Happiest Thoughts:
Great Thought Experiments of Modern Physics[*]

Kent A. Peacock
Department of Philosophy
University of Lethbridge

1. Introduction: Of Chickens and Physicists

A physicist can be defined as a person for whom a chicken is a uniform sphere of mass M. The point of this joke (which this author first heard from a physics professor) is that physicists shamelessly omit a lot of detail when they attempt to model and predict the behaviour of complex physical systems; indeed, one of the important skills that physics students must learn is knowing what to *leave out* when setting up a problem. This penchant for simplification does not necessarily mean that physicists are hopelessly out of touch with reality, however; for one can learn a surprising amount about how real things behave by thinking about apparently simplistic models.

A typical textbook example of a physical model might be the block sliding down an inclined plane. The plane is at a definite angle with respect to the force of gravitation; the block (a rectangular chunk of indefinite stuff) has a given mass, and there will be a certain coefficient of friction between the block and the plane. The assignment might be to calculate the coefficient of friction that would be sufficient to prevent the block from sliding down the plane, as a function of the angle of the plane. Now, no block of material in the world is perfectly uniform, no planar piece of material is perfectly flat and smooth, no actual coefficient of friction is known to arbitrary accuracy, and gravity is never exactly uniform in direction and magnitude. And yet, there are many physical systems in the real world which are sufficiently *like* idealized models such as this, to a definable degree of approximation, that their observable behavior can be predicted using such models. Models are therefore useful not only because they help us to picture how basic physical principles work in a concrete situation, but also as frameworks on which to hang a practical calculation.

Even textbook problems couched in terms of simple models such as the inclined plane amount to thought experiments of a sort. Usually, though, we reserve the honorific *Gedankenexperimente* for idealized scenarios that give us new insights into the meaning or limitations of important physical concepts, usually by testing their implications in extreme or highly simplified settings. Suppose, for example, that I confusedly believe that all objects fall at a rate that is a function of their mass. Galileo has an elegant thought experiment that shows that

---

[*] To appear in James Robert Brown, Yiftach Fehige, and Michael T. Stuart (eds), *The Routledge Companion to Thought Experiments*. Preprint posted to philsci archive, August 6, 2016. Comments, corrections, and objections are welcome; contact the author at kent.peacock@uleth.ca.

my notion is a mistake: all objects in a uniform gravitational field must fall with the same acceleration (ignoring air resistance), on pain of outright contradiction.[1]

Galileo's thought experiment, like most typical textbook models, can be translated into experiments that can actually be performed. But sometimes one can learn a lot even from thought experiments that cannot be done, at least in the simple terms in which they are first described. Mach invited us to rotate the entire universe around Newton's bucket of water. No granting agency will fund that feat, and yet Mach's insights contributed (in a complicated way) to the construction of a theory (general relativity) that has testable consequences (Janssen, 2014).

In this chapter I describe several thought experiments that are important in modern physics, by which I mean the physical theory and practice that developed explosively from the late 19th century onwards. I'm going to mostly skip thought experiments that merely *illustrate* a key feature of physics[2] in favour of those that contributed to the advancement of physics. Some thought experiments (such as Galileo's) provide the basis for rigorous arguments with clear conclusions; others seem to work simply by *drawing attention* to an important question that otherwise might not have been apparent. Many of the most interesting thought experiments have ramifications far beyond what their creators intended. I'll pay special attention to one particular thought experiment, defined by Einstein and his collaborators Boris Podolsky and Nathan Rosen (1935). The Einstein-Podolsky-Rosen (EPR) thought experiment dominates investigations of the foundations of quantum mechanics and plays a defining role in quantum information theory; I will suggest that it may even help us understand one of the central problems of cosmology. In the form in which Einstein and his young colleagues first described it, the EPR experiment was another idealization that probably cannot be performed. Despite this, it has evolved into practicable technology. While the thought experiments to be discussed here are key turning points in the history of modern physical theory, in several cases (including the EPR experiment) their full implications remain to be plumbed.

It is an extraordinary fact that most of the definitive thought experiments in twentieth century physics were born from the fertile imagination of one person, Albert Einstein. This forces us to ponder the importance of individual creativity in the advancement of science. Music would be very different, and much diminished, had Beethoven died young. If Einstein had not lived, would others have made equivalent discoveries? It seems likely that many of his advances would have been arrived at by other competent physicists sooner or later—except perhaps for general relativity, for the very conception of the possibility of, and need for, such a theory was due to the foresight and imagination of Einstein alone.

2. Sorting Molecules: A Thought Experiment in Thermodynamics

We'll start with Maxwell's Demon, a thought experiment that bridges 19th and 20th century physics.

---

[1] For a nice analysis of Galileo's thought experiment, see (Arthur, 1999).

[2] Such as the double-slit experiment. For a lucid exposition, see (Feynman, Leighton, & Sands, 1965 Ch. 1).

James Clerk Maxwell (1831–79) is best known for his eponymous equations for the electromagnetic field. He also made important contributions to statistical mechanics; in particular, he was the first to write the Maxwell-Boltzmann probability distribution which describes the statistics of particles in a Newtonian gas. Although Maxwell did not originate the concept of entropy (which was due to Rudolf Clausius, 1822–88) he was well aware of the Second Law of Thermodynamics, which, in the form relevant to our discussion here, states that no process can create a temperature difference without doing work. Maxwell fancifully imagined a box containing a gas at equilibrium, with its temperature and pressure uniform throughout apart from small, random fluctuations (Norton (2013a)). A barrier is inserted in the middle of the box, and there is a door in the barrier. A very small graduate student with unusually good eyesight is given the task of tracking the individual gas molecules and opening or closing the door so as to sort the faster molecules into (say) the left side, and the slower molecules into the right. Apparently, then, a temperature difference can be created between the two sides by means of a negligible expenditure of energy. The problem is to say precisely why such a violation of the Second Law of Thermodynamics would not be possible.

There is a large literature on Maxwell's demon, which we can't hope to do justice to here. (For entry points, see (Maroney, 2009; Norton, 2013a).) It is well understood that there is a sense in which statistical mechanics would, in principle, allow us to beat the Second Law—albeit, in general, only for extremely brief periods of time. The easiest way to create a temperature difference between the two boxes is simply to leave the door open for a very, very long time. Eventually enough fast molecules will, by pure chance, wander into one side and enough slow molecules, again by pure chance, will wander into another, to create a measurable temperature difference between the gasses in the two boxes, at least until another fluctuation erases the gains made by the first. Now imagine that the hole has a spring-loaded door which could snap shut as soon as a specified difference in temperature ΔT was detected between the two partitions. The thermometer and door mechanism will have some definite energy requirement, but this can be made independent of ΔT. If we want ΔT to be large enough that it implies an energy transfer greater than the energy requirements of the door mechanism, all we have to do is wait long enough and eventually a large enough fluctuation will probably (not certainly) come along—although the larger we want it to be, the longer we (again, probably) have to wait. (Let's call this process "fishing for fluctuations"; like ordinary fishing the result can never be guaranteed.) As soon as the desired temperature difference is detected, the door snaps shut and we would have "trapped a fluctuation" in a way that apparently violates the Second Law. This example underscores the point made by Ludwig Boltzmann (and apparently well understood by Maxwell), which is that the Second Law is a statistical statement. Violations of the Law by pure chance are possible. In trying to exorcise Maxwell's demonic assistant we are dealing with a question of what is overwhelmingly probable, not what is certain in a law-like way. The question is not whether energetically-free sorting against entropic gradients (such as temperature or concentration) could be done at all, but whether it can be done reliably, repeatedly, and in a time span shorter than the life of the observable universe.

In 1914 Marian Smoluchowski presented a critique of the Demon in terms of the statistics of fluctuations (see Norton (2013a)), which Smoluchowski argued would *almost* always wipe out any gains against equilibrium that any conceivable demon could make. Leo Szilard argued in 1929 (1972) that information-theoretic constraints would prevent the demon from beating the Second Law. The acquisition and manipulation of the information that the Demon would need in order to track the particles would, Szilard argued, inevitably dissipate more waste heat than could be gained by sorting the molecules. John Norton (2013a) champions Smoluchowski's analysis, and argues that the information-theoretic approaches to refuting the Demon popular from Szilard onward are all more or less circular in that they *presume* the validity of the Second Law. More recently, Norton argues (Forthcoming) that the question of information is irrelevant and that the Demon can be ruled out on the basis of Liouville's Theorem of statistical mechanics, which shows that the operation of a Maxwell's Demon is strictly impossible in any system that undergoes Hamiltonian evolution (i.e., virtually every conceivable classical system).[3] I will not attempt here to decide upon the correct theoretical analysis of the Demon. Instead, let us invoke the other towering figure of 19th century science—Darwin—and sketch a view of the Demon from an evolutionary point of view.

In biology it is well known that cell membranes regularly perform a process called *active transport*. This involves the pumping of a wide variety of molecules or ions through tiny pores in a cell membrane against entropic gradients. The intricate molecular machines that perform active transport in the walls of virtually all kinds of cells are the closest things in the biological world to Maxwell's Demon, although they tend to create concentration differences, not temperature differences. The crucial point is that any sort of active transport that has so far been observed and studied by cell biologists requires the expenditure of energy. As R. N. Robertson puts it,

> Systems which can transport molecules against their concentration gradients or ions against their electrochemical potential gradients are called *active transport* systems… Such systems use energy provided by the cell to work against the tendency for everything to reach chemical equilibrium… Formally, active transport is a reversal of the decrease in free energy which occurs when concentration or electrochemical systems tend toward equilibrium (1983, pp. 134–135).

Is it conceivable that natural selection could arrive at a form of active transport that does not require the expenditure of some of a cell's budget of metabolic energy?

There is growing evidence that natural selection tends to act so as to minimize the use of available energy. Damian Moran and co-authors (2014) studied a species of eyeless fish which live in underground caves entirely devoid of light. These fish use dramatically less metabolic energy than their surface-dwelling cousins, not only because they have no eyes, but because they do not partake in the energy-intensive circadian rhythm typical of animals exposed to the cycles of night and day. As Moran et al. (2014) put it,

---

[3] Norton also develops the same result for quantum systems.

> While it is a strange thought for terrestrial vertebrates to entertain, it may be unnecessary for animals living in caves or the deep sea to rouse their metabolism for the onset of a day that will never arrive.

If there is no particular survival advantage to paying the high metabolic cost of sight and all of the activities that go with it, a species that evolves its eyes away will have a survival advantage over one that does not. And it seems likely that this would apply generally: all things being equal, if there is a way to reduce energy expenditure it will tend to be found—and favoured—by natural selection. Cellular life goes back well over three billion years, and natural selection has had all of that time to sample the possibility space for active transport and to converge on the means that are the most economical and efficient in their use of cellular resources, especially energy. If it were indeed possible to micro-sort molecules against entropic gradients without the expenditure of energy, natural selection almost certainly would have found a way to do it by now.[4] Arguably, then, we can take this as a good sign that Maxwell's busy Demon is *practically* impossible, whether or not the last theoretical loopholes in the arguments against it can ever be closed.

3. Thought Experiments in Relativity

We turn now to a series of thought experiments that played important roles in the development of the special and general theories of relativity. There is a vast literature on these thought experiments, and we can do little more here than sketch the most interesting ones and point in some directions in which philosophical or scientific questions about them may still linger.

Einstein used a number of elegant models to illustrate how relativity works. For example, he imagined a railway carriage rolling along at constant velocity with respect to a level embankment, and used this scenario to illustrate the relativity of simultaneity (Einstein, 1961). However, accounts of these illustrative models are widely available and they do not seem to have played a major role in his discovery of the theory.

3.1 To Catch a Light Beam

In his "Autobiographical Notes," Einstein, writing more than fifty years after the fact, claims that he "hit upon" a paradox at the age of sixteen:

> If I pursue a beam of light with velocity *c* … I should observe such a beam of light as an electromagnetic field at rest though spatially oscillating. There seems to be no such thing, however, neither on the basis of experience nor according to Maxwell's equations. … One sees in this paradox the germ of the special theory of relativity… (Einstein, 1951)

As John Norton explains (2013b), both the exact timing and content of Einstein's youthful insight are open to question; quite likely Einstein, as do many of us, had gently revised his recollections decades after the events he describes. In a much earlier account given by Einstein of this thought experiment (Norton, 2013b, pp. 130–31) he denied that he had, at age sixteen, a

---

[4] Norton (Forthcoming) makes essentially the same argument with respect to ribosomes.

clear notion of the constancy of the velocity of light; rather the function of the thought experiment for the young Einstein simply seems to have been that it raised a striking and suggestive question: precisely what would one see if one could catch up with a light ray? In this respect the light ray experiment is similar to Maxwell's Demon—no clear conclusion, just a question. Thus, we see that unlike Galileo's thought experiment refuting the Aristotelian notion of variable rates of fall, which presents a rigorous *reductio* argument with a definite conclusion, thought experiments can contribute to advances in science simply by vividly directing attention to a problem that no one else seems to have worried about.

It is easy for modern commentators to Whiggishly interpret the light ray experiment in the light of what is *now* known about relativity. *Prima facie*, the post-1905 interpretation is as follows: because all motion is relative, there should be no way to detect one's velocity in any universal or absolute sense. If I could catch up to a light beam then I would presumably see a pattern of standing waves; and if light waves move at a universal speed with respect to the hypothetical ether, which was supposed to be grounded in absolute space, then I would know my absolute velocity. But even this line of thought is less clear than it first seems. Einstein in his "Autobiographical Notes" claims that according to Maxwell's equations there is no such thing as a standing light wave. This was an odd thing for Einstein in the late 1940s to have said, since he must have known that there are indeed standing wave solutions of the field equations. It is a question of having the right boundary conditions. For instance, a standing electromagnetic wave can be set up between two mirrors facing each other. Perhaps what Einstein had in mind was a free wave propagating in empty space—a highly idealized conception in itself. Or perhaps what he meant was that there is no such thing as a single electromagnetic wave pattern that can be either standing or moving *depending only upon an observer's state of motion*. For whether or not an electromagnetic wave pattern is standing or travelling depends upon the phase relationships among its components, and phase relations are Lorentz covariant. Thus, a wave pattern that is standing in one frame is standing in all.[5] Post-1905, what the light ray experiment actually says is that there is no such thing as catching up to a travelling electromagnetic wave until it turns into a standing wave—although this is really a logical *consequence* of the constancy postulate for light and not a ground for the latter.

As Norton (2013b) points out, there is good reason to think that the pre-1905 Einstein was not in fact targeting the ether theory of light, but rather the emission theory of light—the view that the speed of light depends upon the speed of its source. If the light ray experiment is to be interpreted as an argument at all, Norton says, it is an argument against the emission theory, not the ether theory of light.

Whatever Einstein may have had in mind at the age of sixteen or seventeen, the upshot is that by 1905 he had arrived at a theory in which Maxwell's equations are taken as laws of nature

---

[5] John Norton (private communication) has cautioned me to speak carefully here. Consider standing waves between two facing mirrors. The nodes are at rest with respect to the mirrors, and so of course they must move with respect to someone moving with respect to the mirrors. However, everyone will agree that the wave pattern is standing *with respect to the mirrors*.

which are the same for all inertial frames of reference. These equations predict the existence of electromagnetic waves that travel with an invariant velocity, for the speed of light is a *constant* in the equations. Therefore, if Maxwell's equations are laws of nature, the speed of light must be, in effect, a law of nature itself. As such, it must be independent of an observer's state of motion, even when one is moving arbitrarily close to the speed of light. As Einstein demonstrated, this assumption leads to the Lorentz transformations. Then if we wish to make kinematics and dynamics consistent with Maxwell, we have to make them Lorentz covariant as well. If the round Maxwellian peg won't go into the square Newtonian hole, the hole must be made round—and Einstein showed precisely how to do this.

The light beam thought experiment illustrates an important feature of the way physical theories develop: the intuitive or heuristic viewpoint that stimulated a new development in theory is sometimes not preserved by the time that the resulting theory is formalized. The presumption that such formative intuitions must always be preserved in the formalized theory that flows historically from them has been dubbed by John Woods (2003) the "heuristic fallacy". That this presumption is indeed often false is all the more clear when we grasp that thought experiments do not always contribute to the advancement of a science by serving as *arguments* for any particular proposition—at least not in ways that can be unambiguously reconstructed decades after the fact. Einstein himself described his light ray thought experiment as "child-like" and said, "Discovery is not a work of logical thought, even if the final product is bound in logical form" (quoted in (Norton, 2013b, p. 130)).

3.2 "The Happiest Thought of My Life"

By 1907 special relativity was consolidated. After another foray into quantum mechanics, in which he produced the first qualitatively correct quantum theory of specific heats and in effect founded modern solid state physics (Pais, 1982, Ch. 20), Einstein turned his attention to the problem of unifying gravitation with the principle of relativity. The obvious barrier to writing a relativistic theory of gravitation was that Newton's law of gravitation contains no dependency on time, and is therefore an action-at-a-distance theory (a fact that Newton himself had deplored (I. Newton, 1692)). There was a subtler but no less fundamental problem with Newton's theory: in his picture it is entirely a coincidence that gravitational mass (the "charge" that appears in the force law) and inertial mass (the resistance of an object to an accelerating force) happen to be precisely the same quantity. Einstein reports that in the course of writing a review article on relativity, he was suddenly struck by "the happiest thought of my life" (Pais, 1982, Ch. 9):

> The gravitational field has only a relative existence in a way similar to the electric field generated by magnetoelectric induction. *Because for an observer falling freely from the roof of a house there exists*—at least in his immediate surroundings—*no gravitational field*. … Indeed, if the observer drops some bodies then these remain relative to him in a state of rest or uniform motion… The observer therefore has the right to interpret his state as 'at rest.'

Galileo’s observation that all bodies in a uniform gravitational field fall with the same acceleration (neglecting air resistance) thus takes on a “deep physical meaning”: gravitation is therefore simply a manifestation of inertia, a “fictitious force” such as the centrifugal and Coriolis forces. Most important for Einstein, this fact is a manifestation of the Principle of Relativity:

> if there were to exist just one single object that falls in the gravitational field in a way different from all the others, then with its help the observer could realize that he is [falling] in a gravitational field…”. [This is] therefore a powerful argument for the fact that the relativity postulate has to be extended to coordinate systems which, relative to each other, are in non-uniform motion (in Pais, 1982, p. 178).

On the basis of this thought experiment, Einstein formulated his Equivalence Principle, which expresses the equivalence of gravitation and acceleration: an accelerated frame is equivalent to an inertial frame experiencing a gravitational field.[6] The Equivalence Principle is often illustrated by the elevator thought experiment, which appears in his first exposition of relativity for the general readership, written and published in 1916 (Einstein, 1961). If I am floating in the midst of in a windowless elevator car in free-fall, I have no way of telling from any measurements I can perform within the car whether it is falling freely in a uniform gravitational field, or moving inertially in deep space, far from all matter. (Out of a commendable concern for the safety of the experimenter, Einstein advises that he “fasten himself with strings to the floor” (Einstein, 1961, p. 66).) If there is a rocket engine attached to the base of the elevator car, and I am held to the floor of the car by a constant force, I have no way of telling whether the rocket is burning with constant thrust, or whether I am sitting on the surface of a large planet whose gravitational field is such as to generate a uniform acceleration equal to that of the rocket.

The idea that gravitation is an inertial force is simple and beautiful; it had to be right. However, it was incompatible with the long-held presumptions (which Kant thought were *a priori*) that space has to be Euclidean and time has to be absolute. This was apparent to Einstein in at least two ways, which again can be illustrated with elegant thought experiments.

First, go back to the elevator car sitting on the surface of a spherical planet. Small test masses will fall toward the centre of mass of the planet; thus, two test masses released side by side will move toward each other as they fall. Now, the gravitational potential is a function of the distance from the centre of mass of the planet. Therefore, two test masses released one above the other will tend to move apart from each other when they are released and allowed to fall freely. (The one closer to the centre of mass falls with greater acceleration.) The tendency for freely falling matter to be stretched radially and squeezed tangentially is called the *toothpaste tube* effect. It can be described as a manifestation of *tidal forces*, which are due to differences in gravitational potential from point to point. A gravitational field can be detected by the tidal

---

[6] There are several readings of the Equivalence Principle; one must in particular distinguish between what Einstein himself seems to have had in mind, which he expressed in more than one way during the years in which he developed General Relativity, and the way it is used in modern formulations of the theory. See (Anderson, 1967; Misner, Thorne, & Wheeler, 1973; Norton, 1986).

accelerations it produces (which, again, become vanishingly small within a small enough region of spacetime). Here is the catch: if we want to follow the Equivalence Principle and insist that the test particles are moving inertially, and if we accept that inertial motion follows the shortest paths (the "geodesics") in a geometry, then we are forced to the conclusion that in the presence of gravity the geometry of space (more precisely, spacetime) cannot be Euclidean—for in a Euclidean geometry the inertial paths would be parallel.[7]

Second (and here is yet another thought experiment), Einstein considered a circular disk rotating with uniform angular velocity about its centre. Those portions of the disk not at the centre will be Lorentz-contracted in the tangential direction with respect to the centre; the Euclidean relationship between the circumference and diameter of the circle will therefore fail.

In this way the Equivalence Principle combined with well-established conclusions from special relativity led Einstein to the realization that in order to fulfill his ambition to create a fully relativistic theory of gravitation, he would have to radically alter the geometry of spacetime. It might have been less intellectually risky to give up on the notion of gravitation as a manifestation of inertia, but Einstein boldly grasped the second horn of the dilemma and (under Marcel Grossman's tutelage) taught himself the requisite mathematics—Riemannian geometry and tensor analysis. Thus it was that a theory that was sparked by beautifully intuitive and simple thought experiments quickly acquired mathematical complexity so daunting that in 1913 Max von Laue (who had made a key contribution to general relativity by defining the 10-component stress-energy tensor) wrote of the "extraordinary, in fact inconceivable complexity" of the nascent theory as a reason for rejecting it (quoted in (Gutfreund & Renn, 2015, p. 115)).

The rest is (complicated) history: after a number of false starts, Einstein perfected his field equations of gravitation in late 1915. They received their first experimental confirmation with Eddington's famous eclipse expedition of 1919 that showed Einstein's prediction of the bending of starlight near the limb of the Sun to be correct (to within, at that time, a rather large margin of error). The theory has since then survived every observational test to which it could be subjected. While few experts doubt that general relativity must eventually be replaced with a quantum theory of spacetime, it remains the limit toward which such theories must converge within its very large realm of applicability, just as general relativity itself had to converge to the Newtonian picture where the latter is applicable.

Einstein devoted the larger part of his later research efforts to formulating a unified field theory that would, in principle, provide a geometric picture of all forces in nature (Sauer, 2014). In very simple terms, the idea of the unified field theory was to see whether every sort of force could be inertial. This implies further leaps in the complexity of spacetime geometry. It is clear by now that it is impossible to fully account for the structure and behaviour of elementary particles without taking quantum mechanics into account. Einstein's later attempted unified field theories were all classical—local, continuous, and deterministic—and are now generally

---

[7] For an exceptionally clear and user-friendly explanation of how these thought experiments imply the curvature of space-*time,* see (Norton, 2015a).

considered to be magnificent failures (Pais, 1982, Ch. 17). Perhaps the fact that they were not guided by any clear, intuitive thought experiments had something to do with this.

3.3 Holes in Spacetime

We need to also take a quick look at Einstein's "hole" argument, which played a key role in Einstein's tortuous route to the final form of his field equations for gravitation. There is a large literature on this subject, much of it dealing with issues beyond the scope of this paper. (See (Janssen, 2014; Norton, 2015b; Stachel, 2014).) What we need to do here is, again, see the simple, intuitive picture that guided Einstein—or in this case, almost misguided him.

By 1912 Einstein had nearly succeeded in formulating the gravitational field equations that he would publish three years later. However, he was stymied by two difficulties, one technical, the other conceptual. The technical problem was that the field equations he and Grossman had constructed did not seem to reduce to the Newtonian picture when they should have. This glitch disappeared by the time Einstein arrived at the correct equations, but it led him to question whether it would be possible to find field equations that were generally covariant. He formulated his "hole" argument in order to show that the field equations could not be expected to be generally covariant, but the argument instead helped him to clarify the meaning of the concept of covariance.

General covariance was intended by Einstein to be an extension of the special-relativistic Principle of Relativity to all possible states of relative motion. Philosophically, general covariance is an expression of Einstein's realism, which implies among other things that physical realities are not affected merely by how we choose to describe them. In practice, Einstein thought, this would mean that a mere coordinate transformation should make no difference to the observable *predictions* that we should be able to extract from the theory; therefore, the *form* of the equations of the theory should be preserved by any smooth (continuous and differentiable) mathematical transformation. Transforming from one coordinate system to another—for instance, transforming from Cartesian to polar coordinates—should make no difference to what it is that the coordinates are being used to describe. The methods of differential geometry pioneered by Gauss and Riemann were thus ideal for Einstein's purposes, since they enable one to make a distinction between intrinsic properties of a geometric structure (such as curvature) and extrinsic properties which are purely an artifact of the choice of reference frame or coordinate system.

But would it be possible to write a generally covariant theory of gravitation that satisfied the Equivalence Principle, which implies that the mass-energy structure of spacetime determines the inertial paths of matter? Einstein asked what would happen if there were a *hole in spacetime,* a bounded vacuole containing nothing that could serve as a *source* of the gravitational field; i.e., no matter, energy, fields, or particles whatsoever. If determinism holds, the field inside the hole should be fixed in a unique way by the matter outside it. Presumably we can extend our coordinate grid to cover the hole. Now, apply a smooth transformation to the grid-points inside the hole (but not those outside it). This should transform the gravitational field inside the hole, since the field is a function of the metric structure. By general covariance, both the original and the transformed coordinates should be equally acceptable descriptions of the situation inside the hole. And yet, one seems to have two distinct field structures that are somehow a consequence of the *same* mass-energy distribution. Given the stark choice between abandoning determinism or general covariance, Einstein chose the latter.

Einstein seems to have been briefly satisfied with his hobbled field equations of 1913, but it became increasingly apparent that they were not observationally adequate. Spurred by competition from David Hilbert, Einstein returned to general covariance and in late 1915 arrived at the set of field equations that so far remain by far the best theory of spacetime structure we have. In 1916 Einstein set aside worries about holes in spacetime by arguing that the only things that can actually be observed are coincidences in space and time between "material points":

> All our space-time verifications invariably amount to a determination of space-time coincidences (2015, pp. 187–188).

These are the only things that we actually observe and so our theory should be built out of them; how we paint coordinates onto those observable point-coincidence events should make no difference to the physics we describe. Carlo Rovelli puts it in more modern terms:

> Reality is not made up of particles and fields on a spacetime: it is made up of particles and fields … that can only be localized with respect to one another. No more fields on spacetime: just fields on fields (2004, p. 71).

In short, the best answer to Einstein's "hole" argument is that the notion of an entirely matter-free hole in spacetime has no physical meaning. If a cosmological constant is admitted (as it must be) then even empty space itself has a gossamer "dark energy", but one need not invoke the cosmological constant to see the point. By about 1920 Einstein himself acknowledged (Einstein, 1922) that the old ether of the 19th century is reborn in general relativity as a dynamic substance ultimately indistinguishable from the matter that lives within it. As the distinguished relativist Bryce DeWitt explained,

> General relativity not only restores dynamical properties to empty space but also ascribes to it energy, momentum and angular momentum. In principle, gravitational radiation could be used as a propellant. Since gravitational waves are merely ripples on the curvature of spacetime, an anti-etherist would have to describe a spaceship using this propellant as getting something for nothing—achieving acceleration simply by ejecting one hard vacuum into another. This example is not as absurd as it sounds. It is not difficult to estimate that a star undergoing asymmetric (octopole) collapse may achieve a net velocity change of the order of 100 to 200 km s^{-1} by this means (DeWitt, 1979, p. 681).[8]

Will the warp-drive spacecraft of the future climb to the stars on jets of pure ether?

---

[8] Note that DeWitt's account of space as capable of possessing momentum is in interesting tension with what Einstein said in 1920: "[A]ccording to the general theory of relativity space is endowed with physical qualities; in this sense, therefore, there exists an ether. … But this ether may not be thought of as endowed with the quality characteristic of ponderable media, as consisting of parts which may be tracked through time. The idea of motion may not be applied to it" (1922, pp. 23–24).

4. The Role of Thought Experiments in Einstein's Discoveries

Obviously, Einstein's discoveries were not based purely on thought experiments. By 1905 he was almost certainly aware of the negative result of the real experiment done by Michelson and Morley[9] though he does not mention it by name in "On the Electrodynamics of Moving Bodies" (Einstein, 1905). Rather, he begins his great paper by objecting to the redundancy and arbitrariness in the way that Maxwell's electrodynamics was "usually understood"[10] at the time. In particular, he disliked the fact that the induction of a current in a conductor by a magnetic field was described in different ways depending upon whether the conductor or the magnet were presumed to be at rest, even though the observable effects depend only upon the relative motion of the two. (As several authors have pointed out, this simple consideration in itself constitutes an elegant thought experiment.) Einstein was guided by his acute sense of logical economy as much as his physical intuition. But of course, one should not under-rate Einstein's "muscles of intuition"[11] when he was in the "prime of his age for invention."[12] Einstein trained those muscles during his years in the patent office, where he was constantly required to analyse how alleged inventions were actually supposed to work.[13] As Peter Galison documents (2003), many of these inventions were concerned with the measurement of time or the determination of synchrony.

The role of many of Einstein's thought experiments such as the light beam chase in the formation of his theories seems to have been essentially suggestive, not logical. A comparison can be made with August Kekulé's apocryphal vision of a snake swallowing its tail, which he claimed led him to grasp that benzene has a ring structure.[14] Just as Kekulé envisioned a physical structure, thought experiments such as the light beam chase lead us to picture an argument structure; that is, they don't always constitute an argument, but they suggest one. Consider Einstein's "happiest thought"; in itself, it was simply an observation that a person falling freely in a uniform gravitational field would not feel a force due to gravitation. Much of the impact of this observation for Einstein, from his personal accounts of the event, seems to have been its *affect*; seeing the point of the experiment (that gravitation is remarkably *like* an inertial force) was like grasping the punch-line of a good joke. (Einstein was known for his often-raffish sense of humour.) There is a sense of pleasant surprise, of immediate certainty, like Martin Gardner's *aha!* moment when you *get* the trick that solves a problem (1978). A good thought experiment has the effect of an instantaneous paradigm shift: staring at the duck for a long time, you finally

---

[9] Pais (1982, Ch. 6) marshals evidence to this effect.

[10] Quotation from (Stachel, 2005, p. 123)

[11] Keynes' famous phrase (Keynes, 1978) applies to the young Einstein as much as Newton.

[12] This is adapted from the phrase that Newton applied to himself during that fertile period when he created the differential calculus and laid down the elements of his mechanics (S. I. Newton, 1888).

[13] A. Fölsing: "…for young Albert Einstein, examining patents was more than just a livelihood… His virtuosity with 'mental experiments' was not all that far removed from intellectual penetration of an invention…" (1997, p. 103).

[14] See (Martin, 1997, pp. 165–167). Kekulé proclaimed (perhaps with tongue in cheek), "Let us learn to dream, gentlemen, then perhaps we shall find the truth."

see the rabbit (Kuhn, 1970). It is, in itself, not an argument but a perception that suggests an argument. Almost everyone has fallen or jumped off of something at some time; only Einstein noticed the implications. Pasteur famously said that chance favours the prepared mind; so, also, do thought experiments.

5. Thought Experiments in Quantum Mechanics

With the advent of quantum mechanics, physics moved away from the intuitive and the visualizable. Nevertheless, some notable thought experiments played key roles in the growth of quantum theory. There is not space to analyse all of them here in detail. I'll make just a few comments about the virtual oscillators that Planck used to justify his blackbody radiation law, another elegant model by Einstein which he used in 1909 to demonstrate the wave-particle duality of light, and Heisenberg's microscope and the role it did or did not play in his derivation of the Uncertainty Relations in 1927. I'll have a lot more to say about the Einstein-Podolsky-Rosen (EPR) thought experiment and the lead-up to it.

5.1 Planck's Resonators

In late 1900 Max Planck finally accepted that he had to adopt Ludwig Boltzmann's discrete counting methods in order to find a derivation for the radiation formula that he had stumbled upon by a combination of interpolation and inspired guesswork a few weeks earlier. But he needed the most general possible model he could devise for the interaction of matter and the radiation field. In 1900, of course, virtually nothing was known of the detailed structure of matter; furthermore, the whole point of the calculation was to arrive at a formula that would be valid for any kind of material whatsoever (for the only defining characteristic of a perfect black body is that it absorbs all the radiation that hits it). So Planck imagined that radiation exchanges energy with "virtual" harmonic oscillators or resonators in the walls of the radiation cavity. Planck's virtual oscillators were "spherical chickens," shorn of all detail except the physically plausible assumption that they could somehow come into resonance with incident radiation. But the model was not enough; in order to get the right formula, Planck had to commit what he called an "act of desperation" (Stone, 2013, p. 59) and assume that radiation exchanged energy with the resonators in discrete chunks of magnitude $E=h\nu$, where ν is the frequency of the incident light and h is a new constant of nature.[15] Thus was quantum mechanics born, though it would take over twenty years more (and much hard work by the "valiant Swabian" and others) before it was widely accepted that the radiation field itself is quantized.

Should Planck's story about his virtual resonators count as a thought experiment, or just as an apt model on which a calculation could be hung? Models can be of specialized systems of narrow interest, while a game-changing thought experiment focusses attention on a key feature

---

[15] John Norton (private communication): "The amazing result in thermodynamics is that if systems A and B are in equilibrium, all that matters for A about B is B's temperature. … Whatever properties the matter may have, the equilibrium state of the radiation will be the same." So Planck needed only the sketchiest picture of matter, combined with the assumption of quantization, to get his result.

of wide generality and suggests what calculations must be done in a large class of cases. Planck used his model to ask, what are the most general features of the way in which radiation must interact with any conceivable kind of matter? It was simply mathematics that forced Planck to introduce energy quantization, because that was the only way he could get the right answer. Quantization for Planck was simply a formal step, for which he could not see any independent physical justification. The defining features of Planck's model were not by themselves sufficient to point to the solution of the problem he set it, but the simplicity and generality of the model clarified the question to the extent that it left Planck only one mathematical option. For this reason I'm happy to count Planck's resonator model as one of the most consequential thought experiments of the new physics.

5.2 Of Fluctuations and Mirrors

Now we turn to another ingenious scenario by Einstein which certainly does count as a thought experiment of the first order (although it requires a good knowledge of statistical mechanics to fully appreciate it). In 1909 Einstein imagined a mirror inside a Planckian cavity, able to travel back and forth freely on a rail perpendicular to its face (Einstein, 1909; Pais, 1982, pp. 408–409; Stone, 2013, pp. 136–140). In the cavity is a quantity of ideal gas and radiation, all in equilibrium at a definite temperature. What we do next is Einsteinian simplicity at its best: ask what is required for the mirror to be in equilibrium with the gas and radiation—for it must be. But this thought experiment is no mere dazzling *aperçu*. Einstein applied his mastery of the statistics of fluctuations, which he had exhibited in his work in 1905 on Brownian motion, and derived a key formula for the radiation fluctuations that the mirror must experience in order to remain in equilibrium. The mirror is subject to pressure fluctuations both from the gas and from the radiation, and the gas and radiation fluctuations must be in equilibrium with each other. The new thing Einstein did was to use Planck's blackbody energy distribution formula to compute the fluctuations due to radiation. Planck gives the energy distribution at the given temperature in terms of frequency, which shows the energy flux to which the mirror is subject. Then the crowning touch: Einstein imagined the mirror to be transparent to all frequencies except a narrow band; using Planck's Law for the energy density in that band, he arrived at an expression for the root-mean square fluctuations in the position of the mirror due to what he called "radiative friction." The key result was that this expression contains two terms. One represents wave-like fluctuations caused by constructive and destructive interference, due to small variations in phase, polarization, and frequency. (In modern terms, the light in the cavity would be said to be incoherent.) The other has the form of fluctuations due to impacts from discrete bundles of energy of magnitude $h\nu$. At low radiation densities, the particle-like fluctuations strongly dominate. Einstein took this result, which is a direct consequence of Planck's well-verified distribution law, as good evidence for his view that "the next state of theoretical physics will bring us a theory of light that can be understood as a kind of fusion of the wave and emission [particle] theories of light" (in (Stone, 2013, p. 137)).

5.3 Heisenberg's Microscope

This is not the place to recount the whole story of how modern quantum mechanics burst on the scene from 1925 to 1927. (For a short version see (Peacock, 2008); for exhaustive detail, see (Mehra & Rechenberg, 1982).) By 1927, it was clear that the new quantum theory challenged classical intuitions in a number of related ways. Schrödinger had created wave mechanics in 1926, hoping that it would give a realistic, classical underpinning to quantum statistics. And yet, it soon became apparent that QM only gives us probabilities, which are calculated indirectly from the (complex-valued) wave function by means of the Born Rule, $P(x) = |\Upsilon(x)^2|$. What, then, did $\Upsilon(x)$, which Einstein sarcastically dubbed the *Gerspensterfelder* (ghost field), actually represent? The one thing that Schrödinger's theory did not do was get rid of what he called the "damned quantum jumps" (Stone, 2013, p. 268); rather, it simply gave a remarkably efficient set of algorithms for using the "ghost field" to calculate the probabilities that those jumps would occur. Why these algorithms work so well remained (and remains) a mystery.

Another profound mystery was the appearance of non-factorability (or non-factorizability, as it is sometimes more awkwardly called). As soon as two or more particles undergo some sort of dynamical interaction, the wave function for the combined multiparticle system has cross-terms which imply much stronger statistical interdependencies between the particles than seemed to be possible. These cross-terms are in general algebraically irreducible (except for the special and limited case of so-called *product states*); once the particles have interacted their observable properties remain closely correlated (or anti-correlated) even when the particles have separated to arbitrary distances. Schrödinger (1935b) coined the term "entanglement" (from the German, *Verschränkung*) to describe this mysterious interdependency of non-factorable systems, and famously stated,

> When two systems, of which we know the states by their respective representatives [wave functions], enter into a temporary physical interaction … and when after a time the systems separate again, then they can no longer be described … by endowing each of them with a representative of its own. I would call that not *one* but rather *the* characteristic trait of quantum mechanics. (1935, p. 555)

A remarkable property of entangled states is that, as Schrödinger suggested, the individual particles in an entangled state cannot be represented as pure states, only as mixtures (classical ensembles of quantum states). They are not "things in themselves"! Einstein hated this feature of quantum mechanics, and he was convinced that it marked a fundamental flaw in the theory.

In 1927 Einstein attempted to construct his own version of wave mechanics (Howard, 2007). His aim was to produce a wave mechanics without the pesky cross-terms, and he failed because it cannot be done. The short paper he produced was presented at a meeting but remained unpublished (Einstein, 1927). It was his last attempt to make a constructive contribution to quantum mechanics—and definitely not up to his usual standard.[16]

In the same year Werner Heisenberg introduced his indeterminacy relations. Heisenberg was a complex and contradictory character. His ethically dubious participation in the Nazi

---

[16] See Peter Holland (2005) for a detailed analysis of Einstein's abortive 1927 wave mechanics.

atomic project during WWII (Rose, 1998) sadly tarnishes the brilliance of his contributions to theoretical physics in his dazzling youth, when he laid down the essential principles of modern quantum mechanics at the age of 23.

The central epistemological problem that quantum mechanics will not permit us to ignore is that it is impossible to observe and measure the properties of a particle without physically interacting with it. Einstein based special relativity on operational definitions of quantities such as position and time, constructed so that it would be possible to clearly distinguish between these apparent properties of objects that are partially due to the way they are observed and those that are intrinsic to the objects. Quantum mechanics forces us to ask whether the notion of intrinsic properties has any physical meaning at all.

To investigate this problem, Heisenberg imagined a microscope designed to detect the position and velocity of an electron by scattering gamma rays off it (Heisenberg, 1927, 1930). By the wave-particle duality both electrons and gamma rays have both a wave and a particle aspect. It is necessary to use high-energy gamma rays since there is an inverse relationship between wavelength and energy: the lower the wavelength the higher the resolving power of a microscope, so that only high energy electromagnetic radiation has a short enough wavelength to detect an electron within any reasonable range of error. The key idea of the thought experiment was to apply basic laws of optics to show that the more accurately one could resolve the momentum of the electron, the less accurately one could resolve its position, and *vice versa*. Heisenberg arrived at the now-familiar inequality,

$$\Delta x \Delta p_x \geq h,$$

where the deltas are the uncertainties in position x and momentum p_x, and h is Planck's constant of action. We can't measure one of the deltas with full precision without rendering the other completely indefinite.

Are the indeterminacy relations merely epistemic, a reflection of our *practical* inability to know the precise values of quantities that are, in fact, pre-existent? The microscope experiment itself suggests that these endemic uncertainties are merely the product of the fact that we cannot avoid using very short wavelength radiation to "see" an electron; because of the inverse relation between wavelength and energy, the more accurately we want to detect the position of the target particle, the more we must change its momentum. This epistemic reading of the thought experiment seems to leave open the possibility that the electron may still have well-defined values of both position and momentum even if we can never hope to simultaneously measure them. The problem, one might think, is only that there is no procedure in which we could reduce our "jiggling" of the observed system to zero.

Heisenberg realized that there are both mathematical and philosophical grounds for rejecting this naïve interpretation of the indeterminacy relations. Mathematically, observable quantities come in *conjugate pairs* defined by commutation relations which show precisely the extent to which the commutative law fails for the linear operators representing those observables. Shortly after Heisenberg's publication of the indeterminacy rules, Schrödinger and others showed that they are simply a mathematical consequence of the commutation relations between position and momentum (Cohen-Tannoudji, Diu, & Laloë, 1977, pp. 286–287). According to the mathematical formalism that emerged in the years 1926–30, this inability to simultaneously tie down both values of a conjugate pair is not merely due to unavoidable experimental clumsiness due to the finitude of the quantum of action. Rather, in the mathematics of wave mechanics, asking for simultaneous, exact values of position and momentum is a mathematical contradiction in terms, like asking for a square circle (because position and momentum are Fourier transforms of each other). Philosophically, Niels Bohr argued that this was a manifestation of what he called *complementarity*: the types of experiments in which one can measure position are simply incompatible with the types of experiments in which one can measure momentum. We need both wave and particle viewpoints to fully describe physics, and yet at the quantum level these two modalities cannot be applied simultaneously—where "simultaneously" does not necessarily mean "at the same time coordinate" but "in the same procedure". Thus, the microscope thought experiment is not a way of rigorously deriving the indeterminacy rules from the formal principles of quantum theory, but rather a highly suggestive semi-classical approximation.[17]

5.4 Einstein Challenges Quantum Mechanics

Einstein remained stubbornly convinced that the properties of physical systems could not depend upon the types of measurements we choose to perform on them, especially if those properties could be inferred from measurements performed at a distance. He referred to Bohrian complementarity as the "tranquillizing philosophy—or religion?" which, he said, "is so delicately contrived that, for the time being, it provides a gentle pillow for the true believer from which he cannot very easily be aroused" (Fine, 1986, p. 19).

In 1927–30, Einstein brought the full force of his ingenuity to bear on quantum mechanics and the indeterminacy relations in particular, trying to devise thought experiments in which they could be shown to fail. For once his inventiveness failed him; Bohr and others were always able to find loophole in Einstein's arguments that saved Heisenberg. (See Bohr et al., "The Bohr-Einstein Dialogue," in (Wheeler & Zurek, 1983, pp. 1–50).)

By 1935 Einstein had given up on trying to beat the indeterminacy relations and tried a much subtler approach. In that year he and two younger collaborators, Boris Podolsky and Nathan Rosen, published a short, difficult paper (Einstein et al., 1935) in which they outlined an enigmatic thought experiment aimed at showing that quantum mechanics cannot provide a complete description of the entities for which it purports to account. The Einstein-Podolsky-Rosen (EPR) *gedankenexperiment* has evolved from a hypothetical scenario to become a defining paradigm of modern quantum mechanics.[18]

---

[17] Heisenberg's rules are also often called the uncertainty relations. However, as Richard Arthur (private communication) has pointed out to me, the latter term is potentially misleading since it suggests that the Heisenberg rules merely represent epistemic uncertainty about quantities that do, in fact, have definite values.

[18] There is not space in this paper to consider yet another thought experiment of Einstein's, the box experiment, which makes essentially the same point as the EPR experiment. See (Norsen, 2005).

We have already mentioned two features of quantum mechanics that were the most troubling to Einstein, noncommutativity and the indeterminacy relations that follow from them. Non-commutativity plays an important role in the EPR experiment. However, as Don Howard says (2007), "[E]ntanglement, not indeterminacy, was the chief source of Einstein's misgivings about quantum mechanics… Indeterminacy was but a symptom; entanglement was the underlying disease."

The aim of the EPR paper was not to show that QM is incorrect, but rather that it is *incomplete* in the sense that it does not give a description of every "element of physical reality" ("EPR" again) belonging to an entangled state. By this time, Einstein had decided that the indeterminacy relations of Heisenberg point to nothing more than the incompleteness of quantum mechanics itself (incompleteness in the sense that the theory fails to represent properties that particles presumably *do* have). Part of the aim of the EPR paper was to make this notion of completeness precise. A necessary condition for completeness, according to EPR, is that "every element of the physical reality must have a counterpart in the physical theory" (777). And their sufficient condition for reality was this: "If, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a quantity, then there exists an element of physical reality corresponding to this physical quantity" (777). The argumentative strategy for the paper would be to use the *sufficiency* condition for reality to show that quantum mechanics fails the *necessary* condition for completeness.

Here is the gist of the argument. Consider a composite system $U + V$ comprised of two particles which interacted dynamically at one point and then separated far from each other in space. Because they interacted they will possess quantities (such as total momentum or difference in position) that must be conserved globally. EPR take pains to show that these quantities commute on the system as a whole. This fact is crucial to the argument, since the mutually consistent global conservation requirements give us a basis for comparing the results of apparently incompatible measurement procedures on the individual particles.

Now, measure (say) position on *U* at time *t*. This collapses the entangled state into a product state with *V* in an eigenstate of position, allowing us to predict its position with certainty. (As Schrödinger put it, this "steers" the *V*-system into a definite state.) But we could have also measured momentum on *U* at time *t*, and this would collapse the entangled state into a product with *V* in an eigenstate of momentum at time *t*, allowing us to predict its momentum with certainty. We can't measure both position and momentum on *U* in the same procedure, but we are entirely at liberty at time *t* to choose which of the two types of procedures to apply. Thus, measurements we can perform on *U* enable us to predict presumably non-commuting properties of *V* with certainty. Therefore, there seems to be only two possibilities: either *V* was *already* in definite states of both position and momentum (despite Heisenberg), or our choice of measurement strategy on *U* at time *t* spookily influenced the state of *V*—at time *t*! But EPR say (780), "no reasonable definition of reality could be expected to permit this," precisely because *U* and *V* are spatially distant at time *t*. Therefore, quantum mechanics must be incomplete since *by its own admission* it cannot represent properties that system *V must have already had* at time *t*.

The EPR argument thus establishes (validly) a disjunction: either QM is incomplete (in the sense they specify) or there is spooky action (or perhaps both!).

EPR illustrated their argument by applying it in detail to an entangled wave function which has some very interesting properties which I won't attempt to analyze here except to note that there seems to be no practical way to prepare particles in this particular state; thus the EPR experiment, as they described it in 1935, is a pure *Gedankenexperiment*. Later, practicable versions of the experiment would be defined.

Niels Bohr rushed to publish a response (1935). He agreed with EPR that "of course there is in a case like that just considered there is no question of a mechanical disturbance of the system under investigation during the last critical stage of the measuring procedure" (699). However, one cannot hope for a single complete description of the system. Rather, there are complementary descriptions of the system; in one, we can infer the momentum of *V* from momentum measurements on *U*, and in the other we can infer the position of *V* from position measurements on *U*. But from these facts we cannot infer a pre-existent reality in which *V* possessed sharp values of both position and momentum, for one cannot infer sharp values of both position and momentum for particle *V* with a *single* procedure. To ask for anything else would be to ask a foolish question, as Feshbach and Weisskopf put it (1988): "If you ask an inappropriate question, you get a probability distribution as a response." Thus, Bohr's cryptic and complicated response to some extent does clarify what the quantum mechanics of 1935 actually *says* about the EPR scenario, but it is very difficult to avoid the suspicion that a deeper level of analysis is possible—even if it would not be precisely what either EPR or Bohr themselves likely had in mind.[19] Indeed, J. S. Bell would later show that this is the case.

It is often said that EPR argued for a "hidden variable" or "hidden parameter" account of QM (although they did not use those terms themselves). Einstein in later years stated that he thought that there should be an *ensemble* interpretation of quantum mechanics (1969, p. 668), which he apparently understood as a probability distribution over possible *local* states of the particles. The idea would be that the particles have some sort of complex internal coding (perhaps still far beyond the ken of our present physical theories) which is capable of telling them how to behave in order to obey the observed predictions of quantum mechanics, for all of the possible experimental questions they could be asked.

This notion of hidden variables can be compared to heredity. Why do siblings resemble each other? Not because of anything "spooky," but because of a common *local* cause, namely the shared DNA they got from their parents. Although the role of DNA in heredity was not understood in 1935, EPR in effect implied (although they did not explicitly state) that there has to be some sort of "quantum DNA" encoded in the entangled particles when they interacted at their common source, sufficient to explain how the particles react when they are measured. In 1935 there was no obvious way to test this proposal. And so there matters stayed until after WWII.

---

[19] See (Howard, 2007) for an insightful analysis of what was at stake between Bohr and EPR.

The EPR thought experiment is widely misunderstood. Its import was not to demonstrate the existence of entanglement, but rather to use entanglement as evidence that quantum mechanics could not be telling the whole story about the structure of particles. The non-factorability of wave functions for multi-particle systems had been well-known as early as 1927 (Howard, 2007), although Schrödinger did not introduce the term "entanglement" until 1935. Einstein himself had grasped that light quanta were suspiciously too-well correlated from the time of his pre-1910 pioneering papers on the statistics of electromagnetic radiation. The EPR paper did have the valuable effect of *drawing attention* to the phenomenon of entanglement in a way that its challenge to conventional notions of locality and causality could not be ignored. *Via* the work of Bell (shortly to be described), the EPR thought experiment was one of the key sparks for the modern flowering of quantum information theory. Thus, even though the paper did not accomplish what its authors hoped it would accomplish, it has proven to be one of the most provocative and unexpectedly fruitful of Einstein's great papers.

5.5 How the EPR Debate Might Have Gone

It seems hard to imagine that anything useful could be added to the reams of analysis of the EPR paper that have already appeared, and yet one essential question has received very little attention. Is it actually the case that there is no dynamical interaction between the two particles "during the last critical stage of the measuring procedure"? Both Bohr and Einstein themselves thought that this question did not even merit discussion, and numerous papers published since their time purporting to demonstrate "signal locality" in entangled states simply follow their lead and assume without argument that entangled systems are dynamically local.[20] The bald assumption that spacelike separation guarantees dynamical independence (Einstein's *Trennungsprinzip*, his Separation Principle) (Einstein, 1948) seems so utterly reasonable to most authors that very few have thought seriously to question it. And yet, whether or not the separation principle is correct is one of the first questions that should have been examined, not the last.

Had Einstein and other protagonists in the mid-1930s noticed a short remark by Pauli, published in 1933, the debate over the EPR scenario could have taken a different direction.[21] In a review article on quantum mechanics Pauli includes a discussion of many-particle systems. He notes that when there is no mutual interaction between the particles the system is represented by a wave function which is simply the product of individual wave functions belonging to the individual particles. Pauli seems to suggest that in order for such a wave function to be a solution of the Schrödinger Equation for the system, the algebraic structure of the Hamiltonian (energy operator) must be parallel to the algebraic structure of the state function. For a product state, the total Hamiltonian must be additive (i.e., simply the sum of the Hamiltonians for the individual particles):

---

[20] E.g., (Ghirardi, Rimini, & Weber, 1980; Shimony, 1983). For critiques of the orthodox approach to signal locality, see (Kennedy, 1995; Mittelstaedt, 1998; Peacock, 1992).

[21] My attention was drawn to this important passage in Pauli's book by an unpublished presentation by Don Howard (2006).

> An additive decomposition of the Hamiltonian into independent summands corresponds [*entsprich*] therefore, to a product decomposition of the wave function into independent factors.[22]

Pauli does not directly comment on the Hamiltonian structure for non-product (non-factorable) states, but this pregnant remark raises questions that should have been obvious even in 1935. Does Pauli's *entsprich* mean that a state is a product state *if and only if* its Hamiltonian is additive? If so, then many-particle systems represented by *non-factorable* wave functions (such as the special entangled wave-function used as an example by EPR) would have to have *non-additive* Hamiltonians with algebraically irreducible cross-terms. If that is the case, then for any entangled state there must in general be eigenvalues (energy states) of the total system Hamiltonian that are not simply sums of local energy states for the individual particles; let's call this the *energy of entanglement*. Such nonlocal energy eigenstates would be properties of the entangled system as a whole, rather like the way in which the energy of an atomic orbital is a property of the orbital as a whole and cannot be spoken of as localized to the electrons associated with the orbital. It is unclear whether the existence of such non-localized energies can be sensibly described as implying any sort of "action" at a distance, but their existence would be a clear challenge to Einstein's dynamically-local realism because they it would mean that the particles in an entangled state are dynamically entailed no matter how far apart they are.

Nothing I've said here is meant to suggest that Pauli himself would have been an advocate of such a flagrant challenge to relativistic orthodoxy. One can surmise from his own dismissive remarks about the EPR paper[23] that he did not think that the EPR argument merited close analysis. Whatever Pauli may have thought, however, there is increasing evidence, both theoretical and experimental, for dynamical nonlocality in entangled states. This is not the place to review that evidence in detail; suffice to say that the question of dynamic nonseparability in entangled states—and thus the ultimate interpretation of the EPR thought experiment—remains open.[24]

Could it be, then, that both horns of the dilemma offered by EPR must be grasped? Is quantum mechanics both endemically incomplete and dynamically nonlocal? It is the case that, precisely as Pauli indicated in 1933, the one implies the other—something that perhaps should have been obvious a very long time ago? To answer questions like these a systematic study of the dynamics of the EPR and other entangled states needs to be carried out, a task that remains to be done. It is notable that entangled (non-additive) Hamiltonians are commonplace in the

---

[22] (Pauli, 1933); quotation from English translation, (Pauli, 1980).

[23] "Einstein has once again expressed himself publicly on quantum mechanics … every time that happens it is a catastrophe" (Aczel, 2002, p. 117).

[24] Here in brief are two of the many pieces of evidence for nonlocal dynamics to be considered. Theoretical: although many presentations of "Bohmian mechanics" carefully guard themselves from putting it this way, Bohm's quantum potential, which is implicit in the mathematics of wave mechanics, is a manifestly nonlocal contribution to the total energy of multiparticle systems (Bohm, 1952a). Observational: the recent and very important experiment of Lee et al. (2011) *prima facie* seems to show that two physically distant chips of diamond can be put into the same phonon energy state.

literature on quantum information; see, e.g., (Dür, Vidal, Linden, & Popescu, 2001). No one seems yet to have studied the structure of the Hamiltonian for the special wave function used by Einstein, Podolsky, and Rosen in 1935, although this would be of great foundational interest.

I'll conclude the discussion of the EPR thought experiment by making a brief observation that some may find outrageous. Modern cosmology is built upon the Big Bang model, according to which the universe that we observe expanded rapidly (and indeed is still expanding at an increasing pace) from a highly compressed state some 14 billion years ago. (I need not review the very strong evidence for this picture here; see (Kirshner, 2002).) But these are precisely the conditions that lead to quantum entanglement: particles at close quarters interact dynamically and remain statistically entailed thereafter even when they separate in space and time. It would seem that the entire universe is, in effect, a vast EPR apparatus. And if so, what are its dynamics? Is dark energy the entanglement energy of the universe?

5.6 Schrödinger's 'Hellish Device'

Shortly after the EPR paper appeared, Schrödinger published his famous "cat" paradox which was intended to expose the contradictions inherent in quantum mechanics (Schrödinger, 1935a, 1983). The key idea of the thought experiment was in part due, again, to Einstein, and was hammered out in an extensive correspondence between the two physicists in 1935. (For detailed analysis see Fine, (1986).) Einstein's first version of the experiment was entirely at the macroscopic level:

> The system is a substance in chemically unstable equilibrium, perhaps a charge of gunpowder that … can spontaneously combust, and where the average lifespan of the whole setup is a year. … In the beginning, the ψ-function characterizes a reasonably well-defined macroscopic state. But, according to your equation, after the course of a year … the ψ-function then describes a sort of blend of not-yet and of already-exploded systems … in reality there is just no intermediary between exploded and non-exploded. (Fine, 1986, p. 78)

Schrödinger soon hit upon the idea of coupling a macroscopic system (an unfortunate cat) to a quantum-level system. A bit of radioactive material has a half-life of an hour; if it decays the resulting alpha-particle triggers the release of a deadly poison that instantly kills the cat. If this cruel apparatus is described in the language of quantum mechanics, the wave function for the radioactive atoms is entangled with the wave function for the cat, and the cat is presumably in a superposition of states, either alive or dead—until we open the box to see what has happened to it. Schrödinger's version of the experiment is in one respect cleverer than Einstein's, since it shows that macrostates can be infected with quantum uncertainty if they are coupled to microstates in just the right way. Since the idea of the cat being in a superposition of states is presumably absurd, and since the dividing line between the quantum and the macroscopic (classical) level is arbitrary, the idea of quantum systems being in superpositions must also be absurd. Einstein's simpler version of the experiment has the conceptual advantage that it exposes the contradictions that seem to follow from assuming that quantum mechanics (which after all is advertised as a universal theory) is applied to ordinary macrosystems.

A catalogue of responses to the cat paradox is equivalent to a catalogue of proposed solutions to the measurement problem, which is to show how it is that measurements on superposed quantum systems can apparently produce definite, classical results. I can't do justice to this literature here and I will only make a few general observations.

Einstein and Schrödinger's point can be seen to follow from the EPR sufficient condition for reality. Since presumably opening the box presumably does not disturb the system (here we run into a problem similar to Maxwell's demon), and since we know that when we open the box the cat will definitely be in an alive exclusive-or dead state, it must have been in precisely that state before the box is opened. (That is, if we can think of "alive" as a state, then we can think of "alive XOR dead" as a state.) Thus, one way of challenging Schrödinger's thought experiment is to challenge EPR's reality condition by pointing out that while in quantum mechanics the probability may be unity that a system will be found to be in a certain eigenstate when subject to a certain measurement procedure, that fact does not entitle us to say that the system was in that state before it was measured.

Another obvious problem with the cat experiment is that a real feline, alive or dead, is a complex macrosystem comprised of an enormous number of particles in an incoherent state. In quantum mechanical language it can only be described usefully as a mixture which will behave classically (to an extremely good approximation) even if it is coupled to a quantum-coherent system. To this extent the cat was a poor example, even though it drew attention to the problem in an almost poignant way. However, it is now possible to create "Schrödinger cat states"—macroscopic coherent states (for there is in principle no limit to the size of a quantum coherent state) that can be in a superposition (Yam, 2012).

There are at least two lessons to take from the Einstein-Schrödinger cat. First, it is still not fully understood how quantum processes lead to definite or apparently definite results at the macroscale (this is the measurement problem). Second, the thought experiment emphasizes the key fact that the non-classical features of quantum mechanics cannot be safely sequestered to the micro-level. (Here we have another case of an important thought experiment that does not so much provide a conclusive argument as it memorably draws attention to a problem.) As an illustration of this point, another important thing that happened in physics in the 1930s was the discovery of superfluidity and superconductivity, macro-scale phenomena that are entirely manifestations of quantum statistics. If cost were no barrier, it would be possible to create an Olympic swimming pool full of superfluid helium—and recent work in observational cosmology shows that the entire universe is a Planckian cavity (Smoot & Davidson, 1993). If physics is quantum all the way down, it is also quantum all the way up.

6. After Einstein

6.1 Neglected Potential

In 1951 the young American physicist David Bohm published an illuminating analysis of the EPR thought experiment in his text on quantum theory (1951). Bohm reformulated EPR's experiment in terms of spin observables. His version of the experiment had the great virtue that it could in principle be performed, opening up the possibility of an experimental test of the "quantum DNA" hypothesis.

Bohm then created a whole new version of (non-relativistic) QM, based on the quantum potential, a nonlocal potential field which is a function of the shape of the envelope of the wave packet (and thus, in effect, of the phase relationships within the wave function) (Bohm, 1952a, 1952b; Cushing, 1994). Bohm's "interpretation" successfully reproduces the predictions of ordinary quantum mechanics and resolves some challenges (regarding scattering) to a similar theory that had been proposed by de Broglie in 1927 (see discussion in (Cushing, 1994)). His quantum potential contributes to the energy of a composite quantum system as a whole; in general it is distance-independent and it can't be localized to individual particles. Most physicists were horrified; J. R. Oppenheimer (disgracefully) said, "If we cannot refute Bohm, we must choose to ignore him" (Peat, 1997, p. 133).[25] All horror aside, Bohm had apparently done what John von Neumann had argued could not be done (1955), which was to construct a hidden variable theory that apparently could underpin quantum statistics—although in a way that is explicitly nonlocal.

6.2 "The Most Profound Discovery of Science"

John Stewart Bell, who evidently was unworried about whether his career would be irrevocably damaged if he were known to have read Bohm's papers, noted that Bohm had done what von Neumann had claimed was impossible—namely, constructed a hidden variable underpinning for quantum statistics. As noted, Bohm's approach is explicitly nonlocal, and Bell set out to determine whether *any* completion of QM had to be nonlocal.

Bell used Bohm's version of the EPR experiment, and considered correlations between spin measurements taken on the entangled particles (Bell, 1964). He took the novel step of considering measurements taken in different directions (which allowed comparison between different spin components), and he showed that if there were local hidden variables—"quantum DNA"—then the correlations must obey certain mathematical inequalities. Bell then showed that according to quantum mechanics, the expectation values for these correlations violates the inequalities for a wide range of relative detector angles—they can be more strongly correlated (or anti-correlated) than quantum DNA would allow for. By the 1980s (Aspect, Dalibard, & Roger, 1982), experiments showed that Bell's Theorem (the statement that QM violates "local realism") is almost certainly correct, and recent results are closing the last conceivable loopholes (Miller, 2016). There can be no such thing as quantum DNA! Bell's discovery was called (by H. P. Stapp) the "most profound discovery of science" (1975), and well it might be. It would have

---

[25] Bohm was *persona non grata* not only because of his unorthodox physics but also due to his refusal to testify against his friends who were suspected of left-wing sympathies (Peat, 1997). I have heard from a reliable witness that at Princeton in the 1960s it was a career-ender to *mention* Bohm's name.

been fitting if Bohm and Bell had shared the Nobel Prize sometime in the 1980s—but the academic community did not quite have the courage to make such a radical move.

The experimental devices that have been used to test Bell's Theorem are real-world versions of the hypothetical EPR apparatus. Particles are emitted from a source and sent to remote locations where they interact with measurement devices such as polarizers or Stern-Gerlach devices (which detect spin). A key feature of these modern EPR apparatuses is that they employ *delayed-choice*: the decision about which parameter of the particles to measure (such as spin in various directions) is made (automatically, of course, by a randomizing process) *after* the particles are emitted. The timing is thus such that it would be impossible for information about the detector choice on one side of the apparatus to be transmitted to the detector or particle on the other side at any speed less than or equal to the speed of light. If Einstein's separation principle is correct, then the distant particles should exhibit no stronger correlations than those that could be built into them at the source (by the fact, for instance, that their total spins must add up in certain definite ways). But in fact, the Bell-EPR correlations violate the expectations of separability. Does this mean that there really is "spooky action"? The debate continues.[26]

Bell's Theorem is a special case of a more general result, the Kochen-Specker (KS) Theorem (Bub, 1997; Kochen & Specker, 1967; Redhead, 1987): Quantum statistics cannot in general be under-pinned by a Boolean property distribution. The notion of a Boolean structure can be defined precisely in terms of lattice theory (Bub, 1997) but it can be grasped intuitively by thinking of every possible measurement on a quantum system as asking the system a question (which can always be formulated so as to yield a yes or no answer). (For instance, "is your spin-*x* up?") Classical physics presumed that it would always be possible (in principle) to ask such questions in a non-invasive way and that the amount of information to be gathered by asking more and more questions would monotonically increase. But for a quantum system, the list of possible experimental questions must include questions about non-commuting observables. The Bell-Kochen-Specker results states that if we could answer every possible experimental question that could be asked of a quantum system, the set of answers would be logically inconsistent. Bell's Theorem is essentially a special case of the more general Kochen-Specker result, applied to a spatially extended system. As Demopoulos (2004) emphasizes, descriptions of quantum systems are *incompleteable* because the presumption of completeability entails a mathematical contradiction.

This had been anticipated by Schrödinger in 1935:

> … if I wish to ascribe to the model [of a quantum mechanical oscillator] at each moment a definite (merely not known exactly to me) state, or (which is the same) to *all* determining parts definite (merely not known exactly to me) numerical values, then there is no supposition as to these numerical values *to be imagined* that would not conflict with

---

[26] Current orthodoxy states that because of Bell's Theorem, quantum mechanics violates "kinematic" locality but not "dynamic" locality; that is, orthodoxy holds that the dynamics of entangled particles is still local (additive) despite the endemic violation of Bell's Inequalities in a wide variety of entangled systems. In my view this position is hopelessly inconsistent, but this question is beyond the scope of the present paper.

> some portion of quantum theoretical assertions ((Schrödinger, 1935a); trans. J. D. Trimmer, (1983)).

Fitting quantum mechanical predictions to a Boolean substrate is like trying to smooth out a carpet molded to the surface of a sphere onto a flat floor. There will be a lump! We can move it around and even hide it under furniture, but we can't make it go away. Thus, it is not entirely accurate to call "no-go" results such as Bell's Theorem "no hidden variable" theorems; more accurately, they are *no Boolean variable* theories. Even more precisely, they are *not enough Boolean variable theorems*, since non-Boolean quantum systems can have Boolean subspaces defined by complete sets of commuting observables.

Bell's Theorem is still not well-understood, even in the professional community. Here is Nobel-winner Frank Wilczek on entanglement:

> Measuring the spin of the first qubit tells you about the result you'll get by measuring the second bit, even though they might be physically separated by a large distance. On the face of it, this "spooky action at a distance" to use Einstein's phrase, seems capable of transmitting information (telling the second spin which way it must point) faster than the speed of light. But that's an illusion, because to get two qubits into a definite [entangled] state we had to start with them close together. Later we can take them far apart, but if the qubits can't travel faster than the speed of light, neither can any message they can carry with them (Wilczek, 2008, pp. 117–118).

Wilczek's reasoning is unclear, but he seems to suggest that whatever leads to the correlations manifested in entangled states must have been built into the particles when they were emitted. If so, the correlations of quantum mechanics would be no more mysterious than the fact that many copies of an issue of *Physical Review Letters* contain the same information because they were all printed on one press before they were mailed out to various subscribers. It is distressing that a winner of a Nobel in Physics is seemingly unaware that there is a result called Bell's Theorem whose import is precisely to rule out such "reasonable" explanations.[27] The entire point of Bell's Theorem is this: the assumption that entangled particles are encoded at their source with instructions sufficient to satisfy the predictions of quantum mechanics is (in general) mathematically inconsistent with the correlations predicted by the theory (and observed in many kinds of experiments). For an elementary but rigorous demonstration of this fact, see (Maudlin, 2002, Ch. 1).

Bell's momentous result itself is negative: it rules out a certain class of explanations of quantum correlations, but does not by itself say what actually accounts for these correlations (beyond the quantum mechanical algorithms with which one calculates them). The prima facie explanation, if there is one, is that there is indeed some sort of spooky action (faster than light

---

[27] J. S. Bell: "The discomfort that I feel is associated with the fact that the observed perfect quantum correlations seem to demand something like the 'genetic' hypothesis … For me, it is so reasonable to assume that the photons in those experiments carry with them programs, which have been correlated in advance, telling them how to behave. This is so rational that I think that when Einstein saw that, and the others refused to see it, *he* was the rational man. … So for me, it is a pity that Einstein's idea doesn't work. The reasonable thing just doesn't work." (In Bernstein, 1991, p. 84.)

dynamics) going on, precisely as Einstein had feared. An enormous amount of intellectual energy has been expended trying to find some way of explaining or interpreting quantum mechanics so as to avoid this conclusion, which Bell himself and so many others have found so distasteful.[28] It is this author's opinion that the dogged efforts to explain away the appearance of spooky action have become what Imre Lakatos called a "degenerating research programme" (Lakatos, 1976) but it is beyond the scope of this paper to defend this claim. It is enough here to say that Bell's momentous result remains poorly understood more than fifty years after its publication.

6.3 Entangled Paths

We'll conclude our (incomplete) list of important thought experiments in modern physics with a brief look at interferometry, one of the most powerful tools of modern physics. The essential idea of an interferometer is that particle or light waves are emitted from or collected from a common source, directed through different pathways, and brought together and allowed to interfere. The interferometric Michelson-Morley experiment of 1887, no thought experiment, showed that it is impossible to detect the motion of the Earth with respect to the hypothetical luminiferous ether (Taylor & Wheeler, 1966, pp. 76–78). Interferometry plays an increasing role in modern quantum information theory. Nielsen and Chuang remark,

> We can now see what an actual quantum computer might look like in the laboratory (if only sufficiently good components were available to construct it), and a striking feature is that it is constructed nearly completely from optical interferometers (2000, p. 296).

Given that it is still technically impossible to construct most types of quantum computers that have been envisioned, one must say that so far most of the very active field of quantum computing is still in the realm of the thought experiment.

John A. Wheeler, like Mach, was not afraid to think on a cosmological scale. Imagine a quasi-stellar object billions of light years from Earth with a massive galaxy roughly half-way between (Wheeler, 1983, pp. 190–195). The galaxy will act as a gravitational lens (an effect predicted by Einstein), and can focus the light from the distant quasar onto detectors in an Earthly observatory. Light emitted from the quasar can take either path on its route to the lab on Earth. Gravitational lensing thus permits interferometry on a cosmological scale. The light is passed through a filter and then through a lens which focusses the light on the input faces of two optical fibres. The experimenters have a choice: they can either interpose a half-silvered mirror at the point at which the two light beams converge, or leave the mirror out. Omitting technical details, the key point is that with the mirror in place the experimenters will see interference between the light waves from the quasar, which is only possible if the waves had travelled through both paths; whereas with the mirror omitted, the experimenters will detect individual photons in one detector or the other and thus be able to tell which path the photons took. It is

---

[28] In an interview in 1988, Bell stated that according to his theorem, "maybe there must be something happening faster than light, although it pains me even to say that much" (Mann & Crease, 1988, p. 90).

precisely *as if* the choice of measurement procedure here on Earth determines (determined?) which path the photons took when they were emitted from the quasar billions of years earlier.

Does this literally mean that the past has an indeterminate ontology? Wheeler himself suggests that it does:

> …we are dealing with an elementary act of creation. It reaches into the present from billions of years in the past. It is wrong to think of that past as "already existing" in all detail. The "past" is theory. The past has no existence except as it is recorded in the present. By deciding what questions our quantum registering equipment shall put in the present we have an undeniable choice in what we have the right to say about the past (1983, p. 194).

As with Bell's Theorem, one could cautiously interpret Wheeler's Cosmological Interferometry experiment in a purely negative way. We have to concede that it is contradictory to say that the particle had a trajectory before we made our detector choice, but we could refuse to say more. In particular, we might stubbornly refuse to say that our experimental choice here on Earth today *creates* something in the past. But even if we take this cautiously agnostic stance, we are committed to the position that the past is ontologically "gappy." On pain of contradiction, there are some claims about the past that we just can't make; Wheeler's cosmic delayed choice experiment thus may well amount to an instance of the Kochen-Specker Theorem.[29] Arguably it tells against the block universe theory, according to which the universe is a complete four-dimensional, Riemannian plenum, and it may well provide support for the retrocausal interpretation of quantum mechanics, according to which amplitudes from future to past must be included in quantum-mechanical calculations (Cramer, 1986). Like many of the thought experiments sketched in this review, there is still much to be learned from Wheeler's grand interferometer.

7 Have Thought Experiments a Future in Physics?

Einstein himself had a very unusual ability to visualize—or, more accurately, kinaesthetically to *feel* how things work (Einstein, 1945). An ordinary competent physicist may well believe that it is hopeless to attempt to intuitively grasp the workings of nature as fluently as Einstein, any more than an ordinary musician can hope to duplicate the cognitive feats of Mozart. However, one can learn from those with extraordinary skills—one can at least *try* to do what they do. Einstein did one thing that can be done by anyone with sufficient intellectual courage: he deliberately sought out the simple, the obvious, the perception that was right under everyone's nose. One quality that all effective thought experiments have is that the essential insight is both simple and obvious—once you see it. The willingness to seek out the obvious that is not yet obvious to most people is as much a matter of temperament as raw cognitive ability, because it requires one to be unconventional (as was Einstein)—a risk that sometimes even exceptionally intelligent people are not willing to take.

---

[29] This was pointed out to me by Jesse Supina (private communication).

It is reasonable to ask whether there is still a creative role for thought experiments to play in physics as it grapples with the ever-increasing abstrusity of quantum gravity, particle physics, and string theory. It could be argued that the frontlines of theoretical physics now operate on a level of abstraction that is so far from common experience that the kind of ordinary mechanical and spatiotemporal intuitions at which Einstein excelled may no longer have much relevance. I have great faith in the flexibility and adaptability of the human imagination, and I think it is too soon to draw such a pessimistic conclusion. But even if visualization comes to play a decreasing role in the physics of the future, it will always be good methodology to seek out the obvious—and to question the conventional wisdom that too often prevents us from seeing it.

Acknowledgements

For helpful discussions or advice the author is grateful to Richard T.A. Arthur, Bryson Brown, Saurya Das, John Norton, David Siminovitch, and Jesse Supina. The author also thanks the editors of this volume for the opportunity to write this paper, and James Robert Brown in particular for encouragement at many stages of this author's career. Thanks are also due to the University of Lethbridge and the Social Sciences and Humanities Research Council of Canada for essential financial and material support. Of course, none of these fine persons or institutions are responsible for any errors or misconceptions that may have found their way into this work.

References

Aczel, A. D. (2002). *Entanglement: The Greatest Mystery in Physics*. Vancouver: Raincoast Books.

Anderson, J. L. (1967). *Principles of relativity physics*. New York: Academic Press.

Arthur, R. (1999). On thought experiments as a priori science. *International Studies in the Philosophy of Science*, *13*(3), 215–229. http://doi.org/10.1080/02698599908573622

Aspect, A., Dalibard, J., & Roger, G. (1982). Experimental Tests of Bell's Inequality Using Time-Varying Analyzers. *Physical Review Letters, 49*, 1804–1807.

Bell, J. S. (1964). On the Einstein Podolsky Rosen Paradox. *Physics*, *1*(3), 195–200.

Bernstein, J. (1991). *Quantum Profiles*. Princeton, NJ: Princeton University Press.

Bohm, D. (1951). *Quantum Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bohm, D. (1952a). A Suggested Interpretation of the Quantum Theory in Terms of "Hidden Variables". I. *Physical Review*, *85*(2), 166–179.

Bohm, D. (1952b). A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. II. *Physical Review*, *85*(2), 180–193.

Bohr, N. (1935). Can Quantum-Mechanical Description of Physical Reality be Considered Complete? *Physical Review*, *48*, 696–702.

Bub, J. (1997). *Interpreting the Quantum World*. Cambridge: Cambridge University Press.

Cohen-Tannoudji, C., Diu, B., & Laloë, F. (1977). *Quantum Mechanics* (Vol. I). New York: John Wiley and Sons.

Cramer, J. G. (1986). The Transactional Interpretation of Quantum Mechanics. *Reviews of Modern Physics*, *58*(July), 647–688.

Cushing, J. T. (1994). *Quantum Mechanics: Historical Contingency and the Copenhagen Hegemony*. Chicago & London: University of Chicago Press.

Demopoulos, W. (2004). Elementary Propositions and Essentially Incomplete Knowledge: A Framework for the Interpretation of Quantum Mechanics. *Noûs*, *38*(1), 86–109.

DeWitt, B. S. (1979). Quantum gravity: the new synthesis. In S. W. Hawking & W. Israel (Eds.), *General Relativity: An Einstein Centenary Survey* (pp. 680–745). Cambridge: Cambridge University Press.

Dür, W., Vidal, G., Linden, N., & Popescu, S. (2001). Entanglement Capabilities of Nonlocal Hamiltonians. *Physical Review Letters*, *87*, 137901.

Einstein, A. (1905). Zur Elektrodynamik bewegter Körper. *Annalen der Physik*, *17*, 891–921.

Einstein, A. (1909). Entwicklung unserer Anschauungen über das Wesen und die Konstitution der Strahlung. *Physikalische Zeitschrift*, *10*, 817–825.

Einstein, A. (1922). *Sidelights on Relativity*. (G. B. Jeffery & W. Perrett, Trans.). New York: E. P. Dutton (Dover Reprint).

Einstein, A. (1927). Bestimmt Schrödinger's Wellenmechanik die Bewegung eines Systems vollständig oder nur im Sinne der Statistik? Presented at the Prussian Academy of Sciences.

Einstein, A. (1945). A Testimonial from Professor Einstein. In J. Hadamard, *The Mathematician's Mind: The Psychology of Invention in the Mathematical Field* (pp. 142–143). Princeton, NJ: Princeton University Press.

Einstein, A. (1948). Quanten-Mechanik und Wirklichkeit. *Dialectica*, *2*(3–4), 320–324.

Einstein, A. (1951). Autobiographical Notes. In P. A. Schilpp (Ed.), *Albert Einstein: Philosopher-Scientist* (pp. 2–95). New York: Tudor.

Einstein, A. (1961). *Relativity: The Special and the General Theory*. (R. W. Lawson, Trans.). New York: Crown.

Einstein, A. (1969). Remarks to the Essays Appearing in this Collective Volume. In P. A. Schilpp (Ed.), *Albert Einstein: Philosopher-Scientist* (pp. 663–688). La Salle, IL: Open Court.

Einstein, A., Podolsky, B., & Rosen, N. (1935). Can Quantum-Mechanical Description of Physical Reality be Considered Complete? *Physical Review*, *47*, 777–780.

Feshbach, H., & Weisskopf, V. F. (1988). Ask a Foolish Question... *Physics Today*, (October), 9, 11.

Feynman, R. P., Leighton, R. B., & Sands, M. (1965). *The Feynman Lectures on Physics Vol. III: Quantum Mechanics*. Reading, MA: Addison-Wesley.

Fine, A. (1986). *The Shaky Game: Einstein, Realism, and the Quantum Theory*. Chicago and London: University of Chicago Press.

Fölsing, A. (1997). *Albert Einstein: A Biography*. (E. Osers, Trans.). New York: Viking.

Galison, P. (2003). *Einstein's Clocks, Poincaré's Maps: Empires of Time*. New York & London: W. W. Norton.

Gardner, M. (1978). *Aha! Aha! insight*. New York: Scientific American.

Ghirardi, G. C., Rimini, A., & Weber, T. (1980). A general argument against superluminal transmission through the quantum mechanical measurement process. *Lettere Al Nuovo Cimento*, *27*(10), 293–298.

Gutfreund, H., & Renn, J. (2015). *The Road to Relativity: The History and Meaning of Einstein's "The Foundation of General Relativity."* Princeton and Oxford: Princeton University Press.

Heisenberg, W. (1927). Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik*, *43*, 172–198.

Heisenberg, W. (1930). *The Physical Principles of the Quantum Theory*. Chicago: University of Chicago Press.

Holland, P. (2005). What's wrong with Einstein's 1927 hidden-variable interpretation of quantum mechanics? *Foundations of Physics*, *35*, 177–196.

Howard, D. (2006). Early History of Quantum Entanglement. Retrieved May 9, 2016, from http://www3.nd.edu/~dhoward1/Early%20History%20of%20Entanglement/sld001.html

Howard, D. (2007). Revisiting the Einstein-Bohr Dialogue. *Iyyun: The Jerusalem Philosophical Quarterly*, *56*(January), 57–90.

Janssen, M. (2014). "No Success Like Failure...": Einstein's Quest for General Relativity, 1907-1920. In M. Janssen & C. Lehner (Eds.), *The Cambridge Companion to Einstein*. New York: Cambridge University Press.

Kennedy, J. B. (1995). On the Empirical Foundations of the Quantum No-signalling Proofs. *Philosophy of Science*, *62*, 543–560.

Keynes, J. M. (1978). Newton the Man. In *The Collected Writings of John Maynard Keynes* (pp. 363–374). Royal Economic Society.

Kirshner, R. P. (2002). *The Extravagant Universe: Exploding Stars, Dark Energy and the Accelerating Cosmos*. Princeton and Oxford: Princeton University Press.

Kochen, S., & Specker, E. P. (1967). The Problem of Hidden Variables in Quantum Mechanics. *Journal of Mathematics and Mechanics*, *17*, 59–87.

Kuhn, T. S. (1970). *The structure of scientific revolutions* ([2nd ed.,). Chicago: University of Chicago Press.

Lakatos, I. (1976). *Proofs and Refutations: The Logic of Mathematical Discovery*. Cambridge: Cambridge University Press.

Lee, K. C., Sprague, M. R., Sussman, B. J., Nunn, J., Langford, N. K., Jin, X.-M., … Walmsley, I. A. (2011). Entangling Macroscopic Diamonds at Room Temperature. *Science*, *334*, 1253–1256. http://doi.org/10.1126/science.1211914

Mann, C., & Crease, R. (1988). Interview with J. S. Bell. *Omni*, (May), 84+.

Maroney, O. (2009). Information Processing and Thermodynamic Entropy. *Stanford Encyclopedia of Philosophy*. Retrieved from http://plato.stanford.edu/entries/information-entropy/#MaxDem

Martin, R. M. (1997). *Scientific Thinking*. Peterborough, ON: Broadview Press.

Maudlin, T. (2002). *Quantum Non-Locality and Relativity* (Second Edition). Oxford: Blackwell.

Mehra, J., & Rechenberg, H. (1982). *The Historical Development of Quantum Theory* (Vols. 1–6). New York: Springer-Verlag.

Miller, J. L. (2016). Three groups close the loopholes in tests of Bell's theorem. *Physics Today*, *69*(1). http://doi.org/10.1063/PT.3.3039

Misner, C. W., Thorne, K. S., & Wheeler, J. A. (1973). *Gravitation*. San Francisco: W. H. Freeman & Co.

Mittelstaedt, P. (1998). Can EPR-correlations be used for the transmission of superluminal signals? *Annalen Der Physik*, *7*(7–8), 710–715.

Moran, D., Softley, R., & Warrant, E. J. (2014). Eyeless Mexican Cavefish Save Energy by Eliminating the Circadian Rhythm in Metabolism. *PLOS ONE*, *9*(9), e107877. http://doi.org/10.1371/journal.pone.0107877

Newton, I. (1692). Original letter from Isaac Newton to Richard Bentley. Retrieved March 26, 2016, from http://www.newtonproject.sussex.ac.uk/view/texts/normalized/THEM00258

Newton, S. I. (1888). *A Catalogue of the Portsmouth Collection of Books and Papers Written by Or Belonging to Sir Isaac Newton: The Scientific Portion of which Has Been Presented by the Earl of Portsmouth to the University of Cambridge*. University Press.

Nielsen, M. A., & Chuang, I. L. (2000). *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press.

Norsen, T. (2005). Einstein's Boxes. *American Journal of Physics*, *73*(2), 164–176.

Norton, J. D. (Forthcoming). Maxwell's Demon Does Not Compute. In M. E. Cuffaro & S. C. Fletcher (Eds.), *Physical Perspectives on Computation, Computational Perspectives on Physics*. Cambridge: Cambridge University Press.

Norton, J. D. (1986). What Was Einstein's Principle of Equivalence? *Studies in History and Philosophy of Science*, *16*, 203–246.

Norton, J. D. (2013a). All Shook Up: Fluctuations, Maxwell's Demon and the Thermodynamics of Computation. *Entropy*, *15*, 4432–4483. http://doi.org/10.3390/e15104432

Norton, J. D. (2013b). Chasing the Light: Einstein's Most Famous Thought Experiment. In M. Frappier, L. Meynell, & J. R. Brown (Eds.), *Thought Experiments in Philosophy, Science, and the Arts* (pp. 123–140). New York & London: Routledge.

Norton, J. D. (2015a). General Relativity. Retrieved August 6, 2016, from http://www.pitt.edu/~jdnorton/teaching/HPS_0410/chapters/general_relativity/index.html

Norton, J. D. (2015b). The Hole Argument. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2015). Retrieved from http://plato.stanford.edu/archives/fall2015/entries/spacetime-holearg/

Pais, A. (1982). *"Subtle is the Lord...": The Science and the Life of Albert Einstein*. Oxford: Oxford University Press.

Pauli, W. (1933). *Die allgemeinen Principien der Wellenmechanik*. (H. Geiger & K. Scheel, Eds.) (2nd ed., Vol. 24). Berlin: Julius Springer.

Pauli, W. (1980). *General Principles of Quantum Mechanics*. (P. Achuthan & K. Venkatesan, Trans.). Berlin: Springer-Verlag.

Peacock, K. A. (1992). Comment on "Tests of Signal Locality and Einstein-Bell Locality for Multiparticle Systems." *Physical Review Letters*, *69*(18), 2733.

Peacock, K. A. (2008). *The Quantum Revolution: A Historical Perspective*. Westport, CN: Greenwood Press.

Peat, D. (1997). *Infinite Potential: The Life and Times of David Bohm*. Reading, MA: Addison-Wesley.

Pössel, M. (2009). The elevator, the rocket, and gravity: the equivalence principle. Retrieved March 26, 2016, from http://www.einstein-online.info/spotlights/equivalence_principle

Redhead, M. (1987). *Incompleteness, Nonlocality, and Realism: A Prolegomenon to the Philosophy of Quantum Mechanics*. Oxford: Oxford University Press.

Robertson, R. N. (1983). *The Lively Membranes*. Cambridge: Cambridge University Press.

Rose, P. L. (1998). *Heisenberg and the Nazi atomic bomb project: a study in German culture*. Berkeley, Calif: University of California Press.

Rovelli, C. (2004). *Quantum Gravity*. Cambridge: Cambridge University Press.

Sauer, T. (2014). Einstein's Unified Field Theory Program. In M. Janssen & C. Lehner (Eds.), *The Cambridge Companion to Einstein* (pp. 281–305). New York: Cambridge University Press.

Schrödinger, E. (1935a). Die gegenwärtige Situation in der Quantenmechanik. *Naturwissenschaften*, *23*, 807–12–28–49.

Schrödinger, E. (1935b). Discussion of Probability Relations Between Separated Systems. *Proceedings of the Cambridge Philosophical Society*, *31*, 555–563.

Schrödinger, E. (1983). The Present Situation in Quantum Mechanics. In J. A. Wheeler & W. H. Zurek (Eds.), J. D. Trimmer (Trans.), *Quantum Theory and Measurement* (pp. 152–167). Princeton, NJ: Princeton University Press.

Shimony, A. (1983). Controllable and Uncontrollable Nonlocality. In S. Kamefuchi (Ed.), *Foundations of Quantum Mechanics in the Light of New Techology* (pp. 225–230). Tokyo: Physical Society of Japan.

Smoot, G., & Davidson, K. (1993). *Wrinkles in Time: Witness to the Birth of the Universe*. New York: William Morrow and Co.

Stachel, J. (Ed.). (2005). *Einstein's Miraculous Year: Five Papers that Changed the Face of Physics*. Princeton and Oxford: Princeton University Press.

Stachel, J. (2014). The Hole Argument and Some Physical and Philosophical Implications. *Living Reviews in Relativity*, *17*(1). http://doi.org/10.12942/lrr-2014-1

Stapp, H. P. (1975). Bell's Theorem and World Process. *Il Nuovo Cimento B*, *29 B*(2), 270–276.

Stone, A. D. (2013). *Einstein and the Quantum: The Quest of the Valiant Swabian*. Princeton and Oxford: Princeton University Press.

Szilard, L. (1972). On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. In *The Collected Works of Leo Szilard: Scientific Papers*. Cambridge, MA: MIT Press.

Taylor, E. F., & Wheeler, J. A. (1966). *Spacetime Physics*. San Francisco: W. H. Freeman & Co.

Von Neumann, J. (1955). *Mathematical foundations of quantum mechanics*. Princeton, N.J: Princeton University Press.

Wheeler, J. A. (1983). Law Without Law. In J. A. Wheeler & W. H. Zurek (Eds.), *Quantum Theory and Measurement* (pp. 182–213). Princeton, N.J: Princeton University Press.

Wheeler, J. A., & Zurek, W. H. (Eds.). (1983). *Quantum Theory and Measurement*. Princeton, NJ: Princeton University Press.

Wilczek, F. (2008). *The Lightness of Being: Mass, Ether, and the Unification of Forces*. New York: Basic Books.

Woods, J. (2003). *Paradox and Paraconsistency: Conflict Resolution in the Abstract Sciences*. Cambridge: Cambridge University Press.

Yam, P. (2012, October 9). Bringing Schrödinger's Cat to Life. Retrieved May 9, 2016, from http://www.scientificamerican.com/article/bringing-schrodingers-quantum-cat-to-life/