THE PHILOSOPHY OF PHYSICS

The Philosophy of Physics studies the logical and conceptual foundations of the physical sciences, as well as the metaphysics, ontology, and epistemology of the world that those sciences describe.

It is not so much a distinct and established academic discipline as it is a sort of boundary, a sort of frontier, across which theoretical physics and modern western philosophy have been interrogating and informing and unsettling one another, for something on the order of four hundred years now, about the character of matter, the nature of space and time, the question of determinism, meaning of chance, the possibility of knowledge, and much else besides.

Many of the profoundest intellectual commitments of the west have been vividly and famously in play or at risk or under pressure from the very inception - in the work of Galileo - of the modern scientific investigation of the physical universe. And by the time of Newton, a lively conversation between physics and a new and distinctly modern western philosophical tradition was well underway - and has flourished up to the present day.

That conversation is the topic of this article.

THE PHILOSOPHY OF SPACE AND TIME

1) The Newtonian Conception of the World

According to the Newtonian conception of the world, the physical furniture of the universe consists entirely of infinitesimal material points - of what are referred to in the physical literature as *classical particles*.

Extended objects - objects (that is) that take up finite volumes of space, objects like tables and chairs and baseballs and specks of dust and people and buildings and planets and galaxies - are treated in Newtonian Mechanics as *assemblages* of such particles; and their behaviors are determined (in principle, at any rate) by the behaviors of the particles of which they are made up.

Classical particles come in a potentially infinite variety of different *masses* and *electric charges.* and so on, but the only physical attribute of a classical particle that can change with time is its *position*.

A list of what particles exist (then) and of what their masses are electric charges and their other intrinsic properties may happen to be, and of what their positions are at all times, is a list of absolutely everything there is to say - on this conception - about the physical history of the world. One could in principle *read off* of a list like that (that is) where all the tables and chairs and baseballs and galaxies and specks of dust are, and how they got to be where they are, and who did what to whom, and who *wrote* what to whom, and who *said* what to whom, and so on.

And Newtonian Mechanics is deterministic. Given a list of the positions of all of the particles in the world at any particular time, and of how those positions are changing, at that time, as time flows foreword, and of what the intrinsic properties of those particles are, the universe's entire history, in every detail, from that time foreword, is fixed and determined (if this theory is true) with absolute certainty.

This determinism (of which much more later) has seemed to many thinkers, from the outset, as at odds with some of the deepest and most indispensable ideas we have of what sorts of beings we are, and what sorts of lives we lead - ideas that go under names like 'freedom' and 'autonomy' and 'spontaneity' and 'creativity' and 'the openness of the future'..

2) The Logical Structure of Newtonian Mechanics

The rate at which some particular particle's position is changing, at some particular time, as time flows foreword, is called it's velocity at that time. And the rate at which such a particle's velocity is changing, at some particular time, as time flows foreword, is called it's acceleration at that time.

And the Newtonian picture stipulates that forces arise exclusively between pairs of particles, and that the forces which any two particles are exerting on one another at any particular instant depend only on what sorts of particles they are and on their relative positions. And so (on this picture) a specification of the positions of all of the particles in the world at some particular time, and of what sorts of particles they are, amounts to a specification of what the forces on each of those particles are at that time as well.

And the entirety of what Newtonian Mechanics has to say about the way particles move is that a certain very simple mathematical relation - F=ma - invariably holds between the total force on any particle at any particular instant, and its acceleration at that instant, and its mass.

The theory works like this:

Call the "initial" time, the time we shall want to calculate foreword from, t=0.

And suppose that what we're given at the outset are the positions of all the particles in the world (or in some isolated sub-system of the world) at t=0 (call those x_0^i), and their velocities at t=0 (call those v_0^i), and their masses (m^i), and their electric charges (c^i), and all of their other intrinsic properties.

And let's say that what we should like to calculate is the positions of all these particles at t=T.

The most illuminating way of doing that - for our present purposes - will be by means of a succession of progressively better and better approximations.

The first goes like this: Calculate the positions of all the particles at t=T by supposing that their velocities are constant - and equal to their above-mentioned values (v_0^i) at t=0 - throughout the interval between t=0 and t=T.

This calculation will place particle i at $x_0^i + v_0^i T$ at t=T; but it hardly needs saying that this calculation is not a particularly accurate one, because (unless it happens that no forces are at work on any of the particles here) the velocities of these particles will in fact not remain constant throughout that interval.

Here's a somewhat better one:

Divide the time-interval in question into two, one extending from t=0 to t=T/2 and the other extending from t=T/2 to t=T. Then calculate the positions of all the particles at T/2 by supposing that their velocities are constant - and equal to their values at t=0 - throughout the interval between t=0 and t=T/2 (this will place particle i at $x_0^i + v_0^i(T/2)$ at T/2).

Then calculate the forces acting on each of the particles at t=0 (what those forces are, remember, will follow from the positions of those particles at t=0 together with their masses and their charges and their other internal properties - all of which we are given at the outset).

Then calculate each particle's velocity at T/2 by plugging those forces into the above-mentioned law of motion (plugging them, that is, into F=ma), and assuming that the particles' accelerations are constant throughout the interval from t=0 to t=T/2 - and are equal to their values at t=0 (this will put the velocity of particle i at $v_0^i + a_0^i(T/2)$, where a_0^i is equal to the force on particle i at t=0 divided by particle i's mass).

Then, finally, calculate position of particle i at t=T (which is what we're after here) by supposing that this particle maintains this new velocity throughout the interval between t=T/2 and t=T.

This calculation isn't going to be perfect either, but (since the intervals during which the velocities of the particles are erroneously presumed to be constant are shorter here than in the previous calculation) it amounts to a clear improvement.

And of course this improvement can itself be improved upon by dividing the interval further, into four intervals. That calculation will proceed as follows:

To begin with, the approximate positions of all of the particles at t=T/4 can be calculated (just as we did above) from the positions and velocities at t=0 alone. Moreover, the forces on all of the particles at t=0 can now be read off (as we did above) from their intrinsic properties and their positions at t=0, and thus (with the aid of F=ma) the approximate velocities of all the particles at t=T/4 can be deduced as well. And so what we now have in hand is a list of approximate positions and approximate velocities and particle-types at T/4, and of course those approximate positions and particle-types can now be used to determine the approximate forces on all the particles at that time as well, and that will in turn allow us to determine the positions and velocities and forces at T/2, and so on....

And then we can go on to eight intervals, and then to sixteen. And as the number of intervals approaches infinity, the calculation of the particles' positions at t=T patently approaches perfection. And it happens that there is a trick (and the name of that trick is the calculus) whereby - given a simple enough specification of the dependence of the forces to which the particles are subjected on their relative positions - that perfect calculation can actually and straightforwardly be carried out.

And of course T can be chosen to have any positive value we like. And so the positions of all of the particles in the system in question at any time between t=0 and t=infinity (and with that the velocities of all of those particles between those times, and their energies, and their angular momenta, and everything else about them) can in principle be calculated, exactly and with certainty, from the positions and velocities and intrinsic properties of all those particles at t=0.

3) What is Space?

Newtonian mechanics, then, dictates the motions of particles - it dictates (that is) how the positions of particles in space change with time. And of course the very possibility of there being a theory which dictates how the positions of particles in space change with time would seem to require, to begin with, that there be clear and concrete and determinate matters of fact, at every particular time, about what the position of every particular particle in space happens to be. The possibility of there being a theory which dictates how the positions of particles in space change with time would seem to require (to put it another way) that space itself be a separate, free-standing, autonomous, fully existent thing - the sort of thing that this or that particular particle might occupy this or that particular part of, the sort of thing relative to which a particle might move. And there happens to be a long and distinguished philosophical tradition of doubting that anything like that can be the case, that anything like that can even make sense.

The doubt is connected with the fact that we cannot imagine how, that we cannot even imagine what it would amount to, to carry out a measurement of the absolute position of any particular particle, or any particular assemblage of particles, in space. The doubt (to put it slightly differently) is connected with the fact that we cannot imagine how we might even begin to determine by observational means whether or not (say) every single particle in the world had somehow been transported, last night as we lay sleeping, by exactly five hundred thousand miles, to the left. And there is a school of philosophical thought called 'Empiricism', according to which it is either unjustified or mistaken or even incoherent to suppose that there are any matters of fact about the world to which we can even in principle have no empirical access. The only facts there are, on the empiricist way of thinking, are the ones that can in principle be established by means of some imaginable sort of measurement. And so the only spatial facts there are, on this way of thinking, are relational ones. The only facts there are (that is) are the ones about the distances between particles - and talk about absolute positions, talk about which particular part of space, about which particular spatial point, some particular particle at some particular time happens to occupy, is just so much nonsense.

On this view (which is referred to in the context of discussions of the philosophy of space as 'relationalism') two hypothetical worlds which differ only in terms of the 'absolute positions in space' of the particles that make them up, and not at all in terms of the distances between any two of the particles that make them up, do not in fact differ at all. On this view, there is no such separate, autonomous, free-standing, fully existent thing as 'space' - what we call 'space' on the relationalist view, is nothing but a mathematical representation of the infinity of different spatial relations which particles may potentially have to one another. And it is a redundant and misleading representation at that, it is a way of representing things that makes it look as if there are vastly more physical possibilities than there actually are, since (as we have just seen) there will invariably be an infinity of different and equally legitimate ways of representing any complete set of genuine relationalist spatial facts, any complete set (that is) of spatial relations between particles, in terms of the absolute spatial locations of each of the particles involved.

On the opposing view (which is referred to as 'absolutism') what there are facts about does not necessarily coincide with what we might even in principle be able to establish by measurement. On this view, there are very compelling reasons (of which more in a minute) for believing that the two hypothetical worlds described in the above paragraph do differ, notwithstanding the fact that we may well never have any chance of determining which one of them (if either) is ours.

And on the face of it, the Newtonian system of the world is manifestly committed to an absolutist idea of space. Newtonian mechanics (after all) makes claims about how the positions of particles, and not merely their relative positions, change with time. Newtonian Mechanics (to put it another way) makes perfectly explicit claims about (say) what the laws of the motion of a particle entirely alone in the universe would be, but (since a single particle alone in the universe lies at no particular spatial distance from anything) relationalism is committed to the proposition that it is altogether nonsensical even in inquire about what the laws of the motion of a single particle, all alone in the universe, might be.

And so there is a relationalist critique of Newtonian Mechanics, which goes back to Leibnitz, as a redundant and misleading and philosophically defective scientific theory And there is a defense against that critique, which goes back to Newton himself, to the effect that there are features of our own empirical experience of nature - features which were first made explicit, features which were first expressed in mathematical formulae, in Newtonian Mechanics

- which can have no intelligible explanation except by reference to a substantive, free-standing, absolute space. And there has been a lively debate between defenders of these two positions ever since, which has gone through a number of different stages and taken a number of different forms and had a number of different sorts of ramifications, about which we shall have a good deal more to say in what follows.

4) Kant on Incongruent Counterparts

Let_s start with something very simple, something which is in a conceptual sense prior even to the seminal and original and paradigmatic debate about the character of motion in Newtonian mechanics (although, as a strictly historical matter, it was first thought through some years after that debate was already well underway) but which brings out the salient features of that debate in a particularly stark and accessible form.

There is a famous argument of Kant to the effect that the relationalist position on space cannot be right; an argument (that is) to the effect that relationalism recognizes fewer spatial facts about the world than there manifestly are; which goes like this: Consider a pair of possible universes, in one of which the only material object is a right-handed glove and in the other of which the only material object is an (otherwise identical) left-handed one. The two universes (so says Kant) manifestly differ, but they do not differ in terms of any of the complete list of spatial facts recognized by the relationalist. And so (the argument goes) relationalism is false.

And the standard relationalist response at this point is to stand firm and bite the bullet - to simply deny that there *is* any difference whatever (notwithstanding immediate intuitions to the contrary) between the two universes described above.

Let_s spell that out in just a bit more detail.

Note (to begin with) that among the set of all mathematically possible material shapes (which is to say: among the set of all mathematically possible relationalist arrangements of particles, among the set of all mathematically possible sets of interparticle distances) there are those which can, and those which can not, be made to perfectly coincide with their mirror images by means of ordinary three-dimensional translations and rotations. Call the second sort handed. Gloves are handed, then, and shoes are, and pants and hats (for example) aren_t.

While right-handedness and left-handedness are (of course) not legitimate relationalist predicates, handedness simpliciter surely is. Whether or not a certain shape is handed (again) depends exclusively on the distances between it_s constituent particles; and whether the handednesses of any two relationally identical handed shapes are the same or the opposites of one another (whether, that is, any two relationally identical handed objects - any two gloves, say, or any two shoes - can be translated and rotated in such a way as to perfectly coincide with one another in space) can be read off from nothing over and above the distances between corresponding elements of their two sets of constituent particles. For any two relationally identical handed objects (to put it a bit more formally) there will necessarily exist some two-valued mathematical function - which is a function exclusively of the distances between corresponding elements of their two sets of constituent particles - and which takes on the value (say) 1 if and only if the handedness of the two objects are the same and -1 if and only if their handednesses are opposite.

And the relationalist strategy is to exploit presicely that. The relationalist idea (that is) is that relative handedness itself, relative handedness in its essence, is nothing over and above the peculiar exclusively spatial relation picked out by precisely this relaltive-handedness function, and that our impression to the contrary, our impression that (say) two oppositely handed gloves are somehow intrinsically different from one another, can be traced to the fact that the particular sort of relation in question here - notwithstanding that it is a perfectly and exclusively spatial relation - is a relation which (as it happens) no combination of three-dimensional rotations and translations can ever alter. And it will follow from all this (as promised above) that there just can_t be any matter of fact as to whether two relationally identical gloves in different possible universes have the same handedness or not - precisely because there can be no matters of fact about the spatial distance between any particular particle in one glove and any particular particle in the other.

And this was more or less how things stood all the way up until the middle of the middle of the twentieth century, when new fundamental physical laws were discovered which (it seems to have been thought) simply cannot be written down in legitimate relationalist language, and so decide the question in favor of absolutism.

The laws in question concern the decays of certain elementary particles, they concern (that is) processes in which certain individual such particles fall apart into collections of others, of so-called products. And the spatial configurations in which those decay-products, appear, are sometimes handed, and (moreover) it turns out to be a law that certain sorts of particles are more likely to fall apart into (say) the right-handed version of the configuration in question than the left-handed one, and this last phrase is (of course) just not sayable in the vocabulary of the relationalist.

But the relationalist can now - of course - bite the bullet yet again, and say that the law should rightly read just that the configuration of the decay products of the sort of particle in question here is invariably the same, single, handed one, and that the configurations of the decay products of any large group of such particles is likely to fall into two oppositely handed classes (which, according to the relationalist, is to say nothing other than that those decay products are going to be at certain particular sets of spatial distances form one another) and that those two classes are likely to be unequal in size.

And at this point the whole business is manifestly beginning to harden into a stalemate: The internal consistency and the empirical adequacy of the relationalist position is unassailable, but it comes at a conceptual price, it comes along with a certain new offense to intuition, which is that the laws of the decays of the sorts of particles we_ve been talking about here now look curiously non-local, both in space and (even less familiarly) in time. On this construal of the world (that is) what those laws apparently require of each new decay-event is that it probabilistically have the same handedness as, that it probabilistically line itself up with, the majority of the decays of the sort of particle in question which happen to have taken place elsewhere and before. And the entirety of what any relationalist reading of those laws can have to say about (say) the first such decay which happens to have occurred in the history of the world, or about any particular such decay considered in isolation, is (of course) that it deterministically assumes the handed configuration in question.

And the reason for going through all this is that it contains, in microcosm, much of the structure of the larger and more complicated and more seminal dispute - the one we started off with a few pages back - about the character of motion in classical mechanics, the dispute (more particularly) about whether or not there can be a satisfactory relationalist account of the dynamical effects of rotation.

5) The Question of Motion

There is a famous thought-experiment of Newton - which involves a rotating bucket of water - and which is designed to show (just as Kant's story of the gloves was) that relationalism must be false. What Newton hoped to establish (more particularly) is that there is a sense in which relationalism ineluctably defeats itself - that there can be no relationalist account of the dynamics of even those properties of the world which relationalsim itself seeks to describe.

The experiment - in a somewhat cleaner and more streamlined version than Newton's original one - goes like this: Consider a universe which consists of nothing whatever over and above (say) two balls affixed to opposite ends of a spring. And stipulate that the length of that spring, in its relaxed configuration, in its unstretched and uncompressed configuration, is L. And imagine that there is some particular temporal instant in the history of that universe at which it happens to be the case that 1) the length of the spring is greater than L, and 2) there are no two material components of this contraption whose distance from one another is changing with time - there are no two material components of this contraption (that is) whose relative velocity is anything other than zero. And suppose that we should like to know something about the dynamical evolution of this universe in the future of that instant; suppose that we should like to know (say) whether or not this spring is fated, in the future, to oscillate.

On the conventional way of understanding Newtonian Mechanics, whether or not this spring is fated to oscillate will depend on whether or not, and to what extent, at the instant in question, the contraption is rotating with respect to absolute space. In the event that the spring is stationary, for example, the spring will oscillate - but if the spring is rotating, at just the right speed, it will remain stretched. And the trouble with all that (of course) is that on a relationalist picture of the world there can simply not be any fact of the matter about rotations (or the absence of them) relative to absolute space. And so (the argument goes) relationalism is false - the ralationalist vocabulary of physical properties isn't rich enough even to accommodate an explanation of why it is that some such springs will eventually begin to oscillate and others won't.

The standard relationalist responses to all this start out with the observation that the universe of our actual experience has a great deal more in it than the hypothetical universes we have just now been considering. And the idea is that that other stuff, or some part of it ('the fixed stars', say, or 'the bulk mass of the universe'), might be fit to serve as a concrete material stand-in for absolute space, a concrete material system of reference on which a fully relationalist analysis of rotation could be based.

Ernst Mach (for example) is famous for pointing out that - speaking now in absolutist language - the total angular momentum of the universe of our actual experience appears to be zero. And so - in so far (at least) as the universe of our actual experience is concerned - rotation with respect to absolute space amounts to precisely the same thing as rotation with respect to the universe's own center of gravity, or to its 'bulk mass', or to its 'fixed stars' (which were thought, in Mach's time, to make up the overwhelming majority of that mass). And rotation with respect those latter structures is (of course) a perfectly acceptable relationalist physical quantity.

Mach's proposal, then, is that 'rotation' is to be understood - by definition - as rotation with respect to the bulk mass of the universe; that motion in general (for that matter) is to be understood as motion with respect to the bulk mass of the universe; that the correct theory of the motions of material bodies - the relationalist theory of the motions of material bodies - is that they take place in accord with $F = ma$, where 'a' is to be understood as nothing over and above acceleration with respect to the bulk mass of the universe.

On Mach's theory, the total angular momentum of the universe is necessarily zero - it is zero by definition. And so (for example) on Mach's theory, a two-mass-and-spring contraption of the sort we were talking about above, if it should happen to be all there is of the universe, and if there should happen to be any temporal instant in the history of that world at which the string is stretched, will necessarily oscillate.

Note too that the cost of relationalism in this case, just as in the case of the incongruent counterparts, and in the case of the much more elegant and penetrating formulations of essentially Machian theories of classical mechanics in the 20th century due to Barbor and Boretti, and (in so far as we can tell at present) in every case whatsoever, is non-locality. The standard Newtonian laws of motion (remember) govern the motions of particles across the face of an absolute space which is always and everywhere right where those particles themselves are - but what these Machian laws govern are the rates of change of distances between particles, particles which may in principle be arbitrarily far apart.

There is at least one other way of realizing relationalist aspirations in the context of a classical mechanics of the motions of particles. The idea here is not to look for a concrete material stand-in for absolute space, but (rather) to look for a way of systematically discarding those commitments of the original Newtonian mechanics of the motions of particles with respect of absolute space that don't directly bear on the time-evolutions of inter-particle distances, and keeping all and only those that do.

And once problem is set up in these terms, it's solution turns out to be perfectly straightforward. The complete relationalist theory of the motions of material particles - on this way of doing things - will read "All and only those temporal histories of inter-particle distances are physically possible which can be embedded in a full absolutist Newtonian space - which can be consistently imagined as taking place (that is) within a full absolutist Newtonian space - in such a way as to satisfy F = ma"

This theory manifestly satisfies all of the standard relationalist desiderata - it is exclusively concerned with the time-evolutions of inter-particle distances, it makes no assertions whatever about the motions of a single particle alone in the universe, or about the motion of the universe's bulk mass, it is invariant under all transformations that leave the time-evolutions of inter-particle distances invariant, and so on - just as the Machian theories do. But unlike the Machian theories we discussed above - this one as reproduces the entirety of the consequences of absolutist Newtonian mechanics for the time-evolutions of inter-particle distances. On this theory (for example) the total angular momentum of the universe need not be zero; and a two-mass-and-spring contraption of the sort we were talking about above, if it should happen to be all there is of the universe, and if there should happen to be any temporal instant in the history of that world at which the string is stretched, might oscillate and might not. This theory is no less non-local than the Machian theories are, but on this theory the laws of the behaviors of dynamically isolated sub-systems of the universe make no reference whatever to what's going on in the rest of the world.

The dispute about the nature of space - as the foregoing pages show - has from the very beginning been structured by and around an intuition to the effect that all that we can ever directly observe of the mechanical properties of particles are the distances between them. And this intuition turns out to be worth the trouble of interrogating somewhat further.

How directly do we observe such distances? It goes without saying that all sorts of old-fashioned skeptical doubts about the veridicality of such observations can certainly, in principle, be entertained. The business of deciding how seriously such doubts ought to be taken, in the context of discussions of the metaphysics of space and time, will depend (among other things) on the extent to which they can be incorporated into a bona-fide, predictive, confirmable scientific theory of the time-evolution of the world. If they can, if (for example) it should turn out to be possible to write down a version of classical mechanics in which all talk of inter-particle geometrical distances is somehow systematically reduced to talk of the mathematical structure of inter-particle dynamical interactions, then the way would clearly be open for an important new step forwards in the empiricist re-interpretation of space. As of this writing - however - the detailed study of such matters has only just gotten underway.

6) Time

Thus far, we have left the question of time - the question of what sorts of temporal facts there are, the question of what sorts of temporal claims make sense - altogether to one side. But it's clear that precisely the same sorts of empiricist considerations as we have thus far been bringing to bear on questions of space will have consequences for time as well. Those will be looked into in the present section.

The first thing to note, the starting-point of the discussion, is that we can no more detect our absolute location in time than we can in space. And so there can manifestly be no facts - from an empiricist perspective - about what time it presently is. Mach reasoned (moreover) that we can have no direct observational access to the sizes of time-intervals, or even to ratios of the sizes of time-intervals, or to anything about the way the world unfolds in time over and above the facts about which events occur after, and which occur before, and which occur simultaneously with which others.

Now, Newtonian mechanics identifies physical systems with certain particular sorts of dynamical structures as 'clocks', or (rather) as 'good' clocks, or (rather) as 'accurate' clocks. But what this sort of identification amounts to, from an empiricist perspective, has nothing whatever to do with correlations between the configurations of clock-faces and 'what time it is', or between changes in the configurations of clock-faces and 'how much time has passed' - since, for an empiricist, there are not going to be any facts of the matter as to what time it is, or how much time a certain process takes up, or anything of the sort. What it is, from this point of view, to be a hand on an accurate clock, is to be the sort of physical system whose position (relative to the other hands, or to the rest of the clock, or to the laboratory, or whatever) is correlated by means of a particularly simple and powerful and general law to the physical properties of the rest of the world. To the extent that time-intervals are the sorts of things that can even intelligibly be talked about, on this view, they are not measured - they are, rather, defined - by changes in the faces of clocks. What the mechanics of classical particles is about, all there is for the mechanics of classical particles to be about, on this way of thinking, are the laws of the dependancies of inter-particle distances on clock-readings.

Let's put it a bit more generally: What the classical mechanics of particles is about, and all there is for the classical mechanics of particles to be about - on an empiricist way of thinking - is the laws which determine which sequences of inter-particle distances are physically possible, and which are not.

And there is a straightforward technique for cooking up a theory like that - just as there was in the case of the purely spatial relationalism that we considered in the previous section - by systematically discarding those commitments of the original Newtonian mechanics of the motions of particles with respect to absolute space and absolute time that don't directly bear on those sequences, and keeping only those that do. What that will yield, in this case, is "All and only those sequences of inter-particle distances are physically possible which can be embedded in a full absolutist Newtonian space-time - which can be consistently imagined as taking place (that is) within a full absolutist Newtonian space-time - in such a way as to satisfy F = ma".

An explicitly Machian theory of physically possible sequences of inter-particle positions has recently been devised by Julian Barbour and Boretti. Barbour and Boretti start by defining a distance-function in the abstract space of relational particle-configurations - and then they write down a law of motion to the effect that only those trajectories through configuration-space (that is: only those sequences of inter-particle distances) are allowed which satisfy a certain simple and very elegant mathematical condition. And this law turns out to reproduce the relational consequences of Newtonian mechanics - but only those for universes with zero total angular momentum and (in this case) zero total energy as well.

All of what was said in the preceding section about the comparison between the Machian and the embedding theories - in connection questions of locality, and questions of the applicability of the laws to isolated sub-systems of the universe, and questions of the empirical equivalence of these theories to absolutist Newtonian mechanics - applies here, more or less unchanged, as well.

7) The Special Theory of Relativity

Consider Newtonian Mechanics in its original, absolutist, version.

And think of two observers, one of which is at rest with respect to absolute space and the other of which is moving along a straight line with a constant velocity. Observers such as these, whose accelerations with respect to the absolute background space are zero, are referred to in the physical literature as inertial. Now, imagine that each of these observers sets up a comprehensive frame of reference for the purpose of describing how it is that the world is laid out, and how it unfolds, with respect to herself - a frame of reference of which she herself is the spatial origin, and relative to which any spatio-temporally localized event can be assigned a unique triplet of spatial co-ordinates and a time. Call one of these observers (and her associated frame of reference) K and call the other (and her associated frame of reference) K'. Call the spatio-temporal co-ordinate axes of K x, y, z, and t, and call the spatio-temporal co-ordinate axes of K' x', y', z', and t'. And suppose that K is in motion relative to K', in the positive x direction, with velocity v; and suppose that K and K' coincide at the time t = t' = 0.

Then it follows from what seem at first like altogether elementary and ineluctable geometrical considerations that the relationship between the spatio-temporal address that K assigns to any event and the spatio-temporal address that K' assigns to it is given by the so-called Galilean Transformations: $x = x' - vt$, $y = y'$, $z = z'$, $t = t'$. And there are two trivial consequences of these transformations that will particularly concern us in what follows: one is that if a body is traveling in the x direction with velocity j as judged from the perspective of K, then it is traveling in the x' direction with velocity j-v as judged from the perspective of K'; and the other is that the acceleration of any body, under any circumstances, as judged from the perspective of K, is always identical with its acceleration as judged from the perspective of K', or (for that matter) from the perspective of any observer whatever who in not herself accelerating with respect to K.

Now, if K measures time-evolutions of the positions and velocities and accelerations of particles relative to herself, what she will find (according to Newtonian mechanics) is that those quantities all evolve in time in accord with the Newtonian equation of motion, F = ma. Note, however, that this equation makes no explicit mention whatever of either the position or the velocity of any particular particle. And recall our conclusion above that another observer, any other observer, so long as they are moving with a constant velocity, so long (that is) as they are not accelerating, with respect to this first one, will measure the same acceleration of any particular particle at any particular time as the first observer will. And of course all such observers will agree on the mass-value for each particular particle, and on the magnitude and the direction of the force to which every particular particle, at any particular time, is being subjected. And so, if the motions of all of the particles in the world are such that they obey F = ma with respect to absolute space, if (that is) the motions of all of the particles in the world are such that they obey F = ma with respect to this first observer, then those motions will necessarily obey F = ma with respect to all observers moving with a constant velocity with respect to that first one as well! Thus, F = ma represents the law of the motions of particles not only with respect to absolute space, but with respect to any observer whatever who is not accelerating with respect to absolute space - and the Newtonian law of motion is therefore referred to in the physical literature as invariant under transformations between different inertial frames of reference.

It follows that motion with a constant velocity with respect to absolute Newtonian space is completely undetectable, in a Newtonian universe, by means of any sort of physical experiment. And it is of course for precisely this reason that the debate between absolutists and relationalists about the nature of motion is entirely taken up with cases of acceleration, and (more particularly) with cases of rotation. And by the middle of the 19^{th} century, the very general thesis that all of the fundamental laws of physics must share this invariance under transformations between different inertial frames of reference - this invariance (as it is referred to in the physical literature) under 'boosts' - had become a profound article of faith in theoretical physics.

By the end of the 19^{th} century, however, there was for the first time a serious candidate for a modern fundamental physical theory - the magisterial and spectacularly successful theory of the behaviors of electric and magnetic fields due to Maxwell - according to which there is, as a matter of fundamental physical law, a certain definite velocity at which light propagates through empty space. And a law like that is manifestly not invariant under Galilean transformations.

Moreover - and more puzzling still - a variety of experimental attempts at measuring the velocity of light from the perspectives of different inertial frames of reference all invariably yielded the same results! All of those experiments - which (again) were conducted in frames of reference which were in motion relative to one another - invariably found that the velocity of light as measured relative to the particular frame of reference in which they were being conducted had precisely the value predicted by Maxwell's theory.

The leading attempt at coming to grips with all of this - prior to the revolutionary work of Einstein in 1905 - was due primarily to Lorentz. Lorentz's picture involved (to begin with) explicit violations of the invariance of the fundamental laws of physics under boost-transformations, and he proposed to account for the puzzling outcomes of the experiments described in the previous paragraph by means of a theory of the systematic and lawlike malfunctioning of clocks and measuring-rods which are in motion with respect to absolute space.

On Lorentz's approach, there are real and physically significant facts about the velocities of bodies with respect to absolute space - but it also turns out (as a consequence of the systematic malfunctionings of measuring-devices referred to above) that those velocities can as a matter of fundamental principle not be measured. But to Einstein - whose thinking at that time was very much under the influence of the Empiricism of Mach - this sort of thing was unacceptable. Einstein's approach was to see what might follow - in the light of Maxwell's laws - from the resolute denial that there are any physical facts about the velocities of bodies with respect to absolute space, to see what might follow (that is) from the resolute insistence that the Maxwellian laws were true and invariant under transformations between frames of reference in motion with respect to one another with constant velocities, to see what might follow (that is) from the resolute insistence that the speed of light is indeed somehow the same with respect to all so-called 'inertial' frames of reference.

And the only way of satisfying the requirements of Einstein's program is manifestly to reject the 'altogether elementary and ineluctable geometrical considerations' that lead to the Galilean transformations - which is nothing less than to abandon every previously entertained idea about the structure of space and time. And precisely that is what Einstein proceeded to do.

Einstein was able to show by means of a number of very straightforward thought experiments that it must follow from a law, which is invariant under transformations between relatively moving frames, to the effect that light invariably propagates through empty space with a certain particular velocity, that the facts about which events are simultaneous with which others must vary between such frames, and that (more generally) that facts about the time-intervals and the spatial separations between given events must vary as well. On Einstein's program - on The Special Theory of Relativity, that is - judgements about the simultaneity of two events, or about the time-interval or the spatial distance between them - are altogether on a par with judgements about (say) which objects are to the right or to the left of which others: they are matters about which there are simply not any absolute facts, they depend altogether on one's perspective, on one's physical point of view.

Moreover, the above results can be parlayed, without much further trouble, into a new and complete and somewhat more complicated set of special-relativistic equations for transforming between frames of reference which are in motion relative to one another with uniform velocities, a special-relativistic replacement (that is) for the Galilean transformations mentioned above, the so-called Lorentz transformations; and the physical content of the special theory of relativity essentially comes down to the demand that the fundamental laws be invariant under the Lorentz - rather than the Galilean - transformations.

The well-established historical practice of referring to all this as a theory of 'relativity' turns out - on reflection - to be more appropriate in some respects than it is in others.

The theory does entail that a great deal of what had been taken throughout the previous history of physics to have an absolute or invariant or categorical sort of significance - the facts (say) about which events in the history of the world are simultaneous with which others, or about the time-intervals or the spatial distances between events - turns out to be relative, turns out (that is) to depend, in a way not previously dreamed of, on one's point of view. The relativity of simultaneity, in particular, came as a shock to philosophical traditions according to which (say) only the present is real, or time passes through a continuous succession of universal nows, or the past (but not the future) is metaphysically settled. The relativity of simultaneity (after all) is just the thesis that there is no such absolute and categorical and perspective-independent thing as 'the present', the relativity of simultaneity is just the thesis that there fail to be any absolute and categorical and perspective-independent facts about what is happening precisely now (as opposed to sometime before now, or sometime after it) on the surface of Jupiter, or at a certain particular point in China, or even across the room.

On the other hand, the popular notion that the upshot of Einstein's great achievement is that, somehow, 'everything is relative', is probably not even intelligible, and is certainly not true. The Special Theory of Relativity was explicitly designed (after all) in such a way as to guarantee that it is an absolute and categorical and perspective-independent law of nature that the velocity of light in empty space is (approximately) 186,000 miles per second - and the theory entails (more generally) that there is a certain algebraic combination of the spatial and the temporal distances between any pair of events - the so-called spatio-temporal interval between the two events - on which all inertial observers necessarily do agree. This is why the Special Theory of Relativity is often described as the discovery that what had previously been referred to as 'space' and 'time' are both parts of a single, novel, geometrical structure called 'space-time'.

Moreover, the special theory of relativity is much less accommodating to relationalist aspirations - and in that sense (at least) it is much less a theory of 'relativity' - than the Newtonian picture of the world. All of the standard relationalist strategies turn out to be straightforward non-starters in the context of Special Relativity: the business of replacing talk about the time-evolutions of positions in absolute space with talk about the time-evolutions of inter-particle distances goes nowhere, since in the relativistic context, no such purely spatial distances - or at any rate none with the requisite sort of independence of one's point of view - even exist; and the project of formulating a fundamental physical theory which is invariant under transformations between relatively accelerating frames of reference fares no better, since in the relativistic context, global frames of reference for accelerating observers can not even be coherently defined.

8) The General Theory of Relativity

One of the towering achievements of the mathematics of the 19^{th} century was the discovery of the conceptual possibility of space itself being curved.

Consider a society of perfectly two-dimensional ants (say) whose lives and measuring-instruments and empirical experiences are confined entirely to a certain perfectly two-dimensional surface. And suppose that this surface is not flat. Suppose (that is) that this surface, as viewed from the perspective of a larger three-dimensional space which contains it, as viewed from the perspective of a larger three-dimensional space in which we can imagine it being embedded, has a bump in it, as depicted in figure x .

The ants will manifestly have no trouble in confirming, entirely by means of measurements carried out with their two-dimensional rulers within their two-dimensional surface, that all of the lines labeled R have the same length, that (in other words) all of the points on the line L are equidistant - in so far as any idea of distance that can make sense to these ants is concerned - from the point P, that (in other words) L is a circle. Moreover, they can easily carry out a measurement of L's circumference with their rulers as well. But the ratio between the circumference of this circle and its radius, as the ants measure it, will clearly be smaller than 2(pi).

And so - as Gauss was the first to understand - ants like these will be confronted in their investigations with a direct empirical violation of the laws of Euclidian geometry. Ants like these (to put it slightly differently) will be capable of detecting - by means of experiments entirely confined to within the surface that constitutes their universe of their experience - that that surface is curved, that (as it were) the space they live in is curved!

Nobody at the time entertained the slightest suspicion that considerations like these might actually apply to the three-dimensional space of our experience. But the mathematical possibility of spaces of three and higher dimensions which are curved, and whose curvature could (moreover) in principle be discovered by observers confined entirely to within them, was articulated in magnificent and profoundly illuminating detail in the latter half of the 19th century by Reimann and Lobachevsky and others. And (in particular) a very powerful and intuitive generalization to these non-Euclidian geometries of the Euclidian notion of a ‘straight line’ - a line which is (as it were) precisely as straight as the space it traverses will accommodate - was developed. These ‘generalized straight lines’ are referred to in the mathematical literature as geodesics.

At this point Einstein enters the picture.

Einstein was intrigued by the fact that the ‘mass’ which makes an appearance in F=ma - the mass, that is, which measures the resistance of material bodies to being accelerated by an impressed force, the so-called inertial mass - is invariably exactly the same as the ‘mass’ which determines the extent to which any material body exerts an attractive gravitational force on any other one. The two seem to have nothing whatever - on a conceptual level - to do with one another. That fact that they are invariably identical to one another amounts - in the Newtonian context - to an astonishing and altogether inexplicable coincidence.

It’s this equivalence that entails that any given gravitational field will accelerate any two material bodies, whatever their weights or their constitutions, to precisely the same degree. It’s this equivalence (that is) that entails that any two material bodies, whatever their weights or their constitutions, will share precisely the same set of physically possible trajectories - just as they do in empty field-free space - in the presence of a gravitational field.

And this last way of putting it - this observation that the effects of gravitational fields on the motions of material bodies are completely independent of the particular physical properties of the material bodies on which those fields happen to be acting - positively cries out for something along the lines of a geometrical understanding of gravitation. The idea would be that there is - in the end - only a single, simple, breathtakingly beautiful law of the motions of material bodies both in free space and in the presence of gravitational fields - which is that the trajectories of such bodies are invariably straight lines, or (rather) that the trajectories of such bodies are invariably generalized straight lines, or (more precisely) that the trajectories of such bodies are invariably geodesics; and that gravitation is to be understood not as a force, but (rather) as a departure from the Euclidian laws of geometry.

Such an understanding is out of the question (however) so long as the geometry in question is a geometry of three-dimensional physical space, since (for example) all of the paths in figure y represent physically possible trajectories along which a projectile might travel from point A to point B, but at most one of them - no matter what the geometry of the three-dimensional space in which A and B may happen to lie, can possibly represent a geodesic. But note that in the four-dimensional space-time of the Special Theory of Relativity, it will become a geometrically significant consideration that no two of those paths through space can be traversed in the same amount of time. Looked at in space-time, the trajectories in figure y - if they are all taken to start out from the spatial point A at the same time - take the form depicted in figure z.

And it turns out that there is indeed a beautiful and straightforward non-Euclidian four-dimensional geometry of which all of <u>those</u> are geodesics.

Einstein began by writing down a far more powerful and more general version of the equivalence of inertial and gravitational mass - the so-called principle of the local equivalence of inertial and gravitational <u>fields</u> - according to which the laws which govern the time-evolutions of <u>all physical phenomena whatsoever</u> relative to a frame of reference which is freely falling in a gravitational field - in the immediate vicinity of the origin of that frame - are precisely the same as the laws which govern the time-evolutions of those phenomena relative to an <u>inertial</u> frame. And he was able - after years of prodigious effort - to parlay that principle, together with the Special Theory of Relativity, together with the results of the Newtonian theory of the gravitational 'force', into a fully relativistic, fully geometrical, theory of gravitation - the so-called 'General Theory of Relativity'.

What Einstein produced, in the end, was a set of differential equations - the so-called Einstein Field Equations - relating the geometry of space-time to the distribution of mass and energy <u>within</u> it; and the general theory of relativity consists of a law to the effect that the four-dimensional geometry of space-time and the four-dimensional distribution of mass and energy <u>within</u> that space-time must together amount to a <u>solution</u> to those equations.

On this theory, space-time no longer functions as a fixed backdrop against which the physical history of the world plays itself out, but <u>as an active player in its own right</u>, <u>as dynamical entity on a par with all the others</u>. On this theory (more particularly) the geometrical structure of space-time - the structure (that is) that grounds the distinction between constant rectilinear motion and acceleration - is at least in part determined by <u>the universal configuration of mass</u>. And this sounds, on the face of it, very Machian, very relationalist, very empiricist in spirit; and indeed Einstein's original conception of what it was that this theory might <u>amount</u> to - which was a <u>general theory of the relativity of motion</u> - was very much in that spirit as well.

But that project is now widely agreed to fail. General relativity turns out to be committed (for example) to claims about which motions are and are not possible for a single particle entirely alone in the universe, and it turns out to allow for the existence of universes whose total angular momentum is non-zero, and (moreover) it inherits all of the structural hostility to relationalism that we earlier descried in the <u>special</u> theory.

Solutions to Einstein's field equations exist in which the universe as a whole, although it is finite, has no end or boundary or 'outside' - solutions in which (that is) a generalized straight line which is extended far enough in any direction is space will eventually return its starting-point. Moreover, these solutions generally entail that the gross structural features of the universe as a whole <u>change with time</u>, in accord with definite <u>cosmological laws of evolution</u>.

The theory has also been shown - over the past half-century or so - to allow for the existence of a breathtaking array of different space-time geometries which are philosophically intriguing or unsettling or in one way or another in need of elucidation. One of many such examples is the fact that there are solutions to the field equations which can in principle accommodate <u>time travel</u> in to <u>one's own past</u> - and this in particular has been the focus of intense scientific and philosophical attention in recent years, as it seems to threaten outright logical contradiction in cases where (say) one travels back into the past and murders one's own grandmother prior to the birth of one's parents.

THE DIRECTION OF TIME, AND THE FOUNDATIONS OF STATISTICAL MECHANICS

1) The Problem of the Direction of Time

There is a perennial tension between fundamental microscopic physical theory and everyday macroscopic human experience about the question of precisely how the past is different from the future.

We shall consider that tension here in its original, straightforward, Newtonian, version - but it persists in very much the same form, in the contexts of very different fundamental theories of physics, up to the present day.

Newtonian Mechanics has a number of what are referred to in the literature as fundamental symmetries; and what that means is that Newtonian Mechanics entails - that F=ma entails - that there are certain sorts of facts about the world which - as a matter of absolutely general principle - don't make any dynamical difference. The reader will recall (for example) that much of the first part of this article was taken up with the fact that absolute positions play no dynamical role in Newtonian Mechanics - although relative positions surely do - and the section on the special theory of relativity contained a discussion of the fact that velocities play no such role either.

And neither - as it turns out - does the direction of time.

Think (say) of watching a film of a baseball which is thrown directly upwards, and which is subject to the influence of the gravitational force of the earth; and then imagine watching the same film in reverse. The film run forwards will depict the baseball moving more and more slowly upwards; and the film run in reverse will depict the baseball moving more and more quickly downward. What both films will depict, though, is a baseball which (whatever it's velocity) is accelerating, constantly, at the rate of 32 feet per second per second, in the direction of the ground.

And this is an absolutely general phenomenon: The apparent velocity of any particular material particle at any particular frame of any film of any classical physical process run forwards will be equal and opposite to the apparent velocity of that particle at that frame of that film run in reverse; but the apparent acceleration of any particular particle any particular frame of the film run forwards will be identical, both in magnitude and in direction, to the apparent acceleration of that particle at that frame of the film run in reverse.

The Newtonian law of motion (which is, remember, the entirety of what the Newtonian picture of the world has to say about the motions of particles) is that a certain mathematical relation holds, at every instant, between mass and force and acceleration. And of course the mass of any particular particle at any particular frame of the sort of movie we've been talking about depends on nothing other than what particular particle it is; and the force on any particular particle at any particular frame of the sort of movie we've been talking about depends on nothing other than what particular set of particles happen to exist, and what their spatial distances from one another at that frame happen to be; and what we've just seen is that the acceleration of any particular particle, at any particular frame of such a movie, will be entirely independent of the direction in which the film is run. And so if a certain film, run forwards, depicts a process which is in accord with Newtonian mechanics, then, necessarily, the same film run in reverse will depict a process which is in accord with Newtonian mechanics as well.

This is a point worth repeating, worth writing down (that is) in a number of slightly different ways:

1) It is a consequence of Newtonian mechanics that nothing in the laws of nature can be of any help whatever in deciding which way a film is being run.

2) It is a consequence of Newtonian mechanics that whatever can happen can just as easily, just as naturally, happen backwards.

3) The Newtonian-Mechanical instructions for calculating future physical situations of the world from its present physical situation turn out to be identical to the Newtonian-Mechanical instructions for calculating past physical situations of the world from its present physical situation. The instructions for calculating (say) the positions of all of the particles in the world ten minutes from now are that the present positions of all those particles, and the rates at which those positions are changing as time flows forwards, be plugged into a certain algorithm; and the instructions for calculating the positions of those particles ten minutes ago are that their present positions, and the rates at which those positions are changing as time flows backwards, be plugged into precisely the same algorithm.

4) If we are told the positions of all the particles in the world at present, and if we are told the rates at which those positions are changing as time flows towards some other moment M, and if we are told the size of the time-interval that separates M from the present, then we can in principle calculate the positions of all of the particles in the world at M, with certainty, without our ever having been told (and also without our ever learning, as the calculation proceeds) whether M happens to lie after the present or before it.

5) If the laws of Newtonian Mechanics are all the fundamental natural laws there are, then there can be no lawlike asymmetries whatever between past and future.

And it hardly needs to be said that all this is very much at odds with our everyday macroscopic experience of the world. To begin with, the world is full of ordinary physical processes that do not, or don't regularly, or don't naturally, or don't familiarly, happen backwards - think (for example) of the melting of ice, or the cooling of soup, or the breaking of glass. Moreover, our experience of being in the world is characterized by a very profound asymmetry of epistemic access: our capacities to know what happened yesterday - for example - and our methods of finding out what happened yesterday, are as a general matter very different from our capacities to know and our methods of finding out what will happen tomorrow. And finally, there is an asymmetry of intervention: it seems to us (that is) that we can bring it about that certain things occur - or that they don't - in the future, but we feel absolutely incapable of doing anything at all about the past.

This is the tension, this is the problem, that was mentioned at the outset of this section. It has been lurking deep in our picture of the world, in more or less the same form, from the Newtonian beginnings of modern physics, through the great 19th century investigations of analytical mechanics and electromagnetism, through the revolutionary 20th century developments of the special and general theories of relativity and of quantum mechanics, through the relativistic quantum theories of fields, through contemporary speculations about strings and branes.

2) Thermodynamics

A wonderfully concise and powerful and general account of the time-asymmetry of ordinary physical processes - the first of the three asymmetries (that is) which were mentioned in the paragraph before last - was gradually pieced together, over the second half of the 19th century, in the course of the development of the science of Thermodynamics.

The sorts of physical systems in which manifest past-future asymmetries arise are invariably macroscopic ones, and (more particularly) they are invariably systems consisting of enormous numbers of particles. Systems like that apparently have distinctive properties - and so it happened that a number of investigators, around the middle of the century before last, undertook to develop an autonomous science of such systems.

As it happens, these investigators were primarily concerned with making improvements in the design of steam-engines - and so the system of paradigmatic interest for them, and the one which is still routinely appealed to in elementary discussions of Thermodynamics, is a box of gas.

Consider (to begin with) what terms are appropriate for the description of something like a box of gas. Consider what it is to give an account of the physical situation of a system like that. The fullest possible such account is (needless to say) a specification of the positions and velocities and internal properties of all of the particles that make the gas and its box up. From that, and from the Newtonian laws of motion, the positions and velocities of all of those particles at all other times can in principle be calculated. And from the full history of those positions and velocities everything about the history of the gas and its box can in principle be read off. But the calculations involved here are impossibly cumbersome. And there is patently another, simpler, more powerful, more useful, more familiar, altogether different way of talking about such systems, which is to talk about them in a language of the macroscopic: to talk about things like the size and shape and mass and motion of the box as a whole and the temperature and the pressure and the volume of the gas. And there is patently a possible science of these temperatures and pressures and volumes - a science (that is) of macroconditions. We know it to be a lawlike fact, after all, that if we raise the temperature of a box of gas high enough the box will blow up. And we know it to be a lawlike fact that if we squeeze a box of gas from all sides the box will get harder to squeeze as it gets smaller. Never mind (for the moment) that this must all in principle be deducible from Newtonian Mechanics. It must be possible (or at any rate it seems it must be possible, or at any rate it seemed so to these investigators) to systematize all this on its own, it must be possible to discover an autonomous set of so-called thermodynamic laws of such boxes of gas which directly relate volume and temperature and pressure to one another, and which make no reference whatever to the positions and velocities of the particles of which (as it happens) the box and the gas consist. And it turns out that there are laws like that - that there is a science like that. And the essential principles of that science are as follows:

There is, to begin with, a thing called "heat". Things get warmer by absorbing heat and they get cooler by relinquishing it. Heat is something that can be transferred from one body to another. When a cool body is placed next to a warm one (for example) the cool one warms up and the warm one cools down, and this is in virtue of the "flow" of heat from the warmer body to the cooler one. And the above-mentioned investigators were able to establish - by means of a combination of straightforward experimentation and brilliant theoretical argument - that heat must be a form of energy.

There are two ways in which gasses are known to be able to exchange energy with their surroundings: they can exchange energy as heat (which is what happens when bodies at different temperatures are brought into thermal contact with one another) and they can exchange energy in mechanical form, as "work" (which is what happens when, say, the gas lifts a weight by pushing out on a piston). Since total energy is conserved, it must be the case that in the course of anything that might happen to a gas:

$$DU = DQ + DW, \qquad (1)$$

where DU is the increase, in the course of the occurrence in question, of the total energy of the gas, and DQ is the energy the gas absorbs in the course of that occurrence in the form of heat, and DW is the energy the gas absorbs in the course of that occurrence in the form of mechanical work. Equation number (1) - once again - expresses the law of the conservation of total energy, written down in the macro-language of this autonomous science of temperatures and pressures and volumes; and it is referred to as <u>the first law of thermodynamics</u>.

The <u>time-asymmetries</u> of ordinary physical processes are the province of the <u>second</u> law of thermodynamics, which concerns the evolution of a quantity called the <u>entropy</u>. The discoverers of thermodynamics were able to identify a straightforward function of the familiar thermodynamic variables - variables like temperature and volume and pressure and the like - which increases in every single one of the almost unimaginably diverse set of ordinary physical processes that never occur in reverse. This new thermodynamic variable - the entropy - increases when heat spontaneously passes from hot soup to cool air, and when smoke spontaneously spreads out in a room, and when a chair sliding along a floor slows down due to friction, and when paper yellows with age, and when glass breaks, and when a battery runs down, and whenever any of the ordinary sorts of physical precesses that never happen backwards is underway. Everything we know about which ordinary physical processes can and which cannot occur in reverse is concisely summed up by the proposition that <u>the total entropy of an isolated system can never decrease</u> - which is the second law of thermodynamics.

From these two laws together, a comprehensive theory of the behaviors of the thermodynamic properties of macroscopic physical systems - their temperatures and pressures and volumes and energies and chemical compositions and phase-transitions and so on - can be derived. And once these laws are in place, the question of <u>explaining</u> them or <u>understanding</u> them in terms of the Newtonian mechanics of the individual molecules of which macroscopic systems <u>consist</u> very naturally suggests itself. And it was in the course of attempts by Gibbs and Maxwell and Poincare and especially Boltzmann to imagine how such an explanation might look, at the close of the 19th century, that the problem of the direction of time first came to the attention of physicists.

3) The Foundations of Statistical Mechanics

The fact that a certain set of dynamical laws is symmetric under a certain transformation certainly does <u>not</u> entail that all of the <u>individual trajectories</u> which those laws <u>allow</u> have that sort of symmetry <u>as well</u>. A single particle moving in some particular direction in space - for example - is perfectly compatible with the Newtonian laws of motion, notwithstanding that there is <u>no</u> particular direction in space which those laws in and of themselves pick out as in any way <u>special</u>. And the canonical way of coming to grips with the tension we have just been talking about - which has its origins in the pioneering work of the physicist Ludwig Boltzmann - is to exploit precisely <u>that</u>, to see the asymmetry of our macroscopic experience under time-reversal (that is) as entering into the world not by way of the dynamical laws (which, as we have just seen, seem to have no such asymmetry <u>in</u> them), but rather by way of the <u>particular trajectory the</u>

world happens to be on, by way (that is) of the world's initial conditions.

Boltzmann's achievement - very briefly - was to make it plausible that the initial conditions in question here can be put into a strikingly simple and elegant form: that the Newtonian dynamical laws, together with the hypothesis that the initial macroscopic condition of the universe had a very low entropy-value, together with a very natural-looking probability-distribution over the set of exact microscopic conditions which are compatible with that initial macrostate, can generate precisely the sorts of ordinary physical asymmetries under time-reversal that characterize our everyday macroscopic experience. And we have lately seen the beginnings of promising attempts to extend this program into an account of the origins of the other two time-asymmetries - the asymmetries of epistemic access and intervention.

This approach - while it is universally admired as one of the great triumphs of theoretical physics - is also the source of a great deal of perennial uneasiness.

To begin with, the initial condition Boltzmann has in mind involves some statistical elements, some probabilistic elements - and there has since been a century of intense and unresolved philosophical debate about what it can possibly mean, about weather it can mean anything at all, to speak of the 'probability' that the initial exact microscopic condition of the universe as a whole (of which, after all, there is only one) was this or that.

Moreover, there something ineluctably awkward about the whole strategy of explaining the utterly familiar and fundamental and ubiquitous time-asymmetries of our everyday experience in terms of the world's initial conditions: The time-asymmetries of our everyday experience - the fact, say, that if two bodies at different temperatures are put into thermal contact with one another, heat will flow from the hotter body to the cooler one - are (on the one hand) the very models and paradigms of physical law. And (on the other) we have long been used to thinking of initial conditions in physics as falling under the category of the accidental, of the contingent, of what might have been otherwise, of how things merely happen to be.

All this has again and again prompted the investigation of a number of alternative approaches. Ilya Prigogine and his collaborators (for example) have been working for decades, and with prodigious mathematical sophistication, on an altogether different attempt at curing this tension. Prigogine's idea is that the very business of representing the history of the world as a trajectory, the very business (that is) of thinking of the world as having any particular precisely defined set of initial conditions, is somehow oversimplified or misleading or false. The world (so the story goes) does not have an initial condition, but (rather, somehow) a little continuous group of them; and this (so it is said) somehow renders the above-mentioned puzzling statistical features of the canonical picture an entirely natural and transparent and expected sort of thing, and this (if everything works out, and taken together with the chaotic structure of the Newtonian dynamics of certain paradigmatic sorts of physical systems) is supposed to point the way to the ultimate resolution of the tension about the direction of time. The trouble (or rather, one of the troubles) is that it has never been made particularly clear what it might mean, or what it might amount to, to think of the world as having a multiplicity of initial conditions; and (moreover) that all of the arguments that have been offered on behalf of this idea (which invariably seem to come down to the fact that "Whenever we perform an experiment....we are dealing with situations in which the initial conditions are given with a finite precision...") seem to mix up epistemic considerations with metaphysical ones.

There are also approaches that are intimately tied up with various attempts at solving the quantum-mechanical measurement problem - these will be briefly discussed in chapter 4.

QUANTUM MECHANICS

1) The Principle of Superposition

Let's begin by talking about some experiments with electrons.

The experiments all involve measurements of two perpendicular components of the intrinsic angular momenta, of two perpendicular components of what are usually referred to as the the "spins", of electrons. Let's call them the x-spin and the y-spin.

It happens to be an empirical fact that the x-spins of electrons can assume only one of two possible values. Let's call them +1 and -1. The same goes for y-spins.

The measurement of x-spins and y-spins is a routine matter with currently available technologies. The usual sorts of x-spin and y-spin measuring-devices (which are represented in figure 1, and which will henceforth be referred to here as "x-boxes" and "y-boxes") work by altering the direction of motion of the measured electron based on the value of its measured spin-component, so that the value of that spin-component can be determined later on by a simple measurement of the electron's position.

Another empirical fact about electrons is that there are as a rule no correlations whatever between their x-spin values and their y-spin values: Of any large collection of, for example, x-spin = +1 electrons, all of which are fed into the left aperture of a y-box, precisely half (statistically speaking) will emerge through the y-spin = +1 aperture, and half will emerge through the y-spin = -1 aperature; and the same goes for x-spin = -1 electrons fed into the left aperture of a y-box, and the same goes for y-spin = +1 and y-spin =-1 electrons fed into x-boxes.

And another empirical fact about electrons, and an extremely important one for our purposes here, one that it will be worth discussing in some detail, is that a measurement of the x-spin of an electron can disrupt the value of its y-spin, and that a measurement of the y-spin of an electron can disrupt the value of its x-spin, in what appears to be a completely uncontrollable way.

If, for example, a measurement of y-spin is carried out on any large collection of electrons in between two measurements of their x-spins (as in figure 2), what invariably happens is that the y-spin measurement changes the x-spin values of half (statistically speaking, again) of the electrons that pass through it, and leaves the x-spin values of the other half unchanged.

No accurate measurement of y-spin has ever been designed which has anything other than precisely that effect on x-spin values, and no physical property of the individual electrons in such collections has ever been identified which determines which of them get their x-spins changed in the course of having their y-spins measured and which do not.

The received view of these matters in the scientific and philosophical literature has been that there can in principle be no such thing as a y-spin measurement which has anything other than precisely that effect on x-spin values, and that which electrons get their x-spins changed by measurements of their y-spins and which don't is a matter of pure, fundamental, ineliminable, chance.

Moreover, if there can in principle be no such thing as a measurement of x-spin which fails to uncontrollably disrupt the value of y-spin, and if there can in principle be no such thing as a measurement of y-spin which fails to uncontrollably disrupt the value of x-spin, then, patently, there can as a matter of principle be no way of ascertaining both the value of the x-spin and the value of the y-spin of any particular electron at any particular moment. This is an example - but only one among literally infinitely many - of the uncertainty principle: measurable physical properties like x-spin and y-spin are said to be incompatible" with one another, since measurements of one will always (so far as we know) uncontrollably disrupt one another.

Now consider the rather complicated device shown in figure 3. In one corner there is a y-box. Y-spin = +1 electrons emerge from that box along the route labeled "y = +1", and at a certain point on that route there's a "mirror" or a "reflecting wall", which changes the direction of motion of the electron but doesn't change anything else about it (more particularly, it doesn't chance the value of the y-spin of an electron that bounces off it) as shown. And similarly for y-spin = -1 electrons.

And at the point where the two routes re-converge, there is a "black box", which also changes the directions of the motions of electrons, without altering the values of their y-spins in the process, in such a way as to make the two routes coincide after they pass through it.

Suppose, now, that we feed a large collection of x-spin = +1 electrons, one at a time, into the y-box; and then, as they emerge from the apparatus at "y = +1 and y = -1", we measure their x-spins. What sorts of results should we expect? Well, our previous experience informs us that half of such electrons (statistically speaking) will turn out to have y-spin = +1, and so will take route "y = +1" through the apparatus; and half of them will turn out to have y-spin = -1, and so will take route "y = -1" through the apparatus. Consider the first half. Since nothing that those electrons will run into in between the y-box and "y = +1 and y = -1" can have any effect on their y-spin values, they will all emerge from the apparatus as y-spin = +1 electrons, and concequently (our experience informs us here again) 50% of them will turn out to have x-spin = +1 and 50% will turn out to have x-spin = -1. The second half, on the other hand, will all emerge as y-spin = -1 electrons, but of course their x-spin statistics will be presicely the same. Putting all this together, it follows that of any large set of x-spin = +1 electrons which are fed into this apparatus, half should be found at the end to have x-spin = +1 and half should be found at the end to have x-spin = -1. But when such experiments are actually performed, what happens is that exactly 100% of the x-spin = +1 electrons that initially get fed into this apparatus - even if they are fed in in such a way that no more than a single one of them is ever in the apparatus at any particular time - come out with x-spin = +1 at the end.

Suppose that we rig up a small, movable, electron-stopping wall that can be slid, at will, in and out of, say, route "y = +1" (see figure 4). when the wall is "out", we have presicely our earlier apparatus; but when the wall is "in", all electrons moving along "y = +1" get stopped, and only those moving along "y =-1" get through to "y = +1 and y = -1".

What should we expect to happen when we slide the wall in? Well, to begin with, the overall output of electrons at "y = +1 and y = -1" ought to go down by 50%, since the input x-spin = +1 electrons ought to be half y-spin = +1 and half y-spin = -1, and the former shouldn't now be getting through. What about the x-spin statistics of the remaining 50%? Well, when the wall is out, 100% of the x-spin = +1 electrons initially fed in end up as x-spin = +1 electrons. That means that all of the electrons that take route "y =+1" end up with x-spin = +1 and that all of the electrons that take route "y = -1" end up with x-spin = +1, and since we can easily verify that whether the wall is in or out of route "y = +1" can have no effect whatever on the x-spins of electrons travelling along route "y = -1", that implies that those remaining 50% should be all x-spin = +1.

What actually happens when the experiment is performed is this: The output is down by 50%, as we expect. But the remaining 50% is not all x-spin = +1. It's half x-spin = +1 and half x-spin = -1. And the same thing happens, similarly contrary to our expectations, if we insert a wall in the "x = -1" path instead.

Consider, then, an electron which passes through the apparatus when the wall is out. Consider the possibilities as to which route that electron can have taken. Can it have taken "y = +1"? Apparently not, because electrons which take <u>that</u> route (as we've just seen again) are known to have the property that their x-spin statistics are 50-50, whereas an electron passing through our apparatus with the wall out is known to have x-spin = +1, with certianty, at "y = +1 <u>and</u> y = -1". Can it have taken "y = -1", then? No, for the same reasons. Can it somehow have taken <u>both</u> routes? Well, suppose that when a certian electron is in the midst of passing thorugh this apparatus, we stop the experiment and look to see where it is. It turns out that half the time we find it on "y - +1", and find nothing at all on "y = -1", and half the time we find it on "y = -1", and find nothong at all on "y = +1". Can it have taken <u>neither</u> route? Certianly not. If we wall up <u>both</u> routes, nothing gets through at all!

And it has become one of the central dogmas of theoretical physics over the past half-century or so these experiments leave us no alternative but to <u>deny</u> that the very <u>question</u> of which route such an electron takes through such a contraption <u>makes any sense</u>. The idea is that what these sorts of experiments force us to acknowledge is that asking what route such an electron takes is (somehow) like asking about (say) the political convictions of a tuna sandwich, or about the marital status of the number 5.

The idea is that asking such questions amounts to a misapplication of language, that it amounts to what philosophers call a <u>category mistake</u>. What typically gets said of such electrons in physics textbooks is emphatically <u>not</u> that they take either the "y = +1" route or the "y = -1" route or both routes or neither route through the apparatus, but that there is simply <u>not any matter of fact</u> - not merely no <u>known</u> matter of fact, but <u>no matter of fact at all</u> - about which route they take, that they are in what physicists call a <u>superposition</u> of taking the "y = +1" route and the "y = -1" route through the apparatus.

2) The Measurement Problem

Notwithstanding the profound violence all this does to our earlier picture of the world, to the very idea of what it is to be <u>material</u>, to be a <u>particle</u>, a compact set of rules called ‘Quantum Mechanics’ has been cooked up, which has proven extraordinarily successful at predicting all of the thus-far-observed behaviors of electrons under the circumstances we have just been talking about, and which (as a matter of fact) has proven extraordinarily successful at predicting all of the thus-far-observed behaviors of <u>all</u> physical systems under <u>all</u> circumstances, and which has functioned for more than seventy years now as the framework within which virtually the entirety of theoretical physics is carried out.

The mathematical object with which Quantum Mechanics represents the states of physical systems is called the <u>wave-function</u>. And it is a cardinal rule of Quantum Mechanics that representing things that way represents them <u>completely</u>, that <u>absolutely everything there is to be said</u> about any given physical system at any given temporal instant (that is: the value of every single physical property of that system whose value there is at present <u>any matter of fact</u> about, and the <u>probability</u> of any partiuclar <u>outcome</u> of any particular <u>measurement</u> one might choose to <u>carry out</u> on that system, <u>whether there is at present any matter of fact about the value of the property to be measured or not</u>) can be <u>read off</u> from it's wave-function.

In the particularly simple case of a single-particle system of the sort we've been concerned with here, the quantum-mechanical wave-function takes the form of a straightforward function of among other things) position. The wave-function of a particle which is located in some region A, for example, will have the value zero everywhere in space except in A, and will have a non-zero value in A. And the wave-function of a particle which is located in some other region region B will have the value zero everywhere in space except in B, and will have a non-zero value in B. And the wave-function of a particle which is in a superposition of being in region A and in region B (the wave-function, for example, of an initially x-spin = +1 electron which has just passed through a y-box) will have non-zero values in both of those regions, and will be zero everywhere else.

What the laws physics are about, according to quantum mechanics (and indeed all the laws of physics could be about, all there is for the laws of physics to be about, according to quantum mechanics), is how the wave-functions of physical systems evolve in time. And it is an extraordinary perculularity of the standard textbook formulation of quantum mechanics that there are two very different categories of such laws, one of which applies when the physical systems in question are not being directly observed, and the other of which applies when they are.

The laws in the first category are usually written down in the form of linear differential "equations of motion". And they are designed to entail (for example) that an initially x-spin = +1 electron which is fed into a y-box will emerge from that box (just as it actually does) in a superposition of travelling along the "y = +1" route and travelling along the "y = -1" route. And as a matter of fact all of the experimental evidence we presently have suggests that those laws turn out to be the laws which govern the evolutions of the wave-functions of all isolated microscopic physical systems whatever, under all circumstances. And so (since microscopic physical systems are after all what everything else in the world consists of) there would seem on the face of it to be very good reason to suppose that those linear differential equations are the true equations of motion of the entire physical universe.

But there are reasons why (if wave-functions are indeed complete descriptions of physical systems, as quantum mechanics maintains) that can't possibly be quite right.

To begin with, the laws expressed by those equations are completely detirministic, whereas there seems to be an element of pure chance (as I discussed above) in the outcome of a measurement of (say) the position of an electron which is initially in a superposition of being in region A and being in region B.

Moreover, it can be shown that what the linear differential equations of motion would predict about the end of a measuring-process like that (if those equations were indeed the true equations of motion of the whole world) is not that the pointer on the measuring-device would either point at ‘A’ - indicating that the electron was found in A - or that it would point at ‘B’ - indicating that the electron was found in B (which is what happens when you actually go and do measurements like that) - but rather that the pointer on the measuring-device would, with certainty, end up in a superposition of pointing at ‘A’ and pointing at ‘B’.

This analysis can even be expanded to include a human observer, whose role is to look at the pointer so as to ascertain how the measurement comes out - and when this expanded analysis is carried through, it emerges that the linear differential equations of motion predict that when everything is done, the observer will with certainty end up in a superposition of believing that she sees the pointer pointing at A and believing that she sees the pointer pointing at B!

What these equations predict (to put it slightly differently) is that the pointer on a measuring-device like that would end up, with certainty, in a physical state in which there is simply no matter of fact about where it's pointer is pointing - and that a human observer who looks at that pointer will end up, with certainty, in a physical state, a brain state, is which there is simply no matter of fact about where she believes that pointer to be pointing. And it hardly needs saying that that (whatever that is, precisely) is not what happens when you actually go and do such a measurement!

And so (the standard reasoning goes) the first category of laws needs to be supplemented with a second, which will be explicitly probabilistic, and which will entail (for example) that if the position of an electron whose wave-function looks like the one in fig (that is: an electron about whose position there is, at present, according to quantum theory, no matter of fact; an electron which is at present in a superposition of being located in region A and in region B) were to be measured, then there would be a 50% chance of finding that electron in region A (which is to say: there would be a 50% chance of that electron's wave-function being altered, in the course of the measurement, to one whose value is zero everywhere other than in region A) and a 50% chance of finding it in region B (which is to say: there would be a 50% chance of it's wave-function being altered, in the course of the measurement, to one whose value is zero everywhere other than at the point B).

As to the distinction between those circumstances in which the first category of laws applies and those in which the second category of laws applies, all that the received view has to offer is that that has something to do with the distinction between a "measurement" and an "ordinary physical process", or between what observes and what is observed, or between what lies (as it were) in front of measuring-devices and what lies behind them, or between subject and object. And it has for some time now been viewed by many physicists and philosophers as a profoundly unsatisfactory state of affairs that the best existing formulation of the most fundamental laws of nature should depend on distinctions as imprecise and elusive as those.

The problem of what to do about this, which has emerged over the past thirty years or so as the central problem at the foundations of quantum mechanics, has come to be called the measurement problem - and we shall be considering attempts at solving this problem in the section after next.

3) Attempts at Solving the Measurement Problem

There have been two main traditions of thinking about how to solve the measurement problem.

One of those is starts out by denying that the received way of thinking about what it means to be in a superposition is the right way of thinking about it; to deny, for example, that there fails to be any determinate matter of fact, when a superposition of ‘pointer pointing at ‘A’‘ and ‘pointer pointing at ‘B’‘ obtains, about where the pointer is pointing.

The idea (to come at it from a slightly different angle) is to construe quantum-mechanical wave-functions as less than complete descriptions of the world. The idea is that something extra needs to be added to the wave-function description, something that can broadly be thought of as choosing between the two conditions superposed above, something that can be thought of as somehow marking one of those two conditions as the unique, actual, outcome of the measurement that leads up to it.

The most famous and most successful theory we have in this tradition to date is due to

David Bohm.

Bohm's theory starts out with the stipulation that particles are the sorts of things that are invariably located in one or another particular place. Moreover, on Bohm's account, the wave-functions which are at the center of the quantum-mechanical description of the world are no longer merely mathematical objects, but physical ones, physical things. Wave-functions, on Bohm's theory, are somewhat like classical force-fields; and what wave-functions do in Bohm's thoery (just as force-fields do in classical mechanics) is to push the particles around, to guide them (as it were) along thier proper courses.

The laws which govern the evolutions of those wave-functions in time (which, as it happens, are stipulated to be presicely the linear differential quantum-mechanical equations of motion we discussed above; but this time with no exceptions whatever), and the laws which dictate how those wave-functions push their respective particles around (which are unique to Bohm's theory) are all fully detirministic.

Thus, the positions of all of the particles in the world at any time, and the world's complete quantum-mechanical wave-function at that time (which together comprise all there is to say about the world, on Bohm's theory) can in principle be calculated with certainty from the positions of all of the particles in the world and the world's complete quantum-mechanical wave-function at any earlier time; and any incapacity to carry out those calculations, any uncertainty in the results of those calculations, is necessarily (on this theory) an epistemic uncertainty, a matter of ignorance, and not a matter of the operations of any irreducible element of chance in the fundamental laws of the world. Nonetheless, this theory entails that some such ignorance (precisely enough, and of precisely the right kind, to reproduce the familiar statistical predictions of quantum mechanics by means of the kind of averaging over what one doesen't know that goes on in classical statistical mechanics) exists for us as a matter of principle, some such ignorance is unavoidably forced upon us by the laws of evolution of the theory. The dynamics acts so as to prevent us from ever knowing enough about the physical state of the world to make those predictions which the standard irreducibly statistical formalism of Quantum Mechanics can't make for us.

The account that Bohm's theory produces of the experiments with the two-paths contraption on figures 3 and 4 (the experiments, that is, which seemed to imply that electrons can be in states in which there fails to be any matter of fact about where they are) runs roughly like this:

Consider (as we did above) the case of an initially x-spin = +1 electron which is fed into the apparatus. On Bohm's theory, that electron will take either the "y = +1" route or the "y = -1" route, period. Which one of those two routes it takes will be fully detirmined by its initial conditions, by (more particularly) its initial wave-function and its initial position, but of course certian of the details of those conditions will prove impossible, as a matter of principle, to ascertian by measurement. Anyway, the crucial point here is that whichever route the electron happens to take, it's wavefunction will (in accordance with the linear differential equations of motion) split up and take both. And so in the event that the electron in question takes (say) the "y = +1" route, that electron will nonetheless be re-united, at the black box, with that part of its wave-function which took the "y = -1" route; and of course how that other part of the electron's wave-function ends up pushing the electron around, once the two are re-united, may well depend on whether or not it happened to (say) run into a wall along the "y = -1" route; and so that other part of the electron's wave-function can (as it were) inform an electron which travels through the contraption along one route about what's going on along the other one.

Moreover, it turns out that one of the consequences of the laws of Bohm's theory is that only that part of any given particle's wavefunction which is presently occupied by the particle itself can have any effect whatever on the motions of other particles; and so any attempt to detect the 'empty' part of the wave-function of an electron passing through the two-paths contraption (any attempt, that is, to detect some part of such an electron's wave-function travelling through the contraption along the path which the electron itself does not take) will (since the detecting device will of course consist of particles other than that electron) necessarily (notwithstanding the fact that that "empty" part is really, physically, there) fail.

And that accounts (without any recourse to anything like a superposition) for all of the odd behaviours we discussed of electrons which are fed into such contraptions.

Moreover Bohm's theory can manifestly have nothing along the lines of a measurement problem.

Notwithstanding the fact that on Bohm's theory the linear differential equations of motion are invariably the true equations of the time-evolution of the wave-function of the entire universe (measuring-devices, observers, and all!), there are also invariably definite matters of fact about the positions of particles, and (consequently) about the positions of pointers on measuring-devices.

There are a number of other extra-variable responses to the measurement problem which have emerged more recently. These are referred to in the literature as modal interpretations of quantum mechanics - and all of them start off start off (just as Bohm's theory does) by stipulating that the linear dynamical equations of motion are always exactly right, and that there are certain particular properties of physical systems (let's call them the extra properties of those systems) whose values are determinate even in the event that the quantum state of the world is a superposition of states corresponding to several different values of those properties.

On Bohm's theory, those extra properties are the positions of particles. On modal interpretations, things are a bit more complicated: On those interpretations, the identities of the extra properties can vary from moment to moment; and those identities depend on what the overall quantum state of the world is, and the particular way in which they depend on what that overall quantum state is (that is: the explicit rules whereby they depend on what that overall quantum state is) are designed with the aim of guaranteeing that measurements always have outcomes.

Moreover, modal interpretations (unlike Bohm's theory) are not entirely deterministic. The evolution of the quantum state of the world is of course entirely deterministic on these interpretations (just as it is on Bohm's theory), and the rules whereby the identities of the extra properties depend on what the quantum state of the world is are deterministic too, but the probabilities associated with the various possible values of the extra properties, on modal theories, are real physical chances.

The second of the two traditions I mentioned some pages back - the second of the two big ideas, that is, about how to solve the measurement problem - is to affirm that the standard way of thinking about what it means to be in a superposition is the right way of thinking of thinking about it, to affirm that a quantum-mechanical wave-function does amount to a complete description of a physical system, and to deny that the time-evolutions of those wave-functions always occur in strict accordance with the standard linear deterministic equations of motion.

The idea here is to somehow alter the equations of motion (without, of course, altering any of the innumerable empirical consequences of those equations which are now experimentally known to be true) so as to guarantee that superposition of the sort that figure in the measurement problem are made to go away.

There is a long history of speculations in the physical literature (speculations which have notoriously involved terms like "macroscopicness", and "consciousness", and "irreversibility", and "record", and "subject", and "object" and so on) about precisely what sorts of alterations are called for here; but there has to date been only one fully- worked-out, traditionally scientific sort of proposal along these lines, which is due to Ghirardi and Rimini and Weber, and which has been developed somewhat further by Philip Pearle and John Bell.

Ghirardi, Rimini, and Weber's idea goes (roughly) like this: The wave function of any single-particle system almost always evolves in accordance with the linear deterministic equations of motion; but every now and then (once in something like 10^9 years), at random, but with fixed probability per unit time, the wave function is suddenly multiplied by a narrow bell-shaped curve - a curve (more particularly) whose width is something on the order of the diameter of a single atom of one of the lighter elements - which has the effect of localizing it, of setting it's value at zero everywhere in space except within a certain small region. The probability of this bell-curve's being centered at any particular point x depends (in accordance with a precise mathematical rule) on the wave-function of the particle at the moment just prior to that multiplication. Then, until the next such 'jump', everything proceeds as before, in accordance with the deterministic differential equations.

This is the whole theory. No attempt is made to explain the occurrence of these 'jumps'; that such jumps occur, and occur in precisely the way stipulated above, can be thought of as a new fundamental law; a beautifully straightforward and absolutely explicit law of the so-called "collapse of the wave-function", wherein there is no talk at a fundamental level of 'measurements' or 'recordings' or 'macroscopicness' or anything like that.

Note that for isolated microscopic systems (i.e. systems consisting of small numbers of particles) 'jumps' will be so rare as to be completely unobservable in practice. On the other hand it turns out that the effects of these jumps on the evolutions of the wave-functions of macroscopic systems (systems like measuring-devices, for example) can sometimes be dramatic. Indeed, a reasonably good argument can be made to the effect that these jumps will almost instantaneously convert superpositions of macroscopically different states like {particle found in A + particle found in B} into either {particle found in A} or {particle found in B}, and that they will do so in very good accordance with the standard quantum-mechanical probabilities governing the outcomes of measurements like that.

There is a third - very ambitious, very novel, and very influential - tradition of attempts to solve the measurement problem which can be traced back to a 1954 paper of Hugh Everett III, and which encompasses ideas like the ‘Many-Worlds’ interpretation of quantum mechanics, and the ‘Relative States’ interpretation of quantum mechanics, and the ‘Decoherent Histories’ interpretation of quantum mechanics, and still others besides. The defining characteristic of this interpretation is the demand that quantum-mechanical wave-functions in and of themselves be taken as complete descriptions of the world, and that the evolutions of those wave-functions in time are invariably in full accord with the linear dynamical equations of motion - one or the other of which was explicitly denied (you will recall) by all of the proposed solutions to the measurement problem we have considered so far. Although these theories have been the center of a great deal of interest among physicists over the past several years, it remains unclear whether any of them can accommodate a coherent account of the probabilities which lie at the heart of what quantum mechanics has to say about observable behaviors of physical systems.

One of the important consequences of attempts at solving the measurement problem for the philosophy of science in general has to do with the threat of the so-called underdetermination of theory by experiment. It can be shown (that is) that various of the non-collapse proposals, although they radically differ with one another on questions as profound as (say) whether or not the fundamental dynamical laws of physics are deterministic, do not differ in ways that can ever, even in principle, be made available to empirical investigation.

4) Nonlocality

In a famous paper in 1935, Einstein, Podolsky, and Rosen produced an argument to the effect that - if the predictions of Quantum Mechanics about the outcomes of experiments are correct - the Quantum-Mechanical description of the world is necessarily incomplete.

A description of the world is 'complete', for E.P.R., just in case nothing that's true about the world, nothing that's an 'element of the reality' of the world, gets left out of that description. This entails that if we want to find out whether or not a certain description of the world is complete, we need first to find out what all the elements of the reality of the world are; and E.P.R. have nothing to offer in the way of a general prescription for doing that. What they do - which is something much narrower, but which turns out to be enough for their purposes - is to write down a merely sufficient condition for a measurable property of a certain system at a certain moment to be an element of the reality of that system at that moment. The condition is that "if, without in any way disturbing a system, we can predict with certainty (i.e. with probability equal to unity) the value of a physical quantity, then there exists an element of reality corresponding to that physical quantity".

Let's see what that amounts to. Consider a question like this: If a measurement of a certain particular observable (call it O) of a certain particular physical system (call it S) were to be carried out at a certain particular future time (call it T), what would the outcome be? Suppose that there is a method whereby I can put myself in position, prior to T, to answer that question, with certainty. And suppose that the method whereby I can put myself in that position involves no physical disturbance of S whatsoever. Then (according to E.P.R.) there must now already be some matter of fact about what the outcome of such a future O-measurement on S would be; there must now already be some fact about S (since the facts about S aren't going to get tampered with from the outside in the course of my putting myself in a position to answer the question about the O-measurement) in virtue of which that future measurement would come out in that particular way.

So, what E.P.R. want to argue (once again) is that if the empirical predictions of Quantum Mechanics are correct, then there must be elements of the reality of the world which have no corresponding elements in the Quantum-Mechanical description of the world.

The argument involves a certain particular physically possible state of a pair of electrons which has since come to be referred to in the literature as a ‘singlet’ state or an ‘E.P.R.’ state. Whenever an E.P.R. state obtains, there fails to be any determinate matter of fact - on the standard way of thinking, on the hypothesis (that is) that quantum-mechanical wave-functions are complete descriptions of physical systems - about the values of either the x-spin or the y-spin of either of the particles involved. On the other hand, quantum mechanics entails that if measurements of x-spin were to be carried out on each of the particles in an E.P.R. state, the outcomes of those two measurements would with certainty be the same; and quantum mechanics also entails that if measurements of y-spin were to be carried out on each of the particles in an

E.P.R. state, the outcomes of <u>those</u> two measurements would with certainty be the same as well.

Now, we have a certain inestimably deep intuition about how the world works - an intuition which goes under the name of <u>loclaity</u> - an intuition to the effect that nothing that happens <u>here</u> can directly and without mediation cause anything whatever to happen <u>elsewhere</u>. Of course, all sorts of things that happen here <u>can</u> - <u>indirectly</u> - be the cause of something that happens elsewhere. The flip of a switch one room, for example, can certainly be the cause of a light's going on in another; the point is just that no such causes can be <u>direct</u>, that no such causes can operate without <u>mediation</u>. If a flipped switch on one room <u>is</u> the cause of a light's going on in another, then (so this intuition advises us) there must necessarily be some causal sequence of events at contiguous points in space and at contiguous moments in time (the propagation of an electric current through a wire, say) stretching all the way without a break from the flipping of the switch to the light's going on. And of course the <u>capacity</u> of any such sequence of events to occur will necessarily depend on what sorts of physical conditions obtain in the space between the rooms in question - it may, for example, require that there is an <u>electrical conductor</u> stretching all the way, without a break, from the switch to the light. And that sequence of events (in virtue of <u>being</u> a sequence, in virtue of being a string of causes whose effects are new causes whose effects are new causes...) must necessarily require some finite <u>time</u> to completely unfold.

And it is crucial to the argument of E.P.R. that locality is true. E.P.R. assume (that is) that things could in principle be set up in such a way as to guarantee that the measurement of the x-spin or the y-spin of electron number one produces no physical disturbance whatsoever in electron number two. One could, for example, separate the two electrons by some immense distance - since quantum mechanics predicts that none of the above-mentioned characteristics of E.P.R. states depend on how far apart the two electrons happen to be. Or one could insert an impenetrable wall between them or build impenetrable shields around them - since quantum mechanics also predicts that none of what's been said depends on what happens to lie between or around those two electrons. Or one could set up any desired array of detectors in order to verify that no measurable signals ever pass from one of the electrons to the other in the course of the experiment - since Quantum Mechanics predicts that no such array, in such circumstances, whatever sorts of signals it may be designed to detect, will ever register anything.

Thus, whenever the E.P.R. state obtains, there is a means of predicting, with certainty, and - in principle, if locality is true, and if everything has been correctly arranged - without disturbing electron two, what the outcome of any subsequent measurement of the x-spin of electron two will be. The way to do that is to measure the x-spin of particle <u>one</u>; since it's known that the outcome of any measurement of the x-spin of particle two will invariably be the <u>opposite</u> of the outcome of any measurement of the x-spin of particle one. And so it follows from the reality-condition that x-spin must necessarily be an element of the reality of particle two, that there must necessarily be a matter of fact about what the value of the color of electron two <u>is</u>, whenever the E.P.R state obtains. And precisely the same considerations clearly apply to y-spins as well.

So the standard way of thinking must simply be false, since there can be circumstances (like E.P.R. states) in which there are simultaneously determinate matters of fact about the values of both the x-spin and the y-spin of a single electron, even though those two observables are supposed to be incompatible with one another.

As a matter of fact, it turns out that there are an infinite number of distinct and mutually incompatible observables of electrons which have the property that if measurements of any one of them are carried out on each of a pair of particles in the E.P.R. state, then quantum mechanics will entail that the outcomes of those two measurements will with certainty be the same. And so it will follow from the above sorts of arguments whenever an E.P.R. state obtains, there must simultaneously be determinate matters of fact about the values of all of those, for each of the electrons in the pair.

Moreover, the formalism itself (quite apart from any particular way of thinking about it) must necessarily be incomplete, since there are necessarily certain facts, certain elements of the physical reality of the world, which have no corresponding elements in the formalism. There are facts about the x-spin and the y-spin of electron two, for example, when E.P.R. state obtains, but there isn't anything in the mathematical description of the E.P.R. state, in this formalism, from which the values of the color or the hardness of electron two can be read off.

If all this is right, then whenever The E.P.R. state obtains, there are an infinity of distinct and mutually quantum-mechanically incompatible observables of each of the electrons in the pair each of which has a determinate value. And if that's so, then the statement that The E.P.R. state obtains - since that state does not specify a value of any of the above-mentioned observables - necessarily constitutes a very incomplete description of the state of a pair of electrons. The statement that the E.P.R. state obtains must be true of a gigantic collection of different 'true' states of such a pair, in which those observables take on various different combinations of values.

Nonetheless, the information that the E.P.R. state obtains must certainly constrain the 'true' state of a pair of electrons in a number of ways, since the outcomes of spin-measurements on such pairs of electrons are (after all) determined by what their 'true' states are, and since we're assuming that the Quantum-Mechanical predictions about the statistics of the outcomes of such measurements are correct.

Consider what sorts of constraints arise. First of all, if the E.P.R. state obtains, then the outcome of a measurement of any one of the above-mentioned abservables of electron one will necessarily be the same as the outcome of any measurement of the same observable on electron two; and so, whenever the E.P.R. state obtains, the 'true' state of the pair of electrons in question is constrained, with certainty, to be one in which the value of every such observable of electron one is the opposite of the value of that same observable of particle two.

There are statistical sorts of constraints as well. There are, in particular, three observables of these electrons - one of them is the x-spin, and we can call the other two the l-spin and the k-spin - such that if any one of them is measured on electron one and any other one of them is measured on electron two, the outcomes are equal 1/4 of the time, and opposite 3/4 of the time.

And now a well-defined question can be posed as to whether these two constraints (the deterministic constraint about the values of identical observables for the two electrons, and the statistical constraint about the values of different observables for the two electrons) are mathematically consistent with one another. It was John Bell who first clearly posed and answered that question, twenty-nine years after the publication of the E.P.R. argument; and it turns out that the answer to that question is no.

And so the conclusion of the E.P.R. argument is logically impossible; and so either locality must be false or the predictions of Quantum Mechanics about the outcomes of spin-measurements on E.P.R. states must be false - since those are the only two assumptions on which the argument depends. And it happens that those predictions are now experimentally known to be true; and so the assumption that the physical workings of the world are invariably local must - astonishingly - be false.

Here is a concise and useful way of telling the story:

Einstein and Podolsky and Rosen noticed that there was something odd about the collapse-postulate for two-particle systems. They noticed that it was non-local: if the two particles are initially in an E.P.R. state, then a measurement carried out on one of them can bring about changes, instantaneously, in the Quantum-Mechanical description of the other one, no matter how far apart those two particles may happen to be, or what might lie between them.

E.P.R. suspected that this non-locality must merely be a disposable artifact of this particular mathematical formalism, of this particular procedure for calculating the statistics of the outcomes of experiments; and that there must be other (as yet undiscovered) such procedures, which give rise to precisely the same statistical predictions, but which are entirely local.

And it emerged thirty years later, in the work of Bell, that that suspicion was demonstrably wrong.

Bell's work has sometimes taken to amount to a proof that any attempt to escape from the standard way of thinking, any attempt to be realistic about the values of the spin-observables of a pair of electrons for which the E.P.R. state obtains, must necessarily turn out to be non-local. But things are actually a good deal more serious than that. What Bell has given us is a proof that there is as a matter of fact a genuine non-locality in the actual workings of nature, however we attempt to describe it - unless (which at this writing seems unlikely) one or another of the interpretations in the tradition of Hugh Everett III should turn out to be true. That non-locality is, to begin with, a feature of Quantum Mechanics itself, and it turns out (via Bell's theorem), that it is necessarily also a feature of every possible manner of calculating - outside of the tradition of Everett - which produces the same statistical predictions as Quantum Mechanics does. And those predictions are now experimentally known to be correct.

PROSPECTS AND CONNECTIONS

1) Quantum Theory and the Structure of Space-Time

There are a number of quite fundamental tensions between quantum theory and the special theory of relativity. These tensions have been very much in plain sight for more than thirty years at this writing - but it is only of late that the will to resolutely look them in the face has begun to take hold.

To begin with, every understanding of quantum theory we have, every attempt at solving the problem of measurement we know of, is committed to a description of the states of physical systems at least partly in terms of wave-functions. And the wave-functions for systems consisting of more than a single particle are simply not expressible as functions of space and time - they are ineluctably functions of time and position in a much larger-dimensional space, a space in which the fundamental relativistic criterion of lorentz-invariance can apparently not even be unambiguously defined - called configuration-space.

Moreover, there is a very intimate connection - a connection which has been at the center of the canonical understanding of the special theory of relativity from its earliest beginnings - between lorentz-invariance and locality. This connection is now understood not to be a matter straightforward logical implication - indeed, we can now point to a number of explicit models of simple physical theories which are both lorentz-invariant and non-local - but none of these seem to have quite the same sort of non-locality in them as quantum theory does. Both the standard formulation of quantum mechanics and every single one of the proposals we have for solving the measurement problem - with the exception, once again, of the theories in the tradition of Everett - requires that lorentz-invariance be explicitly false. Every one of those proposals - more particularly - requires that there be a preferred, absolute, non-lorentz-invariant standard of simultaneity.

These tensions have already generated a broad and unprecedented revival of interest in the long-neglected approach of Lorentz to the physical phenomena associated with the special theory of relativity, and there can be little doubt that the business of coming fully to grips with all this - and with its further ramifications for the much-discussed project of the reconciliation of quantum theory with the general theory of relativity - will be a central concern of the philosophical foundations of physics over the coming years.

2) Quantum Theory and the Foundations of Statistical Mechanics

There has been a vague and mostly unspoken idea floating just under the surface of theoretical physics for many years - which is only now beginning to receive a detailed and rigorous and quantitative sort of examination - to the effect that the statistical character of our everyday experience, and it's asymmetry under time-reversal, may have something deep to do with the statistical character and asymmetry under time-reversal of a number of proposed solutions to the infamous quantum-mechanical problem of measurement As of this writing (for example) there is reason to hope that if anything along the lines of the GRW theory of the collapses of quantum-mechanical wave-functions should turn out to be true, then the melting of ice and the cooling of soup and the breaking of glass and the passing of youth could be shown to represent genuine stochastic time-asymmetric dynamical laws; there is reason to hope (that is) that if that sort of a theory should turn out to be true then the melting of ice and the cooling of soup and the breaking of glass and the passing of youth could be shown to be the sorts of transitions which ore overwhelmingly likely to occur, and overwhelmingly unlikely to occur in reverse, completely irrespective of what the initial conditions of the universe may happen to have been.

3) Frontiers

The investigation of the philosophical foundations of physics has traditionally been occupied with attempts to clarify the logical structures and philosophical commitments and inter-theoretic relations of various individual physical theories - the general theory of relativity, say, or statistical mechanics, or quantum field theory, or any number of others. But over the past several decades, in some quarters, a more ambitious project - an inquiry into how the entirety of physical science hangs together - has begun to take shape.

Prior to the end of the 19th century, no physical theory so much as invited consideration as a candidate for a complete account of the behaviors of physical systems. All of them manifestly left vast chunks of the story out - none had anything to say about light, for example, or about atomic spectra, or about the rules of chemical combination. By the 1920's, however, all that had changed. By then, for the first time, it began to make sense to inquire into the possibility that quantum theory (say) might provide the framework of a complete and unified mechanical account of the entirety of the physical world.

Moreover, it was discovered at about that same time that substantive and non-trivial and radically counter-intuitive conclusions about the behaviors of macroscopic measuring-instruments and about the inner lives of embodied subjects have been drawn directly out of the mathematical structure of a certain proposed set of microscopic fundamental laws of physics! That (after all) is precisely the content of the measurement problem. And this sort of a move - this sort of a willingness to entertain the possibility of the most radical imaginable sort of completeness of physics, this sort of determination to push the general project of a physical account of the world as far away as possible from its originally intended applications, to push it at exactly those points at which it seems most at risk of collapsing - is precisely what's essential and distinctive about the way the interrogation of the foundations of physics has lately - in some quarters - been pursued.

Over the past several decades, for example, it has begun to seem altogether reasonable to inquire into the influence of the structure of the fundamental laws of physics on the question of what sorts of calculations it might in principle be possible to carry out - and such inquires have led directly to the rapidly growing field of Quantum Computers. And it has lately become a widely accepted condition of the entertainability of any proposal for a fundamental physical theory of the world - a condition which has played a crucial role in a number of recent debates - that the theory in question contain an account of how sentient inhabitants of the universe it describes might come to have reason to believe that that theory is true. And there have been attempts (which were mentioned briefly in chapter 2 of this article) to understand how the structure of the fundamental laws can account for the time-asymmetries of epistemic access and causal intervention, which lie at the core of role of time in human affairs. And there are many other such examples besides.

All of this is regarded in some quarters as the opening of a distinctive new frontier of physics. This (if it exists) is not a frontier of the very big or the very small or the very fast or the very slow, which are the traditional frontiers of physics, which are the parts of nature about which there is no tradition of thinking outside of physics - but (rather) a frontier at which fundamental physics penetrates into the everyday human experience of being in the world, a frontier at which fundamental physics collides in the most direct imaginable way with an entire culture of altogether different ways of thinking.

BIBLIOGRAPHY

Hans Reichenbach's The Philosophy of Space and Time is the book that set the terms of the 20^{th} century debate about the philosophy of space and time, and Lawrence Skalr's Space, Time, and Spacetime is an excellent and very accessible account of the subsequent course of that debate. On a slightly more technical level, John Earman's World Enough and Space-Time is a very careful discussion of the question of the relativity of motion, and of the question of substantivalism, and of a whole host of related topics, including many important and original contributions to those subjects. Bangs, Crunches, Shrieks and Whimpers, by the same author, is a very up-to-date account of work on the philosophical interpretation of the various conceptually puzzling solutions - some of which were mentioned in the text - to the Einstein field equations of General Relativity. David Mermin's Space and Time in Special Relativity a very accessible and yet very profound discussion of the foundations of the special theory of relativity - including a very illuminating discussion of the whole matter in the manner of Lorentz. The earlier sections of Julian Barbor's book The End of Time contain an admirable clear account of his important work in the development of a Machian version of Newtonian mechanics.

The book that set the modern agenda for discussions of the foundations of statistical mechanics is also the work of Reichenbach, and is called The Direction of Time. Sklar's encyclopedic Physics and Chance - with wonderful background chapters on the philosophy of probability - is also very useful. A recent book on this topic, which discusses the connections between the foundations of statistical mechanics and the foundations of quantum mechanics and the recent attempts at extending the Boltzmanian account of the thermodynamic time-asymmetry to the asymmetries of knowledge and intervention is David Albert's Time and Chance.

The best book yet written on the general foundations of quantum mechanics is a collection of John Bell's papers on the subject, entitled Speakable and Unspeakable in Quantum Mechanics. A host of other very serviceable ones have been published in recent years, including Miachel Lockwood's Mind, Brain, and the Quantum, Bas van Fraassen's Quantum Mechanics, and Albert's Quantum Mechanics and Experience. The best account of the collision between quantum mechanics and the special theory of relativity is undoubtedly Tim Maudlin's Relativity and Quantum Non-Locality.

David Z Albert
Columbia University