May 9, 2000; slightly revised September 29, 2000

# What Can We Learn about the Ontology of Space and Time from the Theory of Relativity?

John D. Norton

Department of History and Philosophy of Science

University of Pittsburgh

Pittsburgh PA USA 15260

In the exuberance that followed Einstein's discoveries, philosophers at one time or another have proposed that his theories support virtually every conceivable moral in ontology. I present an opinionated assessment, designed to avoid this overabundance. We learn from Einstein's theories of novel entanglements of categories once held distinct: space with time; space and time with matter; and space and time with causality. We do not learn that all is relative, that time in the fourth dimension in any non-trivial sense, that coordinate systems and even geometry are conventional or that spacetime should be reduced ontologically to causal, spatio-temporal or other relations.

## 1. Introduction

When the first decades of the twentieth century brought Einstein's special and general theories of relativity, philosophy of space and time lost its ancient bearings. The immovable landmarks of centuries and millennia fell. With them went the familiar questions of the old era.

How are we to fit the absolute truths of Euclid and Newton into a serviceable philosophy? How can we assimilate their certainty with the fragility of human experience and learning? To ask was futile. Newton's space and time had become mistaken opinion. Euclid's geometry had slipped from truth *a priori* to mere contingency in the nineteenth century and now to simple falsehood.

An exuberance of philosophy ensued. Treatises needed no longer to be plotted always to yield the familiar ending, the vindication of the old regime. New questions could be asked and imaginative answers given freely. This anarchy of ideas had its price. As viewpoints multiplied, Einstein's theories were invariably summoned in support. Scarcely any declined to call, be they new findings on methods of discovery, on the nature of scientific theories, on the essence of space and time, on matter and cause, on being and bunkum. Einstein's achievement had become the new absolute, the assured ending of every plot, just as there was no agreement on what that ending was.

My purpose in this paper is to bring some order to this surfeit of wisdom. I will seek to answer the question of the title: what can we learn about the ontology of space and time from the theory of relativity? My task is not to survey the many answers on record. Rather it is to find principled grounds for sifting among them and separating out a consistent, supportable view. Where can we find these grounds? This volume celebrates John S. Bell, so it is appropriate to look to him for guidance. He once remarked (1987, p.67)

> I have for long thought that if I had the opportunity to teach this subject, I would emphasize the continuity with earlier ideas. Usually it is the discontinuity which is stressed, the radical break with more primitive notions of space and time. Often the result is to destroy completely the confidence of the student in perfectly sound and useful concepts already acquired.

He was reflecting on how to teach special relativity and not how to draw ontological morals from it. Nevertheless the danger he addressed threatens both teaching and philosophy. It is the temptation to exaggerate the difference between relativity theory and the classical theory of space and time it replaces. We too often claim to find morals in relativity theory that could equally have

been drawn from earlier theories.

## Four Requirements

This concern motivates the first of four requirements that I shall ask all ontological morals to meet:

1. Novelty . The morals we draw should be novel consequences of relativity theory. They should not be results that could have been drawn equally from earlier theories.

The advent of relativity theory has allowed us to discern possibilities we just overlooked in the past. These are not morals of relativity. To say they are confuses message and messenger.

2. Modesty. The morals we draw should be consequences of relativity theory. They should not be results we wish could be drawn from relativity theory, but are only suggested to us by the theory.

Relativity has inspired many programs of research into space and time that are based on ontological themes. The hope is that our next great success in the physics of space and time will verify them. Unless these themes are consequences of relativity theory, however, they are not morals to be drawn here.

3. Realism . Relativity theory is to be construed as literally as possible.

We cannot draw ontological morals from relativity theory at all unless we take a particular attitude to the theory. In so far as is possible, we must take the theory to mean literally what it says; this is my favored formulation of realism. We are not compelled to adopt realism. But without it, there is no rhyme or reason in answering the question of the title. We could choose to be fictionalists. Then we would judge the ontological pronouncements of relativity theory, whatever they might be, as useful mythmaking, devoid of insight into that which exists.

4. Robustness . We should not draw morals in one part of the theory that are contradicted in others. In particular, the morals we draw from examination of special relativity should survive the transition to general

relativity.

Failure to heed “robustness” has caused much unnecessary confusion. Many of the philosophical responses to relativity theory look at the special theory alone and trumpet results that are almost immediately contradicted by the emergence of general relativity.

### Things to Come

The morals that I will discern all pertain to the theme of entanglement and consequently have a Bell-like ring to them. They will be elaborated in the sections to follow. In Section 2, I will describe how special relativity brought a new entanglement of space and time and I will go to some pains to formulate the entanglement in a way that respects Robustness . In Section 3, I will review the most obvious moral that came with the extension of the special theory to the general theory, the entanglement of spacetime the container and the matter it contains. In Section 4, I will turn to the moral that has only been pursued seriously in more recent years, the entanglement of spacetime and causality, which forces any serious philosophical analysis of causation to examine what Einstein wrought. Finally in Section 5 I examine some of the popular claims that have not entered my select compendium and explain why I have spurned them.

## 2. The Entanglement Space and Time

### The Relativity of Simultaneity…

As Einstein grappled with the problems in electrodynamics that gave us the special theory of relativity, the discovery of one key misapprehension about space and time allowed him to reconcile two apparently incompatible notions, the principle of relativity, demanded by experiment, and the constancy of the speed of light, demanded by Maxwell’s electrodynamics. He could assert both if he was willing to suppose something utterly at odds with classical theory, that observers in relative motion may disagree on which spatially separated events are simultaneous. Two events judged to occur at the same time by one observer might be judged to be sequential by

another in relative motion. This result is the relativity of simultaneity. It is described in careful detail in the first section of Einstein's (1905) celebrated "On the Electrodynamics of Moving Bodies," for all that follows in Einstein's paper depends upon it. It expresses a profound entanglement of space and time, a moral worthy of inclusion in our catalog.

This notion lay behind Minkowski's (1908, p. 75) immortal declaration when he introduced the concept of spacetime:

> The views of space and time which I wish to lay before you have sprung from the soil of experimental physics, and therein lies their strength. They are radical. Henceforth space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality.

Prior to relativity theory, space and time were treated separately. One considered space at one instant of time; and then successive spaces as the instants passed. Minkowski combined these into a single four dimensional spacetime manifold of events. That much was not incompatible with classical theory. In it, just as in relativity theory, the set of all events in space and time form a four dimensional manifold. On pain of violation of Novelty we cannot claim it as a moral of relativity theory. It was just a classical possibility that was not exploited. The novelty lies in the way the new spacetime can be decomposed to spaces that persist through time. It reflects a new entanglement of space and time. In classical theory, the decomposition is unique. There is one way to do it. In relativity theory, as shown in Figure 1, each inertially moving observer finds a different way to slice spacetime into spaces. There is a different decomposition associated with each inertial frame of reference.

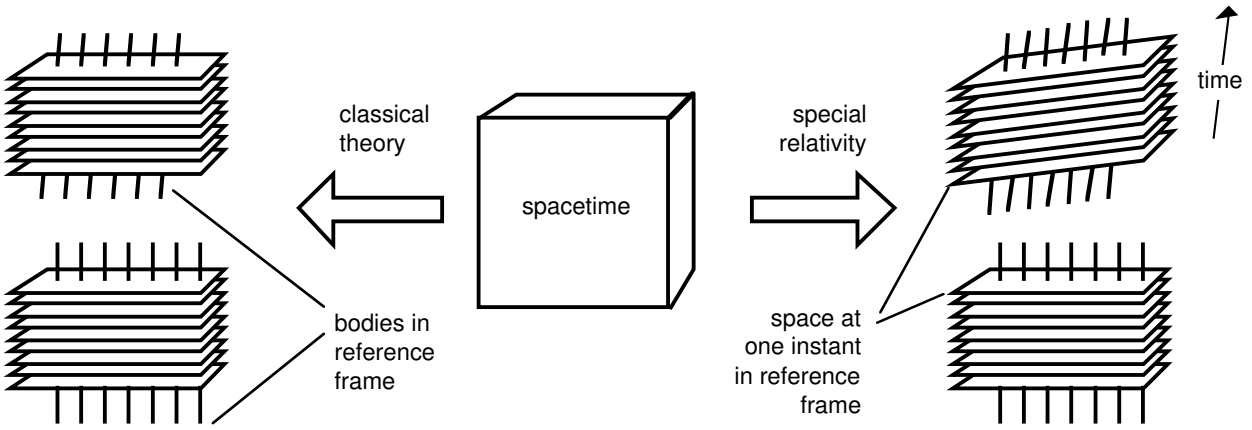



Figure 1. Simultaneity in classical theory and special relativity

Each such space consisted of simultaneous events. Since these observers could not agree on which events are simultaneous, they cannot agree on how to form the spaces. We cannot select one slicing as the correct slicing. Each is geometrically identical and any criterion that would elevate one would do the same to all the rest. It is just like seeking diameters of a perfect circle. There is no one correct diameter that bisects the circle. There are infinitely many and they are all identical in their geometric properties

## …Is not Robust

Important as it is, these traditional presentations of the relativity of simultaneity cannot stand as a moral since they fail the requirement of Robustness . With the advent of the general theory of relativity, spacetime took on a more varied geometrical structure. It could still be sliced in many ways into spaces that persist with time, but, in important cases, just one slicing is preferred geometrically. To visit some familiar examples, consider the Robertson-Walker spacetimes used in standard big bang cosmology. Just one slicing gives spaces filled with a homogeneous matter distribution. Any other slicing mixes events from different epochs with differing densities of matter. Or consider a Schwarzschild spacetime, the idealized spacetime of our sun. Just one natural [1] slicing gives us spaces whose geometric properties remain constant

[1] That is, natural in the sense that the slicing satisfies the technical condition of orthogonality

with time. Analogously many chords might bisect the area of an ellipse, but bisection along the principle axis is geometrically distinct from all the others.

## Infinitesimal Neighborhoods of Events in Classical and Relativity Theory

What of the celebrated entanglement of space and time brought by the special theory. Has the general theory parted what the special theory had joined together? It has not. To find a robust entanglement, we must seek it in a more subtle way. We will find it by exploiting a fundamental fact about the spacetimes of both the special and general theory. They differ in domains of any finite extent. However, if we select just one event and consider the events infinitesimally close to it, then we have found a mini-spacetime that is the same in both special and general relativity. This mini-spacetime mimics the bigger spacetime of special relativity. The entanglement of space and time of the relativity of simultaneity can be found within it. We can use it to formulate this entanglement in a way that is robust under the transition from special to general relativity.

Let us proceed to these mini-spacetimes. In order not to violate Novelty , I will first construct as much of them as I can in a way that is compatible with both classical and relativistic theories. In both, the set of all events forms a four dimensional manifold. That means that we can label events with four real numbers, the spacetime coordinates, and we can then use those numbers to decide which events are near which. This notion of nearness lets us extract the mini-spacetime of events infinitesimally neighboring some arbitrary event O, as shown in Figure 2.

---

with the world lines of the matter of the sun.

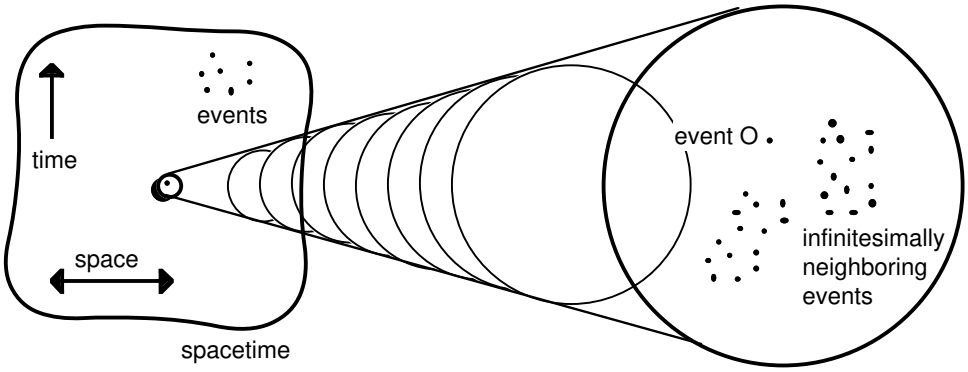



Figure 2. Extracting an infinitesimal neighborhood in both classical and relativistic spacetimes.

Within the mini-spacetime, we can identify events, such as T, which come temporally later than event O. A definite amount of time will elapse between event O and T as we pass along the spacetime trajectory OT. That time is physically measurable, for example, by counting the ticks of a clock that moves along the trajectory OT. Similarly, we can find an event S that occurs simultaneously with O (for at least one observer). The distance along the interval OS is physically measurable. For example, we might contrive a measuring rod to pass through events O and S so that opposite ends occupy events O and S simultaneously (at least for one observer). In the mini-spacetime, the trajectory OT represents an inertially moving body and the interval OS a straight line.

We may also sum intervals represented by OT and OS using the familiar parallelogram rule for vectors. If we add OT and OS, as shown in figure 3, we arrive at OT'. If OT is the trajectory of some body, then OT' will represent the trajectory of a body moving in the direction of OS with respect to the first body. [2]

---

[2] A footnote for experts who suspect a sin against mathematical rigor in the "infinitesimal" talk: the mini-spacetime surrounding event O is really the tangent vector space at event O in the manifold. So displacements t = OT and s = OS are really tangent vectors. The times elapsed and spatial distance along the displacements (squared) are really the norms of the corresponding vectors, using the appropriate geometric structure. In the case of relativity theory, the norms are supplied by the metric tensor g. In the case of the Newtonian theory, I use a Cartan generally

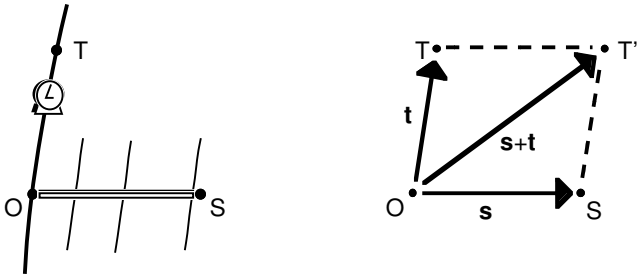



Figure 3. Measuring time elapsed and spatial distance between infinitesimally close events; the addition of displacements between infinitesimally close events.

---

covariant formulation. The norm of t is derived from the absolute time one form dT; the norm of s is derived from the degenerate spatial metric h. Talk of a mini-spacetime of infinitesimally neighboring events is not so misleading, however. The tangent space at O can be mapped onto a neighborhood of O in the manifold of events by such maps as the exponential map. The intervals OT and OS represent inertial motion and spatial straights because the exponential map assigns the vectors t and s to events along geodesics through the event O. By mapping onto neighborhoods of arbitrarily small size, one can come arbitrarily close to the geometric properties claimed for them. The mini-spacetime mimics the full Minkowski spacetime of special relativity in so far as this mapping need not be restricted to arbitrarily small neighborhoods to recover the properties claimed. The map can be from the tangent space to the entire Minkowski spacetime. That is, select a Lorentz normal coordinate system with origin at O in which the metric is g = diag (1, -1, -1, -1). Vectors t = (t, 0, 0, 0) and s = (0, s, 0, 0) are mapped to events T = (t, 0, 0, 0) and S = (0, s, 0, 0) in the manifold, so that the metrical time elapsed along the geodesic OT is t and the metrical distance along the geodesic OS is s. The addition of vectors s + t corresponds to translation from O to event S and then translation by (t, 0, 0, 0) to arrive at event T' = (t, s, 0, 0). The distance OT' corresponds to the $\sqrt{\text{norm}}$ of s + t and is $\sqrt{t^2-s^2}$. This last correspondence exploits the flatness of a Minkowski spacetime and, in general, precludes the vector space mimicking arbitrary spacetimes in general relativity.

## How Classical and Relativistic Spacetimes Differ: The Entanglement

Classical and relativistic spacetimes differ in the disposition of measurable times and distances in these mini-spaces. In the relativistic case, times measured are entangled with distances in a way that they are not in the classical case. The difference is illustrated in Figure 4. We take an arbitrary displacement OT between successive events. We add successively add spatial displacements OS to it to recover new displacements OT', OT'', OT''', etc . In the classical case, the time elapsed along the displacements OT, OT', etc. will all be the same. In the relativistic case, they will differ. One will be the greatest value (labeled "max." in the figure) and the times elapsed will diminish as we proceed in either direction. (This decrease can continue until the time elapsed has dropped to zero in which case the displacement represents the trajectory of a light pulse.)

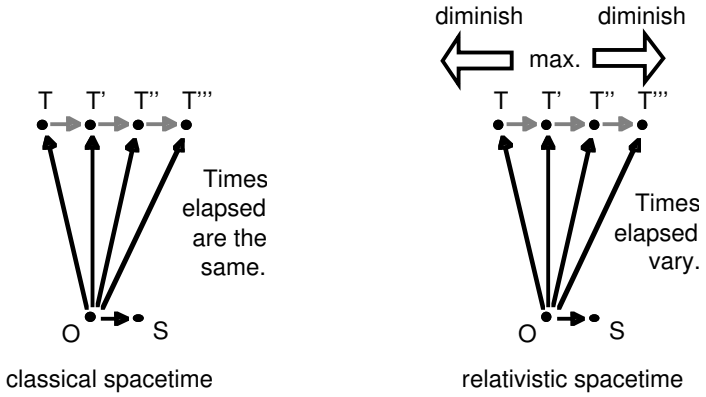



Figure 4. Entanglement of time and space in relativistic spacetimes.

This relativistic effect in the appropriate circumstance is equivalent to the time dilation (slowing) of clocks in relative motion. The clock at rest follows trajectory OT'. The remaining clocks move with respect to it along trajectories OT, OT'', etc. and record a shorter time elapsed. [3] This

---

[3] Another footnote for the experts: The result is really that the Newtonian theory uses separate structures, dT and h, to determine times elapsed and distances, where relativity theory uses a Lorentz signature metric g for both. So taking the vectors t and s above, the Newtonian structures will assign the same norm to t and s + t since dT(t) = dT(s+T). In relativistic spacetimes, the

entanglement is also closely related to a the relativity of simultaneity; it can be used to generate a form of the relativity of simultaneity restricted to the mini-spacetime. For the details, see the Appendix.

## Other Candidates

The analysis above captures as best I can the sense in which space and time are entangled within relativity theory and in a way that respects the requirements of the Introduction. The literature is thick with other proposal that are intended in greater or lesser extent to capture this particular novelty of relativity theory. I explain why I find some of the more prominent wanting.

## Time is the Fourth Dimension

This notion fails the requirement of Novelty . Events, points in space at a particular time, form a four dimensional manifold. That just means that the set of events can be coordinatized by four numbers. This is true in both classical and relativistic theories. In both we can say that the transition from the three dimensional manifold of spatial locations to the four dimensional manifold of events requires an extra dimension associated with time. The assertion is banal.

One m ight try to rescue the notion from banality by urging that there is something more inherently four dimensional about relativity theory. That is true. It arises from the entanglement of space and time in relativity. As we have seen, however, that entanglement involves something more than the four dimensionality of the manifold of events. It involves the measurable times and distances between events and how they become interrelated. The observation that time is the fourth dimension of this spacetime hardly captures this entanglement. Is the tacit claim that time is not just a dimension of spacetime but one that is just like the three spatial dimensions? That is a

---

metric g assigns different norms to them since $g(\quad t,t) \neq g(t+s,t+s)$. The entanglement lies in the metrical structure; the addition of a spacelike vector to a timelike vector alters the norm of the vector.

falsehood. The temporal aspects of spacetime always remain distinct from its spatial aspects. [4]

Indeed “time is the fourth dimension” is a mischievous slogan. It inevitably misleads novices seeking to distill the essence of relativity. Its banal meaning is so obvious that they are drawn to seek a profundity in its connotations. In 1903 the Wright brothers set us free from the two dimensional surface of our earth and allowed us to soar freely in the third dimension, That dimension had always been before us, but we could not exploit it. It controlled us until the Wrights liberated us. Did Einstein in 1905 and Minkowski in 1907 repeat the feat? Did we learn through relativity theory how to free ourselves from the shackles of a three dimensional world and roam freely in fourth dimension, time, that had always stood before us? Of course not.

## The Determinateness of the Future

When Minkowski (1908) introduced the routine use of spacetime into physics, it seemed that this represented the victory of a particular view of time. Minkowski’s spacetime represented all there was: past, present and future, and all at once. Did this finally vindicate an idea whose pedigree traces back to Parmenides in antiquity: time and change are mere illusions? To draw this as a moral of relativity theory, however, violates Novelty . The four dimensionality of the manifold of events is shared with classical theories.

Might there be something special in the nature of the a relativistic spacetime that supports the illusory character of change? An ingenious line of analysis suggests there might be. The argument exploits the spacetime diagram within special relativity shown in Figure 5. Inertial observer A will judge events $A_1$ and $B_1$ to be simultaneous. Inertial observer B moves with respect to A and observer B judges a different set of events to be simultaneous with event $B_1$. It includes

---

[4] Timelike and spacelike vectors remain distinct. We can of course use an imaginary time coordinate in special relativity—$x_4 = ict$—so that the line element becomes $-ds^2 = dx_1^2 + dx_2^2 + dx_3^2 + dx_4^2$. The symmetry of the four coordinates is an illusion. The first three coordinates are reals; the fourth is imaginary.

event A 2 in observer A's future.

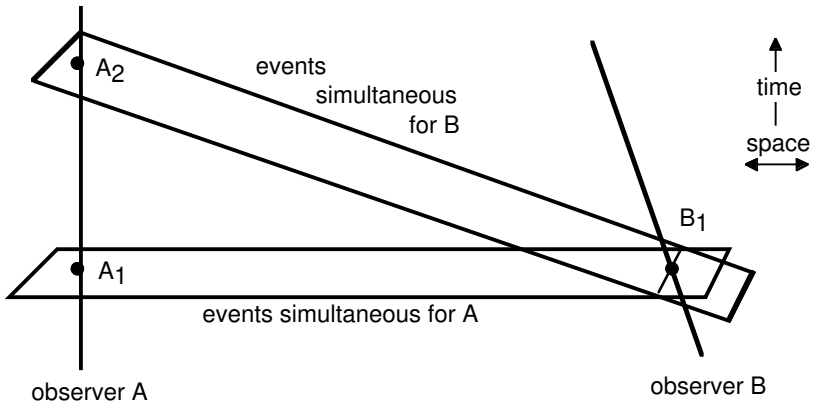



Figure 5. Determinateness of the future?

The argument is then that these judgments of simultaneity allow us to infer that event B 1 is determinate with respect to event A 1 and that event A 2 is determinate with respect to event B 1, so that overall event A 2 is determinate with respect to event A 1. That is, the future event A 2 is determinate with respect to the past A 1.

For our purposes the immediate problem is with Robustness . The argument exploits the relativity of simultaneity and that, we have seen, holds only infinitesimally once we pass to general relativity. Even if that problem could be remedied, we would still face further difficulties. Why, for example, should a judgement of simultaneity be sufficient to allow to us infer determinateness? Why should the relation of determinateness be transitive when we are combining judgments of determinateness from different observers? If these concerns are justified, it would violate Modesty as well. The subject has entered the literature many times. For an entry to this literature see Maxwell (1993), and Stein (1991) and, for broader viewpoints, Capek (1966) and Grünbaum (1971).

## Conventionality of Simultaneity

Einstein used the definition of Figure 13 (See Appendix) to determine which events are simultaneous. Hans Reichbenbach interpreted Einstein's use of a definition as revealing an important convention in the logical structure of relativity theory. If Einstein's definition really is

just a decision on the use of a term, other uses could have been entertained. So, Reichenbach urged, any event between $A_1$ and $A_3$ at A could be deemed simultaneous with event $B_1$; the choice is a matter of convention.

Might this conventionality be the appropriate expression for the entanglement of space and time in relativity theory? The proposal violates Robustness in exactly the same way as the relativity of simultaneity, since the analyses of the conventionality of simultaneity are conducted within special relativity. A version of the conventionality thesis can be created in general relativity by mimicking the analysis within the mini-space. As with the relativity of simultaneity, the conventionality fails if we relate the mini-spacetimes to the larger spacetime in so far as the larger spacetime can host a single preferred relation of simultaneity.

Aside from this problem, the claimed convention has been debated vigorously without a clear decision in favor of either side. The debate has been wide ranging. In my view, its failure to be resolved results from lack of agreement on just what it takes to be a simultaneity relation. What is its physical meaning? Is it synonymous with determinateness—whatever that might be? What are its necessary formal properties? Must it be a transitive relation? Without clear answers, the debate meanders. In one reading that does entail transitivity, defining a simultaneity relation is equivalent to defining a time coordinate in spacetime, where the time coordinate cannot assign equal times to events that can be causally related. That trivializes the convention as merely a part of our broader freedom to choose coordinate systems arbitrarily. It also makes simultaneity conventional in cases in which it manifestly is not, such as in the spacetimes of standard cosmology. Yet the convention does tap into something important and novel in the spacetime structure of relativity theory: there are many more pairs of events that cannot be causally connected than there are in classical theory. (See Section 4 below.) So should we say events are simultaneous just if they are not causally connectible? That violates transitivity and is less useful as a moral because of the ambiguity in the notion of simultaneity. Why not just take the greater freedom in lack of causal connectibility as the moral directly? Its meaning is clearer. If we take the stronger position that simultaneity, whatever it may be, is an inherently causal notion and so

must be definable in terms of causal notions, then Einstein's simultaneity relation turns out to be the only non-trivial, transitive relation so definable. But why should we demand that simultaneity is so definable? For further discussion, see Norton (1992), Janis (1998).

# 3. The Entanglement Spacetime and Matter

## Spacetime Loses its Absoluteness

The g eneral theory of relativity extends the special theory by the incorporation of gravitation. The standard approach had been to treat the gravitation field as a structure contained within spacetime, so that spacetime, the container, and the gravitational field, the contained, remained distinct. Einstein blurred this division of container and contained. The gravitational field became a part of spacetime itself. In the standard approach, we say the earth orbits the sun because the gravitational field of the sun deflects the earth from the natural, uniform, straight line motion dictated by spacetime. In Einstein's theory, we say that the presence of the sun disturbs the geometry of spacetime and that affects what are the natural motions for free bodies. Those natural motions now direct a freely moving earth to orbit the sun.

In adopting this new role, the character of spacetime was altered fundamentally. Formerly spacetime provided an immutable arena in which the processes of the world unfolded. This Einstein (1922, p. 55) characterized as the absoluteness of spacetime, a characteristic special relativity shared with the older classical theory of Newton:

> Just as it was consistent from the Newtonian standpoint to make both the statements, tempus est absolutum, spatium est absolutum, so from the standpoint of the special theory of relativity we must say, continuum spatii et temporis est absolutum. In this latter statement absolutum means not only "physically real," but also "independent in its physical properties, having a physical effect, but not itself influenced by physical conditions."

In Einstein's new theory, this absoluteness was lost. Spacetime is in turn altered by what it

contains. The presence of the sun alters the geometry of spacetime in its vicinity.

In the remainder of this section, I will review two manifestations of this entanglement of container and contained. The first is the failure of a particular view of the nature of spacetime. The second is the now less certain status of energy and momentum.

## Spacetime Substantivalism

There is a natural division within the universes of general relativity. We have a four dimensional manifold of events. And we have a metric field defined on that manifold. Without that metric field, we are unable to specify how much time or space elapses between events. The events of the manifold are like different colors in the rainbow. We can proceed smoothly though the colors: red, orange, yellow, … We can even see that orange is closer to red than yellow. But we cannot assign a distance in meters or a time elapsed in seconds to the passage from red to yellow. It is the same with the events of the manifold. The extra information of the metric field tells us how much space or time lies between events as shown in Figure 6. The aspects of the world that ordinarily we think of as gravitation are also encoded into this metrical information; its disturbance from the familiar Minkowskian disposition of special relativity is associated with the presence of a gravitational field.

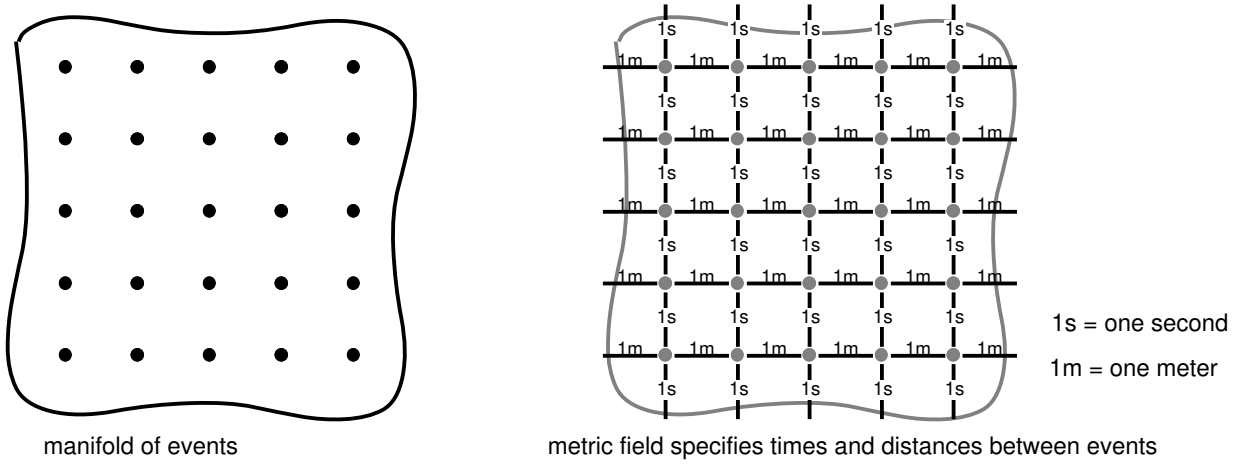


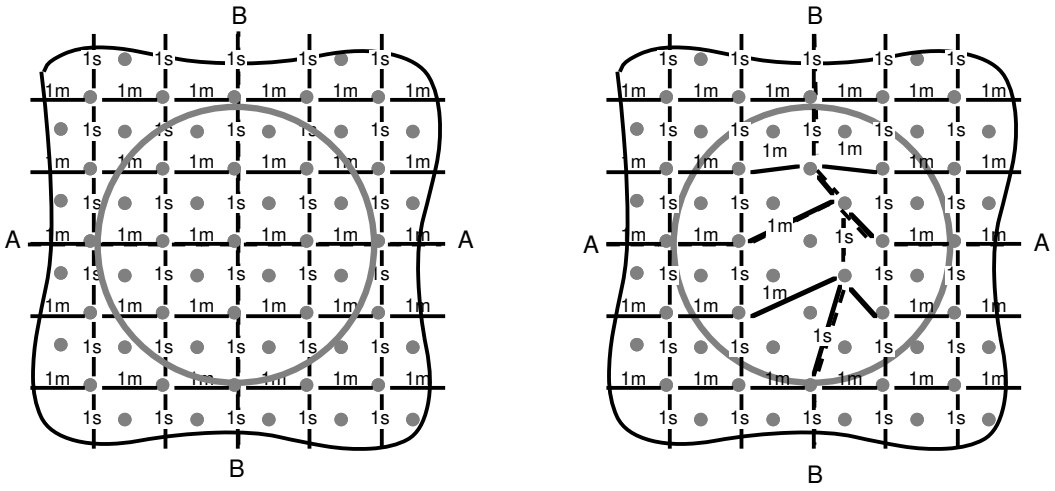

Figure 6. Manifold and Metric

Realism enjoins us to take this division seriously. The division should reflect some

objective aspect of reality. The natural reading that does this is a version of spacetime substantivalism, manifold substantivalism. It identifies the manifold of events as spacetime, the container of the metric field. Moreover it attributes substance properties to the manifold; it has an existence independent of the fields it contains.

## The Hole Argument

The hole argument first appeared in Einstein's work on general relativity towards the end of 1913. John Stachel (1980) recognized its non-trivial importance, bringing it once again to public attention. In its modern form it is used to ends different than Einstein's. (See Earman and Norton, 1987; Norton, 1992, 1999.) It shows that manifold substantivalism leads to some quite unpalatable conclusions that are usually deemed sufficient to warrant its dismissal. The hole argument exploits a property of general relativity, its general covariance. That property allows us to spread the metrical field across the spacetime in different ways. We may take the field and smoothly redistribute the same metrical properties over different events. We may effect this redistribution so that it occurs only within some arbitrarily designated region of spacetime—the "hole." Figure 7 shows the original and transformed metric fields with the hole represented as a large circle.




7. Transformation of the Hole Argument

The two spacetimes are mathematically distinct. For example imagine the straight line

connecting AA picked out by the condition that it have the shortest distance; and the straight line connection BB defined by the analogous condition that it have extremal elapsed time. The two straights will meet inside the hole. But they will meet at different events in the two spacetimes.

How are manifold substantivalists to interpret this difference? They are committed to the notion that the manifold of events has an existence independent of the fields defined on them; the events have their identities no matter what metrical properties we may assign to them. So the difference between the two spacetimes is a physically real difference for them. But it is a difference of a most peculiar type. It turns out that nothing observable distinguishes the two spacetimes. The times elapsed and distances passed to the meeting of AA and BB will be the same in both cases. Worse, everything outside the hole in the two spacetimes is identical; all the differences arise within. This is a failure of determinism of a most serious kind—the hole can be specified to be as small as we like. No specification of spacetime outside the hole can succeed in fixing its properties within. That is, the manifold substantivalist is committed to factual differences between the two spacetimes that are opaque both the observation and to the determining power of the theory.

## Entanglement of the Manifold of Events and the Metric Field

The natural response is simply to assert that the differences between the two spacetime are merely differences in mathematical description; they both describe the same physical reality. This widely accepted escape amounts to the rejection of manifold substantivalism. In particular we say that the meeting points of AA and BB in each case represent the same physical event, even though they are mathematically distinct point events in the manifold.

In rejecting spacetime manifold substantivalism, we see the entanglement of spacetime and its contents. Consider a universe with gravitation but no other contents. We cannot split off the manifold of events as the spacetime container from the gravitational field held within it in the metric field. Which physically real events are identified by which mathematical point events of the manifold cannot be decided without consulting the information of the metrical field. As we

spread that field differently over the mathematical point events of the manifold, we alter their physical identities.

## Is it Novel?

The hole argument entered the literature as a result of Einstein's work on the general theory of relativity and, in its modern guise, we infer from it that manifold substantivalism is untenable. But does it satisfy the requirement of novelty? On this there are differing schools of thought. They divide according to how one understands the requirement of general covariance that allows the transformation of the metric shown in Figure 7. One view holds this to be a special feature of general relativity only. In special relativity, for example, the corresponding metric structure is given once globally and is not subject to transformation. See Stachel (1993). In this view, the hole argument and its consequences satisfies the requirement of novelty. A second view, which I hold, allows that even classical theories may be formulated in a way that permits the transformation of Figure 7. These are called "local spacetime theories" in Earman and Norton (1987). Under that view, the failure of manifold substantivalism is common to all theories, classical and relativistic, if they are appropriately formulated, so the failure does not meet the requirement of Novelty .

## The Problem of Gravitational Field Energy Momentum [5]

In the classical view, space and time are the containers; matter is what is contained. The distinctive property of matter is that it carries energy and momentum, quantities that are conserved over time. A unit of energy cannot just disappear; it transmutes from one form to another, merely changing its outward manifestation. This property of conservation is what licenses the view that energy and momentum are fundamental ontologically. They are the stuff of matter. In the course of many interactions, they are the substances that persist, merely changing their form: the chemical energy of coal is transformed to the heat energy of the fire; to the pressure

---

[5] For a recent discussion of gravitational field energy momentum, see Hoefer (2000).

energy of the steam in the boiler; to energy of electricity in the generator; and so on.

The reality of fields as a type of matter is in part revealed by their carrying of energy and momentum. The electromagnetic field, for example, carries the energy of the sun to us in the form of sunlight. That energy can be put to good use. Vegetation absorbs it and uses it to grow. We can use it to heat water in solar hot water heaters. Sunlight also carries momentum from the sun. As a result, when sunlight blazes down on us, its impact upon us creates a pressure from the momentum imparted to us. As it turns out, that pressure is too small to be noticed by sunbathers. Otherwise, the momentum carried by sunlight would be as familiar as the energy it carries that warms a chilled bather after a swim.

Classically, one would expect the gravitational field to carry energy and momentum as well. When great masses of water run down a mountainside and through a hydro-electric power station, electrical energy is produced and that comes from the energy stored in the classical gravitational field. The effect of the lowering of the water is to intensify slightly the earth's gravitational field. This intensified field has less energy. The lost energy was imparted to the falling water as kinetic energy and then to the generators in the power station. Similarly, the falling water carries momentum that was imparted from the gravitational field. If general relativity is to return a reasonable classical view of gravity in the case of weak gravitational fields, there must be some corresponding provision for gravitational field energy and momentum. The expectation is powerful, but general relativity has shown itself to be most reluctant to give gravitational energy and momentum the homage it draws classically.

## Densities: Bit T and little t

In general relativity, the density of non -gravitational energy and momentum at an event in spacetime is represented by the stress-energy tensor of matter, represented symbolically by T--"big T." It is the structure that encodes the total energy and momentum densities due to all non-gravitational forms of matter, such as fluids, solids and the electromagnetic field. In his original work in general relativity, Einstein defined an analogous quantity, the stress-energy tensor for

the gravitational field, represented symbolically by t—"little t." It was heuristically very important to Einstein because he could use it to convince himself that both non-gravitational and gravitational energy and momentum had the same power to generate a gravitational field. Indeed that the gravitational field's own energy and momentum generates a gravitational field is one way to see how the notorious non-linearity of Einstein's theory arises. However Einstein's little t proved to cause a lot of trouble for the generations of relativists to come. The difficulty is stated most simply in mathematical language: big T is a true tensor, but little t is not; it is a pseudotensor. What this means is that is that big T can be represented independently of the particular coordinate system we use in spacetime, but little t cannot.

This might not seem to be such a problem, until we realize that the choice of different coordinate systems is merely a choice of different ways to describe the same reality. What is real is what is common to all the descriptions. What varies from description to description is merely an artifact of the mode of description. If there is a non-gravitational energy density at an event, big T is non-zero. No change of coordinate system can make big T vanish. It is different for little t. If we have a non-zero little t at some event, we can always choose a new coordinate system in which little t vanishes at that event (and conversely). It is as though we send out many different investigators to inform us of the energy density at some event. All will agree on whether there is a non-gravitational energy density present; they will not agree on whether there is a gravitational energy density present. What are we to think of the density of gravitational energy and momentum at this event when we read these conflicting reports? [6]

---

[6] Analogous problems arise in a formulation of Newtonian theory that represents gravitation as spacetime curvature. ( Wald, 1984, p.286, fn. 4) Does this mean that the effects described in the text violate Novelty . I do not think they do. I take the association of gravitation with spacetime curvature to be the novelty of general relativity and that novelty has been borrowed by the formulation of Newtonian theory at issue.

## The Gravitational Energy and Momentum of Extended Systems

The standard response to this problem in the literature is that "the energy of the gravitational field cannot be localized." (Misner, Thorne and Wheeler, 1973, §20.3, §20.4). We can only talk of gravitational energy and momentum of an extended system and not the density of gravitational energy and momentum at a particular event. In so far as I can understand this response, it really just tells us that little t should be given no physical interpretation. It should merely be used as a mathematical intermediary in computing the gravitational energy and momentum of extended systems.

The difficulty with this response is that the gravitational energy and momentum of extended systems fare only marginally better. (For elaboration of what follows, see Wald, 1984, pp. 285-295). Following the classical model, one would expect that we could take the energy and momentum densities of big T and little t and sum them up over the space occupied by, for example, a galaxy of stars to find the galaxy's total energy and momentum. In general, this cannot be done. One cannot define meaningfully the total energy and momentum of some extended system, where the total energy and momentum is to include both gravitational and non-gravitational contributions. At best, these total quantities can be defined in special cases.

The summation of the information in big T and little t to recover a total energy can be done if there is a rest frame in which the geometry of the spacetime is independent of time. That would arise if we had a completely isolated galaxy not of stars but of passive lumps of matter held apart by sticks such that the whole system just sat there completely motionless. [7] Real systems are not so

---

[7] Technically, the spacetime must admit a timelike Killing field $\xi^a$, which satisfies $\nabla_a\xi_b+\nabla_b\xi_a=0$ and is tangent to the world lines of the frame mentioned. Then the differential law $\nabla_a T^{ab}=0$, can be integrated to give a conserved quantity corresponding to the total energy. If there is a Killing field, then the differential law entails that $\nabla^a(T_{ab}\xi^b) = \nabla^a(T_{ab})\,\xi^b + T_{ab}\nabla^a\,\xi^b = 0$. This quantity $\nabla^a(T_{ab}\xi^b)$ can be integrated over suitable spacelike hypersurfaces $\Sigma$ via Stokes theorem. Following

nicely behaved. Stars radiate and thereby change their mass; stars in galaxies move about relative to each other; gravitational waves impinge upon the galaxy from the outside. All this affects the geometry of spacetime and precludes the summation.

There is another circumstance in which the total energy of gravitational system such as a galaxy can be defined, even if there is considerable internal change in the galaxy. That arises when we presume that the galaxy sits within a spacetime that becomes asymptotically flat as we travel to spatial infinity. That is, if we are far enough away from the galaxy in all directions of space, we find ourselves in spacetime that comes arbitrarily close to the Minkowski spacetime of special relativity. In such a spacetime, we are able to define the energy of a system. We can use those abilities to define the energy of a distant galaxy, since we can treat that distant galaxy in largely the same way as we would a distant object in special relativity.

## Energy, Momentum and Force

In retrospect, we should not have been taken aback by the compromising of energy and momentum in general relativity. The first thing that one learns in approaching general relativity is that the notion of force has been compromised. General relativity no longer offers a precise notion of gravitational force. It has been "geometrized away." But one cannot geometrize away force without other ramifications. Consider how intimately energy, momentum and force are related. In classical theory, if we have a constant FORCE acting on a mass for some TIME during which the mass moves through some SPACE. the ENERGY and MOMENTUM gained by

---

the usual procedure in which a boundary term is contrived to vanish, the energy of the system is recovered as $\int_{\Sigma} T_{ab}\,\xi^{b} n^{a}$, where $n^{a}$ is a unit normal vector to the surface $\Sigma$. This energy is a constant in the frame because of the vanishing of $\nabla^{a}(T_{ab}\,\xi^{b})$. If there are corresponding spacelike Killing vector fields, then the total momentum of the system in the direction of the Killing field can be defined analogously.

the mass is related to FORCE according to:

ENERGY = FORCE x SPACE

MOMENTUM = FORCE x TIME

If gravitational force has somehow been compromised—geometrized away—then we should expect the same to happen to the other dynamical quantities in these two equations, ENERGY and MOMENTUM.

Correspondingly, in those cases in which the classical notion of force is restored, we can define the energy of the gravitational system. The restoration of a Minkowski spacetime in asymptotically flat regions of spacetime allows us to use the resources of special relativity to reintroduce a notion of gravitational force. It is identified with the geometric perturbations of the metrical structure from the exact flatness demanded by a Minkowski spacetime.

## The Infection Cannot Be Restricted to Gravitation Alone

Perhaps our hope is that we might localize the difficulty to gravitational dynamical quantities only. Gravitational force has been geometrized by general relativity, but electrodynamical force has not. So perhaps gravitational energy and momentum are compromised, but not electrodynamical energy and momentum. This hope is quickly dashed once we recall the interconvertibility of all forms of energy and momentum. What classically gives energy and momentum their special status ontologically is that they are conserved through all change, as we saw above. Once we have the transformation of a non-gravitational form of energy into a gravitational form, we are no longer able to assert the conservation of the energy and momentum in all generality. The total energy and momentum of the system is no longer well defined. These conversions are pervasive. We convert energy and momentum between gravitational and non-gravitational forms every time a body rises and falls above the surface of earth; as the earth proceeds in its orbit from aphelion to perihelion to aphelion; as a star undergoes gravitational collapse and releases its gravitational energy as radiation; and so on.

When these conversions proceed, the matter contained by space and time, as captured by its

energy and momentum, loses its unequivocal character as content of space and time. It becomes entangled with its container in the structure of space and time itself.

### All is Geometry?

What moral do we not draw? Once we see how gravitation is geometrized in general relativity, it is easy to expect that all remaining forces will be geometrized as well. Einstein was the most prominent advocate of this view; his legendary quest for a unified field theory amounted to the quest for a theory that geometrized electrodynamical force just as general relativity had geometrized gravitational force. The ontological moral would be that everything is fundamentally a geometrical property of space and time. While one cannot exclude future research establishing such a result, we do not have it now. To conclude this moral would violate the requirement of Modesty.

## 4. Entanglement of Spacetime and Causation

### The Speed of Light

The most prominent fact about special relativity is the constancy of the speed of light. This constancy is somehow built into the essence of space and time according to the theory; spaces and times repeatedly contort themselves in order to preserve its constancy for all inertial frames of reference. We must of course recover from the novice error that there is something special about light that brings all this about. In principle, light—electromagnetic radiation--has nothing to do with it. Special relativity would say the same things about space and time in a completely dark universe. The real result is that there is special speed built into the structure of space and time and, in seeking to go as fast as it can, light happens to travel at that speed.

Are we to recover an ontological moral from the constancy of the speed of light? That might be possible were it not for general relativity. Any such moral would violate Robustness. In general relativity there is no comparable sense of the constancy of the speed of light. The constancy of the

speed of light is a consequence of the perfect homogeneity of spacetime presumed in special relativity. There is a special velocity at each event; homogeneity forces it to be the same velocity everywhere. We lose that homogeneity in the transition to general relativity and with it we lose the constancy of the speed of light. Such was Einstein's conclusion at the earliest moments of his preparation for general relativity. Already in 1907, a mere two years after the completion of the special theory, he had concluded that the speed of light is variable in the presence of a gravitational field; indeed, he concluded, the variable speed of light can be used as a gravitational potential. [8]

## Causal Structure

While the constancy of the speed of light is not preserved in the transition to general

---

[8] See Einstein (1907, 1911). It is not so easy ask if the speed of light is constant in special relativity. At first it looks like the result survives. The metrical norm of any lightlike vector is zero, so that if this zero norm measures the speed then it is always the same—although zero is an unusual measure for the greatest achievable speed. Also, in the neighborhood of any event, one can always set up measuring rods and clocks in free fall and of sufficient smallness so that they measure the same constant for the speed of a light signal. However there seems no general way to extend this constancy to measurements conducted over extended regions, as Einstein realized in 1907. Consider the simplest case of a static spacetime which can be foliated into a family of spacelike hypersurfaces with a time independent geometry. We can use any physical process to assign times to the surfaces, a kind of cosmic clock. But, in the general case, there is no way to do this so that all light signals propagate through the spaces with the same speed on this cosmic clock. For details of the these constructions, see Norton (1985, §3). Einstein's view (as elaborated in Einstein and Fokker, 1914, §2) seems to have been that the constancy of the speed of light entails that the spacetime is conformally flat, such as it was in his reformulation of the Nordström theory of gravitation. The spaces of general relativity are not, in general, conformally flat, so that sense of the constancy of the speed of light fails.

relativity, the existence of a special speed at every event is preserved. To see it, we need to return to the mini-spacetimes surrounding each event that were introduced in Figures 2, 3 and 4. In particular, if we continue to add spatial displacements to the event T, we will eventually arrive at a pair of successive events OL such that the time along OL will be zero, as shown in Figure 8. There is a corresponding displacement OL' of zero time elapsed in the opposite direction.

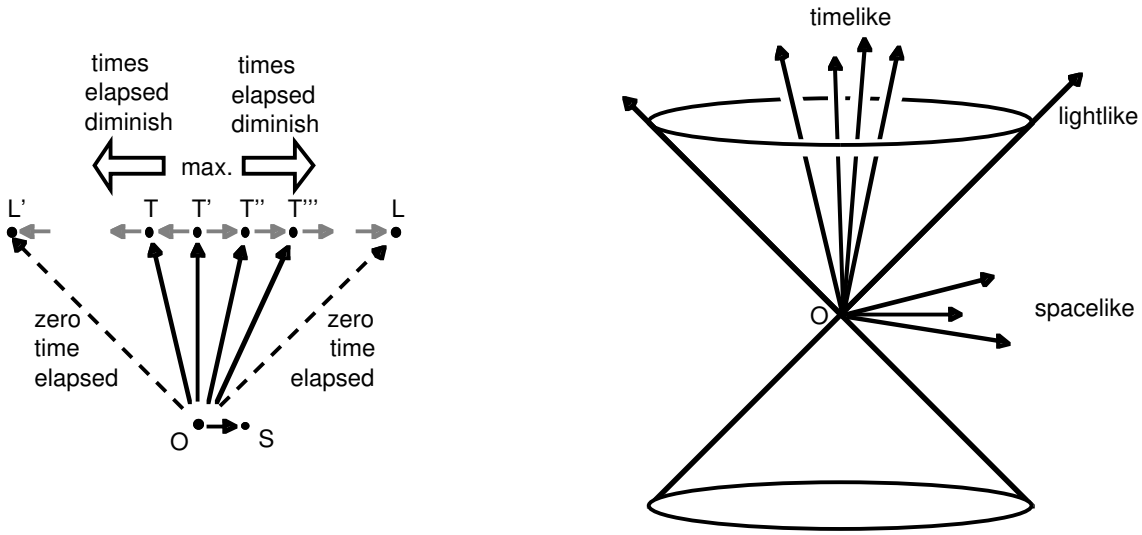



Figure 8 Light Come Structure

The displacements OT, OT', … represent timelike displacements and are instantiated by things moving slower than light. The displacement OL represents lightlike displacements and are instantiated by things moving at the speed of light. The displacements like OS are spacelike; they connect events that some observers will judge to be simultaneous. This three way division of displacements can be effected in the mini-spacetime of every event. If we add in an extra dimension of space, we find that the events lightlike related to the event O lie on a cone as shown in Figure 8. The top half of the cone represents the events in the mini-spacetime surrounding event O passed by an expanding shell of light that originates in O. The bottom half represents the time reverse, the events passed by a spherical shell of light that collapses on O.

So far I have said nothing about whether a body can move so that is passes along a spacelike displacement OS. Relativity theory does not preclude such bodies; they would represent tachyons, particles that travel faster than light. It has long been standard to assume that such propagations

are not possible. This is an additional but natural supposition fundamental to all that follows. What makes it natural is that any spacelike propagation produces temporal anomalies. Even if one observer judges a tachyon to be traveling just slightly faster than light, there will be another observer who will judge its propagation to consist of simultaneous events, so that it moves infinitely fast; and another that will judge its events to progress into the past, so that it moves backwards in time. There is no incoherence in these judgements, as long as we contrive our theory to prevent propagation into the past generating self contradictory closed causal loops. But we are protected from their realization by presuming that nothing moves along spacelike displacements. We are assured that this is not an unreasonable presumption since we have no empirical evidence of tachyons and we have built the highly successful modern quantum field theory on the presumption that there cannot be spacelike propagations in special relativity.

## The Relation to Causality as Normally Understood

It is standard in the physics literature to talk of the light cone structure as the causal structure of the spacetime. That designation can be misleading. General relativity does not have a fully developed metaphysics of causation such as would be expected by a philosopher interested in the nature of causation. Rather we should understand the causal structure of a spacetime in general relativity as laying out necessary conditions that must be satisfied by two events if they are to stand in some sort of causal relation. Just what that relation might be in all its detail can be filled in by your favorite account of causation.

The condition for the possibility of a causal relation between two events is illustrated in Figure 9. Event O and T of the full spacetime are causally connectible if they can be connected by a curve that is everywhere timelike or lightlike. (This means that at the mini-spacetime of every event on the curve, the curve corresponds to timelike or lightlike displacements.) When O and T are so connectible, a body can travel between them without ever exceeding the speed of light. Events O and S are not causally connectible if any curve that connects them must at some event be spacelike; a body that tries to travel between them must somewhere exceed the speed of light.

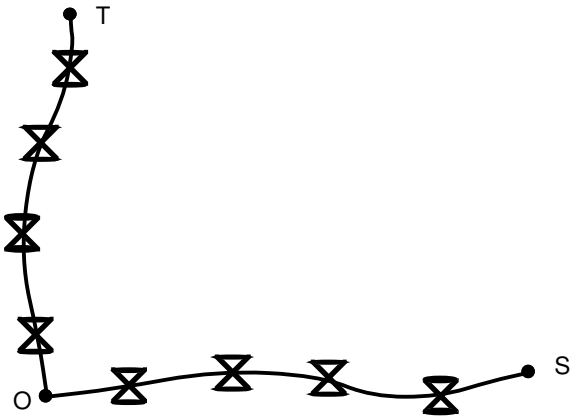



Figure 9. Causal Connectibility

## Causal Isolation

While general relativity only places neces sary conditions on causal connectibility, they prove to be quite powerful. They make possible a far richer repertoire of causal connections, while at the same time thoroughly entangling causal connections with the structure of space and time. In one part of this new repertoire, we have universes that have much less causal connectibility than we would otherwise expect.

The fully extended, matter free Schwarzschild solution is a kind of black hole and one of the simplest spacetimes admitted by general relativity. From outside the black hole it just looks like the gravitational field of our sun. If one were to fall within, one would end one's journey in a singularity as the curvature of spacetime grew without bound. On the other side of the black hole is a second world with a geometry that exactly clones that of the first spacetime. Inhabitants of that new world can also fall within the black hole and even meet those who fell in from the old world, before they come to their end in the singularity. There is a counterpart to that fatal singularity. It looks like black hole in reverse, a singularity that can emit, a white hole. The white hole singularity can causally affect both worlds. These worlds naturally combine into a single universe. However the causally intriguing aspect of this universe is that there is no possibility of direct causal connection between the two worlds. There is no way for inhabitants of one to voyage to the other or even to signal to the other. They are causally isolated.

Another example pertains to the "horizon problem" in cosmology. In standard big bang cosmologies, we can trace back the motion of matter in the universe only finitely far into the past. As we do, the matter of the universe gets compressed to arbitrarily high densities in the approach to the big bang. Take the case in which we idealize the matter as particles—"dust." Because of this unbounded compression, one might expect our piece of dust at some time in the past have had the possibility for causal contact with every other piece of dust. That turns out not to be so. If the expansion of the universe in the cosmology is sufficiently fast, a signal traveling to us at the speed of light from some nominated piece of dust can never arrive. We flee too fast. Our part of the universe is causally isolated from many others. This effect is most easily illustrated in what is known as a conformal diagram such as is shown in Figure 10. The trajectories of the dust particles, really the galaxies, are shown on the left emanating from the big bang. That representation does not allow us to see what can causally connect with what. It is a simple trick to stretch out the diagram so that all lightlike propagation proceeds along lines at 45 degrees. The big bang now appears as a long stretched out band. We read from it that our galaxy G now could have been causally affected by galaxy $G_{near}$, but not by galaxy $G_{far}$.[9] The boundary marked by the furthest galaxies that can affect G now is our "particle horizon."

---

[9] For concreteness I have in mind a Robertson-Walker spacetime filled with pressureless dust at exactly the critical density so that its spatial sections are Euclidean. The invariant interval s is given by the line element $ds^2 = -dt^2 + R(t)\, d\sigma^2$, where $d\sigma^2$ is a Euclidean line element and the time coordinate $t>0$. $R(t)=0$ in the limit as $t \rightarrow 0$, which designates the big bang. In suitable units (Hawking and Ellis, 1973, p. 138) for this most simple of cases, $R(t) = (3t)^{2/3}$. Introducing a new time coordinate $\tau = (3t)^{1/3}$, the line element becomes $ds^2 = R(\tau)\,(-d\tau^2 + d\sigma^2)$, so that the original spacetime is conformal to half a Minkowski spacetime $ds^2 = -d\tau^2 + d\sigma^2$, where $\tau>0$.

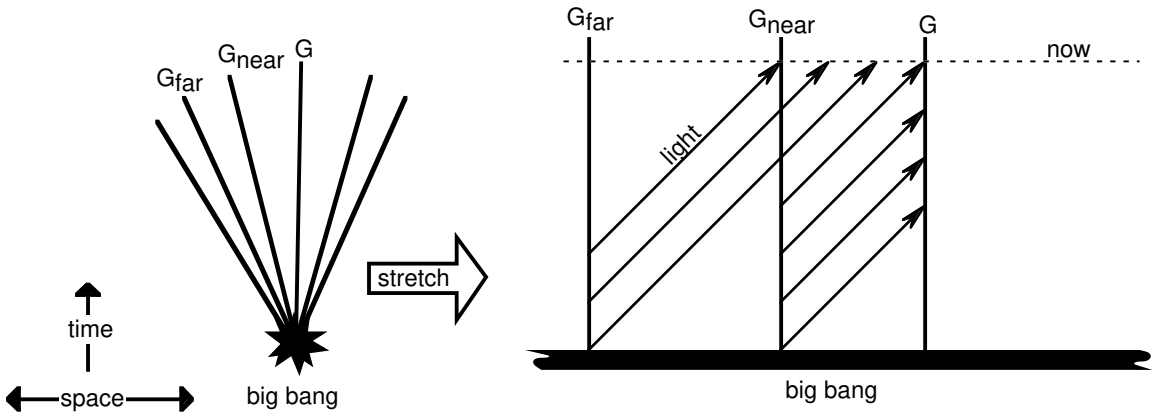



Figure 10. Conformal Diagram of a Big Bang Universe

This is just the simplest example of how horizons can separate off causally inaccessible parts of the universe. In other examples, there remain portions of the universe that are causally inaccessible to us no matter how long we wait on our galaxy, even in the limit of infinite time.

## Causal Abundance

The examples so far have shown us less causal connectibility in general relativity than we might expect. We can also have more and in ways that are traditionally of interest to philosophers.

In ge neral relativity, the trajectory of an observer through spacetime, the observer's world line, is dictated by the spacetime geometry. That geometry admits of quite complicated structure and connections. In particular there proves to be many universes in which an observer's world line can be connected back to meet to meet its own past. This can happen many ways. The simplest just involves a mathematical construction that is essentially identical to what we do when we roll a piece of paper into a cylinder by gluing its opposite edges. We can glue the future of the spacetime to its past and thereby produce a universe in which observers can meet their past selves merely by persisting long enough in time. See Figure 11.

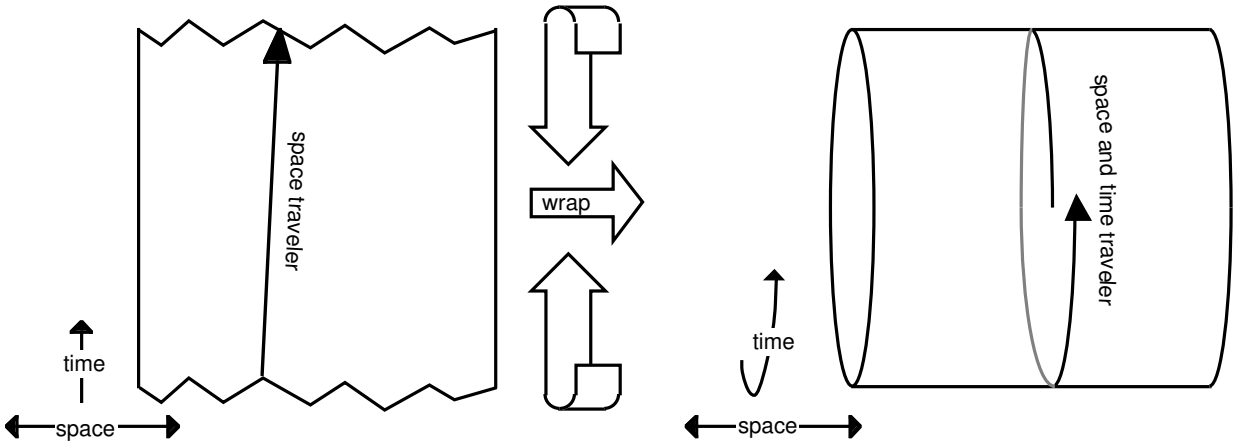



Figure 11 A Universe with Time Travel

There are less contrived but more complicated ways of bringing about this possibility. In a Goedel universe, the cosmic matter rotates and observers who accelerate sufficiently intersect their pasts. In other universes, we need only an infinitely long, dense rod of matter spinning rapidly to achieve the same end. Or in others we open up wormholes that connect one place and time with other places at different times; these are portals through which would be time travelers can pass. See Earman (1995, Ch. 6). If one understands "possible" to mean licensed by our best physical theories, then there can be no doubt that time travel is possible. That does not mean that there is time travel in our universe. Indeed a universe in which time travel actually occurs is likely to be much different from the one we are familiar with. It must be so contrived that present actions can only take place if they will cohere with the interfering machinations of a future time traveler with the past of those actions. See Arntzenius and Maudlin (2000).

In foundations of mathematics and computation, it is often taken as a commonplace that an infinity of discrete actions cannot be completed. Hence what is computable is restricted to what can be calculated in finitely many steps. If one understands "possible" to mean licensed by our best physical theories, then, at least as far as the spatiotemporal aspects are concerned, completing an infinity of computations is possible. In a sense to be explained, general relativity allows systems in which the completion of a quite ordinary infinity of manipulations is allowed. These arise in " Malament-Hogarth" spacetimes. (See Earman and Norton, 1993) The defining

characteristic of such spacetimes is that they admit world lines for a slave master pair. The slave persists for an infinity of time, perhaps fully occupied computing some uncomputable function; the master can be so located in the spacetime that, after finite time has elapsed along the master's world line, the master is able to see the entire infinite history of the slave's world line. If the slave is trying to determine if a given Turing machine halts on a given input by a simple simulation, the master will learn what the slave never learns assuredly at any finite time in the slave's life: whether the machine halts. For example, the slave may be set up to send a light signal to the master just in case the slave's program halts. At no stage of its infinite life will the slave know that the signal was sent; but the master will come to know this assuredly after a finite time of the master's. The slave and master are illustrated in a conformal diagram of a Malament-Hogarth spacetime in Figure 12.

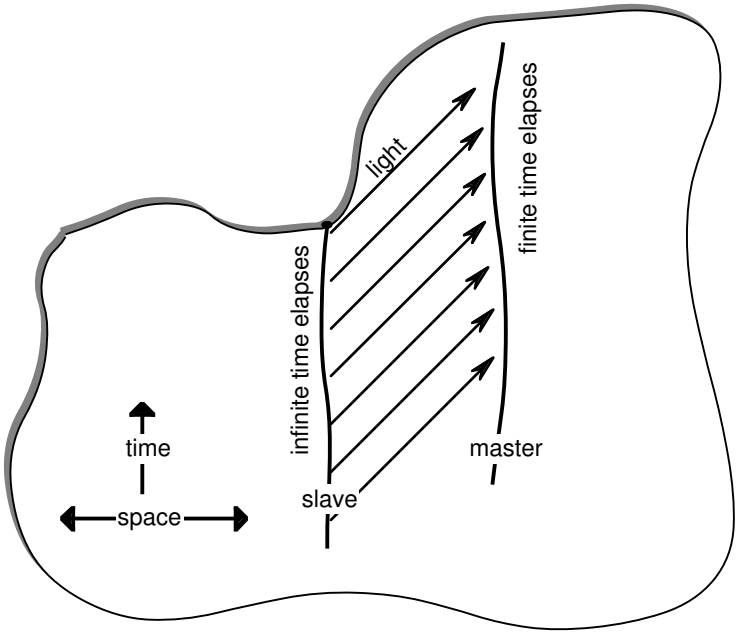



Figure 12: A Malament-Hogarth Spacetime

## Causal Theory of Time

One of the best known attempts to extract a fundamental causal moral from relativity theory is Reichenbach's (1956) "Causal Theory of Time." Its central claim is that the

spatiotemporal relations between events are reducible to causal relations: event A is earlier than B just means that event A could causally affect B. Might we find in this theory an ontological reduction of spatiotemporal structure to causal structure? I do not believe we can read this ontology of cause from relativity theory.

There are two problems. First, Reichenbach's analysis is dependent on a formal result. It is possible to axiomatize the special theory of relativity in terms of causal relations alone, so that the other spatio-temporal relations are derived relations. This formal result holds only special relativity. It fails with the transition to general relativity. The failure is easy to see. Causal connectibility of events is just lightlike or timelike connectibility. In general relativity, one can have many distinct spacetimes with the same relations of causal connectibility.[10] So an axiomatization in terms of causal connectibility alone cannot provide the extra structure needed to distinguish the two cases. As a result, the causal theory of time violates Robustness.

Second, that a spacetime theory can be axiomatized in terms of causal structure does not establish the ontological primacy of the entities taken as primitive in the axiomatization. To think otherwise creates great difficulties. There are many distinct axiomatizations possible for a given theory and we cannot take all the primitives as ontological primary on pain of trivialization. The decision of which axiomatization properly reflects the ontology is quite delicate. It seems natural to axiomatize Newtonian particle mechanics with the mass and velocity of the particles as the primitive notions and their energy and momentum as derived, although the reverse is also possible. This naturalness dissipates once one extends the particle mechanics in almost any way by, for example, adopting a Lagrangian or Hamiltonian formulation, or extending it to a field theory, or relativizing it, or quantizing it. Then energy and momentum appear more fundamental ontologically, with mass and velocity derived quantities. One might look to ontological significance in the simplest of axiomatizations. Such a principle is hard to implement

---

[10] Figure 10 illustrates such a case. The Robertson-Walker spacetime on the left has the same causal structure as a half Minkowski spacetime.

without clear guides on how to assess simplicity. In any case, Reichenbach's and other causal axiomatizations contain large numbers of postulates and informally seem anything but simple.

# 5. Relativistic Morals that Founder

Einstein's theories of relativity are really theories of space, time and gravitation. Einstein did not name them that way as a reflection of how he thought about the theories and how he came to discover them. The result is a special emphasis on relativities of various sorts in attempts to interpret the theories. The tendency has been for these relativities to be overemphasized so that morals derived from them are often unsustainable as novel lessons of relativity theory. I review a few examples.

## "All is Relative"

Need I warn anyone with a modicum of philosophical sophistication that this weary slogan gains no support from relativity theory? The relativity Einstein found in his theories is a relativity of measured quantity to observer. So the length of a measuring rod or the time of a process alters with the motion of the observer. There seems no basis for extending this relativity outside physics to ethics or aesthetics, any more that we would let the wave particle duality of quantum theory license a wave particle character for what is morally good. In any case, this sort of relativity is not novel with relativity theory. In classical physics, the energy and momentum of an object (and many other quantities) vary with the state of motion of the observer. Relativity theory has just increased the number of quantities with this relative character. Moreover the emphasis of relativity was an idiosyncrasy nurtured by Einstein. Minkowski (1908, p.83) saw the same theory quite differently. He deemed the name "relativity postulate" as a "very feeble" way to label the relevant invariance of the theory and preferred the alternative "postulate of the absolute world" in deference to the entanglement of space and time into a single spacetime. Had the coffee table philosophers attended more closely to Minkowski, might we instead be seeking to deflate the slogan "All is absolute"?

## The Relativity of Motion

Einstein discovered the special theory of relativity when seeking to reconcile the experimentally inviolate relativity of inertial motion with Maxwell's theory of electrodynamics. He then found the general theory as part of his efforts to extend this relativity of inertial motion to accelerated motion. While his motive was clear, it remains unclear whether his general theory does extend the relativity of motion to acceleration. The relativity of inertial motion of special relativity is expressed geometrically in a perfect homogeneity of its spacetime, the Minkowski spacetime. It is exactly analogous to the homogeneity of a Euclidean surface. The equivalence of all inertial states of motion is the analog of the equivalence of all directions in the Euclidean surface. In general relativity, the spacetime loses it homogeneity, as does a geometrical surface when it adopts varying curvature, such as the variegated surface of a mountain. We can now pick out preferred directions in this surface of varying curvature by adapting our directions to the curvature of the mountainside. Analogously, the varying curvature of the spacetime of general relativity allows us to pick out preferred states of motion; in a standard big bang cosmology, there is a unique rest state associated with the motion of the galaxies. So, superficially, we cannot draw the moral of the relativity of motion without violating Robustness .

The considerations rehearsed above are just introductory flourishes in a debate of great complexity with many ingenious proposals and counterproposals. For an extended survey, see Norton (1993, 1995).

## Arbitrariness of Coordinate Systems

One of Einstein's favored expressions of the extended principle of relativity was his principle of general covariance. It asserts our freedom to use any spacetime coordinate system, just as his principle of relativity of motion had allowed us to use any inertial frame of reference in our physics. One might be tempted to claim this as a moral of relativity theory, especially since it came to prominence through Einstein's general theory and so might be expected to respect Robustness . It may even have an ontological character in so far as it asserts the insubstantiality of

spacetime coordinate systems. These expectations fail, however.

Einstein's general theory of relativity was the first prominent spacetime theory to employ arbitrary coordinate systems. There is no simple way of formulating the theory without them. Earlier theories of space and time could also be written in arbitrary coordinate systems, although this possibility was obscured by the fact that the theories could be expressed in especially simple forms in specialized coordinate systems. Since all spacetime theories admit formulations that use arbitrary coordinate systems, this purported moral violates Novelty . Since a coordinate system is just a continuous labeling of events with real numbers, we might well wonder how any physical theory could restrict our purely conventional decisions on how we would like to name events. Qualms such as these support the claim, first developed systematically by Kretschmann (1917), that Einstein's principle of general covariance is physically vacuous. See Norton (1993, 1995).

## Relativity of Geometry

Both Einstein (1921) and Reichenbach, one of his earliest and best known philosophical interpreters, advocated what we would now call a conventionality of geometry. Calling it the relativity of geometry, Reichenbach (1928, §8) argues that the geometry of a physical space depends upon a choice on how lengths are compared in different parts of space. The conventionality of the geometry arises from the convention inherent in this last choice.

We cannot accept this claim as a moral of relativity theory on pain of violation of Novelty . Nothing in Einstein or Reichenbach's arguments depends on relativity theory; their arguments can be mounted equally in classical theories. Indeed Poincaré, as both Einstein and Reichenbach acknowledge, had already advocated a version of this conventionality in the form of the claimed conventionality of choice between the geometries of constant curvature. See Friedman (manuscript). I also remain unconvinced that this conventionality is supportable. If the arguments of Einstein and Reichanbach that support it are acceptable, then it seems to me that we must conclude that anything that is not immediately measurable is also conventional. See Norton (1992, §5.2).

### Relational View of Space and Time

Einstein presented his theories of relativity as a part of the relational tradition in theories of space and time. That tradition looks upon space and time as some sort of a construct. The real lies in spatial and temporal relations between bodies; space and time are abstractions from those relations. Or the real lies in relations between events; spacetime is an abstraction from them. In the light of the requirement of Realism , the advent of general relativity would seem not to favor the relationist view. Under a literal reading, general relativity is the theory of a spacetime as a fundamental entity in its own right; it is what endows events with their relational properties, such as the spatial and temporal distances between us. However too strict a realist reading of general relativity can cause trouble, as we saw in the context of the hole argument above. So we might retreat somewhat from the strongest realist reading. However that retreat is still far from what a relationist needs. To extract a relationist moral from relativity theory still seems to extract more that can be read uncontroversially in the theory. It seems to violate Modesty. See Earman (1989).

The most energetica lly developed relational approach lies in the tradition of Machian theories. Einstein originally saw his general theory of relativity as implementing a demand he saw in the writings of Ernst Mach: the inertial properties of a body do not derive from spacetime, but from an interaction with all other bodies in the universe. In spite of his early enthusiasm, Einstein came to abandon the demand that his theory of gravity satisfy this requirement. There is a flourishing tradition in Machian theories, but since it generally seeks to augment Einstein's theories in order to realize its brand of relationism, its Machian inspiration cannot be admitted as a moral of relativity theory, on pain of violation of Modesty. See Barbour and Pfister (1995).

# 6. Conclusion

---

The advent of relativity theory unsettled and energized philosophy of space and time. When Einstein overthrew the rule of Newton, it was as if philosophy was released from an unbending tyranny. In the enthusiasm to partake of the rebellion, it was easy to lose sight of the

philosophical morals that were properly to be learned from the relativity theory. They were readily confused with theses that could equally have been advanced and supported prior to Einstein's theories; or those that were appropriate only at an intermediate stage of the development of the theories; or those that Einstein himself found attractive and heuristically useful in his work, even though they failed to be implemented in his celebrated discoveries. Yet Einstein's endorsement became as sought after as did Newton's in his time.

"To punish me for my contempt for authority, Fate made me an authority myself." [11]

# Appendix: A Robust Version of the Relativity of Simultaneity in the Mini-Spacetimes

The entanglement of measured times and spaces of Section 2 is sufficient to return a version of the relativity of simultaneity that is robust as long as we remain in the mini-spacetimes. This is important since it shows that the entanglement has captured whatever is essential to the relativity of simultaneity. We can generate this version of the relativity of simultaneity by replicating Einstein's procedure of 1905 in the mini-spacetime. Einstein's procedure was based upon a simple definition as illustrated in the spacetime of Figure 13. We have two positions A and B in space. We send a light signal from A to B and it is immediately reflected back to A. By Einstein's definition, the event $B_1$ of the reflection at B is simultaneous with an event $A_2$ temporally half way between the emission and reception of the light signal at A.

[11] Attributed to Einstein in Calaprice (1996, p.8), where the remarks is identified as "Aphorism for a friend, September 18, 1930; Einstein Archive 36-598…"

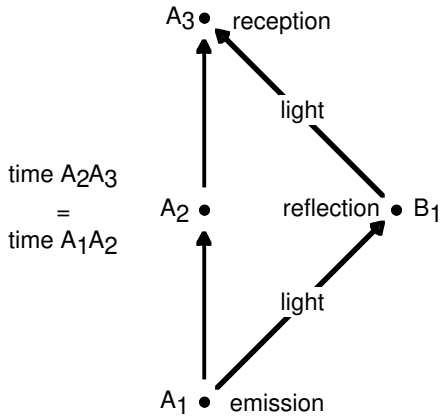



Figure 13. Events A $_2$ and B $_1$ are simultaneous by Einstein's Definition

The definition does not depend on light being used for the signal sent from A to B. It gives the same results with any signal, as long as we are assured that the speed of the signal in each direction is the same; that is, it takes the same time to go from A to B as from B to A. We will used this relaxed definition below.

## …in a Classical Spacetime

If we replicate Einstein's procedure in a classical spacetime, we immediately recover the result that two events, simultaneous for one inertial observer, will be simultaneous for all. Figure 14 shows an inertial observer following trajectory A $_1$A$_2$A$_3$ with A $_2$ the event at the temporal midpoint. Let us suppose that we have found an event B $_1$ that the inertial observer judges as simultaneous with A $_2$. That means that the two transit times of the signals locating B $_1$ are equal.

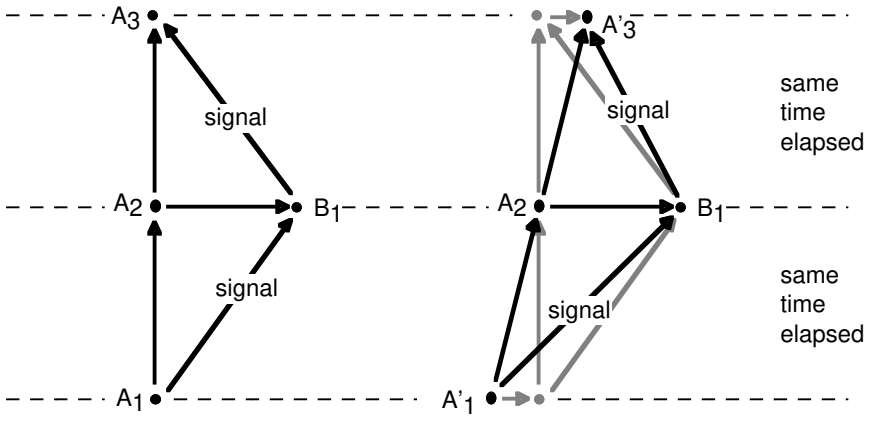



Figure 14. In a classical spacetime, if one inertial observer judges events $A_2$ and $B_1$ to be simultaneous, then so will all other inertial observers.

Now consider a second inertial observer who moves in the direction $A_2B_1$ relative to the first. That observer's trajectory is $A'_1A_2A'_3$, where $A_3A'_3$, $A_2B_1$ and $A'_1A_1$ are all parallel. We assume signals $A'_1B_1$ and $B_1A'_3$ are used to locate $B_1$. We now repeatedly invoke the lack of entanglement of elapsed times for classical spacetimes illustrated in Figure 4. That lack of entanglement will assure us that the same time passes for all the time intervals in the top half of the figure; and similarly for the bottom. First we find from it that the time for $A'_1A_2$ and for $A_1A_2$ are equal, as are those for $A_2A_3$ and $A_2A'_3$. Hence all four of these times are equal. By continuing in this way, we quickly find that the transit times for the two new signals $A'_1B_1$ and $B_1A'_3$ are equal. The conditions of Einstein's revised definition are met and the new observer judges $A_2$ and $B_1$ to be simultaneous.

## …in a Relativistic Spacetime

Once we take into account the entanglement of measured spaces and times of relativistic spacetimes, we find this simple classical result about simultaneity fails. Consider again the two observers $A_1A_2A_3$ and $A'_1A_2A'_3$ as shown in Figure 14. As before, we locate the event $B_1$ by requiring that the transit times of the signals $A_1B_2$ and $B_2A_3$ be the same. The entanglement of measured times and spaces of Figure 4 now precludes the same time elapsing along the many

intervals, unlike the classical case. In particular, it turns out that the transit times for the signals reflected at event $B_1$ are unequal, even though $A_2$ is the temporal midpoint of $A'_1A_2A'_3$. The result is that the new observer does not judge events $A_2$ and $B_1$ simultaneous, unlike the original observer. The new observer must select a new event $B'_1$ as shown in Figure 15 to satisfy the requirement that the signal transit times be the same. [12]

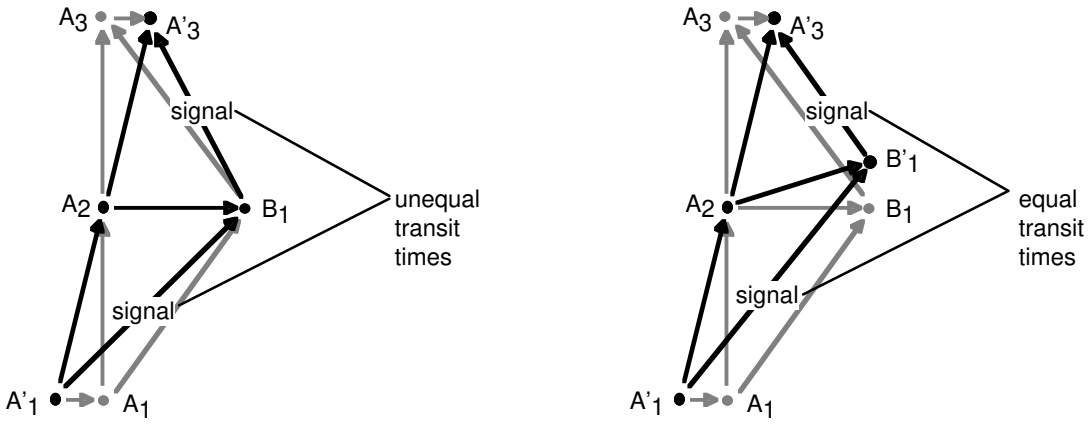



Figure 15. In Relativistic spacetimes, different inertial observers can disagree on which pairs of

---

[12] If a timelike vector $t$ is orthogonal to a spacelike vector $s$ so that $g(s,t) = 0$, then a distinct timelike vector $t'$ will not in general also be orthogonal to $s$. This variability is the robust form of the relativity of simultaneity in the tangent space. To see how it arises in the construction of Figures 14 and 15, let the two observer vectors $A_1A_2$ and $A_2A_3$ be the same timelike vector $t$. Let the vector indicating simultaneous spacelike separation $A_2B_1$ be $s$. The two signals $A_1B_1$ and $B_1A_3$ are $t + s$ and $t - s$. The condition that the time elapsed along both signals is the same is given as

$$g(t + s, t + s) = g(t - s, t - s).$$

Using the linearity of g, this equality becomes

$$g(t,t) + 2\,g(s,t) + g(s,s) = g(t,t) - 2\,g(s,t) + g(s,s)$$

which is satisfied when $g(s,t) = 0$; that is, when $s$ and $t$ are orthogonal. The above demonstration does not require the signals to be timelike vectors. They could be lightlike, in analogy with Einstein's original derivation, or even spacelike (tachyonic).

event are simultaneous.

This version of the relativity of simultaneity survives only as long as we remain in the mini-spacetimes. Once we relate these mini-spacetimes to the larger spacetime, the richer structure of the larger spacetime may select a preferred simultaneity relation. To use the earlier example, the preferred simultaneity relation of a Robertson-Walker spacetime can be projected into the mini-spacetime. So this version of the relativity of simultaneity is not admissible as a moral that must respect Robustness . The entanglement of space and time shown in Figure 4 does survive when we relate the mini-spacetimes to the larger spacetime. Indeed the entanglement becomes of great importance. Through it, we are able to say that free fall trajectories are those along which the maximum time elapses and this condition can be used as a definition of free fall trajectories. Since classical spacetimes do not have the same entanglement of space and time, no comparable definition is possible in them.

# References


Artnzenius, Frank and Maudlin, Tim (2000) “Time Travel and Modern Physics,” in E. Zalta ed. The Stanford Encyclopedia of Philosophy . (http://plato.stanford.edu/)

Bell, John S. (1987) “How to Teach Special Relativity,” pp. 67-80 in Speakable and Unspeakable in Quantum Mechanics . Cambridge: Cambridge University Press.

Barbour, Julian and Pfister, Herbert (1995) (eds.) Mach's Principle: From Newton's Bucket to Quantum Gravity: Einstein Studies , Vol. 6 Boston: Birkhäuser.

Calaprice, Alice (1996) (ed.) The Quotable Einstein . Princeton: Princeton University Press.

Capek, Milic (1966) “The Inclusion of Becoming in the Physical World,” pp.501-524 in M. Capek (ed.) The Concepts of Space and Time. Boston Studies in the Philosophy of Science. Vol. XXII . Dordrecht, Reidel, 1976.

Earman, John (1989) World Enough and Spacetime: Absolute versus Relational Theories of Space and Time Cambridge, MA: Bradford/ MIT Press.

Earman, John (1995) Bangs, Crunches, Whimpers and Shrieks .. Oxford: Oxford Univ. Press.

Earman, John and Norton, John D. (1987): "What Price Spacetime Substantivalism? The Hole Argument," British Journal for the Philosophy of Science.38 , 515-25.

Earman, John and Norton, John D. (1993) “Forever is a Day: Supertasks in Pitowsky and Malament-Hogarth Spacetimes,” Philosophy of Science , 60 , pp. 22-43.

Einstein, Albert (1905) " Zur Elektrodynamik bewegter Körper," Annalen der Physik , 17 , 891-921; translated as "On the Electrodynamics of Moving Bodies," pp. 37-65 in H.A.Lorentz et al., The Principle of Relativity , Dover 1952.

Einstein, A. (1907) " Über das Relativitätsprinzip und die aus demselben gezogenen Folgerungen," Jahrbuch der Radioaktivität und Elektronik, 4(1907), 411-462; 5(1908), 98-99.

Einstein, A. (1911): " Über den Einfluss der Schwerkraft auf die Ausbreitung des Lichtes," Annalen der Physik , 35 , 898-908; translated as "On the Influence of Gravitation on the Propagation of Light, " pp.99-108 in H.A.Lorentz et al., The Principle of Relativity . Dover, 1952.

Einstein, Albert (1921) “Geometry and Experience” pp. 232-46 in Ideas and Opinions . New York: Bonanza, 1954.

Einstein, Albert (1922) The Meaning of Relativity . Princeton: Princeton University Press. 5th ed., 1956.

Einstein, Albert and Fokker, Adriaan D. (1914) "Die Nordströmsche Gravitationstheorie vom Standpunkt des absoluten Differentialkalküls," Annalen der Physik , 44 , pp.321-28.

Friedman, Michael (manuscript) “Geometry as a Branch of Physics: Background and Context for Einstein’s ‘Geometry and Experience’.”

Grünbaum, Adolf (1971) “The Exclusion of Becoming from the Physical World,” pp. 471-500 in M. Capek (ed.) The Concepts of Space and Time. Boston Studies in the Philosophy of Science. Vol. XXII . Dordrecht, Reidel, 1976.

Hawking, Stephen W. and Ellis, George F. R. (1973) The Large Scale Structure of Space-Time . Cambridge: Cambridge University Press.

Hoefer, Carl (2000) “Energy Conservation in GTR” Studies in History and Philosophy of Modern Physics, 31 , pp.187-99.

Janis, Allen (1998) “Space and Time, Conventionality of Simultaneity,” in E. Zalta ed. The Stanford Encyclopedia of Philosophy. (http://plato.stanford.edu/)

Kretschmann, Erich (1917): " Über den physikalischen Sinn der Relativitätspostulat, A Einsteins neue und seine ursprünglische Relativitätstheorie," Annalen der Physik , 53 , 575-614.

Maxwell, Nicholas (1993) “Discussion: On Relativity Theory and Openness of the Future,” Philosophy of Science , 60 , pp.341-48.

Minkowski, Hermann (1908) " Raum und Zeit," Physikalische Zeitschrift , 10 (1909), 104-111; translated as "Space and Time," pp. 75-91 in H.A.Lorentz et al., Principle of Relativity . 1923; rpt. New York: Dover, 1952.

Misner, Charles W., Thorne, Kip S., and Wheeler, John A. (1973) Gravitation . San Francisco: Freeman.

Norton, John D. (1985): "What was Einstein's Principle of Equivalence?" Studies in History and Philosophy of Science , 16 , 203-246; reprinted in Don Howard and John Stachel (eds.) Einstein and the History of General Relativity: Einstein Studies, Vol. 1 Boston: Birkhäuser, 1989, pp.5-47.

Norton, John D. (1992) “Philosophy of Space and Time,” Ch. 5 in M. H. Salmon et al. Introduction to the Philosophy of Science . Prentice Hall, New Jersey; reprinted Hackett, 2000.

Norton, John D. (1993), "General Covariance and the Foundations of General Relativity: Eight Decades of Dispute," Reports on Progress in Physics , 56 , pp. 791-858.

Norton, John D. (1995) "Did Einstein Stumble: The Debate over General Covariance," Erkenntnis , 42 , pp.223-245; volume reprinted as Reflections on Spacetime: Foundations, Philosophy . History, U. Maier and H,-J Schmidt (eds.), Dordrecht: Kluwer, 1995.

Norton, John D. (1999) “Space and Time, The Hole Argument” in E. Zalta ed. The Stanford Encyclopedia of Philosophy . (http://plato.stanford.edu/)

Reichenbach, Hans (1928). Philosophie der Raum-Zeit-Lehre . Berlin: W. de Gruyter; Maria

Reichenbach and John Freund, trans., Philosophy of Space and Time . New York: Dover, 1957.

Reichenbach, Hans (1956) The Direction of Time . Berkely: University of California Press.

Stachel, John (1980): "Einstein's Search for General Covariance," paper read at the Ninth International Conference on General Relativity and Gravitation, Jena; printed in D. Howard and J. Stachel (eds.) Einstein and the History of General Relativity: Einstein Studies, Vol. 1 Boston: Birkhäuser, 1989, pp.63-100.

Stachel, John (1993) “The Meaning of General Covariance: The Hole Story,” pp. 129-160 in J Earman, A. I Janis, G. J. Massey and N. Rescher (eds.) Philosophical Problems of the Internal and External World: Essays on the :Philosophy of Adolf Grünbaum. University of Pittsburgh Press.

Stein, Howard (1991) “On Relativity Theory and Openness of the Future,” Philosophy of Science , 58 , pp.147-67.

Wald, Robert (1984) General Relativity. Chicago: University of Chicago Press