Contents

1

Hilbert on General Covariance and Causality

Katherine Brading, Thomas Ryckman

Department of Philosophy
University of Notre Dame
Notre Dame, Indiana 46556, USA
Department of Philosophy
Stanford University
Stanford, California 94305, USA

1.1 Introduction

Hilbert's work on generally covariant physics in 1915 led him to diagnose a tension between general covariance and causality, and to seek its resolution. In an earlier paper in this series (Brading and Ryckman, 2009), we presented Hilbert's reconsideration of the status of causality in the light of general covariance as it unfolds in Hilbert's First and Second Communications on the Foundations of Physics (Hilbert 1915a and 1915b; Hilbert 1917). In our paper, we claim that Hilbert's "causality problem" and the resolution he offers differ from the (in)famous "hole argument" and its resolution, due to Einstein.[1] The questions and feedback that we continue to receive when discussing this research have made it clear that a supplementary note would be valuable, giving further details of the differences between Hilbert's "causality problem" and that of Einstein, and also making explicit the relationship between Hilbert's proposed resolution and how we think about general covariance and causality in General Relativity today. The purpose of this paper is to address these two points.

We begin with a review of Einstein's "causality problem" and the solutions that he offers (sections (2) and (3)). We then discuss the evolution of Hilbert's "causality problem" through the First and Second Communications (sections (4) and (5)), before addressing (in section (6)) the resolution that he offers in the Second Communication, including its relationship to how we think about these things today. Hilbert's "causality problem" has both a mathematical and an epistemological face, and while the mathematical problem and its resolution are standard fare in General Relativity today, his epistemological discussions remain largely unknown. We end by comparing Einstein's "causality problem" with that of Hilbert, and here make the case that Hilbert was never a victim of Einstein's "hole argument" (see section (7)).

1.2 Einstein's "causality problem"

When Einstein was lecturing in Göttingen during the summer of 1915, he still believed that generally covariant field equations were not to be had. He had two arguments for this, one is the so-called "hole argument"[2] and the other has to do with energy conservation.

Einstein's hole argument has been the subject of detailed consideration in the history and philosophy of general relativity literature.[3] In the argument, Einstein considers a region of spacetime in which there are no matter fields present (the "hole"). He then shows that, in a generally covariant theory, no amount of data about the values of the matter and gravitational fields (or the metric) *outside* the hole is sufficient such that, when combined with the field equations, the values of the gravitational field *inside* the hole are uniquely determined. This was unacceptable to Einstein: motivated by what he would later refer to as "Mach's principle", he was searching for a theory in which the matter fields plus the field equations would uniquely determine the metric.[4] Thus, the hole argument can be understood as posing a kind of "causality problem" for any generally covariant theory.

Einstein formulated the hole argument as a *post hoc* justification for his failure to find generally covariant field equations. His thinking about energy conservation had led him to conclude that we need to restrict the covariance class of our theory. The conclusion of the hole argument was that no generally covariant theory will be physically acceptable. Using energy-momentum conservation to arrive at four non generally covariant conditions, Einstein restricted the covariance properties of his field equations and thereby restored "causality". This seemed satisfactory: Einstein had an argument for why no generally covariant theory could be physically possible (the hole argument), and he had conditions limiting the covariance class of his field equations that were motivated by physical considerations (energy conservation).[5]

While we do not know for sure what Einstein said in his 1915 Göttingen lectures, they were surely the catalyst for Hilbert's First and Second Communucations (1915a and b; 1917). Hilbert is explicit in attributing the idea of generally covariant physics to Einstein, and it is reasonable to infer that two key features of Einstein's work on gravitational theory (concerning a conflict between general covariance and causality, and the use of conservation of energy to resolve the problem) were included in his lectures. It is also reasonable to infer that they were picked up by Hilbert, then to appear in the December Proofs of his First Communication (Hilbert, 1915a).[6] However, as we will see, Hilbert had a different understanding from Einstein of the problem of causality raised by general covariance, and also, therefore, used the considerations about energy conservation rather differently.

The appeal to energy conservation to address problems of causality was consigned to the scrap heap before the year was out. On the 2nd of December 1915, Einstein published the familiar Einstein Field Equations of General Relativity which are, of course, generally covariant; energy conservation no longer restricts the covariance properties of the field equations.

Einstein later "solved" the "problem of causality" posed in his hole argument via his "point coincidence argument". As we hope to make clear in what follows, it is this solution that helps to pinpoint Einstein's own "causality problem" and to make vivid the differences between this and Hilbert's "causality problem".

1.3 Einstein's "point coincidence argument"

As is now well-documented, Einstein extricated himself from the conclusion of the hole argument by means of his so-called "point coincidence argument".[7] However, this resolution of the difficulty was only obliquely expressed in print in the canonical presentation of the new theory published on 11 May 1916.[8] A passage in section 3 of that paper was first identified as the "point-coincidence argument" by Stachel (1989), and it presents a puzzle: it begins with a declaration that the requirement of general covariance removes "the last remnant of physical objectivity from space and time", but in support of this apparently ontological conclusion offers what seems to be a suspiciously epistemological argument. The first premise states that all of our spacetime observations and measurements ultimately amount to a determination of spacetime coincidences. As illustrative examples of this premise, Einstein cites the meeting of two or more material points, and even the coincidence between a pointer and the marks on a dial. The second premise concerns the role of coordinate systems, and is the suggestion that the introduction of a coordinate system merely facilitates the description of the totality of such coincidences. But, since two coincident point events (described by identical coordinates in a given coordinate system) will remain coincident in a new coordinate system (arrived at from the first by an arbitrary coordinate transformation), we have no reason to prefer one system of coordinates to any other. Thus, we arrive at the requirement of general covariance.

The verificationist flavor of this argument was widely hailed by Machians (such as Phillip Frank) and positivists of various stripes. Following Stachel (and in the light of much later Einstein texts), we suppose a more charitable gloss can be given the point coincidence argument. What Einstein should have said is that the discordant conclusion stemming from the hole argument (that generally covariant field equations lead to indeterminism) no longer goes through once it has been recognized that systems of spacetime coordinates have no metrical or other physical significance, but serve as essentially arbitrary labels for spacetime points. Thus the supposedly distinct solutions generated by given matter sources can now be recognized as being merely different mathematical descriptions of the same physical state of affairs. It is this realization that enables Einstein to evade the causality problem posed by the hole argument.

1.4 Hilbert's "causality problem", 1915

Throughout the First Communication, both proofs and published version,[9] and the Second Communication,[10] Hilbert never wavers from his commitment to general covariance. As we have argued (2008, 2009), Hilbert saw profound epistemological significance in general covariance, and sought to explore the consequences of adopting it as an axiom of fundamental physics.

Already in the proofs, Hilbert makes clear the implications of generally covariant physics for considerations of causality, as he understood them. He claims that *any* generally covariant theory will face a problem of mathematical underdetermination, stating explicitly in his Theorem 1 that for a system of n Euler-Lagrange differential equations in n variables obtained from a generally covariant action integral, there will be only $n - 4$ equations for the n variables. Hilbert then argues as follows (1915a, p. 4):

> Therefore, if we want to preserve the determinate character of the fundamental equations of physics according to Cauchy's theory of differential equations, the requirement of four additional non-invariant equations supplementing [the field equations] is essential.

In this extract we see two things clearly stated: the first is Hilbert's "causality problem", and the second is his proposal for its resolution. Hilbert explicitly states that the causality problem associated with general covariance concerns Cauchy determination. The question is whether a generally covariant theory admits of a well-posed Cauchy problem, and Theorem 1 suggests that it does not. In the context of spacetime theory and a system of second-order partial differential equations on that spacetime, a well-posed Cauchy problem requires that the initial data assignments to the unknown field functions and their first (time) derivatives on a spacelike hypersurface determine the second time derivatives of the given field quantities, and thereby unique solutions off the hypersurface (for appropriate regions). Hilbert's Theorem 1 pinpoints failure of Cauchy determination as a consequence of general covariance. This is Hilbert's "problem of causality" in 1915.

Distinct from this diagnosis of a causality problem is the solution Hilbert offers in the proofs: the addition of four further equations. Hilbert followed Einstein in using energy conservation to provide these additional equations, but his purpose was somewhat different. For Hilbert, general covariance retains its axiomatic status, and the field equations remain generally covariant; but, for the sole purpose of meeting the mathematical requirement of Cauchy determination within this generally covariant structure, additional conditions (deriving from energy conservation) are imposed.[11]

With the publication of the generally covariant Einstein Field Equations (for which, of course, Einstein no longer uses energy conservation to restrict the covariance properties of the field equations) Hilbert had to abandon this solution to his causality problem. And, indeed, when we look at the published version of the First Communication, this whole application of energy conservation has gone. But Theorem 1 is still there in the published version: so the causality problem is still there, but now Hilbert has no solution for it.

1.5 Hilbert's "causality problem", 1917

Hilbert's Second Communication (1917) includes a new treatment of the causality problem originally posed in the First Communication, embedding and developing the original mathematical problem in an explicit epistemological context. As we discuss in what follows, Hilbert presents his causality problem as an apparent conflict between the *axiom* of general covariance and our *experience* of the world as causally ordered and causally determinate. This makes explicit the *epistemological* aspect of the problem, which in the First Communication had appeared under a predominantly *mathematical* guise. For Hilbert the deep problem is the epistemological problem.

It is clear that from the outset Hilbert saw deep epistemological significance in general covariance. He saw the adoption of general covariance as an important step towards removing the contributions of human subjectivity from the conceptual structure of physics; specifically, by making it independent of the way in which world-points are designated (through coordinates). Thus, immediately following the statement of his axiom of general covariance in his First Communication, he writes that this axiom is: "the simplest mathematical expression of the demand that the inter-linking of the potentials $g_{\mu\nu}$, q_s is by itself entirely independent of the way one chooses to label the world's points by means of world parameters." This statement appears in both the proofs and the published version (Hilbert, 1915a, p. 990; 1915b, p. 1004). The point is repeated again in Hilbert (1919-1920, p. 49), and in 1921 Hilbert describes the move to general covariance as an emancipation from "the *subjective* moments of human *intuition* with respect to space and time" and "a radical elimination of *anthropomorphic* slag".[12] This understanding of general covariance leads to a corresponding epistemological aspect of the "causality problem", because the requirement of causality appears to be inconsistent with the emancipation achieved by general covariance.

As in the First Communication, the tension between general covariance and causality is given a precise mathematical characterization: Hilbert points out that general covariance leads to a mathematical problem with respect to Cauchy determination. New to the mathematical discussion is Hilbert's observation that arbitrary point transformations (diffeomorphisms) do not respect the relation of cause and effect among world points lying on the same timelike curve: they allow transformations that reverse the temporal order of "cause" and "effect" or place them in spacelike relation. Thus, the mathematical problem now concerns both causal ordering as well as Cauchy determination.

Hilbert's position is that we need to reconcile the general covariance of the conceptual structure of physics with our experience of the world as causal (both causally ordered, and causally determinate in the sense of Cauchy determination). That is, we must be able to recover the world *as we subjectively experience it* from the generally covariant structure of *objective* physics. This is Hilbert's "causality problem" in 1917.

1.6 Hilbert's 1917 resolution of his "causality problem"

To address his "causality problem", Hilbert begins by introducing the notion of "proper coordinate systems": by definition, transformations among such coordinate systems respect the distinction between spacelike and timelike coordinate axes, preserve the temporal ordering of cause and effect, and ensure Cauchy determination. If we restrict ourselves to the use of proper coordinate systems, we will extract causally determinate structures appropriate for expressing our *experience* of the world, from the generally covariant conceptual structure.

The mathematical aspect of Hilbert's "causality problem", of achieving Cauchy determination, is solved via appeal to "proper coordinate systems". As Stachel (1992) states, Hilbert was the first to discuss the Cauchy problem for the Einstein equations. The solution Hilbert offers in his Second Communication (Hilbert, 1917) makes explicit use of Gaussian coordinates, which put the remaining 10 equations into Cauchy normal form. While Stachel also points out that many of the subtleties and difficulties posed by general covariance escaped Hilbert in 1917, the problem that Hilbert posed and the general method for its solution continue to be a part of standard practice in General Relativity.[13]

The epistemological face of Hilbert's "causality problem" is less familiar, and it is to this that Hilbert turns his attention having addressed the mathematical issue. We noted above that, for Hilbert, general covariance has profound epistemological significance as *the* criterion of physical objectivity, enabling us to eliminate the "anthropomorphic slag" associated with preferred choices of coordinate systems. What Hilbert is looking for is an account of the relationship between the physically objective world as expressed by generally covariant field equations, and our subjective experience of the world as causally ordered and causally determinate. According to our reconstruction (see Brading and Ryckman, 2008, 2009), the approach he takes is to consider the status of the statements that we make reporting our experiences of the world as causally ordered and determinate. In the Second Communication (p. 1024) Hilbert explicitly insists that physically meaningful propositions in physics *must* have a generally covariant formulation. However, this condition is not by itself sufficient for physical meaningfulness. In the light of "the principle of causality" (p. 1024), a second condition must be added, according to which when such a proposition is expressed with respect to a "proper" coordinate system, the truth value of that statement is *uniquely* determined by an appropriate spacelike hypersurface. In this way, Hilbert unites the objectivity achieved by general covariance with the subjectivity of our experience of the world as causal, dissolving the appearance of conflict.

In sum, Hilbert's contribution to the Cauchy problem arises from the mathematical face of his causality problem. The problem is a genuine technical challenge facing generally covariant physics, and Hilbert's proposal for a solution remains familiar in the practice of contemporary General Relativity. For Hilbert, this mathematical problem is embedded in a deeper epistemological problem: that of reconciling our experience of a causally ordered and univocally determinate world with the four-dimensional structure of generally covariant physics. This problem has not gone away, and though Hilbert's proposal may be unfamiliar in contemporary discussions, we submit that it may be worth revisiting in earnest.[14]

1.7 Hilbert and Einstein compared

Einstein abandoned general covariance for over two years, justifying this in part by appeal to his hole argument, and the "causality problem" that this argument poses. Then, with his generally covariant field equations in hand, Einstein restored causality by means of his "point coincidence argument", which is best understood as asserting that systems of spacetime coordinates have no metrical or other physical significance, but serve merely as arbitrary labels for spacetime points.

We maintain that Hilbert's "causality problem" was never that posed by Einstein's hole argument,[15] despite superficial similarities in some of Hilbert's discussions (see, for example, Hilbert 1916). In our opinion, the clearest way to see this is to recognize that Einstein's solution to his own problem, concerning the status of spacetime coordinates, is something that Hilbert emphasized from the outset, and which furthermore does nothing to address the problem that Hilbert was addressing.

On the first point, it is significant that in 1915 Hilbert termed the labels for points in his four-dimensional spacetime "world parameters". This terminology highlights the analogy with the arbitrary character of curve parameterizations in the calculus of variations. As Howard and Norton (1993) point out, the Göttingen mathematical community was thoroughly familiar with the use of arbitrary coordinates in the work of Lagrange, Gauss, and Riemann, and it seems highly unlikely, to say the least, that Hilbert was confused about this issue. This evidence is circumstantial, but compelling, and strongly supported by the statement of the problem that Hilbert gives in his Theorem 1 (Hilbert 1915a and 1915b), where Hilbert makes clear that his "causality problem" arises due to *dependencies* among the field equations. In other words, Hilbert *started* from a position in which coordinates are merely arbitrary labels for spacetime points, and thus his "causality problem" cannot be that which Einstein expresses in the hole argument; rather, it arises when we seek to recover the mathematical property of Cauchy determination.

That Einstein's and Hilbert's problems with causality differ is further emphasized by the second point noted above: Einstein's solution to *his* causality problem (via the point coincidence argument) does nothing to address Hilbert's. Recall that, essentially, Einstein's solution in his point coincidence argument posits a four-dimensional arrangement of events on a spacetime manifold. This four-dimensional arrangement does not come causally ordered, nor does it tell us how to recover univocal determination from our field equations. Both of these can be achieved by picking an appropriate co-ordinate system, and this is exactly what Hilbert points out that we need to do: in order to recover – via the field equations – our experience of the four-dimensional world of point events as univocally causally ordered, we need to make use of "proper" coordinate systems.[16] This does *not* imply that Hilbert attributed physical significance to coordinate systems. On the contrary, as we have emphasized, and as Stachel (1992, p. 412) remarks, Hilbert assumed from the outset that "all solutions related by a coordinate transformation must be regarded as physically equivalent."

1.8 Conclusion

Hilbert's "causality problem" is not that which Einstein expressed in his "hole argument" and resolved with his "point coincidence argument". Rather, Hilbert *begins* from the key premise of the "point coincidence argument", that coordinates are arbitrary labels, and seeks to solve the mathematical and epistemological problems that then arise. The mathematical problem, and Hilbert's proposed solution, remain standard fare in General Relativity today: the imposition of coordinate conditions for arriving at a solution to the Einstein Field Equations. The epistemological problem, of how we reconcile general covariance with our experience of the world as spatiotemporally and causally ordered and causally determinate, remains a puzzle admitting of no uncontroversial solution. Hilbert's proposal piggy-backs on his solution to the mathematical problem and, in our opinion, its implications for the interpretation of General Relativity and associated epistemological issues deserve further exploration.[17]

References

Barbour, J. B, 2005 . *Absolute or relative motion?* Vol. 2: *The deep structure of general relativity.* Oxford University Press.

Brading, K. A., and Ryckman, T. A, 2009 . Brading, K. A., and Ryckman, T. A., 2008, Hilbert's 'Foundations of Physics': Gravitation and electromagnetism within the axiomatic method', Studies in History and Philosophy of Modern Physics 39, 102-153.

– – 2009. Hilbert's Axiomatic Method and his "Foundations of Physics": Reconciling Causality with the Axiom of General Invariance, Einstein Studies...

Corry, L., J. Renn, and J. Stachel (1997). Belated decision in the Hilbert-Einstein priority dispute, Science 278, 1270-73.

Einstein, A., and M. Grossmann, 1914. 'Entwurf einer verallgemeinerten Relativitätstheorie und einer Theorie der Gravitation", *Zeitschrift für Mathematik und Physik* 62, 225-259.

Einstein, A., 1916. Die Grundlage der allgemeinen Relativitätstheorie, Annalen der Physik 49, 769-822. Reprinted in Einstein (1996), 284-337. Translated by W. Parrett and G. Jeffrey as 'The Foundation of the General Theory of Relativity', in H. Lorentz et al. (1923), 109-65.

Hilbert, D., 1915a. Die Grundlagen der Physik (Erste Mitteilung). Annotated "Erste Korrektur meiner erste Note", printer's stamp date "6 Dez. 1915". 13 pages with omissions. Translation as The foundations of physics (Proofs of first communication), in Renn and Schemmel (2007), 989-1001.

Hilbert, D., 1915b. 'Die Grundlagen der Physik (Erste Mitteilung)', Nachrichten. Königliche Gesellschaft der Wissenschaften zu Göttingen. Mathematische-Physikalische Klasse, 395-407. Translation as 'The foundations of physics (first communication)', in Renn and Schemmel (2007), 1003-15.

Hilbert, D., 1916. "Das Kausalitätsprinzip in der Physik". Typescript of lectures, dated 21 and 28 November 1916 by Tilman Sauer. Bibliothek des Mathematisches Institut, Universität Göttingen. 17 pages.

Hilbert, D., 1917. 'Die Grundlagen der Physik (Zweite Mitteilung)', Nachrichten. Königliche Gesellschaft der Wissenschaften zu Göttingen. Mathematische-Phyikalische Klasse, 53-76. Translation as 'The foundations of physics (second communication)', in Renn and Schemmel (2007), 1017-38.

Howard, D. and J. Norton (1993). Out of the labyrinth? Einstein, Hertz, and the Göttingen answer to the hole argument, in J. Earman, M. Janssen, and J. Norton (eds.), The attraction of gravitation: New studies in the history of general relativity (Einstein studies v. 5). Boston/Basel/Berlin: Birkhäuser, 30-62.

Janssen, M. and J. Renn (2007). Untying the knot: How Einstein found his way back to field equations discarded in the Zurich notebook, in M. Janssen, J.D. Norton, J. Renn, T. Sauer, and J. Stachel (eds.), The genesis of general relativity, v. 2. Einstein's Zurich notebook: Commentary and essays. Dordrecht: Springer, 839-925.

Majer, U., 1995. 'Geometry, intuition and experience: from Kant to Husserl. *Erkenntnis*, 42, 261-285.

Norton, J., 1984. 'How Einstein found his field equations: 1912-1915', *Hist. Stud. Phys. Sci.* **14**, pp. 253-316.

Norton, J., 1993. 'General covariance and the foundations of general relativity: eight decades of dispute', *Rep. Prog. Phys.* **56**, pp. 791-858.

Renn, J. and M. Schemmel (eds.), 2007. The genesis of general relativity. vol. 4. Gravitation in the twilight of classical physics: The promise of mathematics. Dordrecht: Springer.

Renn, J. and J. Stachel (1999). Hilbert's foundation of physics: From a theory of everything to a constituent of general relativity, Berlin: Max-Planck-Institut für Wissenschaftsgeschichte, Preprint 118. Reprinted in Renn and Schemmel (eds.), 2007, 858-973.

Ryckman, T., 2005. *The Reign of Relativity*, Oxford University Press.

Sauer, T. (1999). The relativity of discovery: Hilbert's first note on the foundations of physics, Archive for History of Exact Sciences 53, 529-75.

Sauer, T. (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 History of Exact Sciences 59, 577-90.

Stachel, J., 1989. "Einstein's Search for General Covariance, 1912-1915." in D. Howard and J. Stachel (eds.), Einstein and the History of General Relativity. Basel: Birkhäuser, 63-100. Based on a paper circulating privately since 1980.

Stachel, J., 1992. The Cauchy problem in general relativity - The early years, in J. Eisenstaedt and A.J. Kox (eds.), Studies in the history of general relativity. (Einstein studies, v. 3) Boston/Basel/Berlin: Birkhäuser, 407-418.

Notes

[1] See also Brading and Ryckman, 2008, section 7.

[2] See Stachel, 1989.

[3] In addition to Stachel, 1989, see also Norton, 1984, pp. 286-91, and 1993, sections 1-3; Ryckman, 2005, section 2.2.2; and references therein.

[4] For more on Einstein's (mis)appropriation of Mach's principle, see Barbour, 2005.

[5] Einstein and Grossmann, 1914. See Janssen and Renn, 2007, for the details of the story.

[6] The "December Proofs" (Hilbert, 1915a) were recently discovered (see Corry, Renn and Stachel, 1997) and contain significant differences from the published version (Hilbert, 1915b). For discussion of these differences, and differing opinions on their significance, see Renn and Stachel, 1999, and Sauer 1999 and 2005.

[7] See Stachel, 1989, and also Norton, 1993, section 3.5, and Ryckman, 2005, p. 21.

[8] Einstein (1916).

[9] Hilbert, 1915a and 1915b respectively.

[10] Hilbert, 1917.

[11] Some commentators have mistakenly asserted that Hilbert's equations are not generally covariant. On this issue we wholeheartedly support Ohanian's recent statement when he writes (2008, p. 355 (n. 56 to p. 221)): "The fact is that Hilbert's variational equations are covariant, but he supplements them, correctly, by extra, noncovariant, coordinate conditions that are needed to make the solution unique, as is well known to anybody who has ever tried to construct a solution of the Einstein equations."

[12] Cited in Majer, 1995, p. 284.

[13] See, once again, Ohanian (2008), n. 56, cited above.

[14] Indeed, to go further, Hilbert's epistemological analysis of the differing status that should be accorded to general covariance versus causality might perhaps be suggestive to those working on the interpretation of General Relativity as a gauge theory, and the associated "problem of time" in quantum gravity.

[15] Stachel (1992) writes that while Einstein was "always a bit vague about just what he meant by causality" in his hole argument, Hilbert on the other hand "gave a quite precise meaning to the concept", formulating it in the context of whether the field equations can be expressed in Cauchy normal form. Surely it is right that Hilbert was led to think about causality in the context of general covariance by Einstein's concerns in the summer of 1915, and Stachel is of course exactly right that Hilbert's version of the problem is stated in the precise mathematical language of the Cauchy problem. What we wish to emphasize is that the problem Hilbert thus arrives at is importantly *different* from that with which Einstein wrestled in his hole argument.

[16] See section on Hilbert's causality problem, above.

[17] While Hilbert himself addressed the epistemological problem within a Kantian framework (see Brading and Ryckman, 2008, 2009), it is not obvious that the core proposal requires this.