Introductory note to *1946/9

1. The texts (see the textual notes for a full description) from which these two versions have been selected for publication here[a] are all preparatory to the article *1949a*, published by Gödel himself, and were destined for the volume *Schilpp 1949* in which that article appeared. As has been remarked in the introductory note to *1949a* (these *Works*, Volume II, pages 199–201), these papers are in content by no means mere preliminary drafts of that article: they are quite extended essays on a subject rather different from that of *1949a*; and, it seems fair to say, they reveal far more of Gödel's philosophic views on the physical world and our knowledge thereof (his views, that is, on the metaphysics and epistemology of natural knowledge). On the other hand, it must be recognized that Gödel's own decision was *not* to publish a work based directly upon these essays, but to begin again and to write the far shorter and more circumscribed *1949a*. In particular, none of the manuscripts here in question was put into final form for publication;[b] there are certain roughnesses, and—of greater moment—certain questionable passages, that one would expect Gödel to have amended in a final revision. (It will be necessary presently to comment on some of these passages.) The papers should be read, therefore, in clear awareness that they are not finished products, but working drafts. As such (as working drafts, that is, on a subject about which Gödel was reflecting deeply over a period of years) they are of a very special interest.

The subject of these essays, in contrast with the published paper *1949a*, may be characterized briefly thus: The very short piece *1949a* is principally concerned to argue that relativity theory supports the doctrine, attributed to "Parmenides, Kant, and the modern idealists"

[a]The decision to publish versions *B2* and *C1* was based upon the facts that (1) the manuscripts fall into two distinct groups: *A* and the two labeled *B* on the one hand, the two labeled *C* on the other; (2) within the first group, *B2* is undoubtedly the latest; and (3) of the *C* versions, *C2*, which remains incomplete, has all the earmarks of an attempted but abandoned revision of *C1*. (Cf. also §2 of this introductory note.)

However, a passage in *A*, pp. 8–11, seemed to the editors to contain remarks of considerable interest not to be found in the later versions. Therefore it is included in this volume as Appendix A.

[b]The title of these texts appears on one of the lists, made by Gödel, with the title *Was ich publizieren könnte* ("What I might [or "could"] publish"); see the preface to this volume. That certainly encourages the belief that Gödel did not mean to disown the contents of these versions when he decided to publish something different. But it is equally certain that the lists in question *cannot* mean that he regarded any of these as in finally acceptable form; some of the detailed points to be discussed below show this very plainly.

(among these last, McTaggart in particular), that *change is illusory*. The present discussion, as its title indicates, singles out Kant among philosophers; it treats in some detail the relationship Gödel finds between Kant's doctrine of the "transcendental ideality" of time and space, on the one hand, and relativity theory on the other. This relationship, in Gödel's view, is significant in both directions: he suggests (a) that Kant's doctrine is in some ways remarkably confirmed by relativity theory, and that to see this is of importance for our understanding of this theory; and (b) that the turn taken in relativity theory requires, if the Kantian principles he discusses are to be maintained seriously, a quite radical reinterpretation and revision of Kantian philosophy itself. But Gödel's extended train of thought defies summary, and no attempt is made here to summarize further his theses or his arguments.

The contents of this introductory note are as follows: §**2** deals first with the question of the chronological and substantive relations among the five manuscripts *A*, *B1*, *B2*, *C1*, *C2*. §**3** extends that discussion, with some amplification of a point concerning Gödel's discovery in general relativity. §§**4**–**5** are concerned with questions of mathematical interpretation raised by Gödel's essays, and in part with needed corrections of some of his statements. The remaining §§**6**–**9** deal with philosophical questions: §§**6**–**8** with issues raised by Gödel's interpretation of Kant, §**9** with a matter independent of Kant (and previously raised in the introductory note to *1949a*).

It should be noted that the treatment of philosophical issues in §§**6**–**9** is quite independent of what precedes those sections; a reader who wishes to, therefore, can read §§**6**–**9** without (or before) consulting the preceding technical matter.

2. On the question of the chronological placement of these papers, some remarks of David Malament's[c] deserve to be recorded:

> The first thing that struck me after looking through the manuscripts is that they almost certainly were all written *before 1949a*, and that the sequence
>
> $$(\mathrm{A}, \mathrm{B1}, \mathrm{B2}) \rightarrow (\mathrm{C1}, \mathrm{C2}) \rightarrow \mathit{1949a}$$
>
> exhibits an interesting line of development. It can be traced in the references Gödel makes to his own technical work in relativity theory. Here are three pieces of evidence.

[c]In a letter, dated 23 August 1986, to Solomon Feferman, from which I quote with the writer's permission.

a. ... there is no mention in B2 that Gödel spacetime exhibits closed timelike curves, even though this fact significantly strengthens Gödel's argument (for the connection between relativity theory and a Kantian conception of time). It is cited in C1 and C2, and given prominence in *1949a*. I just cannot imagine that Gödel would have undertaken to write on Kant, time, and relativity theory after the composition of *1949a* and have suppressed reference to the fact. Incidentally, it is *not* difficult for me to imagine that some interval of time lapsed between his initial discovery of Gödel spacetime and his subsequent realization that it contains closed timelike curves. (I have in mind a certain technical point here, but shall not bother to elaborate.) Presumably A, B1, B2 were written sometime in that interval, and C1, C2 were written after.

b. In a note added to *1949a* on "2 September 1949" Gödel announced that for *every value* of the "cosmological constant" there exist rotating solutions that do not admit (any) universal time functions. (For spacetime models of the sort Gödel was considering this condition is not only entailed by the existence of closed timelike curves, but is strictly equivalent to it.) The discovery was important because it is customary to take the constant to be zero, and Gödel spacetime only qualifies as a solution to Einstein's equation if one allows it to be non-zero (or reinterprets it as a large "unphysical" negative pressure term in the energy momentum tensor field). The discovery clearly bolstered Gödel's case. He obviously thought it sufficiently important to warrant a late change in his manuscript. (Publication of the Schilpp Einstein volume came in December of 1949.) Yet there is no mention of the discovery in any of the five manuscripts here under consideration. This suggests, at least, that they were written before September 2, 1949.

c. In footnote #53 ⟦in this edition, footnote 15⟧[d] of C1 Gödel anticipates the objection that the idea of "time travel" leads

[d] A concordance is given in the textual notes between the footnote numbers as they appear in Gödel's manuscripts and those in the present edition. (It should be noted that when Gödel interpolated new footnotes, he did not renumber the older notes, but gave higher numbers to the new ones. Therefore the order of his numbering does not always follow the order of the text, but does convey information about the chronological development of his thoughts.)

to absurdity, and hence that Gödel spacetime can be rejected *a priori* as an "unphysical", extraneous solution to Einstein's equation. In response he argues that the objection presupposes the "practical feasibility" of time travel, and that this "*may very well* be precluded by the velocities very close to that of light which would be necessary for it ...". (emphasis added) The statement of the objection is very close in formulation to that in *1949a*. But in the latter Gödel's response is more confident, and is supported by the citation of a calculation. He asserts that "the velocities which would be necessary to complete the [time travel] voyage in a reasonable length of time *are* far beyond everything that can be expected ever to become a practical possibility". (emphasis added) Presumably Gödel did the calculation he cites in the interval between the composition of C1 and *1949a*.

The evidence cited by Malament in (a), as tending to show that the three manuscripts *A*, *B1*, *B2*, were completed before Gödel's discovery of the existence of closed time-like world-lines in the cosmological solutions he calls (in the *C* drafts) "R-worlds", gains still further support from another point of contrast between those versions and that of the *C*s. In each of the former three there occurs the remark—referring to "relativity theory" quite generally—that "for the series of events happening to one material point the 'before' has always an objective meaning", in that "it subsists relative to all observers".[e] In both *C1* and *C2*, on the other hand, we read that "*that* passing of time which is directly experienced has no objective meaning in the R-worlds," because "it is possible in these worlds to travel arbitrarily far into the future or the past and back again". Although the former passage ascribes objective meaning to the "before" of a material point (rather than to that "directly experienced"), the latter stands in essential contradiction to it; for the travel Gödel describes goes (in part of its course) *from* what is "after" for some material point *to* what is "before" for it—thus directly refuting the claim that "it ⟦the 'before'⟧ subsists relative to all observers".

To these remarks there should be added a reference to an illuminating passage in Gödel's lecture **1949b*,[f] in which he comments on the motivation of his own search for rotating solutions to Einstein's field equation. Gödel there notes the equivalence of (a) the vanishing everywhere

[e]See p. 13 of *B2* below, fn. 23 and the text above it.

[f]I wish to thank Malament for calling this passage from **1949b* to my attention at a time when I had not seen that text. Cf. further his introductory note to **1949b* in this volume.

of the angular velocity of matter and (b) the existence of a one-parameter system of three-spaces everywhere orthogonal to the world-lines of matter; and adds (page 12):

> This incidentally also was the way in which I happened to arrive at these rotating solutions. I was working on the relationship ... between Kant and relativistic physics insofar as in both theories the objective existence of a time in the Newtonian sense is denied. On this occasion one is led to observe that in the cosmological solutions known at present there does exist something like an absolute time. This has been pointed out by epistemologists, and it has even been said by the physicist Jeans that this circumstance justifies the retention of the old intuitive concept of an absolute time. So one is led to investigate whether or not this is a necessary property of all possible cosmological solutions.

Thus one sees that it was his reflections on the philosophical bearings of relativity theory—themselves presumably stimulated by the request from Schilpp for a contribution to the Einstein volume—that led Gödel to the discovery of the R-worlds. More particularly, one sees quite explicitly that the starting point of his technical discovery was the realization that (in a matter-filled general relativistic cosmos) the existence of a "natural" cosmic time—defined by slices orthogonal to the world-lines of (cosmic) matter—depends upon the non-rotation of that matter. This is the "technical point" Malament had in mind in his parenthetic remark near the end of (a) in the passage quoted above (written at a time when he had not seen *1949b*): Gödel sought, and found, rotating solutions, because he knew that no "natural" cosmic time can exist in worlds represented by such solutions; only in the further study of the geometry of these rotating solutions did the *stronger* result emerge that these contain closed time-like lines, and therefore *no cosmic time whatever* that is free of anomalies can be defined for them.

3. That perhaps deserves some amplification:

The first, weaker, result—invoked by Gödel, in the passage just quoted, against the contentions of unnamed "epistemologists" and of "the physicist Jeans"—is characterized in the *A* and *B* drafts as follows:[g]

[g]See p. 10 of *B2* below. This paragraph is carried through with very minor alterations of phrasing from *A* through *B2*.

> This view ⟦namely (with attribution to Jeans) that "matter and the curvature of space-time produced by it, if the structure of the world as a whole is taken into account, may enable us to determine some objectively distinguished ordering of all events to which the properties contained in our intuitive idea of time could consistently be attributed"⟧ is supported by the fact that in all known cosmological solutions ... such an "absolute world time" really can be defined. But nevertheless the conclusions drawn above ⟦namely (see Gödel's page 7 *ad finem*), that "what remains of time in relativity theory as an objective reality inherent in the things neither has the structure of a linear ordering nor the character of flowing or allowing of change"⟧ can be maintained because there exist other cosmological solutions for which a definition in terms of physical magnitudes of an absolute world time is demonstrably impossible. If however such a world time were to be introduced in these worlds as a new entity, independent of all observable magnitudes, it would violate the principle of sufficient reason, insofar as one would have to make an arbitrary choice between infinitely many physically completely indistinguishable possibilities, and introduce a perfectly unfounded asymmetry.

Now, the state of affairs asserted here as obtaining in Gödel's new cosmological solutions is—so far as Gödel's explicit statement goes—hardly different from that obtaining in the space-time structure of *special* relativity: that is, here as there, the choice of a particular "world time" is seen as not inherent in the objective structure of the world because it is an *arbitrary* choice among objectively "equivalent" alternatives (introducing "a perfectly unfounded asymmetry"). Indeed, it appears most probable that the train of Gödel's thought in this investigation was one that *did in fact* start from special relativity. For everything Gödel says about relativity theory up to the paragraph running from the end of page 9 to page 10 of *B2* is entirely true of the special theory, and could be understood as referring to that theory alone, except for footnote 4—which is numbered 46 in Gödel's original text (i.e., it is a fairly late addition). And in the paragraph just mentioned, the transition to general relativity is explicitly motivated by the following remark:

> With much more justification ⟦than something just criticized⟧, it is objected ⟦footnote: "See *Jeans 1936*"⟧ that the impossibility of defining any absolute time ... *in the empty space-time scheme of special relativity theory* (upon which the foregoing considerations have been based) does not exclude ⟦an "absolute world time" in the cosmological solutions of general relativity⟧. ⟦Emphasis added.⟧

Thus we see the likelihood of the following remarkable sequence in Gödel's investigation: (a) philosophical reflections on the status of time and change in relativity theory, with primary attention to the special theory, leading him to make his comparisons with Kant; (b) a possible check to this line of thought—namely, the objection (associated with Jeans) that *general* relativity might warrant a *reinstitution* of time as "absolute" and change as correspondingly "real"; (c) the technical investigation, motivated by that objection and by the idea that in a "rotating world" a situation *resembling that of special relativity* might obtain, so far as the problem of time is concerned—an investigation that was successful in finding examples of such "rotating worlds"; (d) finally, the *entirely unexpected* discovery that in these "rotating worlds" a very much more thoroughgoing impossibility holds for anything like the conventional notions of time and change.[h]

The difference between what was envisaged by Gödel in (c), and what was eventually found by him in (d), is this: The principle by which, in the ordinary cosmological models appealed to by Jeans, an "absolute" world-time is distinguished, is that events ought to be regarded as "absolutely simultaneous" when they lie in some one three-dimensional (necessarily space-like) submanifold of space-time that is *everywhere orthogonal to the world-lines of "cosmic" matter.* No other "natural" criterion suggests itself; and in the absence of such a criterion, Gödel says, the choice of one among the "infinitely many physically ... indistinguishable possibilities" would be arbitrary, and would "introduce a perfectly unfounded asymmetry". Still, to put the case a little crudely, a determined "absolutist" might argue from the theological postulate that there is a uniquely distinguished "God's time" ("*Gottes Zeit ist die allerbeste Zeit!*")—even if merely physical considerations are inadequate to tell us which time that is. In the rotating worlds of Gödel, however, the situation is radically different: there are not "infinitely many physically

[h](*Note by the editors.*) Researches of John W. Dawson, Jr., have shown that confirmation of the stages in this sequence is provided by entries in Oskar Morgenstern's diaries (among his papers at the Perkins Memorial Library of Duke University) and by Gödel's own statements in letters to his mother. On 23 September 1947 Morgenstern reported that Gödel had found "a world in which simultaneity cannot be defined", and on 7 November Gödel wrote that his investigations had "led to purely mathematical results" that he intended to publish separately. The following January Gödel told Morgenstern that he was eager to return to work in logic, and the next month declared to him that he expected to have his cosmological results written up "soon". Instead, however, there followed a three-month hiatus, at the end of which (on 10 May 1948) Gödel wrote his mother that he had at last settled "a problem" that had for weeks pushed everything else out of his mind. Two days later Morgenstern then noted that Gödel was again "making good progress with his cosmological work": in particular, "Now in his universe one can travel into the past". (The quotations in this note are Dawson's translations of German originals.)

indistinguishable possibilities" among which to choose; rather, there are *no possibilities at all*: not only are there no three-dimensional slices of space-time *everywhere orthogonal to* the world-lines of cosmic matter, there are simply no three-dimensional space-like slices *whatever* that so much as *intersect* all the world-lines of cosmic matter, each in a single point.

This follows almost immediately from what is surely the most striking characteristic of the rotating worlds: the existence in them of smooth, everywhere time-like, everywhere "future-oriented" world-lines, connecting any point in the space-time to any other point.[i] It is this that leads to the consequence stated by Gödel (version *C1*, page 10, paragraph 2) as follows: "Above all they ⟦the "absolute time-like relations" in the R-worlds⟧ define only a partial ordering, or to be more exact: in whatever way one may introduce an absolute 'before', there always exist either temporally incomparable events or cyclically ordered events."

4. Gödel's statement here is open to some criticism (cf. the warning in the opening paragraph of this note that such criticism would be necessary); it seems likely to me that this is a sign of the haste, and perhaps intellectual tumult, in which these thoughts were penned by him after his unexpected discovery.[j] In the space-time structure of *special* relativity, a "natural" (and in *this* sense "absolute") *partial temporal ordering* does indeed exist: namely, "a strictly precedes b" just in case a and b are not identical, and b lies within or on the mantle of the "forward" light-cone of a (or, equivalently, a lies within or on the mantle of the "backward" light-cone of b). This temporal ordering satisfies the "natural" demand that whenever "influence", or "a signal", can be propagated from a to b—and in particular, whenever *a body* can "travel" from a to b—the relation "a precedes b" holds.[k] That is also the situation in

[i]The consequence holds in view of elementary facts about the R-worlds: There, each world-line of cosmic matter is an "open line" (i.e., is isometric with the real-number line); and these world-lines have a canonical and mutually coherent time-ordering. A slice of the envisaged kind, then, would divide each matter world-line into a "past" and a "future" part; and at the same time would divide the entire manifold into two components—one past, the other future, to the slice. A time-like line starting from a point on the future side and ending at a point on the past side would have to cross the slice, and at the crossing-point would have to be directed future-to-past; thus no everywhere "future-oriented" line could start at a point on the future side and end at one on the past side of the slice.

[j]To be sure, the statement is repeated verbatim in the revised text *C2*; on the other hand, the fact that this revision was left incomplete lends some independent support to the view that Gödel did not thoroughly reconsider and criticize what he had written.

[k]If b belongs to the mantle of the forward cone of a (or, equivalently, a to the mantle of the backward cone of b), a *signal* can be propagated from a to b; if to the

those general-relativistic *cosmological* solutions that were known before Gödel's work (although the existence *in general* of space-time structures in which this "natural" condition cannot be met had been noticed long before—cf. below). But the state of affairs described in the preceding paragraph tells us *directly* (since a smooth, everywhere time-like, everywhere future-oriented world-line is a *possible* world-line of a *body*) that our "natural" condition in a Gödel R-world demands that every point of space-time strictly precede every other; and this contradicts what one means by a (strict) ordering relation. In short, Gödel's results, taken with the most natural conditions one would place upon a temporal ordering, imply that *not even a partial temporal ordering is possible* in an R-world.

It might, however, be thought that Gödel had something else in mind: namely, not to define the temporal ordering by the possibility of travel or influence, but to define it simply by the light cones themselves. What one finds if one pursues this line exhibits in yet another way the strangeness of the Gödel universes. The situation is most easily described in a restriction to two dimensions of space (with of course one of time). If one starts at a point a of the space-time, and defines as "later than a" all points in the set swept out by the future-oriented time-like or null geodesics starting from a, one indeed arrives at a non-trivial relation; but this relation *is not an ordering*: it is not transitive. The region in question presents the following picture: It starts, cone-like as one would expect, expanding on all sides from a, and admits a succession of space-like two-dimensional slices of increasing radius, until a certain critical radius is reached. These slices, taken in order until the critical radius is attained, can thus be regarded, so far as all this is concerned, as representing "successive epochs" in the "development" of a part of the world; and so we do indeed have (again: "so far") what looks like a bit of an "absolute", "temporal" ordering. But at the critical epoch (or phase), the slice has ceased to be entirely space-like: rather, one tangent line to the slice at each point of its extremity is a null line—and is indeed tangent to the world-line of propagation of a light signal emitted from the initial point a. These bounding null tangents envelop a circle at the critical radius (a circle, thus, of a definite space-like radius, but of circumference zero). The light-lines from a, however, although they are all tangent to this circle, do not merge with it; they pass beyond and *reconverge* to a point b on the "cosmic" geodesic (what Gödel calls the

interior of the cone, then a signal can be propagated, and (even) a *body* (of non-zero rest mass) can travel, from a to b. (I here make *transport of a body* a more special notion than *propagation of a signal* only as a convenient way to record a distinction; in a quite reasonable sense, "particles" of zero rest-mass, which are propagated along the mantle of the cone, can of course also be regarded as "bodies".)

"world-line of matter") through a. If we think of the "points of cosmic matter" as the *stars* of this world, then light from a given star always reconverges to it—in fact, pulsates periodically out to the critical radius and back again. And so do all other geodesics ("ballistic trajectories") from a given star: they all reconverge, with the same "period" as the light signals.[l] The non-transitivity of the relation defined by this procedure is clear, for instance, from the fact that whereas, in the account just given, no point in space-like relation to b is "later than" a, if one takes a point c on the geodesic from a to b and halfway between them—a point, thus, "later than" a—and constructs its future, this will contain a space-like slice through b of the maximum possible radius (the critical radius). So the idea fails: it, too, does not define an ordering.[m]

A corollary of the result just described is worthy of special note. First let it be remarked that in the R-worlds here under discussion—those of Gödel's initial publication *1949*, in contrast with the *expanding* rotating universes of *1952*—the cosmic masses maintain constant relations to one another, in the sense of the space-time geometry, throughout the history

[l]In the full four-dimensional case this picture is modified: in the "remaining spatial dimension", omitted from the discussion above, geodesics (whether null or time-like) can be thought of as diverging at a constant rate ("linearly") from their starting point.

[m]One remark should be added to avoid a misunderstanding on a small technical point. I have spoken above of defining a relation "simply by the light cones themselves"; but the construction just described uses time-like as well as null geodesics—and of course that exploits the full metric structure of the space-time, not merely the structure defined by the light cones (the "conformal" structure). This leads to a perspicuous geometrical picture, and its extra details are of interest in themselves. The relation can, however, indeed be characterized entirely with the help of the light cones—that is, of null geodesics alone. For in the construction above (still restricting the account to two dimensions of space), the null geodesics emanating from a have been seen to reconverge to b. In doing so, they sweep out, when one considers just the segments from a to b, a surface (the "mantle" of the forward cone from a). This surface has, from the differential point of view, singularities at a and b; but it is topologically a closed manifold, and in fact a topological *sphere*. It therefore has a well-defined "interior region" in the space-time (which itself has the topology of a three-dimensional Euclidean space). The relation "x is later than a" has in effect been defined to hold, *first*, for all points x other than a itself that belong to that surface or to its interior region. To obtain the full relation defined in the text, one must iterate this construction: Let the relation just defined be expressed as $L(x, a)$; let us set $a_1 = b$; and in general, a_n having been defined, let a_{n+1} be obtained from a_n in the same way b was from a. Then the relation "x is later than a" means: for some n, $L(x, a_n)$. (Adding the suppressed spatial dimension, the construction is slightly more complicated but not essentially different. The mantle of the light cone is now a hypersurface, and topologically a 3-sphere in the now four-dimensional Euclidean space. Instead of the focal point b corresponding to a, there now occurs a focal *space-like line segment*, and the construction has to be iterated from this entire line segment—the successive segments, corresponding to the successive points a_n above, growing longer at each stage. But the successive portions of the mantle continue to be 3-spheres, and everything else is as before.)

of each. The above result means that despite this fact—i.e., despite the fact that the *bodies are in no sense "moving apart"*—observers on "sufficiently distant" cosmic masses *cannot reach one another at all* with direct light signals (nor, by the same token, can they reach one another with direct "ballistic signals"). On the other hand, they *can* (in principle) reach one another with *indirect*, or *relayed*, light signals (or ballistic signals). This is in effect just a more vivid formulation of the "non-transitivity" noted above: a light-cone C that originates within, or on the mantle of, a "preceding" one, A, develops so as to "escape from" A. It is also noteworthy that this situation contradicts the conception of the direct light-signal as a "first-signal", or "fastest" signal, in relativistic physics—cf., e.g., *Reichenbach 1928* (version *1957*), pages 143, 166, 238; and *Reichenbach 1924*, *passim.* (To be sure, it is obvious that there can be no "first-signal" in a world in which *every* space-time point can signal to *every* other.)

The only remaining possibility there seems to be for an interpretation that might, *prima facie*, render Gödel's statement correct as it stands is to take, for the "absolute partial ordering" he refers to, just the relation "a and b are two points on a single world-line of (cosmic) matter, and a precedes b on this line". Then, indeed, one has a partial ordering; but it is not true that every extension of this relation either leaves some events temporally incomparable or makes some events "cyclically ordered". For Gödel himself, in his mathematical treatment of the R-worlds, introduces a global space-time coordinate system, with a global time-coordinate that orders "correctly" the events on each world-line of matter. So the ordering by this time-coordinate extends, to the entire cosmos, the partial ordering along the world-lines of matter, in contradiction to Gödel's professedly "more exact" statement.

Nevertheless, what one is entitled to regard as the essential intended content of that statement is true: the ordering by the global time-coordinate violates the requirement that *events "experienceable" by any observer be ordered as they are in that observer's experience.*[n] This

[n]One incidental remark to this: Gödel refers (*B2*, p. 13) to "the fact that for the series of events happening to one material point the 'before' has always an objective meaning"; and of course one will apply the same proposition to the events happening to one observer. But a certain qualification is necessary. In Gödel's R-worlds, one can conceive a material point—or an observer—as having a world-line with the topology of a *circle*. In this case, there ceases to be, for the world-line as a whole, a well-defined ordering of "before" and "after". However, there is still what one may call an "infinitesimal ordering": a determinate direction "past-to-future" *at each point* of the line (that is, "at each instant" for the body or the observer); and this gives rise to a determinate "local ordering ", in the sense that there is induced a determinate ordering on every "sufficiently small" segment of the line. Translated into terms of the "experience" of such an observer, what one would have is, at each instant

requirement is merely the reformulation in a subjective mode of that earlier stated in terms of the direction of "physical influence"; and since the order of influence, or of experience, can be "from" *any* point of space-time "to" *any* other point (or to *the same* point—hence Gödel's invocation of the possibility of "cyclic order"), the requirement (as we have seen) is incompatible with even a partial ordering.

5. Several further passages present deeper puzzles of mathematical interpretation.

The first of these occurs in *C1* shortly before the statement we have just been concerned with, and, like it, deals specifically with the situation in the R-worlds. Gödel speaks (see pages 8–9) of the procedure of defining time relative to an observer "*by direct measuring with clocks and Einstein's concept of* ⟦*simultaneity*⟧" (emphasis added) and states in a footnote that this concept, originally designed for special relativity only, "can be extended to general relativity theory, at least under certain conditions which are satisfied in the R-worlds". This claim is very hard to make good sense of. Einstein's criterion of simultaneity was intended to be employed by *a collection of observers* who are in *the same state of uniform motion* (that is, by *inertial* observers who are, mutually, in a state of relative rest). Among such observers, the relation of simultaneity—or, what comes to the same thing, the synchronization of clocks—was to be established by light signals. In the Gödel R-worlds, one does not clearly see what classes of observers are to play anything like that role. The only plausible analogue of *inertial* observers, in general relativity, is *gravitationally free* observers (i.e., ones who move solely under the influence of the gravitational field). But if any class of these latter can be regarded as "at mutual relative rest" in the R-worlds, it would appear to be the class of those observers who move with the cosmic matter (since the geometric relations of such observers remain stable). And yet these certainly *cannot* be treated as a class among whom Einstein's simultaneity criterion is to be employed: indeed, we have already

and over each small enough segment, a definite sense of the "order of becoming"—the "order of changes"; whereas over the totality of experiences of such an observer (a totality of experiences that would be *without end*), one would have an *eternal recurrence* of those experiences. (The question might be raised whether, besides the possibility of such a *closed* world-line, one might also envisage *self-intersecting* world-lines: ones that "circle back" [or "spiral back"] to earlier stages of the line and *cross* those earlier stages. The answer is that such lines are indeed geometrically describable; but if one wants such a line to represent the history of a *body* (in which case it must be "thickened out" into a *continuum* of lines), then the self-intersection would represent a *collision* of *two* bodies that would proceed to *move through* one another. In so far as one speaks of a physics of "bodies", such a process is probably to be regarded as physically impossible.)

seen that between two sufficiently distant observers of this class *no direct light-signals at all* can be sent.[o]

The problem having been noted, it need not be further pursued. Gödel *may* have had some construction in mind that would form a suitable analogue in the R-worlds to the Einstein criterion in special relativity; and if so, one would like to understand what it was; but whether he did, or whether his remark represents an oversight, there is no serious consequence for his general argument. We may rephrase his conclusion with a qualification: "The fundamental temporal relation 'A before B by t seconds' obtained in this way, ⟦*if such a relation exists at all,*⟧ is certainly not something inherent in the events ...". The interpolated reservation does not weaken Gödel's assertion.

The perplexities of the next passage are more complicated. In each of the manuscript versions, there occurs a discussion centering on the claim (a *limitation*, according to Gödel, of the parallel he is exploring between relativity theory and Kant's doctrine of space and time) that in general relativity one can define "absolute spatial relations" that "lead to a structure ... not so very different from that of intuitive spatial relations".[p] The detailed description of the generation of this structure is not quite the same in, first, the texts of A and B; second, footnote 20 (footnote 50 in Gödel's numbering, thus a fairly late addition) of those versions; and, third, the C texts. But setting aside the varying details, if we take the latest version to represent Gödel's more considered view, it appears that the alleged space of "absolute spatial relations" is meant to be obtained in the following way:

1. We restrict attention to relativistic worlds in which there is a cosmic "matter-flow" defined everywhere; so there is a distinguished family of time-like world-lines—in fact, geodesics—of which exactly one passes through each space-time point. Each of these distinguished world-lines corresponds to (is the *history* of) one "material point"; and it is the material points—or, in other words, the distinguished world-lines themselves—that are taken to be the "points" of the space we are to describe.[q]

[o]It might be suggested that relayed signals be used; but this would altogether defeat the intent of Einstein's criterion, and would lead to no coherent simultaneity-relation.

[p]*B2*, pp. 11-12, with fn. 20; cf. also *C1*, pp. 17-22.1.

[q]It has already been remarked that there are variants in Gödel's discussion. In *B2*, fn. 20, the question is raised of the possibility of defining "'absolute' points of space persisting in time and identifiable at different times" in some way that does *not* simply identify these with the material points. (Gödel's formulation of this question is a little unclear—its phrasing does not directly imply that the "absolute points" are to be

2. The matter-flow (or the vector-field associated with it) is assumed to be *smooth*;[r] and Gödel seems to infer tacitly that this entails the existence of a well-defined differentiable-manifold structure on the set of material points (i.e., on our "space"). That, however, does not follow without some further assumption. It is indeed "ordinarily" the case—i.e., it holds for "reasonable" space-times—and the simplest procedure is just to make this the further assumption: that is, if we let $\mathcal{M}$ be the space-time manifold, we assume the existence of a three-dimensional differentiable manifold $\mathcal{M}_1$, and a smooth mapping π of $\mathcal{M}$ onto $\mathcal{M}_1$, such that the inverse image under π of any point of $\mathcal{M}_1$ is the entire world-line of a material point. $\mathcal{M}_1$ is, then, the underlying differentiable manifold of the claimed geometry of "absolute spatial relations". (It should be noted that, given the assumption of its existence, the "absolute" character of $\mathcal{M}_1$ itself—that is, its geometrically objective status, grounded in the structure of $\mathcal{M}$ together with its matter-flow field—is secure.)
3. It remains for us to define a *metric* structure on the spatial manifold $\mathcal{M}_1$. For each point x of $\mathcal{M}_1$, and each point ξ in $\pi^{-1}(x)$ (that is to say, *each epoch in the history of the material point* x), the map π induces a projection of the tangent-vector space to $\mathcal{M}$ at ξ onto the tangent-vector space to $\mathcal{M}_1$ at x. If we consider, in the former tangent space, the subspace *orthogonal to the direction of matter-flow* (that is, to the direction of the world-line x itself at the point ξ), this projection is an *isomorphism* of that (three-dimensional) subspace onto the tangent space to $\mathcal{M}_1$ at x. The three-dimensional subspace in question at ξ has a positive-definite metric induced from the metric of space-time; and through the isomorphism induced by π, this metric can be transferred to the tangent space at x. However, as ξ varies over the world-line x,

other than the material points. But the sequel removes that uncertainty, for Gödel goes on to speak of the need for "some requirement" to be imposed on this problem, because "otherwise points of space may simply be identified with material points".) At any rate, having raised the question, Gödel remarks that even its meaning is not quite clear (because the needed further requirement is lacking), and he does not pursue it further, nor does he mention it in the *C* drafts.

[r] Again variants must be noted. In *B2*, fn. 20, Gödel envisages the possibility of a *discontinuous* velocity-distribution of matter. It is hard to see how a construction like the one he envisages could be made in this case; and, indeed, in *C1*, p. 22, he restricts the described construction to the case where "the velocity vector of matter is continuous". It remains hard to see how the construction could be done without the stronger condition of (a suitable order of) differentiability, and it is perhaps allowable to assume that, writing a principally non-technical paper, Gödel used the word "continuous" vaguely to connote "smooth". In any case, as will be explained presently, yet stronger assumptions seem to be needed.

the transferred metric on the tangent space to $\mathcal{M}_1$ at x will in general not remain constant. We have, therefore, on the manifold $\mathcal{M}_1$, not a Riemannian metric structure, which consists in a metric on the tangent space at each point of the manifold (with suitable smoothness conditions as one goes from point to point), but rather a structure that consists in the specification, on each tangent space, of a *one-parameter family* of metrics (where the parameter is to be thought of as the "varying epoch"—say, the "proper time"—along the world-line of $\mathcal{M}$ that *is* the point in question of $\mathcal{M}_1$).

4. Gödel says that we can use this construction to define on $\mathcal{M}_1$ the structure of "an ordinary three-dimensional Riemannian space changing in time"; but in a way that is *not* canonically (or "absolutely") determined by our data: rather, "This space is different for two observers A, B only insofar as the phases of development of its several parts (in consequence of the different meanings of simultaneousness for A and B) are collected together in different ways into one phase of the whole space" (*B2*, footnote 20; cf. *C1*, page 22). This "collecting together of phases" is, however, not unproblematic: it presupposes a further strengthening of the assumptions we have made. What is required for the construction Gödel indicates is the existence of a "cross-section" of the mapping π: that is, a mapping σ from $\mathcal{M}_1$ to $\mathcal{M}$ such that the composite mapping $\pi\sigma$ is the identity mapping of $\mathcal{M}_1$ onto itself. Thus σ assigns to each material point an epoch in its own history—a phase of its development; and these selected epochs "collected together" may be looked upon as forming a single "phase of the whole space". That being done, it is easy to extend the construction to *each* phase of each particle (e.g., if the matter-flow geodesics are metrically infinite "open" lines, by taking σ to define the zero-point of time on each material world-line, and then extending to a "time-coordinate" by measuring the proper time, starting from zero, along each such world-line). In this way, letting $\mathcal{T}$ represent the global "time" so defined, we have a representation of $\mathcal{M}$ as the Cartesian product $\mathcal{T} \times \mathcal{M}_1$; and for each "phase" or "epoch" in $\mathcal{T}$, the construction in (3) above gives us a well-defined Riemannian metric structure on $\mathcal{M}_1$—thus, in accordance with Gödel's claim, the structure of a Riemannian space changing in time.

That, however, is not the end of our problems with this passage. For although the collecting together of phases can indeed be performed (on the *assumption*—by no means entailed by what went before—of the existence of a cross-section), one has to ask what geometrical or physical significance such a representation of "space changing in time" will have for a general-relativistic cosmos. Gödel appears to imply that the "collecting together of phases of development" is well-defined *relative to an observer.* Indeed, as we have just seen, he refers to the difference in this representation "for two observers A, B"; and he attributes this difference to "the different meanings of simultaneousness for A and B". (Thus the present difficulties are after all connected with the previous one—that is, with Gödel's supposition that there is a "natural" notion of simultaneity "relative to an observer".) However, even if there were such a well-defined concept of "relative simultaneity" in general relativity, it would remain unclear what significance a geometry pieced together in the fashion described would have (for it must be noted that, for instance, the "distances", or "lengths of curves", in the space $\mathcal{M}_1$ at some phase of its development are *not*, in general, space-like distances, or lengths, of anything whatever in the space-time $\mathcal{M}$). And the absence of a well-defined concept of even *relative* simultaneity makes this difficulty all the stronger. At the very least, Gödel's characterization of this as a structure of *absolute* spatial relations would be very hard indeed to defend.

But there is a special point that may have contributed to what seems indisputably a confusion here: namely, the construction described, when it is carried out for a Gödel "R-world", *does indeed* lead to an "absolute" three-dimensional metric geometry (although not one "changing in time"). For in the first place our assumptions are satisfied in these worlds: there is a mapping π of space-time onto a three-dimensional manifold such that the inverse images of points are the "world-lines of matter", and this mapping admits cross-sections. And in the second place, because the R-worlds are dynamically "stationary", *the geometries induced on the tangent spaces to the three-dimensional manifold are constant in time* (that is, constant over the development of proper time along each world-line). It follows not only that the Riemannian structure induced is itself constant in time, but—of far greater importance—that this structure is *independent* of the arbitrariness that affects the "collection into phases" (independent, e.g., of the choice of a cross-section). So for the R-worlds, but *not* for the most general cosmological models, Gödel's distinction, that whereas there is no cosmic structure at all like that of classical time there is an ("absolute") structure "not so very different from that of intuitive spatial relations", does indeed hold.

A further discussion of geometrical relations, this time in connection with a possible vindication of Kant's views on the *a priority* of spatial intuition—a discussion that occurs in all the manuscripts except the incomplete draft *C2*[s]—poses additional problems. It is, in fact, hard to

[s]See *B2*, pp. 14–15; *C1*, pp. 23–24.

see this discussion as concerned at all with the situation in the theory of relativity (special *or* general); for Gödel here speaks of defining "concepts satisfying Euclid's axioms ... also in a non-Euclidean world", by means that make no reference at all to *time*; so that we seem to be in a universe of philosophic discourse that altogether antedates Einstein—roughly, that of Poincaré's *Science and Hypothesis* (see *Poincaré 1905*, chapters 3–5). Thus the declaration that "we define the 'coordinates' a, b, c of any material point P by the number of times a measuring rod must be applied along the directions α, β, γ... in order to reach P from O" assumes both that these "directions" are well-defined as persisting through time, and that the result of the measurement is independent of the kinematic details of the "application" of the measuring rod (thus in particular of the time taken up by the process).

One point of detail in this passage demands special comment. In *B2*, footnote 27, Gödel says that his definition of coordinates would, in a closed (e.g., a Riemannian—that is, metrically spherical) world, not give unique coordinate values; that it would "have the consequence ⟦either that⟧ every object is in infinitely many different places simultaneously, or that for every object there exist infinitely many exactly equal ones". This is a very inadequate account of the situation. In the first place, it is only for a manifold that is topologically Euclidean that a global system of coordinates of the sort Gödel envisages can be defined, assigning "unique coordinate values" to each point. In the second place, the alternative interpretations envisaged by Gödel, "each object in infinitely many places" or "each object having infinitely many replicas", appear to presuppose that the "closed world" in question admits a Euclidean world as what is called a "covering space": then the covering map represents each point of the covered manifold by infinitely many points, periodically distributed, of the covering space. But this presupposition is *false* for spherical spaces: a sphere (of any dimension ≥ 2), because it is a simply connected space, admits no proper covering spaces whatever.

6. Turning now to issues respecting Gödel's philosophical interpretation of Kant, it should be noted with emphasis—as Gödel does himself in his opening footnote to both texts below—that he does not regard himself as a follower of Kant; and one may perhaps fairly conclude that Gödel is concerned, not primarily with the scholarly authenticity of his reading, but with what he has found suggestive, in Kant's discussions, of themes that in his view are illuminating for thought about modern physics: "Above all", he tells us, "I wanted to show that the questions arising in such a comparison" (namely, "between relativity theory and the Kantian doctrine about time and space") "are interesting and perhaps even fruitful for the future development of physics."

It might be best to let the matter rest there, and invite the reader to reflect upon Gödel's set of variations on themes from (or suggested by) Kant, without editorial animadversion. There are, however, a few points on which Gödel's remarks about Kant's philosophy appear either extremely enigmatic or seriously astray; and it has seemed best to call attention to some of these—not least in the hope that other commentators may succeed in clarifying the enigmas. There are also passages where an interpretation of Gödel's that may strike a student of Kant as misguided appears to merit discussion aimed at clarifying the rationale of Gödel's point of view, and the relation—in part one of *deliberate* divergence—that he sees between himself and Kant.

One of the most transparently non-Kantian features of the view Gödel puts forward lies in the fact that for him, in clear and conscious opposition to Kant, the character of our "spatial (and temporal) intuition"—if we have any such thing as an "innate intuition representing geometrical entities ⟦of some kind⟧"[t]—is not decisive, *either* as guaranteeing, *or* as implying a limitation upon, features of our knowledge of "objects". Thus Gödel tells us (*B2*, page 16; *C1*, page 25) that "What ... Kant did not take into account is that ... space and its properties express themselves also in the sensations, which we know only a posteriori; namely, in the fact that by projecting the sensations in a certain way into a three-dimensional Euclidean space *the laws connecting them can be stated much more easily*" (emphasis added).

To a Kantian certainly, and to a Kant scholar very possibly, this will seem grotesque; for according to Kant, space is "the form of all appearances of outer sense" (A 26/B 42), and as such certainly does "express [itself and its properties] ... in the sensations"[u]—a circumstance Kant cannot seriously be accused of neglecting. But let me suggest that Gödel has here merely expressed himself badly. What I think he claims Kant

[t]Whether we do or do not have such an "innate intuition" (I place the phrase in quotes because it is itself un-Kantian) is regarded by Gödel as a still open and interesting question; cf. *B2*, pp. 14–17 and fns. 28–29; *C1*, pp. 23–26 and fns. 20–21.

[u]"Sensation" here must not be understood in the rather special sense attached by Kant to the word *Empfindung*; for the latter is called by Kant (A 167/B 209; emphasis added) "the *matter* of perception" (in contrast with space and time, the *forms* of sensibility), and is accordingly said to have "no extensive magnitude". Rather, Gödel must use the word for what Kant calls (e.g. B 162) "sensory intuition" (*sinnliche Anschauung*); and it is in this sense, accordingly, that it is used in these comments on Gödel's remarks.

On citations of Kant, see the Information for the reader, p. x. In particular, the *Kritik der reinen Vernunft* is cited in the standard way. All translations from the German (that is, from the cited works of Kant and Weyl) in this introductory note, except in note h, are by the present author. For the *Kritik*, the version of Norman Kemp Smith (*Kant 1933*) has been a most helpful guide, but has not been followed on all points.

has neglected is not simply the fact that space expresses itself in the sensations, but something about *how* it does so; in particular, Gödel claims it does so in a way that, in his opinion, we do and *can* know "only a posteriori". What way this is, is what the passage italicized in the preceding quotation from Gödel tells us. But this in turn requires expansion—and contrast with the doctrine of Kant.

For Kant—and this is a point that in my own opinion Gödel has not sufficiently appreciated (the question will be discussed presently)—space does not *merely* "express itself" in the moment-by-moment character of our (pure and empirical) "outer intuition". That it does so express itself is something Kant claims to establish in the Transcendental Aesthetic, which is short and rather simple; but it is only after the famously difficult (Transcendental) Deduction, and Schematism, of the Pure Concepts of the Understanding that he claims to establish what he calls "the principle of the Axioms of Intuition"; and of this he tells us: "This transcendental principle of the mathematics of appearance greatly enlarges our *a priori* knowledge. For it alone can make pure mathematics, in its complete precision, *applicable to objects of experience*" (A 165/B 206; emphasis added). The implications of this claim, for Kant, are momentous and very far-reaching; they can be conveniently summed up here by saying that he believes that the mathematical structures cognized by us through the "forms" of outer and inner intuition—space and time—affect, necessarily, not only our momentary sensations, but the ordering of these sensations, at each moment and successively, in regular patterns that make "experience of objects" (and, so, "objects of experience") possible. And even further: on the basis afforded by the *a priori* forms of intuition, and the bringing to bear, upon what those forms provide, of the "pure concepts of the understanding" (the *categories*), Kant believes one can ground the possibility of "mathematical constructions"—and, in consequence, *a priori* knowledge—not only of spatio-temporal configurations (figures; abstractly conceived motions), but of the objects of "pure natural science".[v] In short, in the view of Kant, the *laws governing the interrelations, and succession, of our sensory perceptions*, in so far as these can *possibly* be regarded as perceptions "*of objects*"—specifically, of *empirical* objects (*Gegenstände der Erfahrung*)—are necessarily conditioned by the forms of intuition.

[v]Cf. *Metaphysische Anfangsgründe der Naturwissenschaft* (*1786*), Preface: "A rational doctrine of nature (*Naturlehre*) ... deserves the name of a natural science only then, when the natural laws that are fundamental to it are known *a priori* and are not mere laws of experience. One calls a natural cognition of the first kind *pure*; that of the second kind on the other hand will be called *applied* ....

"All *genuine* natural science thus requires a *pure* part, upon which can be grounded the apodictic certainty that reason seeks in ⟦such a science⟧" (pp. 468–469).

That is an aspect of the Kantian philosophy that Gödel does not take seriously. Indeed, as we shall see presently, he goes to some lengths to attempt to exonerate Kant from having held such views at all deeply. Once more deferring that point, what is quite clear is that Gödel is writing from a perspective that knows, that is steeped in, the existence of possibilities for a science of nature not dreamt of by Kant—most notably, the general theory of relativity, with its possibilities for spatio-temporal structure of such a radically new kind as Gödel himself discovered. Gödel's view of this might be expressed as follows: The development of mathematics since Kant has revealed the possibility of "constructions" of a kind that vastly extend the scope of mathematical structures and mathematical reasoning—so that the latter is seen not to depend upon the rather rigid basis that Kant thought indispensable. This more flexible mathematics has led in turn not only to the bare possibility, but, in conjunction with what we have come to know *a posteriori* about "the sensations" (cf. the quotation from Gödel that we are still discussing), to the *fact* that "the laws connecting them" (the sensations) *cannot* be "stated much more easily" by "projecting [them] into a three-dimensional Euclidean space". What Kant can be said, then, to have "overlooked" is simply the possibility of such a development.

But now Gödel makes an interesting un-Kantian move that can be said to have a genuine Kantian ingredient. According to Kant, the "transcendental" *guarantee* of our knowledge in "pure" mathematics and science is correlative to a transcendental *restriction*: The guarantee is strictly dependent upon the forms of intuition as providing the basis of such knowledge (the possibility of "constructing" the objects of knowledge). But the forms of intuition are (Kant thinks) peculiar to *us*; they do not characterize the "things [in] themselves", nor do they necessarily characterize the way the things would appear to other percipient and thinking beings. Therefore the knowledge we obtain through them, and indeed *any* (theoretical) *knowledge we could possibly have*,[w] is restricted in its applicability to objects *as we experience them*; of the "things [in] themselves" we know and can know nothing. If, however, the whole theory of pure science as grounded in the forms of intuition is wrong, and if we have been able to *infer* (or surmise) *from the sensations* a system of laws that do not rest upon a basis determined by those forms, then we may have some grounds for considering the "objects" to which these laws apply to be "the things [in] themselves". And this is exactly the point of view that Gödel suggests. He does not wish to maintain that physics as it exists today has reached true knowledge of what Kant calls *noumena*;

[w]Kant does of course partially exempt our "practical"—that is, ethical—knowledge from this restriction.

but he argues (*B2*, page 13, with footnote 24, and pages 18–21; *C1*, pages 27–30, with the corresponding footnote 27) that it is plausible to consider physics as advancing in steps "beyond the appearances and towards the things".

This intimation that the structure discerned in the general theory of relativity may be *closer to what characterizes the "things in themselves"* than is the more "apparent" structure we encounter in our simpler sensory experience is related—so far as it bears upon Kant's doctrines—to Gödel's contention that, *according to Kant*, although spatio-temporal properties as such are not affections of things in themselves, those properties "correspond" to certain "objective relation[s] of the things to us" (*B2*, page 3; cf. *C1*, pages 6–7). I think it is clear that in this statement Gödel uses the words "objective" and "thing" in what for Kant would be their *noumenal* sense: as referring to "the things in themselves", not to what Kant calls *Gegenstände der Erfahrung.* And it is of some interest that in the incomplete *C2*, presumably the latest of Gödel's drafts, there occurs a slight addition to this discussion, with a further quotation from Kant. The passage in question occurs in what corresponds to the second paragraph under numeral '1' in *B2*, at the beginning of page 4. This whole discussion, with its distinction of "two questions" in numbered sections, is compressed greatly in *C1* (pages 6–7, with no distinction into two questions and no numbered sections), but it is restored in *C2* to something close to its original form. At the end of the paragraph referred to, after the comment that the cited passages seem to imply that "the relations under consideration ... cannot consist solely in the act or disposition of representing", there occurs a new sentence—unfortunately elliptical, because in this version the references to previously cited statements of Kant are left blank: "Finally the passage referred to in ______ seems to indicate that the relations in question are something in the nature of physical influence." As to the quotation meant, it seems possible to identify it with certainty; for the list of relevant Kantian passages (corresponding to (1)–(7) in *B2*, pages 2–3, and to the almost identical paragraph—lacking the identifying numbers—in *C1*, pages 4–5) includes one new item, and it is of obvious pertinence: after quoting Kant (from *Prolegomena* §13, page 286) as saying that appearances are something "whose possibility is based on the relation of certain things unknown in themselves ... to our sensibility", Gödel adds: "and that 'sensual cognition represents ... the way in which the things affect our senses'".[x] To be sure, for Kant, to speak of "physical influence" upon

[x]Gödel does not name the locus of this passage; it occurs in *Prolegomena* §13, Anmerkung III, p. 290.

us of things in themselves may appear quite impossible. The notions of the "physical", or "natural", have for Kant a necessary restriction to what is *given* (in some kind of "intuition", for us necessarily *sensible*);[y] these notions cannot, therefore, be applied to noumena. But to say this is only to say that Kant was a Kantian, on precisely the point at which Gödel *states* that he himself is not. Remove the problematic adjective: Gödel is pointing out to us that Kant explicitly speaks of "the things"—presumably the same that he has called "unknown in themselves"—as *affecting* our senses; and in a way that is "represented" by "sensual cognition". His non-Kantian gloss is that when we have learned that our theoretical physics is able to make constructions of a kind unforeseen by Kant, and not related in the way Kant thought necessary to the character of our "sensible intuition", the possibility is opened of our extending the notion of the "physical" after all into the domain Kant thought inescapably closed to it.

7. The foregoing rather lengthy discussion has been concerned with a matter on which Gödel differs critically—and, I have said, consciously and deliberately—with Kant. And yet, as also intimated above, in some respects—in some *places*—Gödel seems to minimize this difference, even to the point of denying it altogether. These must now, briefly, be pointed out.

In *B2*, page 14, one finds the statement that "the Kantian a priori is incompatible with relativity theory only in one minor point (made by Kant in only one rather obscure passage) and ... Kant's thesis of the apriority of Newtonian time and Euclidean space as he meant it does not contradict relativity theory at all." What follows makes it clear that the sense in which Gödel takes Kant to have "meant" his thesis is one that restricts its application to the structure of what is given at each moment in sense perception; thus, that his reading of Kant here is the one that has already been criticized above: that space "expresses itself" *only* in the moment-by-moment character of our "outer intuition". But what in the world can Gödel mean by the "one minor point" in which Kant's view *is* incompatible with relativity theory; and in what "rather obscure passage" does he say Kant makes that point?

On page 18 we find: "A real contradiction between relativity theory and Kantian philosophy seems to me to exist only in one point, namely, as to Kant's opinion that *natural science in the description it gives of the world must necessarily retain the forms of our sense perception and*

[y]Cf. A 845/B 873: "... *nature*, that is, the totality of *given* objects (whether they are given to the senses, or, if one will, to another sort of intuition)".

can do nothing else but set up relations between appearances within this frame" (emphasis added). That is admirably put, and seems exactly correct. But if this is the single point at which a contradiction occurs between Kant and relativity theory, it must be the point referred to on page 14. How can this point be deemed "minor"? It is not only a point of extremely far-reaching consequences for the philosophy, and the practice, of natural science; it is also absolutely central to Kant's critical philosophy: it is the heart of his doctrine of "transcendental idealism". It does, however, seem to follow from the two passages just cited that this is the point Gödel has called minor. As to the "obscure passage" that he refers to as the only place in which the point is made, Gödel has not told us which place he means, and I am completely unable to guess.

But this is certainly a matter in which we must bear in mind what has been emphasized in §**1** of this introduction—that we do not have from Gödel's hand a final considered version of this essay: indeed, it should be noted that the passage cited from *B2*, page 14, has been revised in *C1* (page 23) so as to *eliminate* reference to the "minor point" and the "obscure passage".

There is another place where version *B2* as here published contains a statement about Kant that is both drastically incorrect and, in my opinion, most likely *not* representative of Gödel's own considered reading of Kant. On pages 16–17, we read:

> This insight ⟦namely, that projection of our sensations into a three-dimensional Euclidean space might be "of no use for stating the laws governing them"⟧ does not necessitate any weakening of Kant's thesis of the subjectivity of space, since *the sensations also* are something subjective (based on the actions of the things on us) and therefore any of their properties may as well be due to subjective as to objective conditions, i.e., to properties of the senses as well as to properties of the objects. Nor does it necessitate a weakening of his thesis of the apriority of space, since Kant did not assert that the "adequacy of our representation of space to the relation of the objects to our sensibility" can be known a priori. [That some particular relations of the things to us are adequately represented by Euclidean geometry in every world ... certainly does not mean yet "a complete adequacy to *the* relation of the things to our sensibility".]

The history of this passage through the successive drafts is rather complex, and need not be discussed here in detail. But it is crucial to remark that in *B1*—which in this substantially agrees with *A*—the second sentence of the quoted paragraph reads:

> It does however necessitate a weakening of his thesis of the a priority of space, in that ⟦*sic*⟧ sense that the "adequacy of our representation of space to the relation of the objects to our sensibility" can only be known by experience, provided a reasonable meaning is given to this phrase.

The sentence thus says the opposite of what it says in the version printed below (and cited above). Now, what one finds in the *B2* typescript is, first, the *same* text as in *B1*: nothing in the sentence as just quoted from *B1* is canceled. *Then*, however, there are written in, above the *uncanceled* text, the following things: above the word 'It', the word 'Nor'; above 'however', 'it'; above the sequence 'sense that the "adequacy', the phrase 'since Kant did not assert'; and after the sequence 'to our sensibility" '—above a parenthesis that had already been canceled in *B1*—the phrase 'can be known a priori' (with no mark of punctuation—no period to stop the sentence). The remainder of the original sentence, from 'can only be known' to the end, is left on the following lines with no modification.

It will be seen from this that Gödel wrote in what can only be regarded as a *tentative* revision (since he did not cancel the original, or indicate at all which words the new ones were to replace). To be sure, the substance of the new words and phrases and the grammatical structure of the sentence determine uniquely how the revised sentence must read; but the manuscript does not give a decisive sign that the revision was judged favorably by Gödel in the end. The corresponding passage in *C1* (page 25) agrees with *B1* (and *A*).

It should be added that, on the substantive issue of Kantian interpretation, the passage as revised in *B2* is quite indefensible. For instance, in the *Prolegomena*, shortly after the passage cited by Gödel on the complete adequacy of our representation of space to the relation which our sensibility has to the objects, Kant says (page 292; emphasis added):

> Thus it is so far from true that my doctrine of the ideality of space and time makes the whole sensible world a mere illusion, that it is rather the only means of rendering secure the application of one of the most important ⟦kinds of⟧ knowledge—namely, that which mathematics propounds *a priori*—to actual objects, ... because without this observation it would be entirely impossible to tell whether the intuitions of space and time ... were not mere phantoms of the brain, made by ourselves, to which there corresponds no object whatever, *at least not adequately*, so that geometry itself should be a bare illusion ....

In Kant's own opinion, therefore, the only possible grounds of assurance of the adequacy of our spatial representation are the grounds provided by his transcendental idealism; and these, he says, are *a priori* grounds.

It seems fair to conclude that the same impulse to minimize Kant's differences with him—i.e., his with Kant—that we have already seen operating, has here led Gödel *at one moment* to venture an interpretation that is really impossible, and that he himself may in all likelihood have come to think better of.

8. There is another issue of Kantian interpretation that should be noted, although it is of no importance for Gödel's general considerations. Certain remarks of Gödel's assume (and appear to take for granted as uncontroversial and well-known characteristics of Kant's doctrine) that what Kant calls "sensibility" is divided into two faculties: the "faculty of sensation" and the "faculty of representation"; and that the latter of these is "something intellectual".[z] But this is quite inauthentic; and the question arises, just what Gödel himself means here and to what in Kant this distinction of Gödel's may be taken to correspond.

Two possibilities suggest themselves. The first is that he is thinking of Kant's distinction of the *matter* of intuition and its *form*. This reading is supported by the considerations (a) that Gödel calls the "first part" the faculty of *sensation*, and Kant says that "in the appearance I call that which corresponds to sensation ⟦*Empfindung*⟧ its *matter*" (A 20/B 34); and (b) that space and time are indeed, of course, according to Kant, *forms* of appearances, or of intuition. But it seems very strained, on this interpretation, that Gödel calls these two parts two "faculties"; that he calls the second part—the *form* of intuition—"the faculty of *representation*"; and that he characterizes this as "something *intellectual*". The alternative appears to be to identify Gödel's "faculty of representation" with the *productive imagination* of Kant. This is at least, in a certain sense, on the way towards being "intellectual"—i.e., a function of the *understanding*. Thus Kant says (B 151–2): "The imagination ... belongs to *sensibility*; yet to the extent that its synthesis is an exercise of spontaneity, which is determining and not merely like the sense determinable, ... to this extent imagination is ... a faculty to determine the sensibility *a priori*; and its synthesis of intuitions, conformable with the categories, must be the transcendental synthesis of *imagination*, which is an action of the understanding on the sensibility ... ." The supposition that what Gödel meant by the "second part of sensibility" was the (productive) imagination has the advantage of making it clearer how he can speak of it as a distinct faculty, and as something intellectual. But if this *is* what Gödel meant, it seems all the stranger that he offered no

[z]See *B2*, p. 4b; *C1*, p. 17.

fuller discussion and no textual references whatever on what is surely a rather subtle part of Kant's doctrine.[aa]

This second way of taking Gödel perhaps gains some support from the only passage that can be seen as *introducing* the distinction to which he later refers as if it were already understood. This is, however, not a passage in Gödel's text, but only in a note; and is itself rather cryptic as to there being "two parts" to the faculty: "'Sensibility' is, according to Kant, the faculty of having sensations under the influence of objects and happenings outside or inside ourselves, *and forming images* of outer objects and psychic processes out of these sensations." (*B2*, page 1, footnote 3; slightly revised version in *C1*, page 2, footnote 8; emphasis added.)

9. Several matters of interest present themselves for comment concerning the philosophical position taken by Gödel himself in these essays, independently of its relation to Kant; but the length to which this introduction has already grown suggests restraint at this point. One such matter, however, must be briefly noted, because it involves a bit of textual evidence that deserves to be put on record.

In the introductory note to *Gödel 1949a*, a question is raised as to why Gödel should have thought that it was necessary to defend the physical non-absurdity of his R-worlds by establishing that the "time travel" *physically* possible in them would not be *practically* possible (see these *Works*, Volume II, page 199, footnote a). The argument given by Gödel in *1949a*, page 561, appears in nearly the same words (but in a weaker form—cf. remark (c) of Malament quoted in §**2** above) in *C1*, page 12, footnote 15. However, Malament has informed me of an earlier version of that note, in which Gödel himself presents exactly the case that is made in the cited introductory note. (This is in one respect very reassuring to the commentator; in another, disconcerting: why did Gödel change his mind?) I quote Malament's minute of this text (I have made one emendation, in brackets, of a typographical error; the initial and final passages in brackets are Malament's own, bracketed in his typescript):

[aa]Since this section was written, a new piece of evidence has been found that lends support to the second alternative discussed, and provides a possible source for the phrase used by Gödel. *Kowalewski 1924* (a volume that Gödel could well have seen) contains the notes on Kant's lectures during a sequence of three semesters (Winter 1791/92 – Winter 1792/93) by one of his auditors. On p. 106 of that volume, corresponding to p. 40 of the manuscript notes on Kant's course "Anthropologie", one finds the boldface heading "Von der Einbildungskraft und Phantasie" ("On the imagination and phantasy"); and immediately following that heading, the sentence (within brackets, indicating—*ibid.*, p. 52—that in the manuscript it appears as a marginal note): "Sie ist das 2te Stück der Sinnlichkeit" ("It is the second part of sensibility").

[The original version of the footnote includes a passage that was subsequently crossed out. It is here restored and typed in boldface. The footnote addresses the standard objection that "time travel" is absurd: if it were possible, one could go back in time and undo the past.]

This state of affairs seems to imply an absurdity. For it enables one e.g. to travel into the near past of those places where he has himself lived. There he would encounter a person which would be himself so and so many years ago. Now he could do something to this person which he knows by his own memory has not happened to him. This and similar contradictions, however, **presuppose, not only the practical feasibility of the trip into the past (velocities very close to that of light would be necessary for it) but also certain decision⟦s⟧ on the part of the traveler; whose possibility one concludes only from vague conviction of the freedom of the will. Practically the same inconsistencies (again by neglecting certain "practical" difficulties) can be derived from the assumption of strict causality and the freedom of the will in the sense just indicated. Hence, as far as the paradoxical situation under consideration is concerned, an R-world is not any more absurd than any world subject to strict causality.**

[Gödel's response (in boldface) seems exactly right. Why did he cross it out and emphasize instead the "practical difficulties of time travel"?]

It seems fitting to conclude with another quotation, bearing on the same issue, containing an early recognition of the possibility of such results as Gödel's, written by another great mathematician and philosopher (and colleague of Gödel and Einstein). The following occurs in *Weyl 1918a*, page 220:

From each world-point there proceeds the double cone of active future and passive past. Whereas in the special theory of relativity these are separated by an intermediate region, it is here ⟦in the general theory⟧ intrinsically quite possible that the cone of the active future encroaches upon that of the passive past; it can therefore happen, in principle, that I now experience events which arise, in part, only as effects of my future decisions and actions. Neither is it excluded that a world-line, though it has time-like direction at each point (in particular, the world-line of my body), returns to the neighborhood of a world-point that it has already once passed. From this there would result a more radical sort of Doppelgängerdom than ever an E. T. A. Hoffmann has thought up. In point of fact, such considerable variations of the g_{ik} as would be required for these effects do not occur in the world-region in which we live; yet to think through these possibilities, as they bear upon the philosophical problem of the relation of cosmic and phenomenal time, has a certain interest. However paradoxical what thus comes to light, a genuine contradiction to the facts given immediately in our experience nowhere arises.

Howard Stein[bb]

Kant's *Critique of Pure Reason* is cited below in the second edition (*Kant 1787*) in the standard way, i.e., the page number preceded by "B". (Gödel does not cite or refer to the first edition *1781*.) Where convenient, these page references are inserted into the text. Page references to Kant's *Prolegomena* (*Kant 1783*) have been supplied by the editors and refer to volume IV of the Academy edition of Kant's works (*Kant 1902–*).

All translations of quotations from Kant are Gödel's own.

The editors wish to acknowledge Howard Stein's considerable contribution to the editing of these texts, in addition to his writing the introductory note. We also relied on an earlier examination and classification of Gödel's drafts by David Malament and Stein.

[bb]I am indebted to Charles Parsons for helpful discussion both of textual points and issues of Kantian interpretation.

David Malament's contribution to this commentary is indicated only in part by the references to him in the foregoing. Numerous conversations with him concerning mathematically obscure points in Gödel's essay have been an invaluable aid to my own attempts to clarify them; I here record my gratitude.