The twins and the bucket: How Einstein made gravity rather than motion relative in general relativity

Michel Janssen

May 14, 2011

Abstract

In publications in 1914 and 1918, Einstein claimed that his new theory of gravity somehow relativizes the rotation of a body with respect to the distant stars (a stripped-down version of Newton’s rotating bucket experiment) and the acceleration of the traveler with respect to the stay-at-home in the twin paradox. What he showed was that phenomena seen as inertial effects in a space-time coordinate system in which the non-accelerating body is at rest can be seen as a combination of inertial and gravitational effects in a (suitably chosen) space-time coordinate system in which the accelerating body is at rest. Two different relativity principles play a role in these accounts: (a) the relativity of non-uniform motion, in the weak sense that the laws of physics are the same in the two space-time coordinate systems involved; (b) what Einstein in 1920 called the relativity of the gravitational field, the notion that there is a unified inertio-gravitational field that splits differently into inertial and gravitational components in different coordinate systems. I provide a detailed reconstruction of Einstein’s rather sketchy accounts of the twins and the bucket and examine the role of these two relativity principles. I argue that we can hold on to (b) but that (a) is either false or trivial.

Keywords: relativity of the gravitational field, twin paradox, rotating bucket, equivalence principle, kinematical relativity, general covariance.

1 Der glücklichste Gedanke meines Lebens

It is so obvious today that the general theory of relativity does not extend the relativity principle from uniform to arbitrary motion that it has become something of a puzzle how Einstein could ever have claimed it did.\footnote{See Norton (1993) for an illuminating discussion of the history of the debate over this issue.} I have argued in several places that, to a large extent, the solution of this puzzle is simply that what Einstein called the relativity of non-uniform motion is more appropriately called the relativity of the gravitational field.
Two observers in non-uniform motion with respect to one another can both claim to be at rest as long as they agree to disagree about whether or not there is a gravitational field. This requires a coordinate-dependent definition of the gravitational field, but, unlike modern relativists, Einstein opted for just such a definition, representing the gravitational field by the so-called Christoffel symbols.

The relativity of the gravitational field is illustrated in Einstein’s admittedly sketchy accounts on the basis of his metric theory of gravity of two celebrated exhibits in the case for absolute motion, the twin paradox (Einstein, 1918e) and Newton’s rotating bucket experiment (Einstein, 1914, pp. 1031–1032). Einstein reduced the bucket experiment to its bare essentials. He replaced the bucket of water by some unspecified rotating body and left out the gravitational field of the earth altogether. The account is part of the introduction to the definitive exposition of what is now known as the Entwurf theory, first proposed by Einstein and Grossmann (1913). In hindsight this theory was only a preliminary non-covariant version of Einstein’s metric theory of gravity. To the extent that Einstein’s account of the bucket experiment works at all, it only does in the generally-covariant version of the theory.

What Einstein’s accounts of the twins and the bucket show is that certain phenomena seen as inertial effects (i.e., effects of acceleration) in a space-time coordinate system in which the component of the system moving on a geodesic remains at rest are seen as a combination of inertial and gravitational effects in a (suitably chosen) space-time coordinate system in which the component moving on a non-geodesic remains at rest. In the twin paradox, the stay-at-home is moving on a geodesic, the traveler is not, and they will accordingly offer different explanations for their age difference upon reunion. In Einstein’s version of the bucket experiment, the stars are moving on geodesics, the rotating body is not, and an observer rotating with this body will offer a different explanation for its bulging out—the effect taking the place of the concave shape of the surface of the water in the original version of the bucket experiment—than an observer at rest with respect to the stars.

As these two examples illustrate, the relativity of the gravitational field was intertwined for Einstein with the relativity of non-uniform motion in the sense that the laws of physics are the same in space-time coordinate systems moving non-uniformly with
respect to one another. In this paper, I provide detailed reconstructions of Einstein’s accounts of the twins (sec. 2) and the bucket (sec. 3) and examine the role of the two relativity principles involved. I show that Einstein failed to extend the relativity of motion from uniform to arbitrary motion in any non-trivial sense. This means that his accounts of the twins and the bucket are also problematic as illustrations of the relativity of the gravitational field. I argue, however, that the principle itself can still be rescued.

The relativity of the gravitational field is directly related to the equivalence principle. In a short but important paper on the foundations of general relativity, Einstein (1918b, p. 241) defined the equivalence principle as the fundamental identity [Wesensgleichheit] of gravity and inertia. This means, he elaborated, that inertial effects and gravitational effects must be manifestations of one and the same field, the metric field, $g_{\mu\nu}(x)$. Einstein proposed to call it the “$G$-field” (ibid.). I will use the modern term inertio-gravitational field. Even though he did not say so explicitly, the relativity of the gravitational field is central to this 1918 definition of the equivalence principle. General relativity, according to Einstein, teaches us that inertio-gravitational fields break down differently into inertial and gravitational components for observers in non-uniform motion with respect to one another, just as special relativity taught us that electromagnetic fields break down differently into electric and magnetic components for observers in uniform relative motion. It is this split of the inertio-gravitational field into inertial and gravitational components that, according to Einstein, is relative in general relativity.

Unfortunately, I know of only one text in which Einstein elaborated on the equivalence principle in this way and explicitly talked about the relativity of the gravitational field. To make matters worse, this text, “Fundamental ideas and methods of the theory of relativity, presented in their development,” written in December 1919/January 1920, was never published during his lifetime. It had been intended for Nature, but was withdrawn in the end and replaced by a much shorter and less informative piece (Einstein, 1921a).\(^2\)

The full text of the original article was eventually published in 2002 in Vol. 7 of The Collected Papers of Albert Einstein (Einstein, 1987–2009, Doc. 31, pp. 245–281). Long before the article finally appeared in print, however, one sentence in it had already become famous: “Then came to me the happiest thought of my life” [der glücklichste Gedanke

\(^2\)See Stachel (2007) for key passages and discussion of both pieces.
[meines Lebens]. This sentence on [p. 20] of the article concludes a paragraph in which Einstein recounted how in 1907 he had embarked on the journey that would lead him to general relativity. As he explained in the first sentence of the next paragraph, his “happiest thought”—or, as I prefer, his “most fortunate thought”—was precisely the relativity of the gravitational field:

In an example worth considering, the gravitational field only has a relative existence in a manner similar to the electric field generated by electro-magnetic induction.\(^3\) Because for an observer in free-fall from the roof of a house, there is during the fall—at least in his immediate vicinity—no gravitational field [emphasis in the original]. Namely, if the observer lets go of any bodies, they remain, relative to him, in a state of rest or uniform motion, independent of their special chemical or physical nature. The observer, therefore, is justified in interpreting his state as being “at rest” . . . The experimental fact that the acceleration in free fall is independent of the material, therefore, is a powerful argument in favor of extending the postulate of relativity to coordinate systems moving non-uniformly relative to each other (Einstein, 1987–2009, Vol. 7, p. 265; the translation follows A. Engel in the translation volume accompanying the Einstein edition).

Commenting on this passage before, I endorsed the claim that “the gravitational field only has a relative existence” but dismissed the inference from the uniqueness of free fall to the extension of the relativity principle from uniform to non-uniform motion (Janssen, 2002, 507–508). Both the endorsement and the dismissal may have been overhasty. The extension of the relativity of motion is not quite as nonsensical and the relativity of the gravitational field is not quite as unproblematic as I suggested.

To start with the former, Dicks (2006) has argued persuasively that we should take seriously that the extension of the relativity principle proposed in this passage (and in other Einstein texts) is to coordinate systems in non-uniform motion with respect to one another. Einstein is not proposing what I will call the strong version of the relativity of arbitrary motion. This version requires that, if two bodies \(A\) and \(B\) are accelerating with respect to one another, it makes no difference whether \(A\) or \(B\) is accelerating. Instead, Einstein proposed what I will call the weak version of the relativity of arbitrary motion. This version requires not that both cases—\(A\) accelerating, \(B\) accelerating—are the same

---

\(^3\)Einstein had just covered the magnet-conductor thought experiment with which he opened the 1905 special-relativity paper. He used this same analogy in his 1914 discussion of the bucket experiment (Einstein, 1914, p. 1032).
but only that they can be accounted for on the basis of laws that hold in the same form in a coordinate system moving with \( A \) and in a coordinate system moving with \( B \).

*Prima facie*, the move from the strong to the weak version—from an equivalence between physical systems to an equivalence between coordinate systems used to describe these systems—may seem to take us merely from the patently false to the utterly trivial. To halt this slide into triviality, coordinate systems need to be endowed with some spatio-temporal meaning, enough, at a minimum, to define their motion with respect to one another. If coordinates are stripped of all spatio-temporal meaning, the equivalence of arbitrary coordinate systems is nothing but general covariance. In that case *any* theory that can be recast in generally-covariant form, including Newton’s theory and special relativity, would satisfy the weak version of the relativity of arbitrary motion, though the term would be misleading as it would be impossible even to define motion without reference to additional geometrical structure.

As we will see, it is essential both to his 1914 account of the bucket and to his 1918 account of the twins that the coordinate systems involved retain a good deal of spatio-temporal meaning. Yet, in the same 1918 article in which he gave his account of the twins, Einstein (1918e, p. 699) wrote that in general relativity space-time coordinates can be “parameters devoid of any independent physical meaning” (p. 699), thereby inviting the trivial reading of the weak version of the general relativity principle. That this reading is indeed trivial was, of course, only driven home in the 1920s with simple generally-covariant reformulations of Newtonian theory (Norton, 1993, sec. 5.3). In the paper on the foundations of general relativity mentioned above, Einstein (1918b, p. 242) could still defend himself against Kretschmann’s (1917) charge that general covariance is physically vacuous by arguing that a generally-covariant reformulation of Newtonian theory would look so contrived that one would immediately recognize the superiority of general relativity (Norton, 1993, sec. 5.2).

Yet, Einstein’s insistence that coordinates are just arbitrary labels only got stronger as time went on. His autobiographical notes, written in 1948, provide a good example. After making the same inference from the equality of inertial and gravitational mass to the extension of the relativity of motion (in the weak sense) as in the passage from 1919/1920 quoted above, he wrote:
This happened in 1908 [Einstein misremembered: this should be 1907]. Why were another seven years required for the construction of the general theory of relativity? The main reason lies in the fact that it is not so easy to free oneself from the idea that co-ordinates must have an immediate metrical meaning (Einstein, 1949, p. 67).

Discussing this passage and a similar earlier one from the published version of a lecture Einstein (1933, p. 288) gave in Glasgow, Stachel (2007, pp. 86–87) points out that by 1912 Einstein had already abandoned the idea that coordinates have direct metrical significance, an idea that had been extremely helpful to him in the formulation of special relativity. 1912 is only four years after 1908—not seven. What Einstein only freed himself from in 1915, Stachel suggests, is the idea that coordinates uniquely identify and individuate space-time points. This idea he gave up in the context of the resolution of the so-called ‘hole argument’ that had convinced him earlier that the field equations of his metric theory of gravity could not be generally covariant. That coordinates do not even identify and individuate space-time points can be used to argue against the notion that space-time would be some sort of substance (Janssen, 2011a, sec. 3). The complete loss of physical significance of coordinates, however, does trivialize the extension of the relativity of motion (in the weak sense) that Einstein routinely claimed was suggested by the equality of inertial and gravitational mass.

As we will see in secs. 2 and 3, the coordinates that Einstein actually used in his accounts of the twins and the bucket in the 1910s have essentially the same status as those in special relativity. They have not lost all metrical significance and they still identify and individuate space-time points uniquely. So what Einstein sought to establish in these accounts was a non-trivial version of the relativity of non-uniform motion in the weak sense—although, as we will see, he initially thought that his 1914 account of the bucket even established relativity of rotation in the strong sense. It took him until 1917 to recognize that it does not. As far as I can tell, Einstein never acknowledged the more subtle problems that, I will argue, defeat his attempts to establish the relativity of non-uniform motion even in the weak sense.

It turns out that to obtain the equivalence between space-time coordinate systems that Einstein was after (between that of the traveler and that of the stay-at-home; between that of the rotating body and that of the distant stars) we have to be careful in our choice
of the coordinate system in which the accelerating party remains at rest. Einstein did not show any concern about the element of arbitrariness in the choice of coordinate systems in his accounts of the twins and the bucket. I suggest that this is directly related to his conviction—explicit in a handful of passages (e.g., Einstein, 1933, p. 286), implicit in others (e.g., Einstein, 1916, pp. 112–113)—that something he called *kinematical* (as opposed to *physical*) relativity is trivially satisfied. In any sensible physical theory, Einstein thought, it is easy to come by space-time coordinate systems in arbitrary states of motion that are all perfectly equivalent to one another as far as the *description* of phenomena is concerned (hence the term ‘kinematical’). However, it is only in general relativity, or so Einstein claimed, that such space-time coordinate systems are also equivalent when it comes to providing a dynamical (or, as Einstein preferred, ‘physical’) *explanation* of these phenomena (in the sense that the phenomena can be accounted for by the same laws in the same form in any of these coordinate systems). Only in general relativity, according to Einstein, do we thus have both kinematical and dynamical (or physical) relativity.

If coordinates do not have any spatio-temporal meaning, kinematical relativity does indeed obtain trivially, though the connotation with motion of the term ‘kinematical’ would be misleading in that case. However, to reiterate, the coordinates in Einstein’s accounts of the twins and the bucket do have spatio-temporal meaning. In that case, as Dorling (1978) has pointed out, though he did not use this terminology, kinematical relativity does not hold for non-uniform motion in either special or general relativity (though it does in Newtonian theory). Amplifying Dorling’s argument, which has not received the attention that I think it deserves, I show in detail how it breaks down in the case of the twins. I consider two possible space-time coordinate systems for the traveler. Kinematical relativity ultimately fails for both, but the one Einstein implicitly chose in 1918 at least gave him a fighting chance to show that his new theory of gravity establishes the kind of physical equivalence he envisioned between the coordinate systems of the two twins. The relativity of the gravitational field is central to this attempt. The phenomenon to be explained is the difference in aging between the two twins. One twin will say that the effect is due to a purely inertial field; the other, Einstein argued, will say that it is due in part to an inertial and in part to a gravitational field.

---

4 See the appendix for a translation of a pair of similar passages (Einstein, 1921c,d).
Similar conclusions can be drawn from Einstein’s account of the bucket. With the right choice of a rotating space-time coordinate system in Minkowski space-time and replacing the non-covariant field equations of the Entwurf theory by the generally-covariant ones of its successor, one can demonstrate physical relativity in this case. In terms of Einstein’s stripped-down version of the bucket experiment, one can show that the bulging out of a rotating body can be explained on the basis of the same physical laws in the coordinate system in which the body is at rest and in the coordinate system in which the distant stars are at rest. Once again, the relativity of the gravitational field is key. In the coordinate system in which the stars are at rest, the bulging out is a purely inertial effect; in the coordinate system in which the body is at rest, it is a combination of inertial and gravitational effects. Contrary to what Einstein believed until 1917, this does not mean that general relativity predicts the same bulging out if we could somehow set the stars spinning while keeping the body at rest. What it does mean is that there is a different perspective on the situation of the rotating body in which its bulging out is a combination of inertial and gravitational effects (Janssen, 2011a, sec. 4). In other words, even if we accept Einstein’s account of the bucket, it would still only establish the relativity of rotation in the weak and not in strong sense.

The analysis of Einstein’s accounts of the twins and the bucket presents us with a dilemma. If coordinates are allowed to keep enough of their spatio-temporal meaning to define motion, the equivalence of coordinate systems in non-uniform motion with respect to one another already breaks down at the descriptive level. If coordinates are completely stripped of their spatio-temporal meaning, general covariance guarantees the equivalence of all coordinate systems regardless of their state of motion, but such equivalence is trivial in that it is no longer characteristic of general relativity. Both kinematical and physical relativity, as Einstein defined them, would obtain in generally-covariant reformulations of Newtonian theory and special relativity as well. The way out of the dilemma, I suggest, is to recognize that, while the slide into general covariance turns the relativity of non-uniform motion of space-time coordinate systems into a feature general relativity shares with older theories, it does not so trivialize the relativity of the gravitational field. Even in generally-covariant reformulations of these older theories, there will be an inertial field and a gravitational field existing side by side. The unification of these two fields
into one inertio-gravitational field that splits differently into inertial and gravitational components in different coordinate systems is one of Einstein’s central achievements with general relativity. These coordinates can be given spatio-temporal meaning with the help of the metric field, but the resulting space-time coordinate systems will in general not be equivalent to one another. In this way, it seems to me, we can hang on to the relativity of the gravitational field while abandoning even the weak version of the relativity of arbitrary motion.

The main purpose of this paper, however, is of a historical nature. Much of the literature on the puzzle how Einstein could claim that his new theory of gravity made all motion relative revolves around his struggle with general covariance (Norton, 1993). Ever since Kretschmann (1917) first pointed this out, it has been clear that general covariance should not be conflated with (the strong version of) the relativity of arbitrary motion. Several commentators have tried to make sense of Einstein’s insistence on the importance of general covariance for general relativity without accusing him of this conflation. In the most ingenious attempt along these lines that I am aware of, Norton (1992) argued that Einstein achieved a general relativity of motion of sorts by using general covariance in the spirit of Klein’s projective geometry rather than in the spirit of modern Riemannian differential geometry. As I did above, Norton emphasized that coordinate systems had a much richer geometrical meaning for Einstein as he was working towards general relativity than they have in modern differential geometry. Norton (1999) subsequently dropped the suggestion that Einstein’s project succeeded on its own terms and retreated to the position that Einstein conflated the geometrical traditions associated with Riemann and Klein, respectively, and, as a result, did not recognize that the status of general covariance in general relativity is very different from the status of Lorentz invariance in special relativity. I agree with this analysis (Janssen, 2005, pp. 61–61) but would add that it only captures one of several strands in Einstein’s ultimately unsuccessful quest for a general relativity of motion (Janssen, 2005, 2011a). Einstein’s notions of the relativity of the gravitational field and of kinematical versus physical relativity played important roles in some of these other strands. These notions have received far less attention than general covariance, partly because they get so little emphasis in Einstein’s own writings. Yet we need to come to terms with them if we want to fully understand Einstein’s failed
quest for a general relativity of motion. A good place to start, it seems to me, is to analyze their role in Einstein’s accounts of the twins and the bucket.

2 The twins

2.1 The twins in Einstein’s 1918 Dialogue

In an article entitled “Dialogue about objections to the theory of relativity,” published in the semi-popular science magazine Die Naturwissenschaften in November 1918, Einstein (1918c) offered an analysis of the twin (or clock) paradox purporting to show that both twins can claim to remain at rest from separation to reunion, as long as they agree to disagree about the presence of gravitational fields. What Einstein wanted to demonstrate with this analysis is not that the acceleration of the “traveler” would be relative after all, in the sense that the situations of “traveler” and “stay-at-home” would be fully symmetric, but that, using the same laws of physics, we can account for the difference in aging between the two of them in coordinate systems co-moving with either one. As we will see, Einstein’s argument only works for a natural but ultimately arbitrary and problematic choice of the space-time coordinate system moving with the traveler.

Before examining this account of the twins, I need to say something about the article in which it appeared. In the Einstein edition, the Dialogue is preceded by the editorial note, “Einstein’s encounters with German anti-relativists” (Einstein, 1987–2009, Vol. 7, pp. 101–113). As this note makes clear, the Dialogue should be seen first and foremost in the context of the ‘anti-relativity movement’.5 This does not mean, however, that Einstein did not intend the arguments he advanced in his Dialogue to be taken seriously by experts in the field. One can argue (Renn, 2007, Vol. 1, p. 16) that it was only with some papers he published in the first half of 1918 that Einstein put the finishing touches on his formulation of the generally-covariant 1915 version of his theory. In addition to the paper mentioned in the introduction, which is explicitly devoted to the foundations of general relativity (Einstein, 1918b), these are papers on gravitational waves (Einstein, 1918a), the De Sitter solution (Einstein, 1918c), and energy-momentum conservation

---

(Einstein, 1918d). Given that he wrote it shortly after these papers, I am inclined to read the *Dialogue* as expressing Einstein’s mature views of the foundations of the theory, even though he wrote it for a broad audience.

![Minkowski diagram for the twin paradox](image)

Figure 1: Minkowski diagram for the twin (or clock) paradox with a “stay-at-home” clock/twin $U_1$ and a “traveler” clock/twin $U_2$. The black dots on the worldlines of $U_1$ and $U_2$ mark the years passing for the two clocks/twins.

Fig. 1 shows a standard Minkowski diagram for the twin paradox. At $D$ (for departure), clock $U_2$ takes off to the right with velocity $v$, while clock $U_1$ stays put. The labels $U_1$ and $U_2$ are those of the *Dialogue*. I will use these labels to refer either to the clocks themselves or to the twin observers traveling with them. I trust it will be clear from context whether I am referring to clocks or twins. At $R$ (for reversal), $U_2$ turns around and starts moving back to $U_1$, again with velocity $v$. At $A$ (for arrival), $U_1$ and $U_2$ are reunited. Special relativity tells us that, if $U_1$ ages $2T$ during the journey $DSA$, $U_2$ only ages $2T/\gamma$ during the journey $DRA$ (with $\gamma \equiv 1/\sqrt{1-v^2/c^2}$; in Fig. 1, $\gamma = 1.5$ [so $v \approx .75c$] and $T = 4.5$ years). $U_1$ and $U_2$ are thus clearly not equivalent. The accelerations of $U_2$ at $D$, $R$, and $A$, produce an asymmetry between them.

In between these accelerations, the situation is completely symmetric. For the description of the situation in between accelerations to be symmetric as well, $U_2$ has to

---

6For elementary discussion, see, e.g., Janssen (2011b, sec. 2.7).
switch from the space-time coordinates of the inertial frame moving with him to the
right before $R$ to the coordinates of the inertial frame moving with him to the left after
$R$. Using these instantaneous rest frames, $U_2$ finds that, while the two halves of his own
trip ($DR$ and $RA$) each take an amount of time $T/\gamma$, the corresponding parts of $U_1$’s trip
($DP$ and $QA$) each only take $T/\gamma^2$. In between accelerations, both $U_1$ and $U_2$ thus find
that the other person ages more slowly than he does by the same factor of $\gamma$ (in Fig. 1,
3 years versus 4.5 years for $U_1$; 2 years versus 3 years for $U_2$).

In Fig. 1, $U_2$ turns around instantaneously at $R$. This can be seen as the limiting case
of $U_2$ turning around in a short but finite time interval $\Delta t \ll T$ around $R$. To order $v^2/c^2$
(which, as we will see, is all that matters), it makes no difference whether the clock at $R$
measuring $\Delta t$ is at rest with respect to $U_1$ or with respect to $U_2$. It will be convenient
to assume that, during $\Delta t$, $U_2$ undergoes a constant acceleration $a$ with respect to $U_1$,
changing its velocity from $+v$ to $-v$. The situation shown in Fig. 1 (and in all other
figures in this paper) corresponds to the limit in which $a \to \infty$ and $\Delta t \to 0$ while their
product $a\Delta t = 2v$ remains fixed. In this limit, the line connecting events simultaneous
to $R$ from $U_2$’s point of view sweeps through the shaded area in Fig. 1, from the line
through $R$ and $P$ to the line through $R$ and $Q$. Hence, from the point of view of $U_2$,
using the space-time coordinates of different instantaneous rest frames before and after
$R$, $P$ and $Q$ are simultaneous, which means that a significant amount of time (5 years in
Fig. 1) elapses on $U_1$ as $U_1$ turns around. This is yet another indication that the states
of motion of $U_1$ and $U_2$ are not equivalent.

In the Dialogue, Einstein, appealing to the equivalence of the effects of acceleration
and gravity, nonetheless gave an account of the situation in which $U_2$ rather than $U_1$
remains at rest the whole time. From $U_2$’s point of view, according to this account, a
gravitational field pops in and out of existence throughout space when he has reached
the midway point $R$ of his journey.\footnote{Two complications are passed over in silence here even though Einstein had been aware of them for many years. In 1912 he had realized that the homogeneous gravitational field corresponding to constant (Born) acceleration in Minkowski space-time will not be a solution of the matter-free gravitational field equations if the gravitational field acts as its own source as it must because of the equivalence of mass and energy and the equality of inertial and gravitational mass. In 1908 Einstein had already realized that the homogeneous gravitational field substituted for constant acceleration only gives the right gravitational forces for particles at rest in that field. See Janssen (2011a, sec. 2) for further discussion and references.} This gravitational field, produces an acceleration
equal and opposite to what from $U_1$’s point of view is the acceleration of $U_2$ as he turns
around. The non-gravitational force acting on \(U_2\) in the vicinity of \(R\), responsible for making \(U_2\) reverse course from \(U_1\)'s point of view, balances the gravitational force, so that \(U_2\) stays put. \(U_1\) falls freely in the gravitational field and changes its velocity with respect to \(U_2\) from \(-v\) to \(+v\). There will be similar transient gravitational fields, with the same difference in response between \(U_1\) and \(U_2\), at the moment of departure \(D\) and the moment of arrival \(A\). Since \(U_1\) and \(U_2\) are at the same gravitational potential at \(D\) and \(A\), these fields will not cause any difference in the time that elapses on the two clocks or in the aging of the two twins. This is different during turn-around. At that moment, \(U_1\) is at a higher potential in the gravitational field than \(U_2\). As a result, \(U_1\) experiences a gravitational blueshift compared to \(U_2\) and runs faster than \(U_2\). Einstein (1918e, p. 699) claimed that “calculation shows” that, from \(U_2\)’s point of view, clock \(U_1\) gains “precisely twice” as much time on clock \(U_2\) during turn-around as it loses during the rest of the trip.\(^8\) Einstein did not give this calculation but it can readily be reconstructed.

Ironically, the wording “precisely twice” makes it clear that the calculation Einstein had in mind is an approximative one. The time \(U_1\) gains on \(U_2\) during \(\Delta t\) from \(U_2\)’s perspective in the limit that \(\Delta t \to 0\) is given by the length \(PQ\) of the line segment \(PQ\) in Fig. 1. Since \(PQ = DA - DP - QA = 2T - (2T/\gamma^2)\), the time gained is

\[
2T(v^2/c^2).
\]

The time \(U_1\) loses on \(U_2\) during the rest of the trip from \(U_2\)’s perspective is given by the difference between \(DR + RA = 2T/\gamma\) and \(DP + QA = 2T/\gamma^2\). To order \(v^2/c^2\), this is equal to

\[
2T(1 - \frac{1}{2}v^2/c^2) - 2T(1 - v^2/c^2) = T(v^2/c^2).
\]

Hence, in this approximation (but this is not true for the exact times\(^9\)), the time gained during turn-around is indeed twice the time lost during the rest of the trip, as Einstein

---

\(^8\)If \(U_1\) and \(U_2\) refer to twins rather than clocks, it makes more sense to say that \(U_1\) loses twice the amount time on \(U_2\) during turn-around that he gains during the rest of the trip (as Karel de Vlieger pointed out to me). More accurate, less ambiguous, but also more cumbersome translations of the terms Einstein (1918e, p. 699) himself used are “running ahead” (vorausilen) and “lagging behind” (zurückbleiben).

\(^9\)Strictly speaking, we therefore cannot apply the calculation to the situation drawn in Fig. 1, where \(v/c \approx .75\). However, the figure still brings out the salient features of the situation—they are just not as pronounced for \(v/c \ll 1\) as they are for \(v/c \approx .75\).
claimed.

Einstein wanted to ascribe the time $U_1$ gains during turn-around from $U_2$’s point of view to a gravitational blueshift. The formula Einstein needed for this purpose also only holds to order $v^2/c^2$ (another confirmation that the calculation referred to in the Dialogue is an approximative one). Derivations of the relevant formula can be found in the earliest papers on what Einstein would soon start calling the equivalence principle. In these papers, Einstein (1907, 1911) derived the red/blue-shift formula analyzing constant linear acceleration in Minkowski space-time, but it can also be derived through an analysis of uniform rotation. On [p. 2] of his lecture notes for a course on general relativity in Berlin in 1919, he derived it both ways (Einstein, 1987–2009, Vol. 7, Doc. 19, p. 148 and editorial note 6 on p. 178). In an appendix first added to the first English edition (published in 1920) of his 1917 popular book on relativity, he went with rotation (Einstein, 1959, pp. 129–130).

Figure 2: The gravitational red/blue shift corresponding to uniform rotation (left) and constant linear acceleration (right).
Fig. 2 puts the two situations, uniform rotation and linear acceleration (for the specific case of the twins), next to each other. Consider the rotating disk bottom left in the figure with radius $R$ and angular velocity $\omega$. On the disk are two clocks, $U_1$ at the center and $U_2$ at the circumference. Special relativity tells us that the rate of $U_2$, moving at velocity $v = \omega R$, is lower than the rate of $U_1$ by a factor
\[
\sqrt{1 - \frac{\omega^2 R^2}{c^2}} \approx 1 - \frac{\omega^2 R^2}{2c^2}. \tag{3}
\]
In the spirit of the equivalence principle, the rotating disk can be replaced by a disk at rest with a static centrifugal gravitational field. In principle (see sec. 3), this gravitational field will have additional components corresponding to the Coriolis force, but those only affect particles moving with respect to the rotating disk. For the centrifugal force on a unit mass on the circumference and at rest with respect to the rotating disk elementary Newtonian mechanics gives
\[
F = \omega^2 R = -\frac{d}{dr} \left( -\frac{1}{2} \omega^2 r^2 \right) \bigg|_{r=R}. \tag{4}
\]
The expression in parentheses is the Newtonian potential $\varphi(r)$ for the centrifugal gravitational field substituted for the centripetal acceleration on the basis of the equivalence principle. This potential is plotted top left in Fig. 2. The rate of $U_2$ is thus lower than the rate of $U_1$ by approximately a factor
\[
1 + \frac{\varphi(R)}{c^2}. \tag{5}
\]
Or, equivalently, the rate of $U_1$ is higher than the rate of $U_2$ by about $1 - \varphi(R)/c^2$. In other words, $U_2$ is red-shifted with respect to $U_1$, $U_1$ is blue-shifted with respect to $U_2$.

In the twin-paradox scenario on the right in Fig. 2, $U_1$ (between $P$ and $Q$) and $U_2$ (at $R$) play the roles of the clock at the circumference (at the lower potential) and the clock at the center (at the higher potential) of the rotating disk, respectively. The rate of $U_1$ is higher than the rate of $U_2$ by a factor
\[
1 - \frac{\Delta \varphi}{c^2}, \tag{6}
\]
where
\[ \Delta \varphi = \varphi(U_2) - \varphi(U_1) = -a(x(U_2) - x(U_1)) \] (7)
is the potential difference between \( U_1 \) and \( U_2 \), the product of the absolute value \( a \) of the acceleration and the distance between the two clocks when \( U_1 \) is turning around. The potential is plotted top right in Fig. 2.

From \( U_2 \)'s point of view (and since there is no gravitational field from \( U_1 \)’s point of view, this is the point of view that matters here), this distance varies. As the Minkowski diagram bottom right in Fig. 2 shows, it initially increases (from \( RP = L/\gamma \) to \( RS = L = vT \)) and then decreases again (to \( RQ = L/\gamma \)).\(^{10,11}\) This corresponds to the gravitational field decelerating \( U_1 \) from \( v \) to the left, its constant velocity between \( D \) and \( P \), to \( 0 \), its velocity at \( S \), and then accelerating it again to \( v \) to the right, its constant velocity between \( Q \) and \( A \). Since \( \gamma = 1 \) to order \( v^2/c^2 \), we can neglect the difference between \( L \) and \( L/\gamma \) when evaluating the expression \( \Delta \varphi/c^2 \) in Eq. (6) and set the distance between \( U_1 \) and \( U_2 \) throughout \( U_1 \)’s turn-around equal to \( L = vT \).

Substituting \( x(U_2) - x(U_1) \approx vT \) and \( a = 2v/\Delta t \) into Eq. (7), and substituting the resulting expression for \( \Delta \varphi \) into Eq. (6), we arrive at the result that the rate of \( U_1 \) is higher than the rate of \( U_2 \) by approximately a factor
\[ 1 + \frac{2v^2T}{\Delta t c^2}. \] (8)

Hence, what only takes \( \Delta t \) on \( U_2 \) takes longer on \( U_1 \). In the limit \( \Delta t \to 0 \) and \( a \to \infty \) with \( a\Delta t = 2v \), \( U_1 \)'s turn-around, instantaneous on \( U_2 \), takes approximately
\[ \lim_{\Delta t \to 0} \left( 1 + \frac{2v^2T}{\Delta t c^2} \right) \Delta t = 2T(v^2/c^2) \] (9)
on \( U_1 \), which is just the length of \( PQ \) in Fig. 1 (cf. Eq. (1)). So, at least to order \( v^2/c^2 \),

\(^{10}\)The hyperbola through the points \( V \), \( S \), and \( W \) connects all points at fixed spacelike distance \( L \) from \( R \). So the space-time line segments \( RP \) and \( RQ \) are indeed shorter than \( RS \) even though they are represented by longer line segments in the Minkowski diagram.

\(^{11}\)That \( RP = RQ = L/\gamma \) follows directly from the observation that for \( U_2 \) the spatial distance between \( S \) and \( R \) is a factor \( \gamma \) smaller than for \( U_1 \). It can be verified through a simple calculation. The Minkowski norm of the interval \( RP \) is \( RP^2 = RS^2 - c^2PS^2 \) (if \( RS \) is measured in meters and \( PS \) in seconds). Using that \( RS = L = vT \) and \( PS = \frac{1}{2}PQ = (v/c)^2T \) (see Eq. (1)), we find that \( RP^2 = v^2T^2(1 - (v/c)^2) = L^2/\gamma^2 \).
Einstein could indeed explain the time that elapses on $U_1$ during turn-around from $U_2$’s point of view in terms of a gravitational blueshift.\footnote{It is routinely pointed out in discussions of the twin paradox that the acceleration of the traveler while turning around cannot be the cause of the traveler being younger than the stay-at-home when they reunite. After all, the age difference would be twice as big if the trip were made twice as long, even if the accelerations involved were kept exactly the same. In Einstein’s 1918 account of the twins, it is only during the acceleration at the turn-around point that the traveler is aging more slowly than the stay-at-home. This raises the question why the age difference at reunion is proportional to the length of the trip. The answer is that it is not the acceleration that causes the traveler to stay younger but the potential difference between the twins at the turn-around point. This potential difference is proportional to the distance between them at that point and thus to the length of the trip (see Eqs. (6)–(8)).}

Once the metric field $g_{\mu\nu}(x)$ is introduced, as Einstein did in 1912, the situation can also be described as follows. In the system of space-time coordinates $x^\mu = (ct, x, y, z)$ of $U_1$, $g_{\mu\nu}(x)$ is the Minkowski metric with constant components $\eta_{\mu\nu} = \text{diag}(1, -1, -1, -1)$ everywhere ($\mu, \nu = 0, 1, 2, 3$). Outside the shaded area in Fig. 1, this is also the metric in $U_2$’s coordinate system. If we take $R$ as the origin of this coordinate system and consider small but finite values of $\Delta t$, this shaded area will correspond to a narrow band around the $x$-axis of $U_2$’s coordinate system. In this region, the components of the metric will not be constants. The 00-component, for instance, will be (to order $v^2/c^2$; cf. note 20):

$$g_{00} \approx 1 + \frac{2\varphi}{c^2} = 1 - \frac{2ax}{c^2}.$$  \hspace{1cm} (10)

The conversion factor for coordinate times to proper times is $\sqrt{g_{00}} \approx 1 - ax/c^2$. Hence, at $S$, the coordinate time $\Delta t$ corresponds to the proper time

$$\sqrt{g_{00}} \Delta t \approx \Delta t + 2T(v^2/c^2)$$  \hspace{1cm} (11)

(recall that $a = 2v/\Delta t$ and $x_S \approx -vT$).

The association of a component of the metric with the gravitational potential suggests that gradients of components of the metric be associated with the gravitational field. Einstein did indeed take that step but modern relativists have not followed his lead. The criterion for the absence or presence of a gravitational field in general relativity today is the vanishing or non-vanishing of the Riemann curvature tensor. Minkowski space-time is flat, which means that the Riemann tensor vanishes everywhere regardless of the coordinate system used. So in modern relativity there is no gravitational field for either
twin. By contrast, Einstein, from 1915 onward, represented the gravitational field by the so-called Christoffel symbols, a combination of derivatives of components of the metric that does not transform as a tensor. This fits with the basic idea of the equivalence principle that, at least locally, acceleration without a gravitational field is equivalent to rest in a gravitational field. It also provides the mathematical expression of the relativity of the gravitational field.

The metric field can be seen as consisting of two component fields, an inertial and a gravitational one, which is why it is also called the inertio-gravitational field. The inertio-gravitational field as a whole “tells” particles (figuratively speaking) to move on (timelike) geodesics (i.e., the straight[est] lines in [curved] space-time) whenever there are no additional non-gravitational forces acting on them. This statement is true in all coordinate systems. In any given coordinate system, however, we can distinguish between an inertial field telling particles to keep moving on a trajectory that in that coordinate system is represented by a straight line and a gravitational field telling particles to deviate from such lines (in a way that, unlike the “marching orders” from non-gravitational forces, is completely independent of the nature of the particle).

The metric field of Minkowski space-time is one particular configuration of the inertio-gravitational field. Once again consider the twins. In $U_1$’s coordinate system, the components of the metric are constants everywhere, so all Christoffel symbols vanish and there is no gravitational field anywhere. In this coordinate system, the inertio-gravitational field only has an inertial component. In $U_2$’s coordinate system, however, there are regions where the components of the metric are not constant. So the Christoffel symbols do not all vanish everywhere and there are regions where there is a gravitational field. In those regions the same inertio-gravitational field that in $U_1$’s coordinate system only has an inertial component has both an inertial and a gravitational component in $U_2$’s coordinate system.

---

13In 1913–1915, Einstein defined the gravitational field simply as the four-gradient of the components of the metric tensor. For discussion of Einstein’s switch from the gradient of the metric to the Christoffel symbols, see Janssen and Renn (2007) and Janssen (2005)
2.2 The twins and the breakdown of “kinematical relativity”

Einstein believed that the equivalence principle not only supported the relativity of the gravitational field but also the relativity of non-uniform motion in the sense that, as emphasized by Dieks (2006), all frames of reference, inertial as well as non-inertial, are physically equivalent. Using the account of the twins in the 1918 Dialogue both as evidence and as an instructive example, I offer a reconstruction and then a critique of how Einstein saw the equivalence of arbitrary frames of reference realized in his general theory of relativity.

Central to both reconstruction and critique is a distinction Einstein made between “kinematical” and “physical” relativity of motion. The only explicit statements of this distinction that I am aware of occur in a stenographer’s transcript of one of his lectures in Princeton in 1921 (Einstein, 1921b, p. 591), in two aborted drafts of the written version of these lectures,14 and in the published text of a lecture in Glasgow in 1933. However, as shown by Dorling (1978), it is implicit, for instance, in Einstein’s (1916, pp. 112–113) discussion of a thought experiment involving two rotating globes in his first systematic exposition of general relativity. In the Glasgow lecture, “Origins of the General Theory of Relativity,” we find:

From the purely kinematical point of view there was no doubt about the relativity of all motions whatsoever; but physically speaking, the inertial system seemed to occupy a privileged position (Einstein, 1933, p. 286).

So Einstein thought that “kinematical” relativity of arbitrary motion was automatically satisfied not just in general relativity but in special relativity and Newtonian theory as well. What is distinctive about general relativity, Einstein suggested in this 1933 passage (as well as in the 1921 passages), is that it also establishes a “physical” relativity of arbitrary motion.

What could this mean in the case of the twin-paradox scenario considered here? Appealing to “kinematical” relativity, both twins, Einstein would presumably argue, would describe the situation the exact same way. $U_1$ will say that he stayed put while $U_2$ moved away at constant speed $v$, turned around, and moved back with that same speed $v$;

14See the appendix for a translation of these two passages—(Einstein, 1921c, pp. 449–450) and (Einstein, 1921d, pp. 459–460).
$U_2$ will say that he stayed put and that $U_1$ moved away, turned around, and moved back. At this purely descriptive level, Einstein suggests, the situation is completely symmetric. At the physical (or, as I prefer to call it, the dynamical) level, it clearly is not. For one thing, $U_1$ and $U_2$ age differently. So whatever Einstein meant by “physical” relativity, it cannot be that the states of motion of $U_1$ and $U_2$ are physically (or dynamically) as well as descriptively (or kinematically) equivalent. What he did mean, as Dieks (2006) persuasively argues, is that general relativity, unlike special relativity and Newtonian theory, should allow us to account for the dynamical difference between $U_1$ and $U_2$ by appealing to the same laws in the same form using either the frame of reference of $U_1$ or that of $U_2$. In both frames, clocks moving at a constant velocity $v$ run slow by a factor $\gamma$ (an effect of the inertial field). In both frames, gravitational fields affect the rate of clocks the same way. And in both frames, gravitational fields are governed by the same laws (courtesy of the general covariance of the geodesic equation and the Einstein field equations, the two basic sets of equations governing the gravitational potential, represented by the metric, and the gravitational field, represented, at least for Einstein, by the Christoffel symbols).

The difference between the two frames, as Dieks (2006, p. 174, p. 176) puts it, is “fact-like” (“de facto”) rather than “law-like” (“de jure”). In $U_1$’s frame the Minkowskian inertio-gravitational field only has an inertial component while in $U_2$’s frame it has both an inertial and a gravitational component. There is also a difference between $U_1$ and $U_2$ (as opposed to a difference between the frames of $U_1$ and $U_2$) in that there are non-gravitational forces acting on $U_2$ while there are none on $U_1$. In the frame of $U_1$, we explain the difference in aging between $U_1$ and $U_2$ appealing to the effect of uniform motion (between $D$ and $R$ and again between $R$ and $A$) on the rate of $U_2$. In the frame of $U_2$, following Einstein’s account in his 1918 Dialogue, we explain the same difference in aging appealing both to the effect of uniform motion on the rate of $U_1$ (between $D$ and $P$ and again between $Q$ and $A$) and to the effect of the gravitational field on the rate of $U_1$ (between $P$ and $Q$). In terms of the distinction I made in the introduction, Einstein thought that his 1918 account of the twins established the relativity of acceleration not

---

15 Likewise, in the example of the rotating globes and as Dorling (1978) points out, the period of revolution of the other globe is different for observers on the two globes (Janssen, 2005, p. 63).
in the strong but in the weak sense.

So much for my reconstruction of Einstein’s position; now for the critique of that position. As Dorling (1978) showed for the example of the rotating globes, it is the assumption that “kinematical” relativity is automatic that is problematic (Janssen, 2005, pp. 62–63). My argument is essentially just an amplification of Dorling’s. According to both special and general relativity, the spatio-temporal descriptions of the twin-paradox scenario in the coordinate systems of $U_1$ and $U_2$ will already be so different that $U_2$, rather than appealing to the same dynamical laws as $U_1$ in his own supposedly equivalent coordinate system to explain why $U_1$ aged more than he did, will recognize that there is a de jure and not just a de facto difference between his coordinate system and that of $U_1$. I will show this for two plausible choices of the space-time coordinate system of $U_2$.

The only escape from this breakdown of kinematical relativity, as far as I can see, is to drop the spatio-temporal from spatio-temporal descriptions. Einstein, to be sure, seems to have been happy to do so. He said so right after his account of the twins in the 1918 Dialogue. And after introducing the distinction between kinematical and physical relativity in the Glasgow lecture, Einstein (1933, pp. 288–289) went on to say that an important step in the development of general relativity was his recognition that coordinates do not have direct metrical significance. I already quoted the relevant passages in the introduction. Unfortunately, stripping coordinates of spatio-temporal meaning trivializes the extension of the (weak) relativity principle to arbitrarily moving frames of reference. It would reduce the combination of kinematical and physical relativity to general covariance.

In Einstein’s 1918 account of the twins, the coordinates used by $U_1$ and $U_2$ are not arbitrary labels but have a definite spatio-temporal meaning. We can make this more explicit by “reifying” the lines of simultaneity for both $U_1$ and $U_2$. This is illustrated in Fig. 3. Imagine two rows of clocks, one moving with $U_1$ and one moving with $U_2$, sliding past one another. The shaded area in Fig. 3 corresponds to the worldlines of all the clocks in the row moving with $U_2$ that meet $U_1$ between $D$ and $A$. The dotted lines are the worldlines of two such clocks. To simplify the analysis, the accelerations at $D$ and $A$ (cf. Fig. 1) have been eliminated. In the situation pictured in Fig. 3, $U_2$’s clocks start to move to the right well before $U_1$ and $U_2$ first meet at $D$ and keep moving to the left until
well after $U_1$ and $U_2$ meet again at $A$. Both $U_1$ and $U_2$ have observers stationed at every one of their clocks.

I will analyze this situation for two different procedures for synchronizing $U_2$’s clocks. In scenario (I), $U_2$, like $U_1$, synchronizes his clocks once and for all some time before $D$ (using, e.g., the standard method of sending light signals back and forth between them). In scenario (II), which Einstein implicitly chose for his 1918 account of the twins, $U_2$ periodically checks the synchronization of his clocks and has the observers stationed at these clocks adjust their settings if it turns out that they are no longer properly synchronized. In this scenario, all clocks moving with $U_2$ will be reset after $R$. In scenario (I), as we will see, kinematical relativity fails during the return leg of $U_2$’s trip. In scenario (II), kinematical relativity appears to hold during both legs of $U_2$’s trip, but the lack of equivalence between the coordinate systems of $U_1$ and $U_2$ even at the purely descriptive level still reveals itself as this scenario requires the clocks moving with $U_2$ but
not those moving with $U_1$ to be reset during the trip. It will be instructive to consider the breakdown of kinematical relativity in these two scenarios in more detail.

For the time being, focus on the lower portion of Fig. 3, the line through $S$ and $R$ and everything below it. The lines of simultaneity for $U_1$ are horizontal lines and are not shown in the figure. The lines of simultaneity for $U_2$ are lines parallel to $MND$ and $PR$. Fig. 3 shows one such line for every year that passes on $U_2$. Suppose $U_1$ and $U_2$ both read 0 at $D$. The observers stationed with the clocks of $U_2$ are instructed to record the time on both $U_1$ and their own clock the moment $U_1$ passes by them. At $P$, for instance, the observer with the dotted worldline through $N$ and $P$ records that his own clock reads $NP = T/\gamma$ while $U_1$ reads $DP = T/\gamma^2$ (use that $DS = T$, $NP = DR = DS/\gamma$, and $DP = DR/\gamma$). Likewise, at $S$, the observer with the dotted worldline through $M$ and $S$ records that $U_1$ reads $DS = T$ while his own clock reads $MS = \gamma T$. Combining such clock readings, $U_2$ determines that $U_1$ runs slow by a factor $\gamma$. Following the exact same protocol, $U_1$ determines that $U_2$ runs slow by that same factor.

Now consider the upper portion of Fig. 3. Once the worldline of one of $U_2$’s clocks crosses the line through $S$ and $R$, it abruptly turns around. In other words, the setup is somehow rigged in such a way that $U_2$’s clocks reverse course instantaneously according to $U_1$ (or, according to $U_2$, the instant that $U_1$ is at rest with respect to $U_2$).

If $U_2$ checks the synchronization of his clocks after reversing course, he will find that they are no longer properly synchronized. It is this loss of synchronization that is behind the breakdown of kinematical relativity. This is true whether or not $U_2$ has the synchronization of his clocks restored. We need to deal with these two scenarios separately.

**Scenario (I)** (cf. Fig. 4). First we compare the description of $U_1$ and $U_2$ of the behavior of each other’s clocks if $U_2$ does not have the synchronization of his clocks restored after $R$. In that case, as shown in Fig. 3, the lines of simultaneity for $U_2$ above the line through $S$ and $R$ are parallel to those below that line. However, as Fig. 3 also shows, the lines for successive years are further apart during the second leg of $U_2$’s trip ($RA$) than they are during the first ($DR$). As a result, $U_2$ will conclude that, during that second leg, the

---

16Similar conclusions will be found if we have $U_2$’s clocks turn around upon crossing the line through $RP$ or the line through $RQ$, i.e., if we have them turn around simultaneously for $U_2$ before $R$ or after $R$, respectively.
Figure 4: Proper time $\tau(U)$ elapsed on clocks $U_1$ and $U_2$ as a function of time $t_1$ of $U_1$’s reference frame (left) and as a function of time $t_2$ of $U_2$’s reference frame (right) if $U_2$ does not resynchronize his clocks (scenario I); for $v < c/\sqrt{2}$ (top) and for $v < c/\sqrt{2}$ (bottom).

rate of $U_1$ is higher than that of his own clocks. Recall that the clock moving with $U_2$ and meeting $U_1$ at $S$ reads $\gamma T$ at $S$ while $U_1$ reads $T$. When $U_2$ itself meets $U_1$ again at $A$, it reads $2T/\gamma$ while $U_1$ reads $2T$. The temporal order of the data points $S$ and $A$ according to $U_2$ depends on the velocity $v$. Note that $2T/\gamma = \gamma T$ for $\gamma^2 = 2$. This will be case for $(v/c)^2 = \frac{1}{2}$ or $v = c/\sqrt{2}$. As long as $v < c/\sqrt{2}$, $\gamma < \sqrt{2}$ and $U_2$ will find that the event $A$ (at time $2T/\gamma > \sqrt{2}T$) comes after the event $S$ (at time $\gamma T < \sqrt{2}T$), as one would expect given that $A$ is a later point in the life of $U_1$ than $S$. As soon as $v > c/\sqrt{2}$, however, the order is reversed. From $U_2$’s point of view, $U_1$ now appears to be aging backwards in time! The problem clearly lies with the time coordinate used in scenario
If \( v > c/\sqrt{2} \), timelike-separated events end up with the same time coordinate (as is evident in the graph bottom right in Fig. 4). This makes for a particularly dramatic failure of kinematical relativity. But it already fails for \( v < c/\sqrt{2} \).

Using the various clock readings listed above, we can plot the proper time \( \tau(U) \) on the clocks \( U_1 \) and \( U_2 \) against the times \( t_1 \) and \( t_2 \) of the twins \( U_1 \) and \( U_2 \). The results are shown qualitatively in Fig. 4, for \( v < c/\sqrt{2} \) in the top half, for \( v > c/\sqrt{2} \) in the bottom half (Janssen, 1988, p. 23). The proper time on either clock plotted against its own time—i.e., \( \tau(U_1) \) against \( t_1 \), \( \tau(U_2) \) against \( t_2 \)—are simply straight lines with unit slope. The proper time \( \tau(U_2) \) plotted against \( t_1 \) will also be a straight line, with a slope of \( 1/\gamma \). However, the proper time \( \tau(U_1) \) plotted against \( t_2 \) will have a discontinuity at \( t_2 = \gamma T \). Up until that time, it will be a straight line with a slope of \( 1/\gamma \) just as \( \tau(U_2) \) plotted against \( t_1 \). Between \( t_2 = \gamma T \) and \( t_2 = 2T/\gamma \), it will still be a straight line but with a slope with an absolute value greater instead of less than 1. Moreover, for \( v > c/\sqrt{2} \), \( U_1 \) appears to be aging backwards in time.

The graphs in Fig. 4 show that kinematical relativity breaks down once \( U_2 \) reverses course. Recall that we started out with two rows of clocks sliding past each other that were both synchronized once and for all by the co-moving observers \( U_1 \) and \( U_2 \). The asymmetry of the situation of these two observers already manifests itself (most glaringly if \( v > c/\sqrt{2} \)) at the purely descriptive or kinematical level. For \( U_1 \), a moving clock (such as \( U_2 \)) always runs slow. For \( U_2 \), a moving clock (such as \( U_1 \)) sometimes runs slow, sometimes runs fast, and sometimes even runs backwards in time. \( U_2 \) has no recourse to any physical mechanism (such as a gravitational blueshift) to explain this different behavior in terms of conditions present in his coordinate system but not in \( U_1 \)'s (such as a gravitational field). \( U_2 \) is thus forced to conclude that his space-time coordinate system is not just de facto but de jure different from that of \( U_1 \).

**Scenario (II)** (cf. Fig. 5). We now compare the description of \( U_1 \) and \( U_2 \) of the behavior of each other's clocks if \( U_2 \) does restore the synchronization of his clocks after \( R \). In this case, the settings of all clocks in \( U_2 \)'s row of clocks are adjusted the moment their worldlines cross the line through \( S \) and \( R \) to make sure that they continue to be properly synchronized according to \( U_2 \) now that they are all moving to the left rather than to the
right. In that case, the lines of simultaneity for \( U_2 \) above the line through \( S \) and \( R \) are parallel to \( RQ \) (cf. Fig. 1).

\[ \text{Figure 5: Proper time } \tau(U) \text{ elapsed on clocks } U_1 \text{ and } U_2 \text{ as a function of time } t_1 \text{ of } U_1 \text{'s reference frame (left) and as a function of time } t_2 \text{ of } U_2 \text{'s reference frame (right) if } U_2 \text{ resynchronizes his clocks (scenario II).} \]

Let \( \hat{U}_2 \) denote the clock in \( U_2 \)’s row of clocks with the dotted line through \( N, P, \) and \( Q \) in Fig. 3 as its worldline. \( \hat{U}_2 \) meets \( U_1 \) at points \( P \) and \( Q \). If two clocks moving with velocity \( v \) are properly synchronized by a co-moving observer, a stationary observer will find that the rear clock is fast compared to the front clock by an amount of \((v/c^2)d\), where \( d \) is the distance between the two clocks for the co-moving observer. So, when \( U_2 \)’s clocks move to the \textit{right} with velocity \( v \), \( \hat{U}_2 \) should be \textit{fast} compared to \( U_2 \) from \( U_1 \)’s point of view by \((v/c^2)d\) where \( d = PR = L/\gamma \) and \( (v/c^2)T/\gamma \). When they all move to the \textit{left} with that same velocity, \( \hat{U}_2 \) should be \textit{slow} compared to \( U_2 \) from \( U_1 \)’s point of view by that exact same amount \( d = QR = PR \). It follows from these observations that the observer moving with \( \hat{U}_2 \) will set back his clock by \( 2(v^2/c^2)T/\gamma \) when he checks the synchronization of \( \hat{U}_2 \) with \( U_2 \) some time after \( \hat{U}_2 \) has turned around. Since from \( U_1 \)’s point of view, \( U_2 \)’s clocks are not only out of sync but also run slow by a factor \( \gamma \), \( 2T(v^2/c^2)/\gamma \) on \( \hat{U}_2 \) corresponds to \( 2T(v^2/c^2) \) on \( U_1 \). This, of course, is precisely the time that passes on \( U_1 \) between \( P \) and \( Q \) (see Eq. (1)). So \( P \) and \( Q \) will be assigned the same value \( t_2 \), which is how it should be as both \( P \) and \( Q \) are simultaneous with \( R \) in the coordinate system used by \( U_2 \) in this scenario. When \( \hat{U}_2 \) first meets \( U_1 \) at \( P \), \( \hat{U}_2 \) reads
$T/\gamma$ and $U_1$ reads $T\left(1 - \left(v^2/c^2\right)\right)$. When $\dot{U}_2$ meets $U_1$ again at $Q$, $\dot{U}_2$ once again reads $T/\gamma$ and $U_1$ reads $T\left(1 + \left(v^2/c^2\right)\right)$. If $U_2$'s clocks do not turn around instantaneously but in a short but finite time $\Delta t$, $\dot{U}_2$ will read $(T/\gamma) + \Delta t$ when it meets $U_1$ again at $Q$. Even in this more realistic scenario, $U_1$ ages $2T\left(v^2/c^2\right)$ from $U_2$’s points of view in the short time $\Delta t$ right after $t_2 = T/\gamma$. This is illustrated in the graph on the right in Fig. 5 (in the limit that $\Delta t \to 0$).

In his account of the twins in the 1918 Dialogue, Einstein argued that the rapid aging of the stay-at-home in the frame of reference of the traveler at the midway point $R$ of his journey can be attributed to a gravitational field that pops in and out of existence at that point and is present only in the frame of reference of the traveler. As we have seen, this argument presupposes a specific choice of the space-time coordinate system of the traveler, namely the one of scenario (II) above. Einstein tacitly assumed that this space-time coordinate system is, to use the term he himself introduced a few years later, kinematically equivalent to the inertial frame of the stay-at-home. To use Dieks’s terms, he assumed that these two space-time coordinate systems differ only de facto and not de jure.

No matter how we phrase it, the assumption fails. We already saw that it fails for scenario (I) in which the traveler and the stay-at-home both synchronized the clocks moving with them once and never adjusted them afterwards. It also fails for scenario (II) in which the stay-at-home and the traveler periodically check whether their clocks are still properly synchronized and adjust them if they are not. The stay-at-home will find that his clocks continue to be properly synchronized, but the traveler will have to adjust his clocks during the return leg of his journey. This will tell him that the rapid aging of the stay-at-home at the midway point is simply an artifact of his space-time coordinates.

More generally, he will recognize that the gravitational blueshift proportional to the distance from $R$ (see Eqs. (6)–(7)) is an artifact of these coordinates. Look back at Fig. 1 and consider a set of vertical lines at increasing distance from $R$. The proper time it takes for a system with such a vertical line as its worldline to cross the shaded region bounded by the line through $R$ and $P$ and the line through $R$ and $Q$ (i.e., the shaded region in Fig. 1, not the shaded region in Fig. 3) is proportional to the distance between
that vertical line and the one through $R$. Yet, in scenario (II), the traveler uses a space-time coordinate system that assigns the same time coordinate (or nearly the same time coordinate if we do not take the limit $\Delta t \to 0$) to the points where these lines enter and exit the shaded region. As with $\hat{U}_2$ and the entry and exit points $P$ and $Q$, the traveler will use the same clock to register the times when a system at rest with respect to $U_1$ enters and exits the shaded region. By the time it exits, the traveler will have arranged for that clock to be set back so that it reads the same time (or the same time plus $\Delta t$) at the exit point as it did at the entry point.

In short, in scenario (II), corresponding to Einstein’s 1918 account of the twins, it will be readily apparent to the traveler that, even at the purely descriptive level, his space-time coordinate system is not equivalent to that of the stay-at-home. Because of that, the traveler will not take the next step and appeal to laws valid in all kinematically equivalent space-time coordinate systems to explain the rapid aging of the stay-at-home during turn-around. Rather than attributing the effect to gravity, the traveler will recognize it for what it is, an artifact of the resetting of clocks involved in setting up the space-time coordinate system he is using.

3 The bucket

Rotation was an especially important target in Einstein’s campaign to make all motion relative (Janssen, 2011a, sec. 4). In the section, “The need for an extension of the postulate of relativity,” in his 1916 review article, “The foundation of the general theory of relativity,” the example of two globes rotating around the line connecting their centers gets pride of place (Einstein, 1916, pp. 112–113). This thought experiment can be seen as a modification of a thought experiments in the celebrated Scholium on space and time in Newton’s Principia (Cohen and Whitman, 1999, p. 408–415). Newton considered two globes rotating around each other, their acceleration revealed by the tension in the cord between them. So, in Newton’s version, the axis of rotation is perpendicular to the cord, whereas in Einstein’s version the axis of rotation coincides with the cord.

As we saw in sec. 1.2, Dorling (1978) put his finger on the problem with Einstein’s analysis, the failure of kinematical relativity.
Einstein’s 1916 article replaced an article with an almost identical title (“The formal foundation of the general theory of relativity”) published about a year and a half earlier (Einstein, 1914). At that point, Einstein was convinced that he had already reached his goal of making all motion relative with the theory that he had originally put forward more modestly as an “outline” [Entwurf] of a “generalized” (my emphasis) theory of relativity and gravity (Einstein and Grossmann, 1913). In the introduction of this premature 1914 review article, Einstein’s Machian account of the more famous thought experiment of Newton’s Scholium, that of the rotating bucket, takes center stage.

Following Einstein’s (1914, pp. 1031–1032) lead, we replace the bucket of water hanging from a slowly unwinding twisted rope considered by Newton by a much simpler system. Instead of a spinning bucket in the gravitational field of the earth, itself spinning on its axis, we consider a globe, held together by non-gravitational forces, rotating with respect to the fixed stars, which, for reasons that will become clear below, we take to form a large spherical shell with the globe at its center. In this case, the centrifugal forces, rather than giving the surface of the water in the bucket its tell-tale concave shape, make the globe bulge out at its equator. This variant of the thought experiment can be seen as a combination of the bucket experiment and Einstein’s 1916 variation on Newton’s rotating globes experiment. Einstein’s brief 1914 discussion of the bucket experiment is in terms of centrifugal forces on a body of unspecified shape rotating with respect to distant masses. Since both Newton and Mach are mentioned by name, it is nevertheless clear that he had the bucket in mind.

We analyze the system of globe and stars, first in terms of ordinary Newtonian mechanics, then in terms of Einstein’s new theory. Let the primed coordinates \((x', y', z')\) be the Cartesian coordinates of an inertial frame at rest with respect to the stars and let the unprimed coordinates \((x, y, z)\) be the Cartesian coordinates of a rotating frame at rest with respect to the globe, our stand-in for the bucket. Let the \(z\)-axis coincide with the \(z'\)-axis and let the rotation be counterclockwise around this axis with angular velocity \(\omega\). The velocity of a point \(x\) of the rotating frame with respect to the non-rotating frame is thus \(\omega \times x\), where \(\omega \equiv (0, 0, \omega)\) is the angular velocity vector. In the rotating frame,
Newton’s second law for the particles making up the globe takes the form

\[ \mathbf{F} = m (\ddot{\mathbf{x}} + 2\omega \times \dot{\mathbf{x}} + \omega \times (\omega \times \mathbf{x})) \],

where dots indicate time derivatives, \( m \) is the mass of the particle under consideration, and \( \mathbf{F} \) is the non-gravitational force keeping the particle in place in the globe (we assume that the gravitational forces between the particles of the globe can be neglected). The last two terms on the right-hand side are (minus) the Coriolis force and the centrifugal force, respectively:

\[
\begin{align*}
\mathbf{F}_{\text{Coriolis}} &\equiv 2m (\dot{\mathbf{x}} \times \omega), \\
\mathbf{F}_{\text{centrifugal}} &\equiv m ((\omega \times \mathbf{x}) \times \omega)
\end{align*}
\]

Moving these terms to the left-hand side, we can rewrite Eq. (12) in the form of \( \mathbf{F} = m \mathbf{a} \):

\[
\mathbf{F} + \mathbf{F}_{\text{Coriolis}} + \mathbf{F}_{\text{centrifugal}} = m\ddot{\mathbf{x}}.
\]

Since the globe is at rest in the rotating frame, the velocity \( \dot{\mathbf{x}} = 0 \) and the Coriolis force vanishes. The Coriolis force, however, does come into play when we consider the stars rather than the globe. In the frame of the globe, the stars are rotating clockwise around the \( z \)-axis with angular velocity \( \omega \). Their angular velocity vector is the opposite of that of the globe. So their velocity is given by \( -\omega \times \mathbf{x} \). They will thus experience a Coriolis force of \( -2m (\omega \times \mathbf{x}) \times \omega \), which is twice the size of the centrifugal force and in the opposite direction. Unsurprisingly, the net inertial force on the stars in the rotating frame is thus a centripetal force of just the right size to keep them in orbit around the globe.

Appealing to the equivalence principle, we reinterpret inertial forces as gravitational forces. An observer rotating with the globe will thus claim to be at rest in a gravitational field that is responsible both for the bulging out of the (stationary) globe and for keeping the stars whirling around her in orbit. Using the same analogy he used in the passage from his 1920 article intended for *Nature* quoted in the introduction, Einstein (1914, p. 1032) noted that this is reminiscent of the situation in electromagnetism where the force on a charged body can be due to a magnetic field for one observer and due to an electric field for another. A rotating body thus illustrates the relativity of the gravitational field.
Of course, the analysis so far has been in Newtonian terms. However, the Newtonian equations of motion in a rotating frame are also the equations of motion according to Einstein’s theory (both in its final form and in its non-covariant Entwurf form) in a rotating frame in Minkowski space-time as long as (A) the time coordinate in the non-rotating frame also serves as the time coordinate in the rotating frame and (B) the velocities of the particles under consideration with respect to the non-rotating frame are small compared to the velocity of light. Before showing this in detail, I want to make a few comments.

1. The main reason for replacing the bucket in the earth’s gravitational field by a globe hovering in space is so that, like Einstein, we can use Minkowski space-time in our analysis of the situation.

2. Since the stars are at rest with respect to the non-rotating frame, condition (B) does not prevent us from using the equations of motion for the stars even if their velocity with respect to the rotating frame, which is given by $\omega d$ (where $d$ is the star’s distance to the axis of rotation), exceeds the velocity of light. This does raise the question, however, whether a rotating observer would still insist to be at rest in a gravitational field if that meant she had to accept that the distant stars are moving at superluminal speeds. Or would she accept in that case that, to use Dieks’s (2006) terms, there is a de jure and not just a de facto difference between the rotating and the non-rotating space-time coordinate systems? Einstein, at least to my knowledge, never addressed these questions, perhaps because, as we will see, he did not expect space-time to remain Minkowskian all the way out to the distant stars.

3. Condition (A) poses the more serious threat to the attempt to establish the de jure equivalence between rotating and non-rotating space-time coordinate systems. As with the twins, Einstein assumed that kinematical relativity would be automatic in this case, whereas it actually requires a specific and ultimately arbitrary choice of a time coordinate. Dorling’s (1978) analysis of Einstein’s (1916) rotating globes example reminds us that a clock on the surface of a rotating body does not keep the same time as a clock at rest (Janssen, 2005, p. 63). An observer rotating
with the body using co-moving clocks would thus recognize the *de jure* difference between the rotating and the non-rotating space-time coordinate systems. As with the twins, we could have the two observers synchronize a set of clocks at rest with respect to them. Have both of them use light signals to synchronize all clocks at some fixed distance from the center with the clock at the center. A little later, have them check whether their clocks are still synchronized. The non-rotating observer will find that they are, the rotating observer will find that the clock at the center is fast compared to the others. If the rotating clocks are to keep the same time as the non-rotating ones, all of them except the one in the center will constantly have to be reset.

4. In addition to bringing out the role of conditions (A) and (B), the elementary derivation of the equations of motion in a rotating frame in Minkowski space-time on which Einstein’s observations about rotation are based will also help us understand how he came to believe that his theory vindicated the Machian idea that, in terms of our example, the bulging out of the globe would be the same no matter whether the globe or the stars were rotating.

Both in general relativity and in the *Entwurf* theory, its non-covariant precursor, the analogue of Eq. (14) for the motion for a test particle of mass \( m \) in a metric field \( g_{\mu\nu}(x^\rho) \) subject to an additional non-gravitational four-force \( F^\alpha \) in some arbitrary space-time coordinate system \( x^\mu \) is

\[
F^\alpha - m \left\{ \alpha_{\rho\sigma} \right\} \frac{dx^\rho}{ds} \frac{dx^\sigma}{ds} = m \frac{d^2x^\alpha}{ds^2},
\]

where \( s \) is the arclength of the particle’s worldline and the quantities \( \left\{ \ldots \right\} \equiv g^{\alpha\beta}[\rho\sigma, \beta] \) are the Christoffel symbols of the second kind. The quantities

\[
[\rho\sigma, \beta] = \frac{1}{2} (g_{\beta\rho,\sigma} + g_{\beta\sigma,\rho} - g_{\rho\sigma,\beta})
\]

are the Christoffel symbols of the first kind (the subscript ‘, \( \mu \)’ denotes differentiation with

\[18\]All Greek indices take on the values 0 through 3 and repeated indices are summed over.
respect to $x^\mu$). The term with the Christoffel symbols in Eq. (15) gives the gravitational force on the particle. Since the Christoffel symbols do not transform as a tensor, they can vanish in one coordinate system and not in another. As mentioned at the end of sec. 2.1, this implements the relativity of the gravitational field. For $F^\alpha = 0$, Eq. (15) reduces to the geodesic equation

$$\frac{d^2 x^\alpha}{ds^2} + \left\{ \begin{array}{c} \alpha \\ \rho \sigma \end{array} \right\} \frac{dx^\rho}{ds} \frac{dx^\sigma}{ds} = 0,$$

(17)

which picks out paths of extremal length in the space-time described by $g_{\mu\nu}(x^\rho)$.

In a Lorentz frame in Minkowski space-time—i.e., an inertial frame with pseudo-Cartesian coordinates $x'^\mu = (ct', x', y', z')$ and components $\eta_{\mu\nu} \equiv \text{diag}(1, -1, -1, -1)$ of the metric—the geodesic equation (17) reduces to $d^2 x^\alpha / ds^2 = 0$ and the equation of motion (15) to $F^\alpha = m (d^2 x^\alpha / ds^2)$. Now consider a frame with pseudo-Cartesian coordinates $x'^\mu = (ct, x, y, z)$ rotating counterclockwise around the $z$-axis with angular velocity $\omega$ with respect to this Lorentz frame. Following Einstein, we use the same time coordinate in the two frames: $t' = t$ [condition A]. To find the equation of motion in this rotating frame, we need to find the components $g_{\mu\nu}$ of the rotation metric—i.e., the metric of Minkowski space-time in the rotating $x^\mu$ coordinate system—and insert them into Eqs. (15)–(17).

In two places in Einstein’s Zurich notebook (pp. 12L-R and p. 42R), we find a quick derivation of the rotation metric (Renn, 2007, Vol. 2, pp. 584–587, 668–669). In the (primed) coordinates of the non-rotating frame, we have

$$ds^2 = \eta_{\mu\nu} dx'^\mu dx'^\nu = (c^2 - v'^2) dt'^2,$$

(18)

where $v' = \dot{x}'$ is the velocity of the particle under consideration with respect to the non-rotating frame. This velocity is the sum of $v = \dot{x}$, the particle’s velocity with respect to the rotating frame, and $\omega \times x$, the velocity of the position of the particle in the rotating frame with respect to the non-rotating frame:

$$v' = v + \omega \times x = (\dot{x} - \omega y, \dot{y} + \omega x, \dot{z}).$$

(19)
Substituting this expression for $v'$ and $dt$ for $dt'$ in Eq. (18), we can write $ds^2$ in terms of the (unprimed) coordinates of the rotating frame:

$$ds^2 = (c^2 - x^2 + 2\omega y \dot{x} - \omega^2 y^2 - y^2 - 2\omega x \dot{y} - \omega^2 x^2 - z^2)dt'^2.$$  

(20)

Regrouping terms, we find:

$$ds^2 = \left(1 - \frac{\omega^2 \rho^2}{c^2}\right)c^2 dt'^2 - dx^2 - dy^2 - dz^2 + 2\frac{\omega y}{c} dx c dt' - 2\frac{\omega x}{c} dy c dt',$$

(21)

where $\rho \equiv \sqrt{x^2 + y^2}$. Using $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$, we can read off the components $g_{\mu\nu}$ of the rotation metric:

$$g_{\mu\nu} = \begin{pmatrix}
1 - (\omega^2 \rho^2 / c^2) & \omega y / c & -\omega x / c & 0 \\
\omega y / c & -1 & 0 & 0 \\
-\omega x / c & 0 & -1 & 0 \\
0 & 0 & 0 & -1
\end{pmatrix}$$

(22)

Contracted with $-g_{\mu\alpha}$, the geodesic equation (17) becomes:

$$-g_{\mu\alpha} \dddot{x}^\alpha - \left(g_{\mu\rho,\sigma} - \frac{1}{2}g_{\rho\sigma,\mu}\right) \frac{dx^\rho}{ds} \frac{dx^\sigma}{ds} = 0,$$

(23)

where we used that $g_{\mu\alpha} g^{\alpha\beta} = \delta^\beta_\mu$; with $\delta^\beta_\mu \equiv \text{diag}(1, 1, 1, 1)$ the Kronecker delta, and that $g_{\mu\nu} = g_{\nu\mu}$. As long as $v' \ll c$ [condition B], $ds \approx c dt$ (see Eq. (18)) and derivatives with respect to $s$ in Eq. (23) can be replaced by time derivatives times $c^{-1}$:

$$-g_{\mu\alpha} \dddot{x}^\alpha - \left(g_{\mu\rho,\sigma} - \frac{1}{2}g_{\rho\sigma,\mu}\right) \dot{x}^\rho \dot{x}^\sigma = 0.$$  

(24)

Inserting the rotating metric (22) into this equation, we find, for $\mu = 1$,

$$\dddot{x}^1 - g_{10,2} \ddot{x}^0 \dot{x}^2 + g_{20,1} \ddot{x}^2 \dot{x}^0 + \frac{1}{2}g_{00,1} \ddot{x}^0 \dot{x}^0 = \dddot{x} - 2\omega \dot{y} - \omega^2 x = 0,$$

(25)

where in the last step we used that $\dot{x}^\mu = (c, \dot{x}, \dot{y}, \dot{z})$, $g_{10,2} = -g_{20,1} = \omega / c$, and $g_{00,1} =$

\footnote{Einstein repeatedly made the mistake of setting $g_{01} = g_{10}$ equal to $2\omega y / c$ and $g_{02} = g_{20}$ to $-2\omega x / c$ (Janssen, 1999, pp. 145–146).}
Similarly, we find, for $\mu = 2$,

$$\ddot{x}^2 - g_{20,1} \dot{x}^0 \dot{x}^1 + g_{10,2} \dot{x}^1 \dot{x}^0 + \frac{1}{2} g_{00,2} \dot{x}^0 \dot{x}^0 = \ddot{y} + 2 \omega \dot{x} - \omega^2 y = 0,$$

(26)

and, for $\mu = 3$, $\ddot{z} = 0$.²⁰ The spatial components of the geodesic equation (23) for a particle moving slowly in a rotating coordinate system in Minkowski space-time can thus be written in vector form as:

$$\ddot{\mathbf{x}} + 2 \mathbf{\omega} \times \dot{\mathbf{x}} + \mathbf{\omega} \times (\mathbf{\omega} \times \mathbf{x}) = 0.$$

(27)

The spatial components of the equation of motion (17) for a particle subject to an additional non-gravitational force $F^\mu = (F^0, \mathbf{F})$ can likewise be written as

$$\mathbf{F} + 2m (\dot{\mathbf{x}} \times \mathbf{\omega}) + m ((\mathbf{\omega} \times \mathbf{x}) \times \mathbf{\omega}) = m \ddot{\mathbf{x}},$$

(28)

which is just the Newtonian equation of motion in a rotating coordinate system (cf. Eqs. (12)–(14)).

Note that the terms of order $(\omega/c)$ in the rotation metric (22) are responsible for the Coriolis force while the terms of order $(\omega/c)^2$ are responsible for the centrifugal force. Before he published the generally-covariant Einstein field equations in November 1915, Einstein repeatedly checked whether the rotation metric is a vacuum solution of the field equations of limited covariance of the Entwurf theory (Janssen, 1999, 2007). One readily verifies that, to first order in $(\omega/c)$, it is. Einstein now inserted this first-order metric into the second-order field equations, solved for the $(\omega/c)^2$ term in $g_{00}$, and checked whether this reproduces the 00-component of the rotation metric. Spurious minus signs and factors of 2 (see note 20) initially confirmed his strong expectation that the Entwurf equations pass this test.²¹ Ignoring strong warning signs that they do not,²² Einstein only accepted

²⁰Using Eq. (4) for the potential $\varphi$ of the centrifugal field, we can write the 00 component of the rotation metric (22) in the form of Eq. (10): $g_{00} = 1 + 2 \varphi/c^2$. The terms $\frac{1}{2} g_{00,2} \dot{x}^0 \dot{x}^0$ ($i = 1, 2$) in Eqs. (25)–(26) thus give components of the gradient of $\varphi$.

²¹See [pp. 41–42] of the so-called Einstein-Besso manuscript on the perihelion motion of Mercury (Einstein, 1987–2009, Vol. 4, pp. 442–447). These pages probably date from May/June 1913. The iterative approximation procedure Einstein used in this calculation is the same as the one he used to calculate the perihelion motion of Mercury (Earman and Janssen, 1993, pp. 141–149).

²²See the so-called Besso memo of August 1913 and a letter from Ehrenfest to Lorentz that same
in October 1915 that the rotation metric is, in fact, \textit{not} a vacuum solution of the \textit{Entwurf} equations.\textsuperscript{23} This precipitated the demise of the \textit{Entwurf} theory. When he wrote his review article on the \textit{Entwurf} theory in October 1914, he still believed that the rotation metric was a vacuum solution. This was crucial to the application of the equivalence principle to rotation in the article. For a rotating observer in Minkowski space-time to be able to reinterpret the inertial effects of rotation as due to a gravitational field, that gravitational field must be allowed by the field equations. The generally-covariant equations with which Einstein replaced the \textit{Entwurf} field equations in November 1915 at least meet that requirement.

However, these generally-covariant equations did nothing to solve another problem with Einstein’s claims about rotation in the 1914 review article. He wrote:

\begin{quote}
We do not necessarily have to trace these centrifugal forces to the motion of $K'$ [a rotating frame]; we might just as well trace them to the \textit{average rotation of ponderable distant masses} in the neighborhood with respect to $K'$, where we treat $K'$ as being ‘at rest’ (Einstein, 1914, p. 1031, my emphasis).
\end{quote}

The italicized clause implies is that the gravitational forces substituted for the centrifugal forces on the basis of the equivalence principle are due to a gravitational field produced by distant matter. So far, however, the stars have been treated as test particles in a given field rather than as sources of that field. The metric field is that of flat empty Minkowski space-time. That we considered this metric field in a rotating coordinate system does nothing to change this fact.

Einstein, however, thought that rotating distant matter—which he pictured, as we did, in the form of a giant spherical shell—would, in the vicinity of its center, give rise to a metric field of the exact same form at the rotation metric. Einstein’s way of trying to prove this is directly related to the way in which he tried to confirm that the rotation metric is a vacuum solution of the \textit{Entwurf} field equations. Solving the first-order field equations for the field near the center of a rotating shell, Einstein suggested, one would

\textsuperscript{23}See Einstein’s letter to Freundlich of September 30, 1915, and the calculation on a sheet of paper subsequently used for the draft of a letter to Naumann of early October 1915 (Einstein, 1987–2009, Vol. 8, Docs. 123 and 124, respectively).
find the rotation metric to first order in \((\omega/c)\) (the Coriolis terms).\(^{24}\) Inserting this first-order metric into the vacuum second-order field equations and solving those, one would find the second-order term of the rotation metric (the centrifugal term).

Strong evidence that this is indeed how Einstein saw the situation comes from the passage in the August 1913 Besso memo referred to above (see note 22) in which Besso warned Einstein that the second step of this procedure fails for the *Entwurf* field equations:

> If through rotation of a hollow sphere one produces a Coriolis field inside of it, then a centrifugal field is produced . . . that is not the same as the one that would occur in a rotating rigid system with the same Coriolis field. One can therefore not think of rotational forces as produced by the rotation of the fixed stars (Janssen, 2007, pp. 785 and 806).

Confirming evidence can be found in letters to Besso and Thirring after 1915 when Einstein thought he had solved the problem by replacing the *Entwurf* equations with the generally-covariant Einstein equations.\(^{25}\)

Even with the generally-covariant field equations, however, Einstein’s scheme does not work. The gravitational field the rotating shell produces near the globe is much weaker than that substituted for the rotation of the globe on the basis of the equivalence principle. Rotation is certainly not relative in the strong sense that rotation of the shell would cause the same bulging out of the globe as the rotation of the globe itself. By the time of his 1921 Princeton lectures, if not earlier, Einstein had recognized this. In the published version of these lectures, Einstein still mentioned the “Coriolis field” (Einstein’s own quotation marks) and the centrifugal field inside a rotating hollow body but now emphasized that their effects are too small to be detected (Einstein, 1922, pp. 64–66; pp. 100–103 in the translation).

Is rotation at least relative in this case in the weak sense that there is only a *de facto* and not a *de jure* equivalence between the non-rotating space-time coordinate system of the shell and the rotating coordinate system of the globe? As we saw above, even this weak equivalence only obtains for a specific choice of a time coordinate in the rotating

---


frame (condition A: $t = t'$) and if no part of the shell is moving with superluminal velocity (condition B: $\omega d < c$).

In Einstein’s own terms, kinematical relativity is not automatic, neither for traveler and stay-at-home in the twin paradox, nor for globe and shell in our simplified version of the rotating bucket experiment. Without kinematical relativity, however, the accelerated observer will not attempt to explain various inertial effects as gravitational effects but accept that they are the result of his or her own acceleration.

4 What is relative in general relativity: motion, gravity, neither?

Einstein’s 1918 account of the twins was meant to illustrate the relativity of non-uniform motion in what I called the weak sense. Einstein showed that the difference in aging between the two twins can be explained on the basis of the same physical laws in a Lorentz frame in Minkowski space-time in which the stay-at-home is at rest and in a suitably chosen space-time coordinate system for the traveler. In Einstein’s (1921b,c,d) own later terms, he showed that, on the (erroneous) assumption that kinematical relativity is automatically satisfied, his new theory of gravity ensures physical relativity as well. In Dieks’ (2006) terms, he tried to show that two space-time coordinate systems in Minkowski space-time, one in which the stay-at-home and one in which the traveler is at rest, differ only de facto and not de jure. The relativity of the gravitational field, as Einstein (1920) called it, was key to his argument. The difference in aging between the two twins is a purely inertial effect in the coordinate system of the stay-at-home but a combination of inertial and gravitational effects in the coordinate system of the traveler.

Einstein’s 1914 account of the bucket in the context of the Entwurf theory proceeds along similar lines and faces similar problems as his account of the twins four years later in the context of the generally-covariant version of his theory. A globe, our proxy for the bucket, and the distant stars take over the role of the traveler and the stay-at-home, respectively. In this case, Einstein concluded that the globe’s rotation is relative both in the weak and in the strong sense. To establish that this rotation is relative in the weak
sense, one has to show that the bulging out of the rotating globe can be explained on the basis of the same physical laws in two space-time coordinate systems, one in which the stars and one in which the globe is at rest. To establish that it is relative in the strong sense, one has to show that rotation of the stars would produce the same the bulging out of the globe as the rotation of the globe itself. Although it took Einstein several years to realize this (see note 25), this second proposition is obviously false.

The analysis of Einstein’s accounts of the twins and the bucket in secs. 2 and 3 shows that even the claim that non-uniform motion is relative in the weak sense is untenable. As Dorling (1978) first noted, although he did not put it in these terms, Einstein’s assumption that kinematical relativity is automatically satisfied is problematic. If we leave the coordinates used enough spatio-temporal meaning to define motion, kinematical relativity fails. If we strip them of all spatio-temporal meaning, kinematical and physical relativity become trivial in that they would obtain in any theory that can be recast in generally-covariant form. Although Einstein often endorsed this trivial reading of kinematical relativity, he used the non-trivial reading in his accounts of the twins and the bucket. Without kinematical relativity, Einstein’s basic argument in these accounts no longer goes through. If the traveler or a person rotating with the globe finds that his or her space-time coordinate system is not equivalent to that of the stay-at-home or a person at rest with respect to the distant stars, he or she will attribute the phenomena to be explained (the age difference between the twins, the bulging out of the globe) to his or her own acceleration rather than to a gravitational field.

The upshot then is that Einstein’s general theory of relativity fails to extend the principle of relativity from uniform to arbitrary motion in any non-trivial sense. Where does that leave the other relativity principle that Einstein sought to illustrate with his accounts of the twins and the bucket—the relativity of the gravitational field? Clearly, these illustrations of the principle, dependent as they are on kinematical relativity, do not work either. Can we give up the illustrations without giving up the principle itself? Can we hold on to Einstein’s idiosyncratic identification of the gravitational field with the Christoffel symbols—the mathematical implementation of the relativity of the gravitational field—or are we forced to adopt the modern invariant definition of the gravitational field (i.e., the non-vanishing of the Riemann curvature tensor)? As I suggested
in the introduction, it seems to me that the relativity of the gravitational field can be separated from the relativity of arbitrary motion with which it got entwined in Einstein’s problematic accounts of the twins and the bucket. In the introduction I already gave a formulation of the former that is entirely independent of the latter. According to general relativity there is one inertio-gravitational field that splits differently into inertial and gravitational components in different coordinates. This is a highly non-trivial statement that captures a key feature distinguishing Einstein’s theory of gravity from older theories. The coordinates by themselves do not have any spatio-temporal meaning. This is not only in keeping with the way coordinates are used by modern relativists, it also avoids the problem that Einstein ran into in his accounts of the twins and the bucket of giving coordinates the kind of direct metrical significance that on other occasions he emphatically denied they have. With the help of the metric field we can give spatio-temporal meaning to these coordinates but the resulting space-time coordinate systems will in general not be equivalent to one another.

Whether or not Einstein’s notion of the relativity of the gravitational field can indeed be salvaged in this or some other way strikes me as an interesting question about the foundations of general relativity. However, as I indicated in the introduction, I am interested first and foremost in the historical question of how exactly Einstein thought about the general relativity of motion that gave his theory of gravity its misleading name. Much of the discussion of this question in the historical and philosophical literature focuses on Einstein’s interpretation of general covariance. This is undeniably an important part of the story, but to get the full picture we need to come to terms with less familiar notions such as the relativity of the gravitational field and the distinction between kinematical and physical relativity. If nothing else, I hope this paper will convince the reader of the importance of various passages in Einstein’s work in which these notions are either defined or (mostly implicitly) put to work. It thus seems appropriate to end this paper with a list of these passages:

- The account of the bucket in the introduction of the review article on the *Entwurf* theory of late 1914 (Einstein, 1914, pp. 1031–1032). See notes 21–25 for references to various other passages dealing with rotation in Einstein’s correspondence and
manuscripts of the period 1913–1917.

- The account of the twins in the “Dialogue about objections to the theory of relativity” of late 1918 (Einstein, 1918e).


- The statement of the difference between kinematical and physical relativity on [pp. 1–2] of the typescript based on a stenographer’s notes of the first of Einstein’s two popular lectures in Princeton on May 9 and 10, 1921 (Einstein, 1921b) and in the opening paragraphs of two aborted drafts of The meaning of relativity, the written version of his three more technical lectures in Princeton on May 11, 12, and 13, 1921 (Einstein, 1921c,d). In the appendix, I present translations of the relevant parts of these last two texts.

**Appendix: Einstein on kinematical versus physical relativity**

The following is my translation of a pair of corresponding passages in two aborted drafts of The meaning of relativity (Einstein, 1922), the published version of Einstein’s lectures at Princeton University of May 1921. These two drafts were published as Docs. 63 and 64 in Vol. 7 of The collected papers of Albert Einstein (pp. 449–469). They were not selected for inclusion in the translation volumes accompanying the documentary edition of Einstein’s papers.


*Kinematical relativity of motion*. That motion by definition can only be conceived of as relative motion of one body with respect to another is a fact that
has long been known to philosophers. One can also express this as follows: from the point of view of pure kinematics there is only relative motion. Out of arbitrarily many bodies moving with respect to each other one can choose an arbitrary one as the body of reference \([\text{Bezugskörper}]\) (in other words, one can look upon that body as “at rest”) and refer the motion of all others to it. The body of reference used in geometry and physics is the Cartesian coordinate system which basically consists of three mutually orthogonal rigid rods, to which (through rod constructions) all points in the world can be referred.

**Physical relativity of motion. This is by no means automatic.** Even though it is a fact that all states of motion are equivalent from a purely kinematical point of view, this equivalence need not exist at all from a physical point of view. A priori it would certainly be possible that there be one state of motion (or a group of states of motion) that would somehow be privileged. This would have the consequence that coordinate systems in a specific state of motion (or specific states of motion) would be especially suited to serve as the body of reference for the description of nature. The laws of nature would take their simplest forms if such coordinate systems were used as the body of reference (justified coordinate systems).

**Physical relativity of motion is the fundamental assumption of relativity theory.** The special theory of relativity is based on the presupposition that there does not exist one physically privileged state of motion, i.e., that nature does not give us any reason to introduce the notion of absolute rest. The general theory of relativity is based on the presupposition that there are no physically privileged states of motion in nature at all, and that, therefore, all coordinate systems are equivalent for the formulation of the laws of nature.


It has been clear for ages that motion can only be thought of as relative motion (motion of one body with respect to another), not as absolute motion (motion of a body not referred to other bodies). It is therefore impossible, as long as one bases oneself solely on the concept of motion, to prefer one state of motion over all other states of motion through some special marks. If one considers coordinate systems \(K, K', K''\) moving with respect to one another, all one can say about the state of motion of any one of them is that or how it is moving with respect to the other systems. By contrast, it would a priori be perfectly possible that some coordinate system be privileged over all others *from a physical point of view.* This would be the case if, in their simplest form, the laws of nature would be valid with respect to one coordinate system \(K\), but invalid with respect to all coordinate systems moving differently. If, however, in the opposite case, the validity of the law of nature is not tied to a specific coordinate system, one can say that the motion (of coordinate systems) is not only relative in the purely (kinematical) conceptual sense (which goes without saying) but also in the physical sense.
Whereas, however, all states of motion are equivalent to one another in the kinematical sense, certain states of motion are privileged according the classical mechanics. For in addition to being valid with respect to a particular coordinate system $K$, the laws of motion are valid only in coordinate systems $K'$ that are in uniform translational motion with respect to $K$ (inertial systems), but not in coordinate systems moving differently with respect to $K$. It is in this sense that one says that mechanics only satisfies the *special* principle of relativity. The theory based on this special principle of relativity is called the special theory of relativity. [This draft breaks off before Einstein gets to general relativity.]

**Acknowledgments.** I am grateful to members of PIG (the Physics Interest Group) at the University of Minnesota (especially Aimee Slaughter) and to the students in the 2010 editions of my honors class on Einstein (especially Alexandra Nelson) for helpful comments. I am grateful to the Einstein Papers Project and the Hebrew University of Jerusalem for permission to publish translations of parts of two Einstein documents in the appendix.

**References**


Dorling, J. (1978). Did Einstein need general relativity to solve the problem of absolute space? Or had the problem already been solved by special relativity? *British Journal for the Philosophy of Science,* 29, 311–323.


