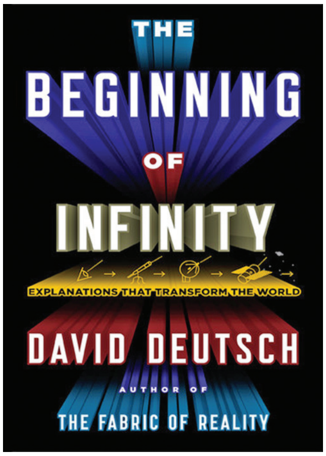
Book Review



The Beginning of Infinity: Explanations That Transform the World

Reviewed by William G. Faris

The Beginning of Infinity: Explanations That Transform the World
David Deutsch
Viking Penguin, New York, 2011
US$30.00, 496 pages
ISBN: 978-0-670-02275-5

The Cosmic Scheme of Things

A number of recent popular scientific books treat a wide range of scientific and philosophical topics, from quantum physics to evolution to the nature of explanation. A previous book by the physicist David Deutsch [2] touched on all these themes. His new work, *The Beginning of Infinity*, is even more ambitious. The goal is to tie these topics together to form a unified world view. The central idea is the revolutionary impact of our ability to make good scientific explanations.

In books written by physicists one expects to find awe of the universe and delight that we have come to understand something about it. There might also be an implicit assumption that people are not particularly significant in the cosmic scheme of things. Deutsch makes the contrary claim (p. 45):

> People *are* significant in the cosmic scheme of things.

He presents the following example. The atomic composition of the universe consists mainly of the light elements hydrogen and helium. Consider a heavy element such as gold. Where can it be found, and where did it come from? There are two very different sources. One is the transmutation of elements in the interior of an exploding supernova. Gold, among other elements, is created and then distributed throughout the universe. A small fraction winds up on planets like Earth, where it can be mined from rock. The other source is intelligent beings who have an explanation of atomic and nuclear matter and are able to transmute other metals into gold in a particle accelerator. So gold is created either in extraordinary violent stellar explosions or through the application of scientific insight.

In the claim there is a potential ambiguity in the use of the word "people". Deutsch characterizes them (p. 416) as "creative, universal explainers". This might well include residents of other planets, which leads to a related point. Most physical effects diminish with distance. However, Deutsch argues (p. 275):

> There is only one known phenomenon which, if it ever occurred, would have effects that did not fall off with distance, and that is the creation of a certain type of knowledge, namely a beginning of infinity. Indeed, knowledge can aim itself at a target, travel vast distances having scarcely any effect, and then utterly transform the destination.

Deutsch is a physicist who believes that matter is strictly governed by the laws of physics. However, he is no reductionist. He believes that atoms are real, but also that abstractions are real. One level of abstraction is knowledge, which he defines as "information which, when it is physically embodied in a suitable environment, tends to cause itself to remain so" (p. 130). This includes both knowledge embedded in DNA and ideas in human brains. Biological knowledge is nonexplanatory, but human knowledge includes explanations that solve

William G. Faris is professor of mathematics at the University of Arizona. His email address is `faris@math.arizona.edu`.

DOI: http://dx.doi.org/10.1090/noti821

unexpected problems. The gold example shows that an explanation can have physical effects just as real as those created by an exploding star.

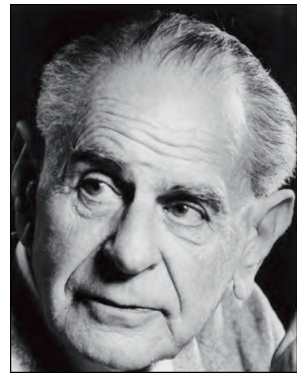
Deutsch's book is also a manifesto in praise of the notion of progress. His claim is that a new kind of progress began with the Enlightenment. In the best case the growth of knowledge may proceed indefinitely in the future without bound. This is the "beginning of infinity" of the title and of the quotation above. The argument has several threads. It begins with a definite position on philosophy of science, centered on an argument that a notion of "good explanation" of reality is the key to progress. It includes a theory of cultural evolution, which explains obstacles to progress and the lucky circumstance that swept them away. There is extensive discussion of the anomalous status of quantum theory, which on some accounts is incompatible with a realistic world view. The argument connects views on topics as varied as mathematical reality, voting systems, aesthetics, and sustainability. The author rejects sustainability as an aspiration or as a constraint on planning; he favors an open-ended journey of creation and exploration.

Good Explanations

Deutsch attributes recent progress to the discovery of how to create good explanations. A "good explanation" is defined as an explanation "that is hard to vary while still accounting for what it purports to account for" (p. 30). This philosophy is inspired by the writings of Karl Popper. In his influential works, Popper argued against empiricism, the notion that we derive all our knowledge from sensory experience. There is no such thing as raw experience; scientific observation is always theory-laden. Furthermore, there is no principle of inductive reasoning that says that patterns of the past will be repeated in the future. In Popper's view the path to scientific progress is to make conjectures and then test them. Deutsch insists that even explanations that are testable are not enough; they should be good explanations in the sense that he defines.

His version of philosophy of science has two complementary strands. One is realism, the notion that there is a physical world about which it is possible to obtain knowledge. The other is fallibilism, the doctrine that there is no sure path to justify the knowledge that has been obtained at any stage. These together support the metaphor of a knowledge base of good explanations that is increasing but bounded above by a reality that is never fully knowable.

One useful feature of the book is the explicit definitions of philosophical terms, explaining how they are used in his system. Thus "instrumentalism" is defined as the "misconception that science cannot describe reality, only predict outcomes of observations" (p. 31). As is clear from this example, definitions often come with a judgment.



Reproduced with the permission of the Library of the London School of Economics.

Karl Popper: philosopher of "theory-laden" science.

Deutsch's main work as a physicist has been on quantum computing, and it is natural that he would test his ideas in the framework of quantum theory. This leads to a problem. Quantum theory is the accepted framework for understanding properties of matter, viewed as consisting of constituents on the molecular, atomic, and subatomic scale. It explains such properties of matter as density and strength and conductivity and color. It also underlies the deeper understanding of the chemical bond. It is so pervasive in fundamental physics that it is difficult to see how to make the smallest modification to it without destroying the whole edifice.

The problem is that the formulation of quantum theory is abstract and mathematical, and, as usually presented, it also has a peculiar dual character. There is a law of time evolution given by the Schrödinger equation: For every time interval there is a transformation U obtained by solving this equation. It maps the initial state of a system to the final state of a system. (The notation U indicates a *unitary* linear transformation.)

There is another law of time evolution, a random transformation R that determines the result of the measurement. Various possibilities for this transformation may occur, according to certain calculated probabilities. The state change under this transformation is called *reduction* or *collapse*.

The origin of the transformation R is that it describes an intervention from the outside. The system is in a certain state. A physicist sets up a measuring apparatus and performs the experiment. The observed outcome occurs. The new state after the experiment is determined by the particular R that corresponds to the experimental outcome. If the experiment is repeated under identical conditions, then the outcomes vary, but the frequencies are predicted by the calculated probabilities.

Such instrumentalist accounts have a long and complicated history. An early version is sometimes referred to as the "Copenhagen interpretation", since it stems from ideas of Niels Bohr and his school. (See [8] for a critical account.) Deutsch will have none of this. For him instrumentalism is defeat; the only acceptable account is realistic. Here is his explicit definition (or dismissal) of Bohr's approach (p. 324):

> *Copenhagen interpretation*
> Niels Bohr's combination of instrumentalism, anthropocentrism and studied ambiguity, used to avoid understanding quantum theory as being about reality.

If the instrumentalist account is not acceptable, then what are the alternatives? One is to develop a theory that includes the deterministic U time evolution but also has additional structure. This structure should introduce randomness in some way; perhaps the reduction transformation R will emerge as a consequence. The other way is to develop a theory with only the deterministic time evolution U. Such a theory must explain why the outcomes of experiments on quantum systems appear random.

The final section of this review gives a more detailed discussion of these alternatives. It will come as no surprise that Deutsch prefers a realistic version of quantum theory with only the dynamical time evolution given by U. This seems the happiest outcome, but we shall see the price he must pay.

Cultural Evolution

Many of the ideas in this book are related to evolutionary theory. In biological evolution genes are replicated, and they change through variation and selection. In cultural evolution the analog of a gene is a "meme". (The term was coined by Richard Dawkins in 1976.) Memes are replicated, and they also change through variation and selection.

Copyright of Max Alexander, STARMUS.

Richard Dawkins: evolutionary theorist.

The copying mechanism is different for genes and memes. A gene exists in a physical form as DNA that may be copied intact through several generations without ever being expressed as behavior. It acts much like a computer program. A meme can be copied only if it is enacted. In fact, a meme has two forms. It may be an idea in a brain. This idea may provoke a behavioral embodiment of the meme, such as action or body language or speech. The idea in the next recipient brain has to be guessed from the observed behavior. The successful meme variant is the one that changes the behavior of its holders in such a way as to make itself best at displacing other memes from the population.

Deutsch contrasts two ways that a meme may successfully replicate itself. A meme may survive as a meme of conformity, because it is never questioned. It relies on disabling the recipients' critical facilities. When such memes dominate, the result is a static society in which dangerous dysfunctional memes are suppressed. The other possibility is a dynamic society. In such a society memes are replicated in a rapidly changing environment. This can take place in a culture of criticism, because the memes embody new truths.

The transition from a static society to a dynamic society depends on an almost accidental shift in how meme transmission is employed. A meme cannot be simply copied from behavior. There has to be a mental capacity to infer the idea from the behavior. This creative ability developed in service to the task of replicating memes of conformity. It was hijacked to the task of creating new knowledge.

Deutsch considers the transition to a society where there is deliberate creation of new knowledge to begin with the Enlightenment. There is no precise date, but the founding of the Royal Society in 1660 is a landmark. There may have been previous transitions that did not survive. Deutsch calls such a period a "mini-enlightenment". He speculates that two such mini-enlightenments may have occurred. One was at the time of the Athenian advances in political freedom and openness to new ideas in philosophy and science. Another was when Florence became a center of creativity in art, accompanied by advances in science, philosophy, and technology. In both cases the initial spark was extinguished. This has major enduring consequences (pp. 220–221):

> The inhabitants of Florence in 1494 or Athens in 404 BCE could be forgiven for concluding that optimism just isn't factually true. For they know nothing of such things as the reach of explanations or the power of science or even laws of nature as we understand them, let alone the moral and technological progress that was to follow when *the* Enlightenment got under way. At the moment of defeat, it must have seemed at least plausible to the formerly optimistic Athenians that the Spartans might be right, and to the formerly optimistic Florentines that Savonarola might be. Like every other destruction of optimism, whether in a whole civilization or in a single individual, these must have been unspeakable catastrophes for those who had dared to expect progress. But we should feel more than sympathy

> for these people. We should take it personally. For if any of those earlier experiments in optimism had succeeded, our species would be exploring the stars by now, and you and I would be immortal.

As we have seen, Deutsch's book combines various threads to create a vision of continuing progress in enlightenment and human flourishing. It is a unified portrait that explains scientific advance and puts it in a moral framework. The vision is supported by a great number of assertions, some of them extravagant. Take, for instance, his claim that "everything that is not forbidden by laws of nature is achievable, given the right knowledge" (p. 76). Here is a supporting argument (p. 56):

> ...every putative physical transformation, to be performed in a given time with given resources or under any other conditions, is either:
> — impossible because it is forbidden by laws of nature; or
> — achievable, given the right knowledge.
>
> That momentous dichotomy exists because if there were transformations that technology could never achieve regardless of what knowledge was brought to bear, then this fact would itself be a testable regularity in nature. But all regularities in nature have explanations, so the explanation of that regularity would itself be a law of nature, or a consequence of one. And so, again, everything that is not forbidden by laws of nature is achievable, given the right knowledge.

A passage like this is tough to decipher.

The question of whether to believe every detail may be beside the point. A manifesto is not a scientific or philosophical treatise; it is an outline of a world view that one can try on for comfort. Some readers will find satisfaction in a systematic view of the world that is compatible with notions of science and progress. Others may find it too simple or too optimistic. In the latter case they are violating yet another of Deutsch's maxims (p. 212):

The Principle of Optimism
All evils are caused by insufficient knowledge.

So much the worse for them.

Quantum Theory

From here on this review concentrates on quantum theory and is more technical. The formulation of quantum theory centers on the wave function, a quantity that is difficult to interpret in terms of a realistic world view. In compensation the theory has an elegant mathematical structure. What follows is a brief review of this structure, followed by a discussion of three possibilities for interpretation. (Deutsch would not agree that interpretation is an issue, since he sees only one reasonable way to think of the theory.)

The state of a quantum system at a given moment in time is described by a complex valued wave function $\psi(x)$. This depends on the positions of all the particles in the system. For instance, say that there are N particles (no spin or statistics). The position of particle i at fixed time is described by a coordinate $\mathbf{x}_i$ in three-dimensional space. The wave function depends on $x = (\mathbf{x}_1, \dots, \mathbf{x}_N)$, which ranges over configuration space of dimension $3N$. This is already an incredible picture of nature: everything is related to everything else through the wave function. In particular, the joint probability density for the system of N particles is proportional to

$$\rho(x) = |\psi(x)|^2. \tag{1}$$

It is possible for the different particles to be highly correlated, even when they are widely separated in space.

The change in the state over a given time interval is ordinarily given by a transformation U. This transformation is deterministic, and it is computed by solving the Schrödinger equation. This is a complicated linear partial differential equation, and much of theoretical physics consists of attempts to solve it for some particular system.[1]

Since the Schrödinger equation is linear, the corresponding transformations U are also linear. This naturally leads to a more abstract point of view. The characteristic feature of vectors is that it is possible to take linear combinations of vectors to produce new vectors. In particular, if ψ_1 and ψ_2 are vectors, then the sum $\psi_1 + \psi_2$ is also a vector. Also, every scalar multiple of a vector is a vector. Since wave functions share these properties, there is a convention of calling a wave function a vector. This in turn leads to the use of geometrical

[1] *In some accounts the complete description of the physical state of a quantum system involves more than its wave function; it brings in additional structure, namely the actual positions of the particles. This leads to the question of how to specify the time evolution of particle positions. This issue is not treated in orthodox quantum mechanics, but Bohmian mechanics and Féynes-Nelson stochastic mechanics each specify a possible continuous time evolution for particle positions. The reviewer thanks Sheldon Goldstein for comments on this topic.*

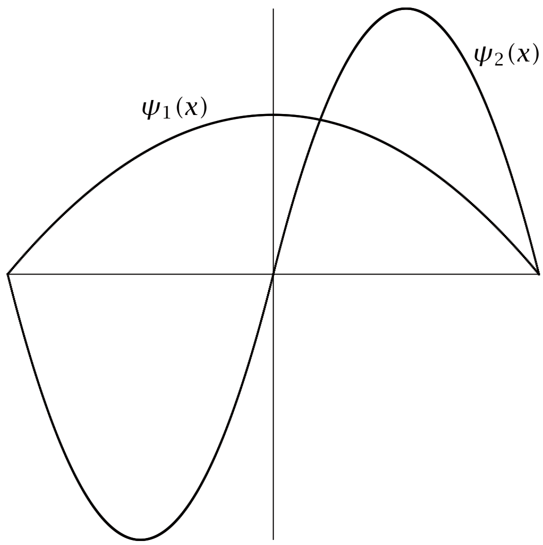


Figure 1. Example of orthogonal functions.

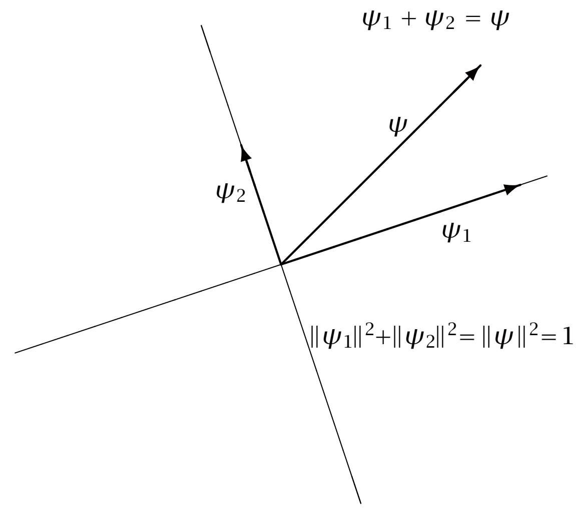


Figure 2. Probability = Pythagoras.

analogies. (The fact that the scalars are complex numbers is only a minor problem.)

Each pair of vectors ψ_1 and ψ_2 has a scalar product $\langle\psi_1, \psi_2\rangle$. This too has an analog for wave functions; the scalar product is defined by the definite integral

$$\langle\psi_1, \psi_2\rangle = \int \overline{\psi_1(x)}\psi_2(x)\,dx. \tag{2}$$

The space of all wave functions is regarded as a vector space of functions, in fact, as a Hilbert space. Each vector ψ in the Hilbert space has a norm (length) $\|\psi\|$ which is determined by the usual formula $\|\psi\|^2 = \langle\psi, \psi\rangle$ for the square of the norm. In the case of wave functions the explicit expression is

$$\|\psi\|^2 = \int |\psi(x)|^2\,dx. \tag{3}$$

Vectors ψ_1 and ψ_2 are orthogonal (perpendicular) if their scalar product has the value zero, that is, $\langle\psi_1, \psi_2\rangle = 0$. (Figure 1 shows an example of orthogonal wave functions. In this example $\psi_1(x)$ is even and $\psi_2(x)$ is odd, so their product has integral zero.) With these definitions various notions of geometry have attractive generalizations. Suppose, for instance, that there is a sequence of orthogonal vectors ψ_j with sum ψ, so

$$\psi = \sum_j \psi_j. \tag{4}$$

In this situation the theorem of Pythagoras takes the form

$$\|\psi\|^2 = \sum_j \|\psi_j\|^2. \tag{5}$$

If $\|\psi\|^2 = 1$, then the theorem of Pythagoras has the special form

$$1 = \sum_j \|\psi_j\|^2. \tag{6}$$

These give positive numbers that sum to one, just what is needed for probability. In fact, this is the standard framework for the interpretation of quantum mechanics. For simplicity consider an observable quantity with discrete values, such as the energy of a bound system. The possible values are indexed by j. The observable quantity determines a decomposition of the Hilbert space into orthogonal subspaces.[2] The state vector is a vector ψ of unit length. It is expressed as a sum (4) of the projected vectors. Different observable quantities define different decompositions. For each observable quantity there are probabilities given by the terms in (6).

This geometrical picture of quantum mechanics, abstract and beautiful, is immensely appealing. The reviewer is tempted to summarize it in a slogan:

Probability = Pythagoras

To see such a surprising and satisfying connection (Figure 2) is to be seduced by perfection. However, as we shall see, there are reasons to resist its allure.

---

[2] *The mathematical structure is an orthogonal direct sum decomposition $\mathcal{H} = \bigoplus_j \mathcal{H}_j$ of the Hilbert space. There are corresponding real numbers λ_j that are possible values for the observable quantity. These data determine a self-adjoint operator A whose action on each $\mathcal{H}_j$ is multiplication by λ_j. If ψ is the state vector and ψ_j is the orthogonal projection of ψ onto $\mathcal{H}_j$, then the expected value of this quantity has the elegant expression $\sum_j \lambda_j \|\psi_j\|^2 = \langle\psi, A\psi\rangle$.*

The trouble begins with a counterexample. Consider a quantum system. Each possible decomposition (4) defines probabilities. It is tempting to think of these probabilities as describing outcomes associated with this system. This is not a consistent view. An early argument to this effect was given by John Bell. A more recent example of Lucian Hardy [5] makes the same point in an even more convincing way. Something else is needed to select a particular decomposition relevant to a given situation. See the appendix for a brief account of Hardy's argument.

What is it that determines which decomposition is relevant? If in a given situation probabilities are to predict frequencies, then which probabilities are to be used, and which are to be discarded? Any serious account of quantum mechanics must consider this problem. There are various ways to address it, but in the reviewer's view they fall into three general classes:

- Instrumentalism
- Quantum theory with additional structure
- Pure quantum theory

The *instrumentalist account* of quantum theory emphasizes its ability to make experimental predictions. This solves the problem, because the outcomes emerge as a result of the particular experimental setup. Deutsch summarizes the usual rule for such predictions (p. 307):

> With hindsight, we can state the rule of thumb like this: whenever a measurement is made, all the histories but one cease to exist. The surviving one is chosen at random, with the probability of each possible outcome being equal to the total measure of all the histories in which that outcome occurs.

He then describes the adoption of the instrumentalist interpretation (p. 307):

> At that point, disaster struck. Instead of trying to improve and integrate these two powerful but slightly flawed explanatory theories [of Schrödinger and Heisenberg], and to explain why the rule of thumb worked, most of the theoretical-physics community retreated rapidly and with remarkable docility into instrumentalism. If the predictions work, they reasoned, then why worry about the explanation? So they tried to regard quantum theory as being *nothing but* a set of rules of thumb for predicting the observed outcomes of experiments, saying nothing (else) about reality. This move is still popular today, and is known to its critics (and even to some of its proponents) as the 'shut-up-and-calculate interpretation of quantum theory'.

In an instrumentalist account the relevant decomposition into orthogonal subspaces is determined by the measurement one performs on the system. This leads to the nontrivial problem: how does one characterize measurement? There is no universally accepted solution. The notion of "measurement" is defined with varying degrees of precision. In some versions a measurement is said to necessarily involve interactions on a macroscopic scale. This is in spite of the fact that the macroscopic world is supposed to be made of atoms.

The mathematical formulation ignores most of this detail. The particular measurement is specified by an orthogonal decomposition, which then determines a decomposition of the state vector ψ as a sum of components ψ_j as in (4). For each j there is a reduction operator R_j. The effect of reduction on the state vector ψ is another vector $\chi_j = R_j\psi$. In some accounts it is required that $\chi_j = \psi_j$, but this is unnecessarily restrictive. All that is needed is that $\|\chi_j\|^2 = \|\psi_j\|^2$. The particular j that is used is random; its probability is given by the squared norm $\|\psi_j\|^2$. Strictly speaking, the vector χ_j is not a state vector, since it does not have norm one. However, it may be multiplied by a scalar to give a normalized state vector, which could be taken as a new state of the system.

In general, a probability describes the statistics of what happens when an experiment is repeated many times. But what happens in a particular experiment? As always in probability, when the experiment is performed, a particular outcome value j' occurs. In this instrumentalist version of quantum mechanics there is an additional postulate: the new reduced wave function is $\chi_{j'} = R_{j'}\psi$. No mechanism is given.

Another direction is *quantum theory with additional structure*. One ingredient might be to try to model the experimental apparatus along with the system of interest. There can also be an explicit mechanism for introducing randomness. Some of the ideas go back to von Neumann; others have been elaborated by various authors. There is no agreement on details. Here is a brief account of a possible view, taken from [4]. It is not universally accepted; in particular, Deutsch presumably would be appalled at the ad hoc introduction of randomness.

The complete description involves the system of interest together with an environment, forming what one might call the total system. The system of interest might be a particle, an atom, or a molecule. To make the description concrete, it will be called an atomic system. The environment could be an experimental apparatus or, more generally, a larger system of some complexity. For brevity call it the apparatus. The combined system is described by a wave function $\Psi(x, y)$, where x describes the positions of particles in the atomic system, and y describes the positions of particles in the apparatus. Typically there is no natural way for the wave function for the combined system to determine a wave function for the atomic subsystem.

One situation where it is meaningful to have a wave function $\psi(x)$ for the atomic subsystem is when the total system wave function is of the product form

$$\Psi_0(x, y) = \psi(x)\phi_0(y). \tag{7}$$

Suppose that this is an initial state that has no immediate interaction between the atomic system and apparatus. This means that the wave function is nonzero only if the $\mathbf{x}_i$ are so far away from the $\mathbf{y}_j$ that the interaction is negligible. The atoms are headed toward the apparatus, but they are not there yet. As long as the combined wave function has this form, the dynamics of the atomic subsystem is described by the Schrödinger time evolution U for the atomic system alone.

Now the system is to interact with the apparatus. The starting point is a decomposition of the wave function $\psi(x)$ of the atomic system as a sum

$$\psi(x) = \sum_j \psi_j(x). \tag{8}$$

The measurement itself is accomplished by the unitary dynamics of the combined system. This deterministic transformation must be appropriate to the decomposition (8) in the following sense. It should map each wave function $\psi_j(x)\phi_0(y)$ to a new wave function $\chi_j(x)\phi_j(y)$ for which the new environment wave function $\phi_j(y)$ factor also depends on j. These wave functions $\phi_j(y)$ should be normalized and form an orthogonal family. This implies that the atomic wave function normalization is preserved, in the sense that $\|\chi_j\|^2 = \|\psi_j\|^2$. For such a transformation to exist, there must be a physical interaction of a suitable type—for instance, electric or magnetic—between the atomic system and apparatus.

Courtesy of Mark Everett.

Hugh Everett: originator of many-worlds quantum theory.

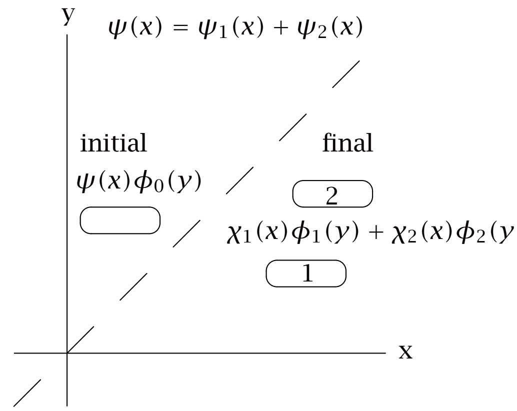


Figure 3. Support of system-apparatus wave function.

Suppose that there is such an interaction. Since the transformation is linear, it maps Ψ_0 to a state Ψ in which the atomic system is coupled to the states of the apparatus. The new wave function is

$$\Psi(x, y) = \sum_j \chi_j(x)\phi_j(y). \tag{9}$$

One can think of $\chi_j(x)$ as the result of applying a reduction operator R_j to the state $\psi(x)$. The probability associated with such an atomic wave function is $\|\chi_j\|^2 = \|\psi_j\|^2$.

The apparatus wave functions $\phi_j(y)$ are functions of the apparatus configuration $y = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_M)$, where M is the number of particles in the apparatus. When M is very large it is plausible that the apparatus wave functions $\phi_j(y)$ with various indices j are macroscopically different. Such effects have been studied under the name *quantum decoherence*. (See [7] for a recent survey.) In the present account the condition that the apparatus states are macroscopically different is interpreted to mean that the wave functions $\phi_j(y)$ with different j are supported on subsets with negligible overlap in apparatus configuration space (Figure 3). The corresponding atomic wave functions $\chi_j(x)$ are the result of dynamical interaction. Each of them is a candidate for the result of the reduction process.

In order to have an actual result, additional structure is required. This is obtained by introducing a new dynamics with random outcome. One of the indices is randomly selected. Say that this is the j' index. Then the corresponding reduced wave function of the atomic system is the $\chi_{j'}(x)$ given by applying $R_{j'}$ to $\psi(x)$. This reduction process does not contradict the unitary dynamics for the atomic subsystem. The wave function for the atomic subsystem is not even defined while the atomic system and apparatus are interacting; it is only defined before the interaction and after the interaction is over. Suppose that for a subsequent time interval a decomposition (9) with nonoverlapping $\phi_j(y)$ persists and there is no immediate interaction between the atomic system and apparatus. Then over this interval of time the wave function of the atomic subsystem continues to be defined, and the deterministic dynamics U describes its time evolution.

The reader will notice that the above account is complicated and artificial. However, it is at least consistent. This is because it is a consequence of a variant of quantum theory [3] that is itself known to be consistent.

The final possibility is *pure quantum theory*. Almost everyone agrees on the deterministic dynamics given by the Schrödinger equation. Perhaps this is all that is needed. This idea is the genesis of the many-worlds theory (see [6] for a recent survey). This theory began with hints by Schrödinger himself as early as 1926 (see [1] for a modern account) and was developed in more detail by Everett in 1957. The rough idea is that when there are macroscopically separated wave functions $\phi_j(y)$, each of these describes a world with its own history. We live in and experience only one such world, but they are all equally real. There is no miraculous reduction of the wave function and no randomness. The apparent randomness is perhaps due to the effect that we experience a history that is typical and hence appears random. Or there may be some other mechanism.

Deutsch is an enthusiastic proponent of such a theory. It does everything he wants, avoiding instrumentalism and ad hoc introduction of randomness. He refers to the resulting picture of physics as the "multiverse," and he is willing to draw the consequences (p. 294):

> ...there exist histories in which any given person, alive in our history at any time, is killed soon afterwards by cancer. There exist other histories in which the course of a battle, or a war, is changed by such an event, or by a lightning bolt at exactly the right place and time, or by any of countless other unlikely, 'random' events.

It is not the case that everything is permitted.

> A great deal of fiction is therefore close to a fact somewhere in the multiverse. But not all fiction. For instance, there are no histories in which my stories of the transporter malfunction are true, because they require different laws of physics. Nor are there histories in which the fundamental constants of nature such as the speed of light or the charge on an electron are different.

As with much of what Deutsch writes, the question is not so much whether to believe it as whether to march under his banner.

References

[1] Valia Allori, Sheldon Goldstein, Roderich Tumulka, and Nino Zanghì, Many-worlds and Schrödinger's first quantum theory, *British Journal for the Philosophy of Science* **62** (2011), 1–27.

[2] David Deutsch, *The Fabric of Reality: The Science of Parallel Universes—and Its Implications*, Allen Lane (Viking Penguin), New York, 1997.

[3] Detlef Dürr, Sheldon Goldstein, and Nino Zanghì, Quantum equilibrium and the role of operators as observables in quantum theory, *J. Statistical Physics* **116** (2004), 959–1055.

[4] William G. Faris, Outline of quantum mechanics, pp. 1–52 in *Entropy and the Quantum*, edited by Robert Sims and Daniel Ueltschi, Contemporary Mathematics, 529, American Mathematical Society, Providence, RI, 2010.

[5] Lucian Hardy, Nonlocality for two particles without inequalities for almost all entangled states, *Phys. Rev. Letters* **71** (1993), 1665–1668.

[6] Simon Saunders, Jonathan Barrett, Adrian Kent, and David Wallace (editors), *Many Worlds? Everett, Quantum Theory, and Reality*, Oxford University Press, New York, 2010.

[7] Maximilian Schlosshauer, *Decoherence and the Quantum-to-Classical Transition*, Springer, Berlin, 2010.

[8] David Wick, *The Infamous Boundary: Seven Decades of Heresy in Quantum Physics*, Copernicus (Springer-Verlag), New York, 1996.

Appendix: The Hardy Example

In quantum mechanics each observable defines a decomposition of the Hilbert space into orthogonal subspaces. Each state vector ψ (length one) is then written as a sum of projections ψ_j onto these subspaces. The probability corresponding to j is $\|\psi_j\|^2$. The Hardy example shows that these probabilities cannot simultaneously predict outcomes for all the observables together.

First consider a single particle. Suppose that for such a particle there are two distinct observable quantities D and U. Each can have either of two values, 1 or 0. The event that $D = 1$ is denoted d, and the event that $D = 0$ is denoted $\bar{d}$. Similarly, the event that $U = 1$ is denoted u, and the event that $U = 0$ is denoted $\bar{u}$.

Now consider two particles, perhaps widely separated in space. For the first particle one can observe either D or U, and for the second particle one can observe either D or U. This gives four possibilities for observation. There are four corresponding decompositions of the state vector describing the two particles:

$$\psi = \psi_{dd} + \psi_{d\bar{d}} + \psi_{\bar{d}d} + \psi_{\bar{d}\bar{d}}, \tag{1}$$
$$\psi = \psi_{du} + \psi_{d\bar{u}} + \psi_{\bar{d}u} + \psi_{\bar{d}\bar{u}}, \tag{2}$$
$$\psi = \psi_{ud} + \psi_{u\bar{d}} + \psi_{\bar{u}d} + \psi_{\bar{u}\bar{d}}, \tag{3}$$
$$\psi = \psi_{uu} + \psi_{u\bar{u}} + \psi_{\bar{u}u} + \psi_{\bar{u}\bar{u}}. \tag{4}$$

Furthermore, Hardy constructs the quantities D and U and the state ψ in such a way that

$$\psi_{dd} \neq 0, \tag{5}$$
$$\psi_{d\bar{u}} = 0, \tag{6}$$
$$\psi_{\bar{u}d} = 0, \tag{7}$$
$$\psi_{uu} = 0. \tag{8}$$

For fixed ψ he chooses the parameters that define D and U to satisfy the last three of these equations. He then optimizes the parameters that specify ψ to maximize the probability $\|\psi_{dd}\|^2$. The maximum value works out to be $\frac{1}{2}(5\sqrt{5} - 11)$, which is about 9 percent.

Suppose that in a given physical situation the observables all have values. According to the first equation above, the probability of dd is greater than zero. So it is possible that the outcome is d for the first particle and d for the second particle. Suppose that this is the case. The second equation says that d for the first particle implies u for the second particle, and the third equation says that d for the second particle implies u for the first particle. It follows that the outcome is u for the first particle and u for the second particle. However, the fourth equation says that this is impossible.

The conclusion is that in a given physical situation the observables cannot all have values. The usual explanation for this is that measurement in quantum mechanics does not reveal preexisting values but in effect creates the values. The four decompositions correspond to four different experiments, and each decomposition provides the correct probabilities for the result of the corresponding experiment. For a given experiment, only one of the four decompositions is relevant to determining what actually happens. The other three decompositions give mathematical probabilities that are not relevant to this context.