will also have *G'(x)* as a solution in that region. *G(x)* and *G'(x)*, together with all other mathematically distinct metric tensor fields that can be transformed into each other by being dragged along with an (active) diffeomorphism, form a equivalence class of solutions. But this equivalence class of mathematical distinct metric tensor fields corresponds to one physical solution to the field equations, that is, to one gravitational field.**

Einstein conclusion…

- In summary, because generally covariant field equations admit non-equivalent solutions for events within *H*, such equations are not acceptable as an appropriate physical theory of gravitation. This is the “hole” argument against general covariance of the field equations.
- So, if we require that the course of events in the gravitational field be determined by the laws to be set up, we must adopt a theory with restricted covariance properties.
- One could think that Einstein’s argument was a sort of excuse to accommodate his “Entwurf” theory with limited covariant properties. But, indeed, at the end of the day, his argument was much deeper than that.

In coordinate-free language

- In trying to explain the line of reasoning behind Einstein's arguments, we kept as close as possible to the mathematical language and methods of his time.

- The modern terminology of differential geometry (coordinate-free language) which distinguishes between coordinate transformations and (active) diffeomorphisms, was not available to Einstein, one may easily clarify these arguments.

- Assume the gravitational field equations are generally covariant. Consider a solution of these equations in which the gravitational field is *g* and there is a region *H* of the universe without matter: the "hole". Assume that inside *H* there is a point *A* where *g* is flat and a point *B* where *g* is not flat.

in coordinate-free …

- A smooth map $\phi : M \rightarrow M$ which reduces to the identity outside H, and such that $\phi(A) = B$, and let $\tilde{g} = \phi^{*} g$ be the pull-back of g under ϕ.

- The two fields $\tilde{g}$ and g have the same past and are both solutions of the field equations but have different properties at the point A. Therefore, the field equations do not determine the physics at the space-time point A. That is, they are not deterministic. However, we know that (classical) gravitational physics is deterministic.

- So, one must pick one of the following:
  (i) the field equations must not be generally covariant;
  (ii) there is no meaning in talking about the physical space-time point A.

  The correct physical conclusion is the second one, that there is no meaning in referring to "the event A" without further specification.

The point-coincidence argument

- By late 1915, after having returned to generally covariant field equations, Einstein introduces the point-coincidence argument: a coordinatization of the manifold is itself not sufficient to determine an individuation of the points (events) of the manifold.

- Einstein then argues: the events of the space-time are implicitly defined and thus individuated only as points of intersection or coincidence of world-lines.

- In regions where no matter is present, the points of a manifold are physically differentiated only by the properties that they inherit from $g_{\mu\nu}(x_\nu)$. So, it is impossible to have two different sets of values of the functions $g_{\mu\nu}(x_\nu)$ assigned to one and the same event of the space-time manifold.

- Therefore, in regions where no matter is present, the points of a manifold are physically differentiated only by the properties that they inherit from the metric field.

At the end of the day: Einstein Got Right

- **The Hole Argument: Points of space-time have no inherent physical properties. They inherit all of these properties from the space-time Structures, including all fields.**

- **Conclusion: No first order space-time structures or fields, no space-time!**

“Relativity and the Problem of Space” (1952)

- On the basis of the general theory of relativity ... space as opposed to ‘what fills space’ ... has no separate existence.

- If we imagine the gravitational field to be removed, there does not remain a space of the type [of SR], but absolutely nothing, not even a ‘topological space’.

SR versus GR: GR is a Background-Independent Theory

- **The contrast between general relativity and all previous theories:**

  Background-dependent theory: Fixed and given space-time stage, on which the drama of physics unfolds, like in Minkowski space-time of SR.

  Background-independent theory: No actors, no stage, no anything. Like GR.

- **Einstein put it this way:**

  “Space-time does not claim existence on its own, but only as a structural quality of the field.”

“That’s all folks…”