Lecture 15: Philosophical Issues Related to Special and General Relativity

(Note: Some issues connected with cosmological applications of general relativity, e.g. the Big Bang, are postponed until later in the course.)

Two general questions: (1) What do special relativity and general relativity have to say about our fundamental notions about space and time? (2) What impact do they have on general issues concerning the meaning of scientific theories, etc.?

(l)(a) The "substantivalist-relationist" debate.

Note that already in the late nineteenth century, the idea of a "field" propagating through "empty" space may be relevant. If anything, this notion favors substantivalism because it suggests there has to be something (ether?) there to "undulate". But if the ether goes, this argument disappears.

Suppose one were a pre-special relativity relationist. What then is "real"? A tempting hypothesis is that whereas the definition of the origin in space is conventional, the "now" of time is not: in that case one would say that events which take place "now", but at any point in space, "really exist", while those in the past and future do not. (Or perhaps: those in the past have real existence, those in the future do not.) But an alternative view is that the "origin" in time is as conventional as that in space, i.e. "now" is merely analogous to "here" (cf. Huw Price, "The View from No-when").

What does special relativity have to say about this? In the first place, "now" cannot be defined unambiguously for distant events. Moreover, the "now" (simultaneity) becomes in a curious sense intransitive. Of course, *any one observer* will regard "now" as a transitive relation: if A is simultaneous with B and B with C, then A is simultaneous with C. However (cf. Sklar p. 72), it is perfectly possible to have 3 events A, B, and C such that A is simultaneous with B for *one* observer, and B is simultaneous with C for a second observer, but for no observer is C simultaneous with A-e.g. it is in the "absolute future" of A. We then have the following paradoxical situation: A says "the state of mind of an observer at B, now, really exists" and the observer at B says "C, now, really exists" but C is in the absolute future of A, so he cannot (apparently) say it "really exists"!

A possible way out: We deny reality to the "elsewhere"! I.e. for an observer at space-time point (xt), all events outside his past and future light-cones are "unreal". Since the light-cone is a Lorentz invariant, this has the advantage that all (inertial) observers at event (xt) will agree on the reality or not of other events. That observers at different space-time points disagree is perhaps no more surprising than that in Newtonian mechanics, observers at different (absolute) times disagree. If one takes this view, then of course we must still decide whether to count as "real" (a) only the present ("here-and-now") (i.e. a single "event"); (b) only the present and the interior of the past light-cone; and (c) the present plus the interiors of the past and future light cones. (But if (c), there

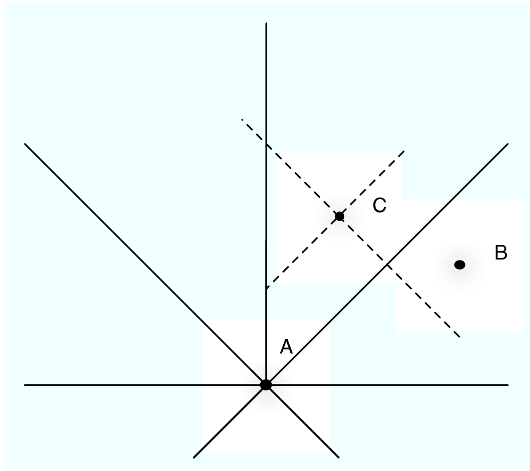


Figure 1: B is in "elsewhere" of A and C, hence we can find a frame in which it is simultaneous with either.

seems little motivation.)*

There is one other aspect of the "substantivalist-relationist" controversy on which *general* relativity sheds an unexpected and curious light. Recall that one of Leibniz's arguments for the relationist point of view was that if absolute space exists (and is infinite) then it is difficult to see why the Universe should be placed at one particular point in it rather than another. Now, one possible answer is of course that the Universe is in fact infinite. But in Newtonian theory this gives rise to severe difficulties (gravitational collapse, etc.).

What does general relativity say about this question? First, while the Universe must be time-dependent (if the cosmological constant is zero: see lecture 26), there is nothing to prevent it being infinite: it may or may not be, depending on the value of the critical "closure" parameter Ω (see last lecture). So if it is indeed "open" or "flat", Leibniz's objection does not arise. More interesting is the case where the closure parameter is such that the Universe is *closed*. In this case, space itself is finite in extent: it just does not make sense to ask what happens beyond it! (Compare the ants on the surface of the football: provided they have no concept of a third dimension, it would simply *make no sense* to them to ask "what happens beyond the surface with which we are familiar"?). Of course, in our (3 + 1 dimensional) case it *would* make sense to ask, if we believed that "in reality" spacetime is more than 4 dimensional as some recent theories in particle physics suggest.

*Note: There are more complications in general relativity because of the possibility of causally anomalous topologies.

(l)(b) The status of inertial frames.

In Newtonian physics, we had the problem of deciding what fixed the inertial frames: Mach's provisional answer was "the (local) mass distribution of the Universe". *Special* relativity does nothing in particular to help with this problem, it needs also simply to posit a class of frames as inertial. Does *general* relativity help here?

In general relativity, the inertial frames are those which are freely falling in the local gravitational field, and it is with respect to these that force-free bodies move in straight lines with uniform velocity. A more general (and reference-frame- independent) statement is that force-free bodies *follow the geodesics of the space* (in particular, light follows a "null geodesic"). What determines the geodesics? The (invariant) geometry of the spacetime, which according to general relativity is (mostly) determined by the distribution *and velocity* of matter. So in a qualitative sense Mach is right: it is the behavior of mass which determines the inertial frames (which choice then in turn determines the behavior of mass!).

However, quantitatively things are more complicated (and perhaps even in 2007 not completely understood!). Suppose, first, we start with a region of spacetime where we can assume that some particular choice of the class of inertial frames is valid (don't, for the moment, ask how we do this!). Let's now sit in one such inertial frame, and introduce into it, say at the origin, a spherical mass, which is at rest but, relative to our coordinate frame, *rotating*. Evidently this would be expected to change the choice of inertial frames, at least locally. The most obvious such effect is of course standard Newtonian-type gravity: however, we want to go beyond this, so we ask: Is a frame freely falling towards the origin with a radial acceleration of $GM/r^2(GM/r^2c^2 << 1)$ a true inertial frame? The detailed formalism of general relativity says no: in the region near the origin, the true inertial frames are not only freely falling but also, to some extent, *rotating* with the mass. However, until one approaches "black-hole" conditions the rotation is never total. Moreover, the effect falls off as r^{-3} (i.e. the apparent angular velocity is proportional to r^{-3}) and so vanishes sufficiently far from the spinning mass. However, let's emphasize that this conclusion only holds because we implicitly assume that the boundary condition for the solution of the Einstein equations "at infinity" is precisely the same as what it was before the mass was introduced-which perhaps begs the question! All we have really shown is that this assumption is at least self-consistent.

We are still, then, forced back to the question: How do we know how to choose the inertial frames "at infinity"? We could try making both of two related (but different) hypotheses: (1) We must choose them in such a way that the total matter of the Universe (or perhaps that part of it which is "sufficiently close") should appear to be at rest or moving uniformly (Mach's principle); (2) If we were to try to make a different choice, then the "dragging" effect already examined would be just such as to make our local inertial frame after all coincide with that chosen in (1) (principle of gravitational induction). Are either or both of these statements correct in standard general relativity?

The answer to (1) appears to be a qualified yes: in most cosmological models currently considered, in particular in the so-called "standard model" (FRW Universe, see lecture

26), it is possible to choose one's reference frame so that the local average mass density of the Universe is at rest, and it is then consistent to take this frame as an inertial frame. Question (2) is a bit trickier: there are sufficient problems in applying the theory to an *external* rotating shell that the answer is not totally clear-cut. However, it appears probable that the principle of gravitational induction will work, as in Newtonian theory, only for a special value of the dimensionless parameter $G\rho R^2$. Since it is precisely this parameter, in general relativity, which determines whether the Universe is open, flat or closed, a belief in the principle of gravitational induction might turn out to imply (e.g.) that the Universe is flat!

(What about an "empty" Universe? Mach's retort: it makes no sense even to discuss this!)

(1)(c) Determinism and predictability.

Laplace in the eighteenth century claimed, on the basis of Newtonian mechanics, that a complete knowledge of the positions and velocities of all particles of the Universe would completely determine its future behavior, and that therefore an "all-seeing being" could in principle predict the whole of the future. Assuming that there is some sense in which this claim is defensible in Newtonian mechanics, does relativity alter the situation?

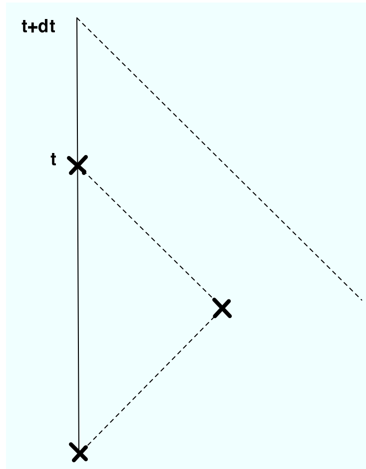
Actually the plausibility of the Laplacean assertion even within Newtonian mechanics rests rather heavily on the implicit assumption that the being is not only "all-seeing" but that he can "see" (i.e. obtain information on) everything *instantaneously*. That this is so is obviously true in an infinite Universe, where the claim has to be that if he can see "instantaneously" a distance R, then he can predict the future up to a time R/v_l, where v_l is the limiting velocity of propagation of physical effects.

In special relativity the whole idea of "seeing" instantaneously does not make sense: the maximum velocity of propagation of *any kind* of information (not just by a light wave!) is c. Thus all the demon can do is to send out light signals and get them reflected back to him with the information. But it is clear that the edge of his backward light cone is continually moving, so at time $t + dt$ there will *automatically* be events in it (which can therefore causally affect him) of which he must be ignorant at time t!

(Does a closed Universe help here? We will have to discuss a bit more cosmology, e.g. the so-called Big Crunch, to determine this.)

(2) The nature and status of scientific theories.

Most of the questions to be raised here are not specific to special or general relativity, and in fact could in principle have equally well been raised in connection with the transition from an Aristotelian-Ptolemaic to a Copernican world view. That they were not, probably has to do with the special relation of general relativity to non-Euclidean geometry: while even the medievals would probably not have argued that we had strong *a priori* reasons (other than semitheological ones) to believe that the Earth was at rest,



a long pre-1900 tradition held, following Kant, that we *do* have a priori reasons to believe that Euclidean geometry is the geometry of the real physical world.

Let us then ask the question: What does it *mean* to say (e.g.) that the Copernican picture of the solar system is "correct" and the Ptolemaic (or perhaps Tychonian) one "wrong"? Or that special relativity is right and the ether theory wrong? Or, if we do not wish to commit ourselves quite that far, what makes the Copernican theory "better" than the Ptolemaic one and special relativity "better" than the ether theory?

At first sight, at least, there are certain statements we can make about rather directly observable events in the physical world whose meaning does not seem problematical, at least in the operational sense that we have well-defined criteria for their truth or not: e.g. "the bus left the bus-stop at 9:01 a.m.", "the sun set directly over University Avenue", "Comet Shoemaker-Levy 9 crashed into Jupiter and broke up into 9 pieces", etc., etc. (Of course, the first two statements do contain certain implicit assumptions about the conventions for measurement of time, etc.) Let's suppose for the moment, for the sake of argument, that these "direct observation-statements" are indeed free of problems. It is clear that statements like "the earth moves around the sun", "the speed of light is the same in all inertial frames", etc., do not have quite the same status. Indeed, as we have seen (sketchily) all the experiments which are traditionally used to support special relativity *can* be explained by the nineteenth-century ether hypothesis, provided that we are prepared to make the appropriate assumptions about length contraction, time dilation, etc. Similarly, the crucial experiments which support (in textbooks!) general relativity, *can* be explained, if we modify our hypotheses concerning e.g. the effect of the gravitational potential on clocks and measuring rods appropriately, without modifying the previously assumed Euclidean geometry of space. In fact the early twentieth-century French mathematician Poincaré constructed a complete philosophy of science along these lines, claiming that the choice between general relativity and such a "neo-Newtonian" theory was no more than a matter of *convention*. (objections to Poincaré's conventionalism, and counter-arguments thereto, are discussed detail by Sklar, pp. 55-69).

A closely related question which is often raised by philosophers of science IS: If two supposedly "alternative" theories explain (and/or predict) exactly the same experimental consequences, are they really "the same" theory simply disguised in different language? In a few rather trivial cases this is certainly plausible: e.g. Sklar's example of revising the theory of electromagnetism by simply exchanging the definitions of positive and negative charge. However, when applied to e.g. the relation between special relativity and ether theories, the answer is much less obvious, and probably depends on what exactly is meant by "really the same theory". Physicists tend to be somewhat impatient with this kind of question, probably because they are conscious that whatever the *logical* equivalence or not in these cases, the *psychological* and *heuristic* consequences are often quite different: While no doubt the ether theorists would eventually have realized that for consistency in their theory they would have to introduce something corresponding to what we know as the "mass-energy relation" $E = mc^2$, they almost certainly would not have come across it as quickly or as simply as Einstein did. Incidentally, if one does take the view that theories with the same empirical consequences are "the same" one is almost inevitably forced to a radically non-realistic view of any concepts (e.g. atoms) which are not directly accessible to observation-something, again, which most practicing physicists find distasteful, because of their consciousness of the heuristic value, if nothing else, of a "realistic" approach. (Example: Mach on atoms).

The situation is further complicated by the fact that not uncommonly a radically new theory implies that what were previously taken to be "directly observable" events/facts actually are not: e.g. in special relativity, the statement that "the bus left the bus-stop at 9:01 a.m." is actually only true for those observers in the inertial frame of the bus-stop. Again, when a Ptolemaic astronomer said "the sun rose at 6:45 a.m." he meant, no doubt, exactly that: nowadays, we regard the statement as in some sense metaphorical!

Ifwe cannot decide on the question of what exactly it means to say a given theory is "correct", can we at least identify the features which make it "preferable" to its rivals (note: a question on the intersection of philosophy, history, and sociology of sciencel). Of course, in principle one should distinguish the question "what features *should* make theory A accepted in preference to theory B?" from "what features do, in practice, get theory A accepted?" (e.g. the dying off of adherents of theory B!), but let's focus on cases (e.g. Ptolemy/Copernicus, ether/special relativity) where most of us would believe that the acceptance was "rational".

Various criteria have been proposed:

(a) "Intrinsic plausibility"-but this may mean no more than coincidence with our common-sense prejudices! (Ptolemy is certainly preferable to Copernicus on this ground!)

(b) Conservatism-a new theory should preserve as much of the old one as possible. This is perhaps reasonable in cases where the old theory has justified itself over some wide domain. In practice, one often tries to keep what appear to be the broadest and most "generic" features of the old theory, e.g. the relation between symmetry and conservation laws. A good example of something most people are extremely reluctant to sacrifice is conservation of total energy.

(c) Simplicity/economy. E.g. one important difference between special relativity and the ether theory is that there is actually not one ether theory but infinitely many, corresponding to the different possible rest frames of the ether! Similarly, in Newtonian gravitational theory one has no means of knowing what gravitational field, if any, one is in, provided that everything around one is subject to it. But the idea of "simplicity" is in general somewhat subjective! (The actual calculations based on general relativity are typically *much* harder than those in Newtonian mechanics: is a picture of the planets which requires them to move in (ugly) ellipses "simpler" than one in terms of circular epicycles?)

It may be interesting at this point to introduce the concept of a "map" and contrast it with that of a "picture"; when we do this, some of the above questions may look a bit different. I quote from my book "Problems of Physics"

> At this point let us step aside for a moment and reflect on what it is we are trying to do when we construct a 'theory' in physics, and what we regard as the criterion for success. Of course, so long as we are working within a given conceptual framework-roughly speaking, what T.S. Kuhn calls a 'paradigm'-such as Langrangian quantum field theory, there is no great problem: our job is simply to find the ingredients which make the theory self-consistent, tractable and in accordance with experiment, and while it goes without saying that this task is highly nontrivial, at least we have fairly well-defined and generally agreed criteria for what constitutes success. But let's ask the question at a broader level: What kind of relationship do we look for between the physical world and our theories of it? Perhaps if we can answer this question, we will be a little closer to understanding just what force some of our Martian colleague's doubts and reservations have.
>
> Needless to say, the question of the relationship between the external world and the scientist's description of it is one which has exercised generations of philosophers of science-it is, perhaps, *the* central question of the discipline-and I have nothing particularly profound or original to contribute to it; I would like merely to present an analogy which, while possibly philosophically naive, may nevertheless I believe, capture the implicit understanding of many working scientists as to what it is that they are trying to do. What I would like to suggest is that we should think of scientific theories as analogous to *maps* of the physical world (notice that I say 'maps', plural, not 'a map'.) What is a map? In the first place, it is not a picture of anything. Rather it is a representation in more or less symbolic form of certain interesting features of a particular geographical region. Think of the various kinds of maps we know: there are Geological Survey maps, real estate agents' maps, road maps put out by the motoring organizations, military maps, maps of the Paris Metro system or of Chicago's O'Hare airport . . . and so on. If we really think about it, they have rather few specific features in common, apart from being two-dimensional representations on pieces of paper.[†] What they do have in

[†]Though I suspect that future generations may use three-dimensional holographic maps for some

common is that each conveys, in the form of a visual Gestalt, that information about the region in question which is relevant for the purpose for which it is designed. Which features of the external physical world are represented in the map depends on its intended function; it need not necessarily preserve metric or even topological relationships.[‡] Mountaineers do not usually complain because their maps fail to show the names of roads, nor travellers on the London Underground because, if you believe the wall maps, all lines run in one of eight discrete directions: each map is perfectly adequate for the purpose for which it is designed. There is not and could not be a 'perfect map'!

One thing about the analogy between maps and scientific theories needs to be specially emphasized: A map is a representation which is constructed *by human beings* with specific human purposes in mind. In so far as the analogy is a good one, therefore, many of the features of our present-day science which puzzle our Martian friend may perhaps be traced not to any 'intrinsic' or 'objective' properties of the world, but rather to the constraints imposed on what for us human beings is an adequate description of it by the capacities and limitations of our own minds. Needless to say, this idea is not exactly a novel one-it goes back at least to Immanuel Kant-but it seems to me that it may be desirable to reemphasize it in the intellectual context of late twentieth-century physics. Perhaps it may, inter alia, shed some light on why we think of 'simplicity' as such a virtue in a scientific theory, and why the 'unification' of apparently disparate natural phenomena has always been such a major goal in physics: it is not so much that there is any a priori reason why Nature should be simple or unified, but rather that it saves us an enormous amount of mental filing space! A 'good' and simple scientific theory, like a good and simple map, gives us the ability to get where we want to, physically, intellectually or technologically, quickly, cleanly and without unnecessary diversions-and that, surely, is a large part of what we mean by 'understanding' a phenomenon in physics. One obvious implication of this point of view is that it may often be more valuable to have a simple theory which gets things approximately right than a complicated one which gets them exactly right (just as no motorist in his right mind would use a series of 6" U.S. Geological Survey maps to get from New York to Los Angeles!) Another intriguing question which it suggests is whether we shall in the end ever really be satisfied with a theory-such as QCD in its current state of development-in which, while the basic postulates can be simply and clearly stated, it is largely impossible to obtain results which can be compared with experiment without recourse to large-scale numerical computation. My guess

purposes.

[‡]If you are a traveller departing from Charles de Gaulle airport in Paris, it will not matter to you if your airport map fails to represent the complicated topology of the connection tubes, provided it gets you to your satellite. (If on the other hand you are an airport security officer, the topology is vital).

is that in the end we shall come to regard such theories as second-best, in that they do not give us any real 'understanding' in the sense in which we have been used to it in physics; but this may no doubt depend on the extent to which we eventually get used to the idea of using computers as a genuinely symbiotic extension of our own mental powers, something which seems to me difficult to predict.