

1 On Learning from the Mistakes of Positivists

1 Aims of the paper

Philosophers have much extended our understanding of the foundations of spacetime physics in this century. They go on doing so. Much of this is the work of positivists and conventionalists. Friedman (1983, Introduction) tells us in his excellent book that we must learn from their mistakes if we hope to go beyond them. I will look at two of their mistakes: one is about simultaneity and the other is about the relativity of motion.

In my view, the most interesting and suggestive arguments take the old-fashioned form that the world could not possibly be as our best scientific theories say it is. One argues that physics, as the scientists hand it down to us, is conceptually awry and must be rewritten, either by amputating parts of it or by interpreting these parts as conventions, not factual claims. They are metaphysical arguments. In this century, Mach, Einstein, Reichenbach and Grünbaum have urged them. Epistemology appears in them only by ruling out concepts from a fact-stating role unless they meet some observational criterion or other.

It would be absurd to claim that this exhausts the normative role of epistemology in the philosophy of science. But I see little value in asking, about these dismembered forms of physical theory, whether they

are better evidenced than the standard forms used by the practising physicist. We philosophers are not likely to choose better than they do, even if we enlarge their view of what they may choose among. Further, a main theme will be that a focus on the observability of ideas and on parsimony of theoretical structure misled us for decades as to what spacetime theories were about. Such preoccupations are still with us. I urge that we need to look, too, at which structures in a theory may help us grasp how it may evolve under the pressures of new evidence and speculation.

I make some large assumptions. No distinction between observation and theory statements will sustain the fashionable thesis that theory is underdetermined by all possible observation statements. I have argued elsewhere (1982) that Quine's pragmatist holism gives no workable account of the deep entrenchment of propositions in theory – the pragmatist surrogate of necessary truth – or of theory change. Also, that Quine confuses the range of observational vocabulary with the range of observable fact (Nerlich 1976). I agree with others who have argued that no strong principle of charity is defensible. I believe that single sentences, as atoms of truth, have been neglected and ill understood by modern pragmatism. I have put these assumptions crudely and briefly, but the aim of my essay will be misunderstood without some brief reference to them.

2 Conventionalism and simultaneity

I turn first, and briefly, to the question of the allegedly conventional determination of simultaneity in SR.

Bridgman's (1962) proposal ought to have closed the debate on the status of simultaneity. Use of slowly transported clocks properly defines a relation which is unique, relative and not arbitrary. This is clear in a spacetime picture; a glance shows that the definition can have no rival. Yet no consensus accepted it. Malament's analogous but more effective presentation of Robb's proof that we can define simultaneity in terms of causality has no advantage of greater transparency, nor of greater fidelity to real causal relations. It succeeded because it did make one, but only one, concession to conventionalism; it was explicitly causal. Thus the core of the objection to conventionalism was perfectly clear.

Let me illustrate how limpid Bridgman's definition is. Choose a ref-

reference frame and two rest clocks, A and B , at different places in it; represent them by two parallel timelike lines in spacetime (fig. 1.1). Choose a point event, e , on one of them – A , for example. Imagine the trajectories of uniformly moving clocks which might lie as close as you please to e and equally close to some point event on B . There are two limits to the trajectories of such clocks. First, there is the light ray which intersects e and some point event on B ; second, there are trajectories which approach arbitrarily close to parallelism with the worldlines of A and B while sharing a point with each. The first limit can offer no satisfactory definition of simultaneity, since different simultaneity relations emerge depending on whether we place e on A or on B (fig. 1.2). We have no way of preferring A over B for the location of the reference event e . But if we choose parallelism with the frame clocks as our limit, then we get the same simultaneity relation whether we place e on A or on B .

Thus we have the ideal situation of a limiting trajectory which approaches as close as you please to A 's, as close as you please to B 's, and which has a point in common with each. So long as we stick to the assumption that our clocks do, indeed, measure the length of their own worldlines, the definition is unique, non-arbitrary and relative, varying quite obviously with the direction of the worldlines (i.e. the frame of reference chosen). Once we accept the definition, then the

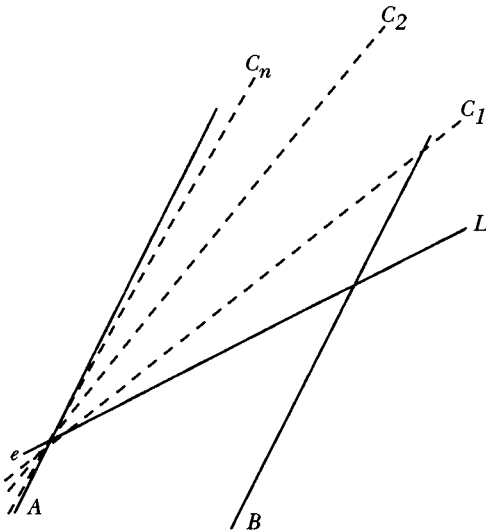


Fig. 1.1.

trajectories in question are clearly those of clocks whose motion is arbitrarily slow relative to the chosen frame. Equally clearly, the resulting relation is satisfactory in being neither absolute nor arbitrary. It is the same as Einstein's alleged convention.

If this is so clear, why was it not seen at once that Bridgman's suggestion settles the issue? Similarly, why was the work of Robb so long neglected? Partly, I suggest, because an argument about the issue was unlikely to be recognised unless it fell within the conventionalist-positivist problematic and style of presentation. Yet then, it was likely to be obscured by other conventionalist preoccupations. Bridgman, Ellis and Bowman were noticed because they argued inside that problematic. But it is foreign to the real thrust of the definition. A structural issue was obscured by too much attention to Ockhamist, reductionist ambitions. There are further reasons for seeing this episode as underlining that our thinking about SR was partly misled by Einstein and Reichenbach. And seriously misled.

A blindness to issues other than parsimony and observationality obscured and confused the issue. It is not a simple matter of the scope of causal definition. Conventions for simultaneity can begin to make sense only in a particular – and imperfectly clear – conceptual setting: one where we speak seriously of the identification of places at different times and thus of the same time at different places. That is the language of frames of reference, not coordinate systems. Frames are distinct only if their rest points and their simultaneity classes differ,

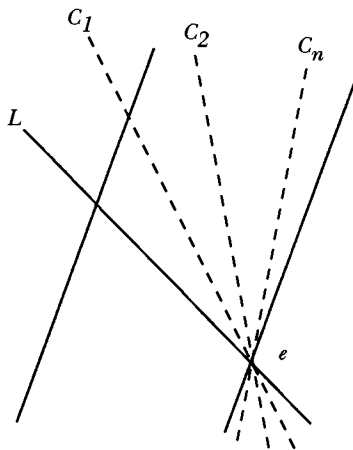


Fig. 1.2.

whereas coordinate systems differ more finely; for example, if we use polar rather than Cartesian coordinates for the same space of rest points. There is hardly an issue of convention or of anything else about simultaneity unless we take frames as somehow significantly different from coordinate systems. Frames of reference give us physics understood not through spacetime, but through space and time. The issue gets to be of significance only if we give some weight to taking space and time by themselves. There are no relevant non-factual, merely conventional statements in a spacetime treatment. Yet any convention in the one case should appear in the other. Relative to a coordinate system, the question whether or not events have different *time* coordinates has just the same status as the question whether they have different *spatial* ones. Neither question is seriously about rest or simultaneity. The space-like size of time-coordinate slices of things is well defined in coordinate spacetime physics, whether or not their spatial coordinates are identical for different time coordinates. But that is not the same as the thing's spatial shape and size. None of these mismatches can license a distinction between fact and convention as to simultaneity in a frame of reference of which there is no trace when we deal with coordinate systems. That is just one of several closely allied confusions.

Another is about the relation of frames of reference to coordinate systems for spacetime. That GR was written covariantly seemed to argue that it also lifts SR's restriction to the privileged set of Lorentz frames. But general covariance among coordinate systems is a much weaker condition than equality among frames of reference and complete symmetry for motion in all spacetimes – that depends on the structure of the spacetime itself. Clearly, inertial frames are preferred for flat spacetime. Someone might think that if skew coordinate systems are usable for a covariantly written theory, deviant relations of simultaneity in corresponding inertial frames must be permissible, too. But it is not so.

Second, the idea of a convention is the idea of a distinction between sentences which report facts and those which do not. It implies that real structure is less than we find in theoretical language. But the use of skew coordinates has nothing to do with that; it can only multiply the stock of things we related factual descriptions to. In fact, use of skew reference frames gives an unacceptable space which is anisotropic in all sorts of ways. Even if it did not, it would not give us a

convention. It is not clear what role conventions can play in simultaneity. Sometimes the literature reads as if it is telling us that moving bodies have no definite shape and size in the frame. It seems to say that, though rest and motion are matters of relative fact, simultaneity has no such factual status. It waits to be determined, even after all relative matters are fixed. That is confusion.

The source of this lies in Einstein's treatment of simultaneity, which *seems* to license SR by purging some classical spatio-temporal structure. In fact, of course, it adds some. This could not have been understood before 1908 and the advent of spacetime. Certainly, SR removes the space stratification of classical spacetime, that is, its absolute partitioning into spaces at times. But SR replaces it with a metric of spacetime, whereas Newtonian spacetime, even if curved, has only affine structure. So SR has more structure, not less. Conventionalism really did obscure this and caused SR to be seriously misunderstood for quite a long time. There is another way to put this misunderstanding: it was not seen that the concept of reference frame, as we find it in classical physics, is inadequate for use in a spacetime with full metric structure. It needs revision. In classical physics, frame-relative descriptions of physical events are completely fixed by a choice of the spacetime curves to represent the frame's rest points. But SR is a more structured theory than classical physics, and choice of rest points does not give enough. Thus conventionalism told us precisely the opposite of the truth about the nature of this theory and its factual richness. It was false that a convention allowed use of familiar sentences, though a less structured world robbed them of factual import. Exactly the opposite is true; a factually richer world demanded a richer concept of reference frame to portray it fully. To describe all the richness of the world in frame-relative facts we need a reference frame which states its simultaneity classes as well as its rest points! The proviso about richness in frame relative facts is intended to cover, for example, the relative shape and size of moving things – their *spatial* characters rather than the *spacelike* dimensions of their coordinate cross sections. See also §4.

3 *The relativity of motion*

The relativity of motion is still imperfectly understood, I think. People thought it ought to get into physics as some sort of axiom, being either

analytic, *a priori* true or, at the very least, self-evidently desirable. It is easy to see why; if motion is change of place, and place is a complex relation of distance and direction among bodies, then motion has to be a symmetrical relation among them. Other grounds were offered too, but epistemological demands on definitions for physical concepts have been among the more important. Many have felt that GR validates the relativity of motion in a way that meets the demands. So GR seems to commend, to scientists and philosophers alike, a positivist approach to spacetime foundations.

The relativity of motion is a big step toward getting space out of our ontology. That looks desirable because space is obnoxious to ontology and epistemology alike. It does enough work in classical mechanics to be indispensable but not enough to be intelligible. It seems featureless and intangible to the point of vacuity. Leibniz's objection, that it could make no difference to have created the universe somewhere else in space with all spatial relations among bodies just the same, looked plainly unanswerable.

But this was delusive. There is nothing *a priori* about the relativity of motion. Only naivety about geometry makes it seem so. It ought never to have figured as an *a priori* requirement on physical theory.

Still, something is wrong in Newtonian physics, and classical physics generally. It has been much discussed. Newton makes an absolute distinction between bodies that accelerate and those that do not. It is a causal distinction within dynamics. The related distinction between rest and uniform motion has no dynamical role; but Newton still wished to draw it. Later classical physicists spoke of acceleration absolutely, but of rest and motion only relative to some privileged frames of reference. They allowed themselves to speak of *changing* velocity absolutely without making sense of *velocity* absolutely. This is not satisfactory.

To call it unsatisfactory is not to call it false. The world could have been as Newton described it. It could also have been as the post-Newtonians described it. Each is unsatisfactory in a different way, and neither way is epistemological. Each description leaves the world incoherent; it gives us distinctions in kinematics which link with no distinctions in dynamics. It does not matter that identity of place is unobservable, but it does matter that it has no causal role in physics, while being a physical distinction. What is physics about if not space, time and matter? To put the ugliness of classical physics another way,

Galilean relativity gives us the sole classical example of change without cause; whether we see it as change of absolute place or change of distance and direction among bodies makes no difference. It is still an ugly affront to the classical presupposition that all changes are caused. The rest of classical science strongly vindicates this presupposition. The scandal of absolute space lies in that affront, not in its unobservability. But it gives a powerful motive for making the relativity of motion an *a priori* thesis.

However, rest and uniform motion are essentially observational concepts in a broadly classical setting; that is, one in which space and time are considered as separate. (If you prefer, think of a spacetime stratified at each time into a unique S_3 space, for instance.) This could not have been understood in the seventeenth century for no one knew about non-Euclidean geometries then. Spaces of constant curvature, either positive or negative, offer free mobility, as Helmholtz called it. That is, an elastic solid which fills any region of the space in a tension-free state, can so occupy any other. But an object which is in a relaxed state while at rest will not be relaxed while it is moving, if it moves in a constantly curved space. It will move under tension and distort its rest shape as a function of its speed and its deformability under a given force. Uniform motion of elastic solids is measurable by strain gauges in non-Euclidean space.

This is easy to visualise in the case of the uniform motion of particles in a dust cloud through a space of constant positive curvature. (I am assuming, here, a classical relation between space and time, not some sort of curvature of a spacetime.) For there, all geodesics eventually intersect and then diverge. In Euclidean space a dust cloud may keep its shape and size just as well in motion as at rest, since the velocity vector of each particle may be the same, in magnitude and direction, as that of every other; so the particles can move along parallel trajectories at the same speed. But this possibility is unique to Euclidean geometry, among spaces of constant curvature, at least. In elliptic space, the geodesics which the bits of dust move along intersect. The result is just as observable if the curvature is negative. We can consistently suppose that Newton's laws still hold for particles in such a space. A cloud of free-fall dust, in elliptic or hyperbolic space, can keep a volume constant in shape and size only if it stays at rest. Absolute rest is kinematically definable. Yet Newtonian laws apply to every point mass, so we here are envisaging worlds in which Galilean

relativity still holds for them. The distinction has been made observable, yet still lacks a role in dynamics; *the breach of the causal principle is not filled*. This is still an incoherent world picture, ripe to create conceptual dissatisfaction.

Even in a world where we can never actually observe them, as in Newton's world, rest and uniform motion are in principle distinguishable by observation. The distinction passes rigid positivist and operationalist constraints. We can always raise the testable conjecture that space has some slight curvature to make practical the verifiability-in-principle of uniform motion. It was simply bad luck that the suggestion could not be made in the seventeenth century.

Variable space curvature yields stronger results. It cripples all those indiscernibility arguments which Leibniz used so persuasively. For in such spaces there is no isotropy and homogeneity from one point to another; spatial points differ in the structures that surround them, in purely spatial ways. So it really might make a difference if all the matter in the universe was one metre to the left of where it actually is. It could be that only that allows all matter to fit into volumes whose geometry lets them be relaxed there. There is no free mobility, even in Helmholtz's sense, if the curvature of space may vary. Note that I am not talking about GR here, where the indiscernibility arguments may be rescued. I mean a classical stratified structure, e.g. a constant S_3 at each time. In GR, the curvature of space accompanies the distribution of matter, so there is always another possible world like the one in question, but with all the matter and all the curvature displaced together in new regions. But if we imagine a possible world where curvature is not tied to matter distribution, this salvation fails (Sklar 1985, p. 14).

4 Non-decisive a priori criteria

Let us consider a Newtonian stratified spacetime, in which we make no appeal to curvature to account for gravitation. It offers a new perspective on the incoherence. This flat Newtonian spacetime is like a pile of hyperspatial sheets, pierced, as if with spokes, by time-extended particles. The sheets may slide across each other, just as a stack of paper sheets would do, taking the spokes with them and retaining the places where they pierce the individual sheets. Each paper sheet has its own

Euclidean spatial metric; the thickness of the sheets gives a time metric. The pile fills a Euclidean 4-space in which all force-free trajectories are straight. Nothing in the model corresponds to the length of a spoke (i.e. of a particle worldline). Nor is there a metric along the curved spokes of accelerating particles, of course. So nothing corresponds, either, to the orthogonality of any straight trajectory to the spacetime sheets they pierce. If we could speak of a spoke as piercing some n seconds of stacked sheets with a minimum spacetime length, then the spacetime would define orthogonal piercing and would define points of absolute rest and thus of absolute uniform motion. These concepts would be definite whether or not our perceptual powers equipped us to detect them. But though Newton presupposed the orthogonality, his physics fails to provide a metric basis for it. The affine structure of Newtonian spacetime has no such determinacy. Nor does it help to embed the pile of sheets in a curved spacetime where all free-fall particles are geodesics and gravity is caught up in geometry. So, again, it is incoherent, conceptually: the product structure of the spacetime yields both metric space and metric time, without any metric in the spacetime itself.

It is not inconceivable that the world should be this way, of course. But, I claim, this is *non-decisively* objectionable on *a priori*, conceptual grounds; that is, in such a world, scientists and philosophers would have legitimate grounds for seeking a more coherent conceptual structure, even if the world were just as this theory describes it. A true account of the world might leave us dissatisfied, with legitimate, specifiable, though not necessitating, grounds for trying to improve the account. In such a world there is a proper case for trying to reconceive it more coherently. We might try and fail, but not because we were obtuse. I am proposing *a priori* reasons which do not necessitate, but which appeal to criteria of coherence. This is something like the converse of Kripke's suggestion that there may be necessary propositions of identity which are not knowable *a priori*. Had Newton been right, I suspect that debates about space would have persisted unresolved in epistemology, physics and metaphysics. For, I think, space is not reducible to spatial relations among objects, yet its role in Newton's physics is obscure and unsatisfactory. That is, Newton described a quite possible, but imperfectly coherent, world.

I call these principles *a priori* because they lead theory choice; they are *reasons* for a choice. Principles of teleology, determinism, continu-

ity and the principle that physics can be truly expressed in laws, are among them. They are not necessary, since teleology, determinism and continuity are either false or presently improbable. The principle that there are true laws is deeply entrenched, despite our belief that we do not yet know any law which is precisely correct. But it is not necessary. That is one way in which these principles are not decisive reasons for choosing theories.

I have tried to illustrate another. Whether such a principle is true or false may depend on the success or failure of some philosophical programme of reduction. I assume that either Newton's version or a post-Newtonian version of classical physics could be true, but not both. I also assume that not both a relationist and a substantialist version of these theories can be true. I do not think that a relationist reduction is possible, but a classical world would always present us with a good motive for trying to make it work, since it is an incoherent world: it disjoins uniform motion and mechanics. We still do not quite know how to settle these factual philosophical debates. Nor are we likely to find out until we learn more about how to identify and describe what I am here calling non-decisive *a priori* reasons round which the debate centres.

Principles of observation and coherence overlap. Any fundamental physical theory applies to every physical entity; thus every basic theory has the same domain as every other, though not the same ideology, of course. It follows that a theory will apply to measuring instruments and to observers and to processes within them involving the properties which the theory is about. The relation between the vocabulary of the theory and that of flashes, bangs, smells and tastes can only be as vague and unstructured as the latter vocabulary itself. This point hardly makes sense save within a realist context in which every subtheory in physics applies to a single domain of theoretical entities.

It may be, of course, that principles about the relation of scientific concepts to observational ones are among these *a priori* reasons which do not necessitate.

For these reasons, the supposed advantages of an unrestricted relativity of motion were illusory. On the one hand, rest and motion are, in principle, quite as observable and absolute as acceleration is. On the other, SR, in providing a metric for spacetime, provided the first metaphysically coherent and intelligible arena for physical events, in which, although absolute rest was not defined, its metaphysics gave no

presumption that it should be well defined. It commits one, not to space and places, but to spacetime. Change of place is not a basic idea of the theory because space itself is not. The crucial distinction lies between linear and curved trajectories in spacetime. No way to identify places across time is needed since space and time, by themselves, have become mere shadows.

5 *The geometry of force*

Another thread woven into the fabric of the relativity of motion is an idea about force; that the spatial relation between the force's source and a test body which it acts on should be rather simple. Classically, forces meet something like the following criteria (Nerlich 1994, p. 177).

- (a) Any force has an identifiable body as its source to which its target is spatially related in a definite way. Sources are force centres.
- (b) The conditions under which the source body acts are specifiable independently of any description of its effect (e.g. a glass rod is electrically charged when rubbed with a silk cloth).
- (c) Each centre of force acts on its target in some definite law-governed way (by contact, inversely proportional to the square of the distance).
- (d) Generalisations describing the action of forces (e.g. all like-charged bodies repel) are defeasible (e.g. unless there is an insulating wall between them) but the defeating conditions must be causal.

Criterion (c) works in classical mechanics like this. In an inertial frame, any test body whose spatial path is not a straight line can be linked to a source object, so that any force vectors which curve the body's path, are related rather simply to the position vectors which point at the source. Even when the vectors of magnetic force round a moving electron are found to be orthogonal to the position vector linking a test body to the source, and in a plane orthogonal to the motion of the source, the relation is simple enough to let us see the field as an emanation of the source. For non-inertial frames, the force vectors for centrifugal and inertial forces have no such simple relation to sources. Indeed, that is plain from the stronger fact that the cen-

trifugal and inertial force vectors will be the same whatever other objects there may or may not be elsewhere. But, I submit, it is only if the underlying geometry is Euclidean that the more general criterion linking force vectors geometrically to source position vectors is swallowed by the stronger criterion that real force vectors must depend *somehow* on the presence of sources. Only Euclidean geometry can realise the stronger, simpler criterion.

This use of criterion (c) is important for relationism and spatial reduction. That is because it gives a prominent, *physical* role to spatial relations among objects. The position vectors are spatial relations based firmly in dynamics; that plausibly contrasts with the wider embedding space which appears as a cloud of merely possible force and position vector links. I do not think this attraction overrides other difficulties which have been found in relationism, but I conjecture that the way it stresses spatial relations plays some role in reductive thinking.

There is some evidence of this in Mach's critique of Newton's rotating bucket experiment. Though Mach complained that Newton's theory conjures with unobservables, the real physics of his proposed answer to Newton's challenge makes the fixed stars sources of the inertial field. The field forces were envisaged as meeting the criteria just set out. Mach's proposal links centrifugal force vectors to position vectors for the stars in a way that was sufficiently simple, given the symmetry of the star shell. The relative motion of the symmetric shell through the non-inertial frame produces forces on the objects at rest in the frame. These force vectors at least *begin* to relate, feasibly, to symmetric position vectors and the direction of their change. The *a priori* complaint about unobservable space cannot be central, because the proposed revision results in more than conceptual changes. I am suggesting that Mach's real wish was to avoid the theoretical incoherencies that I am trying to illustrate and identify.

6 *The relativity of motion in General Relativity*

Whether the unrestricted relativity of motion is a thesis of GR in any philosophically interesting sense depends, as before, on our willingness to take rather seriously a distinction between frames of reference and coordinate systems. The distinction is, arguably, not clear enough

to allow us to do so. If we want to study the physics of SR by the use of reference frames, we must note that 'accelerating' objects can no more function as frames of reference than they can in classical physics. But this says little about our willingness to work within the corresponding spacetime coordinate systems. So it is not perfectly clear that the sort of question that occupied the positivists about the relativity of motion is not simply left behind in GR. But I hope some light may still be shed on philosophical issues by treating frames in a serious, quasi-classical way. (See Jones 1980 for suggestions as to what counts as a frame in GR and for an excellent critique of the general relativism of motion.)

Why does the unrestricted relativity of motion fail for flat spacetimes? Firstly, because the very strong symmetries of flat spacetimes allow us to pick an infinitely large set of linear time-orthogonal coordinate systems, rather simply related to each other by the Lorentz transformation. To these there correspond rigid and global frames of reference, simply transformable into each other through relations which have the dimensions of a velocity. They are privileged frames in flat spacetime, because any other consistent frames of reference are either merely local, have anisotropic spaces, or have spatial and temporal metrics which change from place to place and time to time. So only a restricted relativity of motion holds here.

Nevertheless, though flatness reveals its strong flavour in this way, it produces Euclidean spaces (or everywhere locally Euclidean ones, e.g. in hypertoroid spaces) which, as I already mentioned, strike us as bland to the point of vacuity. It allows free mobility in Helmholtz's sense and it fails to distinguish kinematically between rest and uniform motion. So the very features which make flat spacetime yield a space which looks like a nothing, and thus ripe for reduction, also deliver an unequivocal preference for highly symmetrical and global reference frames. The preference restricts the relativity of motion, thus imperilling the reductionist programme.

We get the opposite state of affairs, in general, when we move to the more complex, variably curved spacetimes of GR. We may lose any or all of the criteria which make Lorentz frames desirable in flat spacetimes. To start with, there may be no possibility of global reference frames. But even where we can use them, we will find, in general, that they are not rigid; clocks will run at different rates in different places and space expand or contract as time goes on. That is to say, we will be

obliged to treat the *coordinate* differences in time and distance seriously if we are taking the idea of frame-relative motion seriously. Of course, arbitrary choice of frame of reference may exaggerate this lack of rigidity wildly. But, in variably curved spacetimes, we lose a general contrast between rigid and non-rigid frames, and between frames that do and those that do not give us isotropic spaces.

Nevertheless, the same variable spacetime geometry, which swallows up preferred frames in its asymmetries, also, in general, gives the spaces of these arbitrary frames an obtrusive and changing geometry which has a very distinctive and prominent kinematic and dynamic character. It is impossible to think of a space as a featureless nothing, if voluminous solids are stressed, or even shattered, simply by moving inertially into regions where the geometry leaves no room for them to exist undistorted.

Consider the problem of an arbitrarily selected local frame in any spacetime, where we treat the space and time of the frame as given by the set of t -constant spacelike hypersurfaces and x_1 - x_3 -constant timelike lines that make up the coordinate lines and planes of some arbitrary coordinate system. The only requirement on the various, arbitrarily curved, space-like hyperplanes is that they must not intersect; similarly for the arbitrary time-like curves. When we project down from this arbitrary coordinate system into the space and time of the frame, the resulting arbitrary spatial geometry will have geodesics of its own which bear no simple relation to the geodesics of the spacetime (as it is easy to see from the special case of flat spacetimes). The motion of any free-fall particle whose spatial path in the frame is not geodesical, will be guided on its curve by vectors of the gravitational field, as determined by projection from the spacetime curvature of the region. These vectors can certainly not be expected to link, in any simple geometrical way, with position vectors, even of nearby sources of the matter tensor and the curvature guidance field, let alone of the array of sources as a whole. This will be especially true, of course, when the geometry of the frame's space is itself complex and changing. But even if it is not, we lose any simple link between force and position vectors. I have suggested that this gives most of its point to the focus on spatial relations, so prominent in reductionist literature. (See Jones 1980 for a rather different perspective on what relativity of motion actually calls for and for a searching critique of it.)

Finally, GR is rather rich in spacetimes which yield preferred refer-

ence frames (or classes of frames), both globally and locally. Nor are such examples among the rarer and more exotic models. Flat spacetimes are examples, as is the spacetime model of our universe preferred in modern cosmology. In fact Rosen has suggested, in several papers (e.g. 1980), that GR might be rewritten in such a way as to make this fact formally more prominent. This has not been widely accepted, but it serves to emphasise the unintended emergence of preferred frames.

It remains true that the basic concepts of GR are relativistically forged in that the theory is written covariantly in a non-trivial way. But that does not mean that the unrestricted relativity of motion is a key concept of the theory, nor that the best way for us to learn about the metaphysics of spacetime structure is to understand the mistakes of the positivists or the philosophical relativists more widely. Their ideas were often confused and baseless.

We might ask whether the nature of the question changes if we approach it in a way less tied to classical and to positivistic ideas about how to answer it. That does give rise to a rather different kind of discussion, but not one which suggests that the positivists were on to something after all. Positivists are strongly inclined to think, like relationists generally, that spacetime relations ought to be analysed as somehow or other causal ones. But in terms of what is generally meant by causality in modern treatments of spacetime (a meaning strongly influenced by positivists themselves) motion is clearly not generally relative in the spacetimes of general relativity. Malament (1985) shows that non-rotation is an invariant property of conformal transformations. The conformal structure of spacetime is essentially the null cone or light cone structure and regarded by positivists as strongly causal. But then the causal structure of spacetime determines the distinction between rotation and lack of it in an absolute fashion. Roger Jones (1980) penetratingly considers a variety of suggestions as to what should constitute a general relativity of motion. He shows clearly that GR is rather restrictive in that regard.

7 Other ways of understanding spacetime

Some think that only conventionalism and positivism offer any systematic picture of the foundations of spacetime theories (Sklar 1985,

p. 303). But this is not so. Conventionalism avowedly fails to find any foundation at all for large parts of the theories, writing them off as mere non-factual conventions. Where it does offer foundations they are narrow, opaque or dubious, as in the case of causality, or confused as in the case of simultaneity. A more pluralistic view of the search for foundations is needed. We need a metaphor of foundations without *foundationalism*.

Let me turn now to describe three ways, quite different from each other, in which we might properly regard a philosophical study (or a quasi-philosophical one) as tracing the foundations of spacetime theory. These by no means exhaust the range of options and are not even intended to represent the traditional main stream of what might be called foundational studies.

Positivist and conventionalist attempts to find the foundations of physical theories are motivated, clearly enough, by the laudably modest aim of saying no more than one must. Yet they also have an immodest tendency to rob theories of legitimate assertive power.

A taste for modesty may lead us to substitute a particular, favoured expression for other extensionally equivalent but intensionally different expressions. Relativity theories provide several striking examples of this, most notably, the very common reference to the null cone as the light cone. The word 'causal' is very widely used for the relation, whatever it is, that divides the surface and interior of the cone from what lies beyond it. But, where positivism hardens its heart against intensional distinctions among coextensive expressions, its drive for economy is no longer simply modest. This is clear when one contrasts it with a modesty in realist attempts to arrive at a foundation for theory. I turn to that in a moment.

In general, positivist investigations search for what I shall call the *restrictive* foundation of theories. That consists of a minimal ideology – *foundationalism* in short – and a minimal set of axioms, which generate a body of theorems previously judged as indispensable. To call this the restrictive foundation stresses its tendency to reduce content and prune ideology. This is a bald account of positivist aims, to be sure, but I hope their familiarity will allow me to be brief and turn to something less familiar.

One motive behind a realist examination of a theory, is to look for what I shall call its *permissive* foundations. Whereas the restrictive foundation presents, in few axioms and a lean stock of predicates, all the

theorems we simply must have, the permissive foundation presents axioms and predicates within which we can speculate most radically on how that same theory may develop. We can reflect on the results of dropping quite deeply entrenched propositions from a theory. Radical speculation can legitimately rest on the theory's broader base for its development. The permissive foundations tell us what this base is, and which theories might yield to correction without the theory's being abandoned. It is not clear that these restrictive and permissive motives for finding foundations must lead in different directions, but in space-time theory they certainly do.

It is very widely believed that a single principle about causality lies at the foundations of the relativity theories. It is the Limit Principle, that nothing outstrips light. The nice things about this Principle are that it is, very likely, true, that it is qualitative, and expressed in lean and intuitive ideology – simply 'particle' and 'outstrips'. If we take this as giving the core of the theory, we are likely to include it among the axioms, and to allow ourselves to speak of the null cone as the light cone, since photons will lie in the conical surface. That is its restrictive foundation.

But there is a price for this. It may cripple speculation within the theory. Questions which might be fruitful cannot be pursued as developments of *that* theory, nor call on its resources, to make speculation clear and definite. The hypotheses that there are tachyons, and that the photon has some finite mass, cannot be pursued within SR if we formalise it according to its restrictive foundation, yet there is a clear sense in which both these suggestions have been investigated consistently within SR (see Feinberg 1967 and Goldhaber and Nieto 1971). Unless each was compatible with the foundations of SR, in some clear sense, neither could have been considered. The idea of permissive foundations for a theory may give to realism a theoretical modesty of its own; for it is clear that there may be a proper diffidence about denying structure just as there is about asserting it. In the present case, suppose that we admit null cones, which are observationally remote structures in an equally remote object, spacetime; then a realistic attitude towards them lets us speculate about epistemically more proximate objects, tachyons and photons, in a tangible and articulate form given by the admission.

When we ask for a permissive foundation, the question is quite differently motivated, and our criteria for a good answer are not at all the same, as for a restrictive foundation. Though the two hypotheses just

mentioned are improbable, we can usefully ask how SR would survive their truth. The answer is quite obvious from a glance at tachyon theory, for all of its equations are Lorentz invariant. So the permissive foundation of SR is the Invariance Principle, not the Limit Principle. The Invariance Principle simply refers to the conical structure within spacetime and to the very powerful symmetries of Minkowski spacetime. The hypothesis of the massive photon may call for some changes in the laws of electromagnetism and optics, but whatever laws are suggested are still required to be Lorentz invariant. So that Principle is the permissive core of SR.

Third and lastly, I want to discuss Ehlers, Pirani and Schild's well-known paper (1972) on the foundations of GR. In this, the authors link various geometrical structures – projective, conformal, affine and so on – material structures such as photon and particle free-fall trajectories. In what way is this paper foundational?

Two physical conditions are needed for it to work at all; particles in free fall and null cone surfaces filled by light. However, GR is not a theory about the constitution of matter and has no ontic commitment to matter in particulate form. It can admit tachyons, and speculation about massive photons. For these reasons, we cannot see Ehlers, Pirani and Schild as presenting an ontology, nor even a likely epistemology, in any standard sense, since neither particles nor the geometry of light propagation are observationally proximate. Yet something like each of these studies is at stake in the paper.

Free-fall *particles* are needed because only point masses can be relied on to trace out the geodesics of spacetime which fall inside the null cone. These trajectories constitute the projective structure of spacetime. The centres of gravity of voluminous, elastic, massive bodies will not do, since the geodesics through various points in the solid will usually not be parallel. The bodies are gravitational multipoles. Internal stresses will tend to force the falling body off geodesical trajectories, even in the absence of external forces. The worldline of neither its geometric centre, nor its centre of gravity will be geodesic. So, in general, it is only freely falling *particles* which inscribe geodesics. Similar reflections apply to the filling of the null cone. Of course the surface of the null cone will still be the boundary between space-like and time-like curves; it will be the overlap of points elsewhere and elsewhere from the apex of the cone whether or not light or matter fills it. So what Ehlers, Pirani and Schild offer us is an elegant and familiar way in which the

geometric structure of GR might be *inscribed*. They single out a subset of worlds, which the theory makes possible, in which the geometric structures at the core of the theory and matter structures, familiar from a range of other theories, come together so that the latter trace out the former. This might be called an investigation of the inscriptional foundations of a theory, presenting one possibility within it.

The idea of an inscriptional foundation suggests an *ideal* way in which our understanding of the structure of spacetime might be traced out by matter. Real matter is not required by the theory to lie in the suggested trajectories. Recent work by Coleman and Korté (1982, 1993) argues that it may be rather less an idealisation than I have just suggested. They argue in detail, first that we can identify paths which are geodesics (and thus the projective structure) of spacetime in terms of a criterion framed in the language just of differential topology. They show how to trace particles from any coordinate system which we can identify by the smoothness and continuity of its curves and to verify whether these particles then follow the geodesics defined. They show how to correct for the effect of forces on charged particles. This, allied to the work just discussed, allows us to show clearly how both projective and conformal structures may be inscribed on spacetime in such a way as to define a Weyl structure yielding an affine connection and, eventually, a metric for spacetime.

I close with the suggestion that the philosophical investigation of physical theories may take several useful forms, that the form of a theory that philosophers find in scientific use is likely to be irreducible, and that time spent in attempting to reduce it to something else may illuminate why the theory is composed as it is, but will seldom result in a justifiable revision of it. Of course investigations into the epistemology of a theory shed light on what goes on in it, but epistemic theories of concept formation have seldom proved constructive or even insightful. One can usefully recognise *a priori* elements in a theory without relegating them to mere convention, and hope, eventually, to see how they are corrigible by observation. In short, there is a variety of ways in which we might look for foundations in physical theory, for many of which it is simply unhelpful to complain that they raise epistemological problems which they make no offer to solve.¹

¹ I am grateful to Graham Hall (Department of Mathematics, University of Aberdeen), Adrian Heathcote (University of Sydney) and Chris Mortensen (University of Adelaide), all of whom read the paper in some form and discussed it with me. They are not responsible for any mistakes.

2 What ontology can be about: a spacetime example

With Andrew Westwell-Roper

At least since Quine's seminal work *Word and Object*, ontology has been dominated by a particular model of ontic commitment, that a theory commits itself to entities through the variables of quantification and what values those variables must take if the theory is to be true. 'To be is to be the value of a variable.' The predicates of a theory do not carry the load of any commitment to what is in the world itself.

The metaphysics of spacetime does not lend itself elegantly to that style of analysis of its picture of the world. Of course there is the central ontic question whether modern physics is committed to an entity, spacetime, or whether this can be reduced or 'parsed away' in a restructured theory. But so much of the metaphysics of spacetime concerns the various levels of structure which it has – whether it is merely a differential manifold, whether it has projective, affine or metrical structure, at what level its structure may cease to be objective and give way to merely conventional determination. This seems likely to bring us to consider some sense in which the predicates of theories rather than their quantificational structure carries some form of ontic consequence.

In this essay, jointly authored with Andrew Westwell-Roper, we explore ontology of structure rather than of quantification and the kind of consequences for methodology in metaphysics is probed. We touch on the work on the foundations of special relativity due to A. Robb and John Winnie

1 Introduction

Quine's original idea about ontology was that ontic questions are referential ones: they ask what *objects* there are. Our best answers to ontic questions tell us what our best theory, in its canonical form, says that there is. We get the answer from the theory by finding what objects must lie in the value range of its quantified variables. So ontology is an objectual matter.

This position evolved somewhat when Quine (1976) argued that *all* the objects quantified over by *any* physical theory could be reduced to nothing but pure sets, although the theory applies non-set-theoretical predicates to those sets. For example, the predicate 'is an electron' might be applied to the set of coordinate quadruples for what were, before the reduction, the spacetime points in the four-dimensional region occupied by an electron. Hence, it was suggested, the ontological question of what objects must lie in the range of the quantified variables is no longer of interest, since those objects can always be pure sets. Rather, the ideological question of what predicates are required by the canonical form of our best physical theory 'is where the metaphysical action is'.

We want to pursue an idea about ontology that would complement Quine's. The objectual questions with which Quine has been concerned are now to be cast as questions about which *monadic* predicates the lexicon of natural science must contain; must it contain the predicates 'is an electron' and 'is a particle', for example? We wish to focus on the polyadic predicates required by natural science, and the *structural* questions they generate. These are questions about the kinds of relations in which the objects of this world stand. (We are not concerned here with *objectual* questions about whether relations exist over and above their relata.) Quine's metaphysics does not tell us much about how to find the structures which the world must have if a given theory of it is a good one, but we contend that an examination of the structural commitment of a given scientific theory can lay bare certain objectual commitments which are not discernible from its quantificational apparatus and monadic predicates.

Philosophical work on spacetime abounds in clear and crisp examples of argument about ontic structure. Our tactic will be to present some of these as briefly and simply as possible in a way that shows the point of treating them as ontic questions which are not objectual

ones. Quine's view of philosophical argument finds a different place for these discussions – the place of how to write a theory in its best canonical form. But we think that this obscures what is most interesting about them. They concern how far one structure of the world may underlie certain other structures as their basis, and they concern just what this basic structure really is. We argue, at the end of the paper, that the answer to the main objectual question of spacetime physics – whether spacetime itself is in the ontology of our best theory – depends on answers to the structural questions.

The main structural questions of spacetime are standardly expressed in this form: can we identify F structure with causal structure? Here F dummies for 'topological', 'conformal', 'metric' and other structural descriptions. Something is wrong with this standard form. The trouble is with 'causal' (as a number of writers agree). We will say why that is not quite right at some length, later. Elsewhere, one of us suggested that the question was about what structures we could get out of the *topology* of spacetime. That is a little misleading, too, for we could obviously get topological structure from topology. Perhaps the flavour of these structural questions might best be given to the general philosopher by phrasing the main question (about metrics) thus: can we identify the quantitative structure of spacetime with a qualitative one?

Conventionalist theories of spacetime give a positive answer to the question, but only in a weak sense. In these theories, all spacetime structure is causal, but this only means that there is, in fact, no metrical structure in spacetime itself. Our theoretical *language* has a metric structure but this is really all talk, so the conventionalist tells us, and which sentences of this language turn out true rests on arbitrarily chosen conventions and reflects no structure in the world. We will not talk about conventionalism in this paper.

According to the theories we want to discuss, the metrical structure of spacetime may be fixed uniquely by its causal (rather, by *some* qualitative) structure. There is no question here of metrical language as something like superfluous baggage. Even in the examples where the programme fails one cannot equate the metric with causality by robbing metrical language of its apparent content. Either causality identifies the complete unique metric congruence structure (the topology, or whatever is in question) or the project fails.

These theories are about structures in spacetime, not about objects

in it. They tell us what relations hold among these structures and which are primary or which independent. They answer questions which are recognisably about scientific realism, in one of its forms, and what is that if not ontology?

2 *Structures in spacetime*

In this section we want to report some results very briefly: successes, failures and prospects. In the next, we comment, equally briskly, on doubts about whether the basic relation really is causality. Later sections get back to some philosophical repercussions of the results and the doubts.

As long ago as 1911 A. A. Robb completed a remarkable account of Minkowski spacetime which treated its whole metric congruence structure in a coordinate free way and which required no use of rods or of clocks (Robb 1914). The theory of these general classes of measuring instruments standardly involves quite complex quantum mechanical ideas to explain their (rough) invariance. This seems unsuitable in the foundations of a theory which need lay down no quantum postulates for its own main development. Robb's treatment takes a single qualitative relation, which he called *after*, and postulates a number of conditions for it. These turn out to be enough to determine all the geometric structure for Minkowski spacetime which it seems reasonable to ask for. Consider these examples.

(a) Select a time-like line in Minkowski spacetime and choose a point on it. Then one can define a causal structure which picks out a hypersurface orthogonal to the line at the point. It is defined (roughly) by the intersection of two light cones from two point events on the line such that the given point is midway between them. The significance of the construction is that orthogonality is thus tied to simultaneity in a reference frame in a direct, not a conventional way; for the (hyperplane) of events orthogonal to the line at the point is just the class of events simultaneous with the point event on the line (with respect to a frame for which the line represents at rest point). The simultaneity-classes definable in this way are just the events simultaneous by standard signal synchrony, those made simultaneous by setting $\varepsilon = \frac{1}{2}$ in short, just the simultaneity classes we want. It has

been shown that more complex, less ‘natural’ simultaneity classes proposed as equally acceptable by conventionalists cannot be defined by any other purely causal construction. Thus the *standard* sets of simultaneous events in any reference frame can be selected without arbitrariness by a purely causal construction (see Malament 1977b).

(b) A series of causal definitions of curves, lines and planes leads on to the whole metric congruence structure of Minkowski spacetime. For any pair of intervals one can find a construction which identifies them as either congruent or not, whether the intervals be space-like, time-like or light-like. Much of the work lies in showing that the construction, based on causal connection or *after*, does actually correspond with the intended geometrical object. Thus it is by a series of definitions that the causal or ‘after’ vocabulary is, in certain constructions, *coextensive* with the geometric one. It shows that we can identify metrical classes with qualitatively specified ones.

This is a surprising and, perhaps, a gratifying result. It appears to carry out Leibniz’s causal programme in a rigorous way, for all spatial and temporal quantities and relations are, indeed, identifiable by means of causal conditions, so long as one sticks to Minkowski spacetime. We can grasp its structure without reliance on the methods of coordinate systems or reference frames and without the use of instruments (rods or clocks) extrinsic to the structure we are probing. In the end, it looks as if we can reduce all this to a matter of concrete relations among concrete events (Winnie 1977).

Unhappily, the result does not easily extend from the case of Minkowski spacetime to Riemannian spacetimes generally. The homogeneity of Minkowski spacetime allows causal constructions to incorporate global constraints on which the success of Robbian definitions depends. But in the full range of spacetimes envisaged in General Relativity one cannot even define the topology of spacetime, in many cases, by means of Robbian constructions. Clearly, this fact weakens the case for saying, even for Minkowski spacetime, that the metric is a causal structure rather than merely that it *coincides* with one (Sklar 1977b).

Nevertheless, the idea that causality – that is to say, the idea that some basic qualitative relation rather like causality – is of fundamental physical importance is strong and apparently getting stronger. Thus Robert Geroch (1971) claims a widespread belief among relativ-

ity theorists that the first, and most basic conditions to consider in any example of a spacetime, are the causal ones.

The situation in the literature, which is by now vast and almost all of it very complex, is a little strange. Exact solutions of the Einstein equations are not easily had and, as we understand it, there is still no synoptic view of the range of examples permitted. Furthermore, it appears that a wide range of spacetime structures that are quite unrealistic, physically, are worth studying, both because they are more easily surveyed in a more or less purely geometrical way, and because some form of spacetime theory may well survive the possible refutation of the Einstein equations. Causal theory seems to be studied as a means of probing not only the physics of GR but also geometric spacetime structures more widely. It is a way of studying what generates the geometry, perhaps in the hope of showing that something physical does (Geroch and Horowitz 1979). We will discuss this hope later.

(c) Last, it has been shown (Malament 1977a) that if we enrich the basic causal structure in an intuitive way then we can get at least the topology of any kind of spacetime out of the enriched structure. The basis is causal *curves*. It can be shown that every transformation of a spacetime which preserves its causal curves preserves all of its topological structure.

3 Is causality really causal?

We have been hinting rather heavily that, although the literature abounds with references to causality and the study is called causal theory, it may not really be causality that is at issue. Some qualitative relation is widely felt to generate spacetime structure, at least locally. Whatever the relation is, it must serve to pick out the class of points inside or on the surface of the null cone at a given point. That is obviously a formal requirement on the constructions causal theorists build. Causality is simply the name for whatever qualitative relation meets this requirement. But some uneasiness is often expressed that this relation may not actually be causality.

A reductive hope is enshrined in the expression 'causality'. It is the hope that Leibniz's programme might succeed, or that something like it may. Perhaps geometric structure can be quite strictly generated out of material structures. If so, that would be significant for the objectual

ontological question whether spacetime is a thing which exists independently of material objects and material relations among them.

There are several grounds for thinking that it is seriously misleading to call this fundamental relation causality. First *actual* causal connections are clearly not adequate to define what is wanted. We would have to include causal connectibility: pairs of events which *could* be causally connected. That no longer sounds like a purely physical relation. There are familiar worries about the use of modal concepts in reductive analysis which need not be rehearsed. Perhaps we should voice a conviction that the use of modalities is standardly a sign that the alleged reduction is actually circular. This case can be made out with particular clarity and certainty in many instances within spacetime metaphysics (Nerlich 1994, pp. 65–67). We take it to apply quite generally in the field.

There are two versions of the non-modal idea of causal connection. First, events are causally connected if they are linked by a particle, either massive or massless. Alternatively two events are causally connected if a signal is sent from one to the other. There are several things wrong with these suggestions but we will not mention them all (see §3). Some comments will make things clearer about what kind of study the ontology of structure is.

(a) At the end of his paper (1977, p. 198) on causal construction John Winnie points out that the basic relation, whatever it is, has to be instantiated in all spacetimes which are solutions for the equations of GR. This includes the so-called vacuum solutions in which the world is empty of matter. In these solutions mass and energy are associated with gravitational waves and with varying and evolving curvatures of spacetime. So something goes on in these universes, matterless though they may be. No particle connections occur in them, however, nor is there any electromagnetic (or indeed any non-gravitational) field through which a light signal, or any other kind of signal, could be sent. Hence there are no causal connections according to either of the above two views of causal connection. Though we might hope to speak here of gravitational influences and, therefore, causal ones between one spacetime point and another, it is not clear how this avoids presupposing that the very structures one hopes to probe are curvatures (ripples etc.) in spacetime.

This objection to causality plainly does not rest on its not having the right *extension* in the actual universe, as a matter of fact. The point

is, rather, that there are models where causal connectibility (causality) rests on no basis of actual causal connection in the model universe. In our world, at least connectibility has a non-modal basis. But the empty spacetime solutions leave the link between the qualitative relation and causal connection unsatisfactorily attenuated.

(b) Several years ago, a number of papers appeared in physics journals raising the question whether or not there might be particles, to be named tachyons, that go faster than light (see e.g. Bilaniuk and Sudarshan 1969; Recami and Mignani 1974). The question whether Special Relativity could consistently incorporate these fast particles was briskly debated but no clear inconsistency was shown. If tachyons are consistent with SR, then they are consistent with GR. No such particle was ever found, though some experimenters thought they had detected them in results that were not replicated. Discussion lapsed somewhat when it was found that the quantum fields of these particles were not stable. The hypothesis is inert, at present, but hardly dead. A revival has been mooted by Michael Redhead (Heywood and Redhead 1983), for it may help to ease the more painful aspects of recent versions of the EPR paradox in quantum theory. But what matters for us, here and now, is that the hypothesis that there are tachyons is not clearly inconsistent with special relativity. It is, however, *quite flatly inconsistent with the claim that causal connectibility is coextensive with the relation we want*. The tachyon hypothesis just is that particle connections fall outside the null cone. It plainly follows from the hypothesis that particle connections would not serve to identify the surface and interior of the null cone. This shows conclusively that there is no *fundamental* postulate of relativity which says that nothing goes faster than light. If there were, one could deduce the inconsistency of tachyons with SR in a one-line proof.

(c) Over the last twenty years there has been speculation among physicists over whether the photon has non-zero rest mass. According to the standard formulation of quantum electrodynamics (which is both a relativistic and a quantum theory), the rest mass of the photon is zero, so that, as a consequence of the relativistic equations relating the mass, energy, momentum and velocity of a particle, its speed is always the frame invariant constant c . However, other formulations have been proposed which retain SR as the relativistic basis of quantum electrodynamics, but reformulate the electrodynamics so that the rest mass of the photon is a constant the value of which is open to

experimental determination (Goldhaber and Nieto 1971; Barnes 1979, §1). If that constant were to have a value greater than zero, then the speed of the photon would be less than c , i.e. light would not travel at a frame-invariant speed in the null cones.

This may seem puzzling since the invariant speed of light plays such a prominent role in most expositions of SR, and it may well seem that to abandon this simply is to abandon SR. It is argued later that this conclusion does not follow. For the moment, we simply record the claims that SR is consistent with light's lying in the interior of the null cone, never on its surface.

Those who postulate a non-zero photonic rest mass sometimes maintain that even if light does not occupy the null cones, nevertheless some other causal process does. Thus Goldhaber and Nieto (1971) suggested that the neutrino, which at that time was thought to be massless, provides the basis of spacetime structure. The neutrino has since been found experimentally to have a small positive rest mass, but the graviton is still thought to be massless.

There are two reasons why the reductionist cannot resort to identifying null cones by means of possible graviton trajectories, however. First, our concern, like that of the reductionist, is with the ontologies of SR and GR, not of quantum electrodynamics or quantum gravity. Neither SR nor GR describes either the photon or the graviton, and we cannot provide ontological foundations for a physical theory in terms of objects which the theory does not explicitly recognize, and could not even implicitly recognize, since it could not describe them adequately. Our discussion of the photon was intended to show only that the theory that the speed of light (conceived as a classical electromagnetic wave) is less than c , might be thought compatible with SR. (The reformulation of quantum electrodynamics which allows for a speed of light less than c involves only a deviation from the classical equations of Maxwell, and does not depend on any quantum aspect of electromagnetism; cf. Barnes 1979, §1.) Similarly, for our purposes the zero rest mass of the graviton in quantum gravity just reflects the fact that the gravitational waves which are solutions of the GR field equation travel at c .

Secondly, and more importantly, the fact that gravitational waves in GR travel at c cannot be used as the basis for identifying the null cones in terms of causal processes in SR or GR. It cannot be so used in SR because SR does not recognize and cannot describe a gravita-

tional field or waves in such a field. It cannot be so used in GR because there gravitational waves are reduced to fluctuations in spacetime structure. Thus, in order to identify a gravitational wave one must first identify the spacetime structure, so that gravitational waves cannot be used to identify some aspect of spacetime structure, such as the null cones. Of course, the reductionist claims that, in one sense, gravitational waves are *not* fluctuations in spacetime structure; he will maintain that spacetime structures are ideal and can be reduced to causal processes of some kind between material bodies. However, to effect this reduction he must be able to reduce at least the conformal structure of the null cones to some system of causal processes, before he can hope to reduce the full metrical structure, including the fluctuations in spacetime curvature involved in gravitational waves, to a system of causal processes. That is, the reduction of gravitational waves to causal processes presupposes the very capacity to identify the null cones which the reductionist is trying to achieve.

(d) Sklar (1977a) suggests that the best reason for an interest in causality lies in the causal curves. That is, curves which are, everywhere, time-like. He claims an epistemic advantage for these. It is that each curve represents what might be the worldline of a consciousness. Seen in that light, it is not the causality of the curve that counts, but rather the fact that we have basic experience of ordering the events which lie on not too large segments of such curves. So Sklar thinks that causality is a misnomer. We have an intuitive understanding of what counts as events being close to one another on such a curve. Once again, it seems questionable whether the basic qualitative relation we want to identify is causality.

4 *What is ontic structure?*

From these examples it would seem that the problems of ontic structure are not extensional ones. It seems likely that causal connectivity and the basic qualitative relation between any point event and any other on or in its null cone are coextensive. If there were tachyons, or if photons had non-zero mass, the relations would not be coextensive, but there probably are no tachyons or massive photons. The questions raised in (b) and (c) of §2.3 concerned only the *consistency* of these hypotheses with relativity theory.

If structural questions are not extensional, then perhaps they are intensional, i.e. questions about which relations-in-intension must be the meanings of those polyadic predicates required by a given scientific theory. However, the attempt to reduce spacetime to a system of relations of causal connectibility discounts this suggestion. For this is an attempt to *identify* causal connectibility with the basic qualitative relation. While there are problems surrounding the identity conditions for relations-in-tension, we may take it that such relations are identical only if they are (logically) necessarily coextensive, and hence only ever identical necessarily. Causal connectibility and the basic qualitative relation clearly are not *necessarily* coextensive, however; the reductionist can only be claiming that these relations are *contingently* identical, so he cannot be talking about relations-in-intension.

Relations which are not relations-in-extension, and admit of contingent identity, should be treated as physical relations instantiated in our physical world, not as linguistic meanings. Their identity conditions may not be completely straightforward (Putnam 1969) but they are not the intensional entities which Quine has rejected for their lack of identity conditions. For our purposes we need only set a necessary condition for the identity of physical relations, viz. that physical relations are identical only if it is *physically* necessary that they be coextensive, where what is physically possible and necessary is determined by what is consistent with the physical theory at hand. The consistency of various non-causal hypotheses with SR or GR is sufficient to show that causal connectibility and the basic qualitative relation are not the same physical relation.

Another structural problem can be put, a bit roughly, like this: we want to order the ideology of the theory or at least its polyadic predicates as any formal system orders its ideology; so that some relations are seen as primitives, others defined, and so on. That ordering tells us which relations the formaliser thinks lie at which levels of basicness.

In dealing with this problem in spacetime we have a huge advantage of precision in that we are dealing with structure *in geometry*. Various levels of structure can be precisely defined and ordered. There are, for example, topological, projective, affine and metric structures in decreasing order of generality. One can define the levels by the group of transformations of space into itself which preserve

structure at these levels. For instance, topology studies those properties of space which are unchanged when it is arbitrarily (but continuously) distorted: these are the properties of separation, neighbourhood, boundary, enclosure, dimension. Roughly, projective geometry studies properties which are preserved when space is distorted in any way which keeps lines straight. Thus each geometric level of properties is precisely defined by the group of transformations which preserve the properties of that level. This means that we can order geometric properties, as to which are the more general and deep-seated, in a very satisfactory way. As a consequence, the constructions on the qualitative relation which aim at capturing topology, affinity or the metric either succeed exactly, or they fail.

Looking for ontic structure involves pondering the ways of some possible worlds. One considers models of GR which are vacuum solutions of the Einstein equation, or models in which the photon is massive or in which there are tachyons. We suggest that this does not befog our structural question because several of the difficulties which cramp the style of possible world metaphysics in general are missing here. First, the candidacy of a model rests on the necessary condition that it be a solution to some well-defined set of equations. Second, we need not be concerned to fix what constitutes the range of all worlds accessible in that they are solutions to the equations. It is frequently enough to look at a few salient or, perhaps, limiting cases to settle main issues. Lastly, there is no evident need to raise questions about the identity of objects in one model with objects of another. We can forget the problem of Trans World Heir Lines (to borrow David Kaplan's engaging phrase).

For all these reasons, then, it seems that the study of ontic structure, need not be less determinate than the study of objectual ontology. It is rather that some physical relations are less tractable than others and that hope in settling questions about physical relations must be tempered by one's sense of how fully the target theory has been articulated.

5 Studying structure

If one branch of ontology is structural rather than objectual, what is the best way of studying it? The discussion so far suggests that we

should formalise our theory and use the quite well-understood resources of formal systems. Choosing axioms, definitions and theorems helps to present just the sort of map we want of a theory's ideology. The suggestion has already been pursued in detail by Sklar (1977a). We add some observations which may help to clarify and support our claims.

(a) *Positivist probes.* Positivist and conventionalist critiques of theories are familiar. They are motivated by the fact that two mutually inconsistent theories may each entail the same body of (actual and possible) observation statements. Theories may be underdetermined by observation, in short. So, in the construction and axiomatising of theories, we face observationally unanswerable questions about which theory to choose. The positivist strategy rejects the questions either by pruning away the vocabulary in which they might be asked, or by so defining the vocabulary in observational language as to make the answers analytically true or false. These recipes tend to yield rather strong medicine. Less stringently, the conventionalist strategy rejects the question by robbing some part of the ideology of its fact- or structure-reporting role. The question gets an answer, though not a structurally significant one. These three styles of treating the problem of underdetermination are motivated by laudably modest wishes to say no more than one can confirm. However, the characteristic fault of such critiques is their strong tendency to evacuate the theory of its assertive power. Its content implodes under positivist pressure.

The somehow laudable distaste for excessive content may lead not just to the three kinds of ideological pruning just mentioned. It also promotes a tendency to substitute for one another extensionally identical expressions which are not intensionally identical. It does not require that, of course, but relativity theory provides clear examples where it has been done: notably, the almost universal reference to the null cone as the light cone. Use of the expression 'causal' seems to have been adopted, also, in the hope that it would be coextensive with the qualitative relation we are pursuing, so as to offer a reduction. In as much as positivism hardens its heart against intensional distinctions among coextensive expressions, its drive for economy ceases to be powered by a seemingly modesty. This is clear once one sees a certain modesty in realist aims as well. We will turn to that in a moment.

In whatever way they work, positivist probes seek to discover the

core of spacetime (or whatever other) theory. The core consists of the minimal ideology and the minimal axioms which use it, so long as they suffice to generate the needed theorems. Thus positivist probes seek what may be called the *restrictive core* of the theory, thus stressing its tendency to reduce content and prune ideology. This is a rather bald account of positivist aims, but we hope the familiarity of the ideas touched upon will allow us to be thus brief.

(b) Realistic probes. Motives which reflect the scientific realism of theorists in choosing formal systems may vary. Clearly, one motive must be to avoid the implosion of theoretical content which positivism is so likely to produce. Among other motives, we wish to stress two which are well illustrated in spacetime metaphysics.

The first is an argument addressed specifically to Reichenbach's (1958) treatment of conventionalism in geometry. Reichenbach argues that two theories underdetermined by the same set of observations have the same factual content. But Clark Glymour (1977) has argued against this. Though the set of observations relevant to some pair of inconsistent theories may be exactly the same, so that they are empirically equivalent in that sense, this body of evidence may still confirm one of these theories better than it confirms the other. This is to say that the theories are distinct and non-synonymous. It is also to say, of course, that our choice between them is not arbitrary. To show this is a matter of articulating in spacetime examples a certain schematic principle of confirmation. It says that a theory with fewer untested hypotheses is better tested by the same body of evidence than a theory with more untested ones. More determinately, let there be two theories P and Q , a set of definitional axioms A , such that $P \& A$ entail Q , but no set of definitional axioms B such that $Q \& B$ entail P . If every test of $P \& A$ from some body of evidence is just a test of some hypothesis in Q , then every body of evidence confirms Q better than it confirms P . We will not pursue this in more detail now.

Glymour's result, whether it is right or wrong, points to something important in motives for scientific realism. One great weakness of positivist critiques of physical theory surely lies in the assumption that we understand epistemic principles better than we do physical ones. But, surely, we do not. It is entirely reasonable to regard the failure of modern epistemology to provide an adequate rationale for the ways we choose our theories as demonstrating the incompleteness of epistemology. At least, that strikes us as quite clear. It is always open

for a realist to say that the current state of epistemology shows, not that science is in bad shape, but rather that some relation between simplicity and truth or some new theorem in confirmation theory such as Glymour's, is a good bet for a research programme in philosophy. That seems overwhelmingly the most plausible and, perhaps, the most fruitful attitude to take to the problem.

A second, quite distinct motive for probing one's theory with hopes for realism lies in a search not just for the restrictive but also for the *permissive core* of the theory. The restrictive core presents, in the fewest axioms, enough to generate all the theorems we think we want. The permissive core presents the axioms without which we could not reconstruct anything recognisable as *this* theory. Put like that, the cores may not sound obviously different and perhaps one might often reach the same axioms in each search. But they can, and in spacetime theory they do, diverge quite widely, motivated, as they are, so differently.

An example built out of material already at hand should make this clearer. It is very widely believed that a single principle about causality lies at the foundations of SR and GR. It is the Limit Principle: nothing outstrips light. We have been rehearsing the results of attempts to make this principle yield the whole metric congruence structure of spacetime in formal systems, attempts which make this the key element of their restrictive cores.

The nice things about the Limit Principle are that it is, very likely, true; that it is suitably qualitative so as to capture (by extension at least) the qualitative relation we have been looking for; and that it uses an ideology tied to agreeably familiar concepts like *particle* and *outstrips*. ('Outstrips' is preferred to 'goes faster than' since the latter is not transparently qualitative.) If we take this principle as a restrictive core (or a main part of it) we are likely to go on in our formalised spacetime theory to define the null cones at each point as the light cones there and, of course, to lay the principle down as an axiom of the formal theory.

We pay a price for this. It may bind theoretical imagination. Questions which might be fruitful speculations cannot arise. The hypothesis that there are tachyons, particles faster than light, is inconsistent with SR thus formalised and so is the hypothesis of the massive photon, since we have tied the conical structure of spacetime to the hypothesis that photons are massless. Yet both hypotheses, and more

notably the former, are compatible with relativity (else they would never have got off the ground). Clearly, the restrictive core has its own modesty, hospitable though it may be to realism. It is no less difficult in denying structure than positivism is in affirming it. In the present case, and surely quite generally, permitting a structure (here the null cones) to belong to an entity epistemically remote (here space-time) is precisely what allows us, within the theory, to speculate about the observable behaviour of epistemically proximate things. It opens the door to the design of new ranges of tests of the theory. What the permissive core permits is speculation *within the theory* as to tachyons and the like. The positivist may speculate about them too, but only by abandoning the theory which both disciplines and guides what he may conjecture and construct. Realism can be the springboard for flights of the empirical imagination.

When we ask for the permissive core, our question is quite differently motivated and our criteria for a good answer are not all the same as for the restrictive core. Though it is unlikely that either hypothesis just considered is true let us ask how SR would survive their truth. The answer is quite obvious once one glances at tachyon theory. The striking fact about this bit of physics is that the equations are all Lorentz invariant. That is they are the same for all inertial coordinate systems, since they remain unchanged, as all equations of physics do, under Lorentz transformations from one inertial frame to any other. So the permissive core of SR appears to be the Invariance Principle: that the laws of physics are invariants of the Lorentz transformations. What one has to say about the massive photon may involve some change in the laws of electromagnetism or optics. But what it most certainly does not demand is a change in the principle that whatever laws there are must be Lorentz invariant. That principle – the Invariance Principle – is the permissive core of SR. When the Invariance Principle is taken as the permissive core the constant ‘ c ’ in the Lorentz transformations cannot be interpreted as anything more than a finite inertial-frame-invariant scalar with the dimensions of speed. To interpret it as the velocity of light would preclude the massive photon hypothesis, and to interpret it as the limiting speed of causal propagation would preclude the tachyon hypothesis. Even to insist at this level that ‘ c ’ has its standard value of 2.9979×10^8 m/sec. would preclude further more precise measurement of the frame-invariant speed.

The distinction between the cores may become active in constructing formal theories (compare Sklar 1977a). We are indebted to this paper, but try here to take a different and in some ways, perhaps an incompatible perspective on the forces at work in constructing formal theories. In the present climate of speculation within physics, we are claiming that no one would seriously consider an equation which is not Lorentz invariant. To express it in the material mode, it may be active in our probing the ontic structure of the world. How? Well, SR is a very highly confirmed theory. It, or something very like it, must be true, locally at least. But *what* is highly confirmed, so virtually certain that speculation inconsistent with it is really somewhat idle? Not the Limit Principle, certainly, strongly recommended though it plainly is by the positivist styles of critique as at the heart of the theory. What our willingness to countenance as speculation with SR – and surely the entertaining of alternatives has to count as a crucial part of creative scientific thought – what this marks out as basic is the Invariance Principle. Not only is this principle quite a different statement from the Limit Principle in being much weaker. It is expressed in an ideology which is richer in what is needed to state the principle itself. It asks for a richer ideology in the further axioms which must complement the Invariance Principle if we are to generate some interesting theorems.

It follows that there are at least two sorts of pressures in structural ontology – toward restrictive cores and limited ideologies and toward permissive cores and speculative ideologies. There appears to be no obvious principle which instructs us how to balance these opposing pressures (Sklar, 1981). But, of course, one wants neither the implosion of content so liable in positivism nor the explosion into pointless structures which results from reflex realism. We point out that a richer ideology does not automatically lead to our including bizarre theorems and a baroque ontology, either objectual or structural. However it does lead to our including theorems which may not differ from their contraries in observational consequences. It also liberates theoretical speculation, and we count this no less important than parsimony.

No mention has been made of the idea that theory structures are governed by choice of conventions. We find it unhelpful to divide the language of our theories into factual and merely conventional, non-factual sentences. This not only gives wrong answers to some ques-

tions but prevents the asking of others. In this context we can ask why we choose this convention rather than that to settle, say, simultaneity in a frame, and we can answer in terms of simplicity, convenience and so on. But we can hardly ask, for want of any idea how to answer, why we have a convention governing a unique choice of simultaneity in a frame but no convention singling out some one frame to be uniquely preferred though, plainly, such a convention is possible. If conventionalism were true surely the answers to such questions would be important and clear principles ought to govern them. But we see no room in conventionalism for such principles. If one choice is seen as an *assumption* about the world, we might at least hope to find a principle which explains how we choose, since assumptions, unlike conventions, are not necessarily arbitrary.

Further, there is nothing whatever arbitrary in the choice of space, time and spacetime, continuous and differentiable, as observationally undetermined arenas for the occurrence of physical things and processes. They are, if you like, virtually inescapable *a priori* concepts, though corrigible ones, which we bring to the organisation of our experience. That they are corrigible is abundantly clear in the physics of this century. The description '*a priori*' wears an antiquated look, to be sure, but not more than the ideas, to which it is no distant kin, of non-factual convention and of the observationally underdetermined as the observationally arbitrary. Just as it was the structure of the laws of classical mechanics that determined our assumptions about spatial congruence and not some observably invariant object, so it is the structure of laws in SR which determines the (local) null cone structure of spacetime and not some observable relation (signalling etc.). Just as an immense diversity of observations confirmed the former by confirming the laws, so it is with the latter.

(c) The Invariance Principle poses a new problem. There is something not quite right about this way of identifying the permissive cores. Something about the structure of spacetime is at the centre. For reasons already glanced at, the Invariance Principle obstructs a clear view of what it identifies because it speaks of extrinsic matters. The Principle is about frames of reference or coordinate systems and about transformations among these. We need to clear those cobwebs away. We can do so by saying something like this: spacetime has a null cone structure at each point within which there are no privileged directions. But reference to null cones imports quantitative ideology.

We want a qualitative principle and expect to find it in the permissive core.

In the light of all that, we suggest the following course. Given any point in spacetime, some set of points count as elsewhere (locally, at least). Also, there is a set of points which are *after* the given point and a set of points which it is *after*, as Robb noted. We do not need a relation which is asymmetrical, in the style of *after*. So let us speak of the *elsewhen* as what is before or after the given point. Then we have a quite elegant vocabulary of symmetrical qualitative relations which allows us to state, qualitatively, the conical structure of spacetime in pre-metrical language. The statement which lies at the permissive core of SR, then, says this:

The elsewhere and elsewhen of any point overlap in a (hyper-conical) boundary.

Any spacetime point in this boundary, which is distinct from the origin point, is both elsewhere and elsewhen from it. There are no privileged directions in the elsewhere or the elsewhen. What we are claiming as the permissive core of SR is not merely something qualitative. It is also topological. The structure given in the statement is virtually the same as the structure of the ε -neighbourhoods in the fine (Zeeman) topology of Minkowski spacetime (Zeeman 1967). The ε -neighbourhood of a point consists of the manifold neighbourhood of the point, with the null cone through the point (though not the point itself) removed. That is to say, it is the union of the elsewhere and elsewhen neighbourhoods minus the boundary in which they overlap. Though these are not a *base* of neighbourhoods of x in the fine topology, they are undoubtedly important for it.

Thus the basic qualitative relation in spacetime is either elsewhere or elsewhen. That it is one rather than another of them has not been argued here. First, and most noticeably, the choice of either relation jars with the whole reductive trend in the positivist critique which aimed to make causality the basic qualitative relation, since each is, essentially, a geometrical relation. Causality is at least cashable in terms of matter, and it is moderately proximate to observation, but elsewhere and elsewhen are among those relations which have always given offence to reductive styles of thinking.

Which is more correct, or more searching, the positivist probe or the realist? That depends (among other things) on what you are will-

ing to grant about what constitutes structural ontology. If structure is not an extensional study, then the positivist theory is incorrect as to questions about physical relations in the theory of spacetime. If structure is extensional, then positivism surely has much to be said for it. We argued earlier that ontic structure is not extensional, but surely those arguments touch the issue rather lightly and perhaps they hinge on special cases. They would repay more critical study. As for which probe is more searching, this depends on what you want to find. Both restrictive and permissive cores are worth finding, and if they are not the same for any theory, then each style of attack on the basis of a theory will reveal something important about structure.

(d) Finally, a rather brief glance at two further issues followed by a general conclusion. What Sklar (1977b) says ‘might be right about the causal theory of time’ is that the continuity or discontinuity of causal paths in spacetime are best taken to be basic, not in being the paths of possible causal connections but in being paths which might constitute the life-history of some consciousness. We have two doubts about this.

First, as Sklar himself mentions, it is not clear why this should give us grounds for preferring the *elsewhen* relation to *elsewhere*. We have a good basic intuitive grasp of local spatial continuity; we are, in fact, just as surely spatially extended as we are enduring in time. Further, it has often been noted that our temporal concepts tend to confuse us in thinking that time flows, that only the present is real and so on. This tendency, is due to the complex structure of temporal perceptions which is not easily described nor easily understood and is still imperfectly grasped (see §11, pp. 268–71). It is, therefore, less than suitable as a justification for our choosing continuous causal curves as basic.

Second, and more importantly, time-like curves are structurally basic in a way that space-like curves are not. In a two-dimension Minkowski spacetime there is a simple *elsewhen* construction dual to each *elsewhere* (or ‘causal’) construction, so that the Robbian programme can be carried out on the basis of *elsewhere* relations just as well. Less obviously, in four-dimensional Minkowski spacetime, where the dual to the timelike curve is the spacelike hypersurface of three dimensions, the programme is also possible on the basis of the relation ‘. . . is elsewhere from —’. (This claim is somewhat conjectural. We have not worked it through.) But there is an asymmetry, for while

no two spacetime points are connectible by a time-like curve, in this spacetime, any two points are connectible by a space-like curve. In spacetime more generally, it is regarded as a pathology of the spacetime if this asymmetry is not preserved. The reason for the asymmetry is simple and basic. Time is one-dimensional and space three-dimensional. It is this structural, non-epistemological feature, which makes the time-like curve so simple and fruitful a means of probing spacetime structure. It is, in particular, a geometrical fact about spacetime and spatiotemporal relations, quite unsuitable for reductive purposes in being just the sort of thing reductionists want to reduce. So the basic qualitative relation that meets (or, perhaps, fails to meet) the formal requirements of Robb's programme is not a causal relation but a qualitative feature of spacetime itself.

If that argument is sufficient, then it illustrates how structural and objectual ontology may be connected. If the basic relation of relativity theory is elsewhere, or if the basic structure is the time-like curve, then a basic object in the ontology of SR and GR physics is spacetime itself. It must exist if the structure exists. Thus objectual ontology may depend, and we incline to the view that it standardly does depend, on structural ontology.¹

¹We are indebted in several places to discussion with Adrian Heathcote.