Introduction to Quantum Theory and Dirac's Course

Home:
The purpose of today's Zoom discussions is to specifically discuss your views on Quantum Theory — your previous encounters with the various problems of quantum theory.

While you appear in a plethora of articles and interviews that cover various aspects of your contributions to physics, this one is specifically devoted to the fundamental aspects of quantum theory. Our discussions will go into some depth and contain popular-level material as well. Kaushik is here to cover the popular angle and I'm here to ask you some detailed questions.

Sengupta:
You took Dirac's course in quantum mechanics[1] at Cambridge during the late 1950s. Will you please elaborate on your experience?

Penrose:
It was, of course, a very elegant set of lectures from which I learned a lot. Many of my friends complained that his lectures were the same as his book [1], but I hadn't read his book at that time. So that didn't matter to me. I was studying pure mathematics at the time. I was doing algebraic geometry and I took three courses that had nothing to do with the research I was supposed to be doing.[2] One of them was by Bondi[3] on general

[1]Paul A. M. Dirac was one of the founders of quantum mechanics. His great treatise, The Principles of Quantum Mechanics, set the stage, tone, and language for this upcoming field in 1930. Many generations of physicists learned quantum mechanics through this classic book.

[2]Penrose did his Ph.D. in algebraic geometry at St John's College Cambridge. Here he attended lectures by Hermann Bondi, Paul Dirac, and S. W. P. Steen, all three of which had substantial influence on Penrose's later work.

[3]Hermann Bondi, the Austrian born British mathematician and cosmologist, is known for developing the steady state model of the Universe with Fred Hoyle and Thomas Gold as an alternative to the Big Bang theory. He was also known for his contributions to general

relativity. Nice course. Dirac's course was (stylistically) very different from Bondi's course. Bondi was very demonstrative, and Dirac's wasn't like that.

Penrose:
I was talking about Bondi. His way of speaking was like that of an actor, whereas Dirac's course was elegantly put together and very precise in what he said. But his tone of voice would be pretty soft, and I did not like it then. The third course was by the Cambridge logician S. W. P. Steen. I learned a lot from him, too. All these courses were very important to me. Bondi's, of course, because I seriously pursued general relativity and Dirac's, as I had learned a lot about quantum mechanics from him. One of the special topics that I learned from Dirac's lecture was about two-component spinors[4] which was very helpful for me because I wanted to learn, in particular, about that topic. Dirac was the one who made it clear to me. The mathematical logic course (given by Steen) was also important because it indicated that there's something different from computational physics involved in human thinking.[5]

Sengupta:
As you had recalled in your lecture/talk, in the first lecture of his course, Dirac had talked about the superposition principle of quantum mechanics— a quantum particle can be here, there, or it can be in a state which is partly here and partly there at the same time. You had then wondered, how can something be

relativity.

[4]Two-component spinors are the basic ingredients for describing fermions in quantum field theory in 3+1 space-time dimensions.

[5]Penrose first heard about the details of Gödel's theorem in a course on mathematical logic (where he also learned about Turing machines) given by the S. W. P. Steen. He was then (1952/53) a first year graduate student studying algebraic geometry at Cambridge University. He was concerned by the possibility that there might be true mathematical propositions that were in principle inaccessible to human reason. When he learned the true form of Gödel's theorem (through Steen's presentation), he was enormously gratified to hear that it asserted no such thing; for it established, instead, that the powers of human reason could not be limited to any accepted preassigned system of formalized rules. What Gödel showed was how to transcend any such system of rules, so long as those rules could themselves be trusted.

at two places at the same time? At some point during that lecture by Dirac, your mind had dozed off to a completely different thought and by the time you were back, Dirac had gone on to the next topic. You had missed his explanation. Thinking back, you now believe that this lapse of concentration was probably a "good thing," because Dirac's explanation would have put your mind at ease, and you would have stopped "worrying" about the problem, which you never did. Were the oddities of superposition principle the main reason behind your interest in the foundational aspects of quantum mechanics?

Infobox 0.1 The superposition principle

The superposition principle is one of the most fundamental concepts of quantum theory whose significance was first clearly brought out by Paul Dirac. According to this principle, any two (or more) quantum states can be added together (superposed), resulting in another valid quantum state. Conversely, every quantum state can be represented as a sum of two or more distinct states.

Mathematically, the superposition principle refers to certain nuances of the Schrödinger equation which is a first-order differential equation in time.

Penrose:
Yes. Well, he did explain, I suppose. But my mind was wandering at that time. He had a piece of chalk, and people told me that he would break it into two pieces. This is what he normally did to demonstrate how you might imagine a piece of chalk in two places at once. But that's all I could remember. Much later I learned that he also believed that quantum mechanics was not a complete theory.

The Measurement Problem and Gravitational Collapse

Home:
Tony Leggett[6] thinks the measurement problem arises because quantum mechanics does not describe the occurrence of a definite outcome, while Abner Shimony[7] says that it arises because, within the Born probabilistic rule, we talk about the probability of an outcome to be observed. But then, if an outcome is not ensured to be distinguishable from others, how can one talk about probability? So, in that sense, the measurement problem shows inconsistency in the formalism. What would be your way of briefly stating what the measurement problem is and why you feel that this indeed is a problem?

Penrose:
It took me a long time to settle on my current view. I don't remember when I first thought it must be a gravitational effect.

Home:
When did you first think that measurement problem is indeed a problem? I ask because the majority of mainstream physicists do not believe it is a problem.

---

[6]Sir Anthony J. Leggett is the British-American physicist who shared the 2003 Nobel prize in physics with two Russian-born physicists, Alexei Abrikosov and Vitaly L. Ginzburg, for their "pioneering contributions to the theory of superconductors and superfluids." Leggett was knighted (KBE) by Queen Elizabeth II in 2004 "for services to physics." Leggett was introduced to the foundational problems in quantum mechanics during the late 1960s by a colleague (Brian Easlea) at the University of Sussex.

[7]Abner Shimony was Professor of philosophy and physics Emeritus at Boston University. He published arguably the most important article of his career when he derived the Clauser–Horne–Shimony–Holt (CHSH) inequalities with his graduate student, Michael Horne, and two other graduate students (John Clauser at Columbia and Richard Holt at Harvard). The CHSH inequality is a new form of Bell's inequality amenable to experimental tests.

Infobox 0.2 Measurement Problem

The Schrödinger equation describes the evolution of a closed quantum system. The evolution is deterministic and reversible (unitary evolution) — given the Hamiltonian of the system whose initial state is known, the final state is uniquely determined by the unitary operator U because the Schrödinger equation in nonrelativistic quantum mechanics is linear. Linear, in this context, means that the sum of any solutions of the Schrödinger equation is also another solution (the principle of superposition). However, an observation — the act of measurement of a quantum mechanical system in a state of superposition with a measuring apparatus — interrupts its unitary evolution. Measurement abandons the superposition state and picks a definite state. It destroys the entanglement in the composite system consisting of the quantum system and the measuring device and collapses it into a definite (single) outcome. This dichotomy is the infamous measurement problem of quantum mechanics. It arises because of the incompatibility between the microscopic world of quantum particles and the macroscopic world of the measuring apparatus.

Penrose:
There's a lot of talk which somehow makes you get used to it and then say it's not a problem. I don't agree with that at all. I think it is a problem. My view is that for a macroscopic object you can have a superposition, but the lifetime is very short. This is because it is a gravitational effect. I can't remember when I first thought it was a gravitational effect. I think it was not immediately, but I came to that view at some point. I certainly had that view when I wrote *The Emperor's New Mind* [2], although, at that time, I didn't have quite the same point of view about this is as I have now. I think the argument that I find most persuasive has to do with the conflict between the two basic principles of physics. One of them is the superposition principle in quantum mechanics, and the other is the Principle of Equivalence in general relativity (Galileo's principle about how you locally get rid of gravity by freefall).

Infobox 0.3 The Principle of Equivalence

The Principle of Equivalence is a fundamental law in physics. It states that gravitational and inertial forces are indistinguishable and, thus, are the same thing. In the Newtonian form, experimenters within a windowless laboratory, freely falling in a uniform gravitational field, would be unaware that the laboratory is in a state of nonuniform motion. An inertial state of uniform motion is unaffected by gravity, implying the universality of free fall: in a uniform gravitational field, all objects fall with the same acceleration, e.g., 9.8 m/s^2 near the surface of the Earth. This was confirmed to a high degree of precision in an experiment conducted by the Hungarian physicist Roland Eötvös. In Einstein's version, the principle asserts that in free fall, the effect of gravity is eliminated in all possible experiments, and general relativity reduces to special relativity, as in the inertial state. Also see Section **??**.

Now I give you this argument (can't remember where I first came up with it, but a version can be found in an article in Foundations of Physics [7]). Consider a tabletop experiment where you want to take into consideration the Earth's gravitational field. But if you are a standard physicist, you would simply add a term in the Hamiltonian for the Earth's gravitational field. But then if you want to be respectful of the Galileo-Einstein Principle of Equivalence, you say the Earth's field is equivalent to an acceleration. So, you do your calculations again. You start with a freely falling frame and you transfer back to the stationary frame on the desktop, and you find that you get almost the same answer.

The key is in "the almost" — they are almost the same. The only difference between the description with the Newtonian way — which would be inserting a term in the Hamiltonian — and the Einsteinian way — which would be considering that it is just an acceleration effect — lies in the phase factor. The wave function is identical, except there is a multiplier, which is the phase factor of modulus unity. So, you might say, well, who cares because if we are only going to work with the probabilities, we are just going to take squared amplitudes.

Hence, it does not replace the phase factor and thus does not make a difference. But then you look carefully at the phase factor and see it's not benign because it involves the time cubed. It basically says that the two perspectives — the Newtonian and the Einsteinian — are equivalent, but they are starting from a different quantum field theory vacuum.[8] Again, you might say, who cares! You stick to your vacuum, and you get the same answer.

Home:
Could you elaborate, please?

Penrose:
Now, let's consider a slightly different situation where your gravitational field is due to a body which is brought into a superposition of two locations. Then you try to use the Einsteinian perspective. You find that you can't do it. Because, when you try to eliminate gravity, the free fall coordinates keep changing as you move around. It's a bit like saying that you're in Pisa dropping stuff from the Leaning Tower. You could say there is no gravity but that doesn't help if you want to talk about the gravitational field in New York — because the direction of the field is different from in Pisa. How do you eliminate them both at once?

[8] Quantum field theory perceives vacuum in space, not as space entirely devoid of matter, but full of fluctuating electromagnetic waves that can never be totally eliminated. These waves appear in all possible wavelengths, and their presence implies that empty space contains a certain amount of energy–an energy that we cannot tap into but which is always there.

Infobox 0.4 The Heisenberg Uncertainty Principle

The Heisenberg uncertainty principle is one of the fundamental postulates of quantum mechanics. This principle is typically expressed in either of two mathematical forms:

- Momentum-position uncertainty: $\Delta p \Delta x \geq \hbar$
- Energy-time uncertainty: $\Delta E \Delta t \geq \hbar$

The momentum (Δp) times the uncertainty in the position (Δx) or alternatively, the uncertainty in the energy (ΔE) times the uncertainty in the time (Δt) is greater or equal to $\hbar$ where $\hbar = h/2\pi = 1.05 \times 10^{-34}$ J s, h being the Planck constant (6.63×10^{-34} J s).

So, what I do at this point is to try taking into consideration a certain error (factor) in the calculations. The Einsteinian perspective is not used because it's hard to see what's going on. Instead, we use the Newtonian perspective and note that we're admitting an error, which we then integrate over the whole of space. Integrating by parts, we find that it causes an uncertainty rather than an error — an uncertainty in the energy of the system. We can now use *Heisenberg's time-energy uncertainty principle*, which tells us that the uncertainty is inversely related to the lifetime of the system.

So, the system behaves just like an unstable particle — it has an energy uncertainty which is inversely proportional to the lifetime. This is what I mean when I say, "there is a lifetime for this superposition," and this is the point of view that I hold at the moment: there is a fundamental uncertainty in the energy which is equivalent to the system having a lifetime. There are various ways of saying this. The most precise way is to say that the energy uncertainty is proportional to a gravitational self-energy of the difference between the two states. Another way of putting it is as follows. Suppose you have an object in a superposition of here and there. Then imagine two instances of this body and ask, how much energy would it cost me to move it away from itself

to that distance by just considering gravity and no other force? That is the quantity that you must consider as the gravitational uncertainty.

The formula is the same as the one calculated by Diösi.[9] He did his calculations several years before I did. I didn't know about it then, but I learned later that he had come to the same conclusions as I did. I think it's one of your questions.

Home:
What is the difference between the Diósi model [8] and yours [7]? Also, please explain the experiment [9] claiming to test the Diósi-Penrose model (or DP model) in this context.

Penrose:
In his model, Diósi predicts that the collapse of the wave function occurs at a specific moment [9]. If you have a body with many particles, there may be continuous collapses happening all the time. It's as if the particles are jumping around, contributing to the heat of the system. This spontaneous heating is what they aimed to measure in the mineshaft. However, in my model, there is no heating because the system does not transition from one state to another. It returns to its original state, so there is no jump and no heating.

Home:
But then how does the outcome get realized? Does it not collapse to one of the outcomes?

Penrose:
Imagine I have two small telescopes in each hand. If I were to put one of the telescopes into a superposition of two locations, in my

[9] A heuristic formula of decay time for the balanced superposition of two mass distributions was originally proposed by Penrose: $t = \hbar/E_\Delta$ were E_Δ is the gravitational self-energy of the difference between the mass distributions in each of the two locations of the object. Its form is very close to the one proposed earlier by Diösi [5], and subsequently modified by Ghirardi, Grassi, and Rimini [29].

picture, the telescope would then collapse to one location or the other as if it had always been there. This means that the object would effectively collapse to a specific history instead of another one. In this scenario, there would be no heating because it involves continuous movements without any sudden jumps that result in the heating.[10]

Home:
Would there be no timescale involved in this picture?

Penrose:
You must consider when the collapse has occurred. However, that's in the realm of quantum reality. In the classical reality, there is a retroactive aspect that it goes back to. On the other hand, in the quantum reality... yes, it does jump.

Home:
Is this an experiment where you plan to test the aspect of timescale or the retroactive aspect?

Penrose:
Time will be a factor. You need to conduct an experiment where you wait for something to happen. The goal is to be able to time when the collapse occurs. The outcome of the experiment can vary depending on when the collapse occurred. The collapse could be influenced by the environment.

Home:
So, this model is similar to spontaneous localization collapse models, such as the GRWP model. It makes testable predictions while going beyond the scope of standard quantum mechanics by modifying or supplementing the unitary formalism.

[10]Penrose's point is that in his particular interpretation of quantum mechanics, the wave function collapse is not a sudden, discontinuous jump from one state to another. Rather, the object collapses into a particular state as though it has been in that state all along. This implies that rather than the particle undergoing a sudden and violent change, it undergoes a more gradual and continuous movement. The lack of sudden jumps also means that there is no energy loss or heating that typically arises from such jumps

Penrose:
Yes.

Home:
What is your opinion on those models? They lack a mechanism because they rely on equations. They incorporate collapse into the equations and make predictions about the lifetime.

Infobox 0.5 Collapse Models

The collapse models are significant among the various concrete schemes proposed to solve the quantum measurement problem. These models tack on a mechanism that forces a quantum system to collapse when it gets too big, explaining why macroscopic objects like cats cannot be in superposition, whereas microscopic objects like electrons can. Collapse models work by modifying the Schrödinger equation, including the collapse of the wave function. The collapse is embedded in the dynamics, not just the measurement processes. The most studied collapse model is the Continuous Spontaneous Localization (CSL) model by Ghirardi, Rimini, Weber, and Pearle [10], [11]. In the GRWP scheme, a universal "background noise" is introduced — one which is not itself describable in quantum mechanical terms but crucial to the viability of the scheme.

Penrose:
That's true. They're brave attempts. I don't think I believe in any of them because they're not relativistic to start with. My considerations of why you need the retroactive aspect come from looking at special relativity. All the models of this nature that I have seen are not relativistic. I don't know of one which is relativistic. A key part is to make them relativistic and since none of these models are relativistic, they're not the answer. They could be brave attempts, and I can certainly take my hat off to them, if you like, and they're worth considering. I think one of them is...

Home:
Philip Pearle?

Penrose:
Philip Pearle, yes [12]. Phillip Pearle's work sparked my interest in this relativistic approach. In his model, he discusses how the amplitude evolves with time, where gamblers seem to have a rule.[11] However, I don't believe that approach works either.

Sengupta:
You go even a step further to say that quantum mechanics is not quite right, which is quite a devastating statement because the theory works perfectly well. You perhaps imply that quantum mechanics will need modifications, particularly in its macroscopic validity. Can you explain in what sense you say this?

Penrose:
In my view, quantum mechanics is incomplete because it lacks a clear explanation for the collapse of the wave function. Unlike others or even the likes of Einstein, Dirac, and Schrödinger, I tend to be more direct in expressing my critique. The inconsistency lies in the fact that quantum mechanics requires the wave function to "collapse" during observations, yet this process doesn't adhere to the Schrödinger equation, which is supposed to govern the behavior of quantum systems when the wave function isn't collapsing. This contradiction suggests that our current understanding of the theory is incomplete.

[11] These papers propose a stochastic spontaneous dynamical collapse model that describes the apparently random spontaneous collapse process in a suitable mathematical way. This is what Penrose may have meant by remarking, "where gamblers seem to have a rule."

Infobox 0.6 Quantum Gravity

Gravity and quantum mechanics were developed and confirmed independently in numerous experiments, but when applied together, they produce nonsense. Physicists believe a working theory of quantum gravity would resolve these contradictions by applying the rules of quantum mechanics to gravity, thereby endowing the gravitational field with the irreducible randomness and uncertainty characteristic of quantization.

For example, a full theory of quantum gravity is needed to understand what happens inside a black hole. The two leading contenders for the quantization of gravity are string theory and loop quantum gravity. While each enjoys many attractive features, neither offers a compelling argument for a correct quantum theory of gravity. Nor there is much observational evidence to guide decisive experiments. After nearly half a century of intense work, none of these efforts have given birth to a unanimously accepted quantum theory of gravity. Thus, quantum gravity remains a theorists' playground, where some ideas will undoubtedly survive while others will be relegated to the backburners of physics as mere historical curiosities [10].

I've suggested that this issue might be addressed by introducing gravity into quantum mechanics. Rather than simply "quantizing gravity" in the same way we might quantize Maxwell's equations, I believe that gravity could potentially provide a solution to the problems within quantum mechanics itself. Quantizing gravity is definitely feasible. I'm not suggesting that you shouldn't attempt to pursue this path for general relativity, as its impact on quantum mechanics will be at least as significant as the effect of quantum mechanics on general relativity. Gravity is unique because it incorporates the Principle of Equivalence—something that is not found in other theories.[12]

Historically, gravity has been foundational in the development of scientific thought, as seen in the work of Kepler and Newton and

[12]Despite almost a century's worth of study, it is still unclear how general relativity (GR) and quantum theory (QT) should be unified into a consistent theory. The conventional approach is to retain the foundational principles of QT, such as the superposition principle, and modify GR. This is referred to as 'quantizing gravity', resulting in a theory of 'quantum gravity'. The opposite approach is 'gravitizing QT' where we attempt to keep the principles of GR, such as the equivalence principle, and consider how this leads to modifications of QT.

even in the formulation of the Lagrangian and Hamiltonian mechanics. I propose that gravity might offer a distinct perspective not just for understanding cosmic scales but also for addressing some of the core challenges in quantum mechanics.

The collapse of the wave function is fundamental to how quantum mechanics operates, and I suspect that gravity's influence could be critical to comprehending this process. Although we don't yet have a well-defined theory that integrates gravity and quantum mechanics to address this issue, I see this as a significant challenge and opportunity for future advances in physics.

Home:
So, if I understand correctly, you're saying that the collapse of the wave function in quantum mechanics is actually a result of the gravitational effect caused by measurement?

Penrose:
Exactly. The idea is that when you measure something, you're interacting with a larger system, and the gravity of that system causes it to collapse. This suggests that all measurements follow this gravitational nature. However, we don't have a detailed theory that fully explains this phenomenon yet. That's where I think our focus should be, rather than solely trying to quantize gravity.

Home:
So, the key point is to develop a theory that explicitly incorporates gravity into quantum measurement because the massive nature of the measuring apparatus and the displacement required during measurement play a crucial role.

Penrose:
Exactly. In quantum mechanics, there is a lack of consistency because it doesn't provide a clear indication of when the wave function should collapse. This collapse of the wave function is also not consistent with unitary evolution. Therefore, we need

to focus on developing a measurement theory that takes into account the gravitational effects and solves these inconsistencies.

Classical vs. Quantum Reality and EPR

Penrose:
I really must discuss two types of reality first. There are, what I call, classical reality and quantum reality. The difference between the two is really explained in the Einstein criterion for reality of a quantum state.

When you have a superposition, you move a body into a superposition of two locations. Then the reduction happens, and it becomes one or the other. Now, in this picture, that effect goes back to where the thing became separated. As though it had been in one or the other all the time. It's retroactive in that sense. It becomes as though it had been all the time in one location or the other. This argument follows from special relativity. if you have a superposition, you think of the world's line of a superposition — the world gets separated into two branches and then, at a certain point, it becomes one or the other.

Now, if you take a non-relativistic view, you could say this happened simultaneously: when one branch becomes 1 all the time, and when the other branch is 0 (when the amplitude becomes completely 1 on one side, and 0 on the other side). If you have relativity, that doesn't make much sense. Because you must think of the simultaneity between one of them becoming the whole state, and the other not being part of the state. To make it relativistically invariant, you must return to the object's initial splitting into the two states.

So, it's really a condition of special relativity that requires you to go back to where the superposition initiated. This is how the

classical reality behaves. Classical reality has this retroactive aspect — when the superposition goes to one way or the other, it's as though it had been that all the time. You must make sure this doesn't lead you into a contradiction — things like grandfather paradoxes and so on. But the argument basically is that the decision must be a random decision.

Home:
What about the EPR situation?

Penrose:
You can consider EPR situations as one way of looking at it. Let's say Alice and Bob share a pair of spin-half particles. When Alice measures her particle, how does that affect Bob's particle? Well, you have to go back along the past light cone of Alice's measurement, and then Bob's particle has a well-defined state resulting from Alice's measurement. However, this doesn't provide a way

Infobox 0.7 The Einstein-Podolsky-Rosen (EPR) Arguments

The Einstein-Podolsky-Rosen (EPR) arguments, formulated in their 1935 paper titled [13] *Can Quantum-Mechanical Description of Physical Reality be Considered Complete?* raised fundamental questions about the nature of quantum mechanics, and in doing so, highlighted the peculiar concept of entanglement where two or more particles become intrinsically linked, regardless of the distance between them.

The main thrust of EPR's argument was to challenge the completeness of quantum mechanics as a description of physical reality. They posited that if quantum mechanics were complete, then every element of physical reality would have a counterpart in the theory.

Infobox 0.8 The Einstein-Podolsky-Rosen (EPR) Arguments. Contd.

EPR introduced a criterion for reality: *If the value of a physical quantity can be predicted with certainty without disturbing the system, then there must be an element corresponding to that physical quantity in the theory.*

The EPR argument is illustrated through a thought experiment involving two particles that have interacted and then separated. According to quantum mechanics, such particles remain entangled, meaning that the measurement of one particle's property instantly determines the corresponding property of the other particle, no matter how far apart they are. This seemed to violate the principle of locality, which states that physical influences cannot travel faster than the speed of light. This non-local interaction seemed paradoxical to Einstein, Podolsky, and Rosen.

EPR favored "local realism," the idea that physical processes occurring at one location should not instantaneously affect outcomes at another location (locality), and that every element of reality is well-defined irrespective of observation (realism). The EPR scenario seemed to violate this concept, suggesting either quantum mechanics was incomplete or it implied faster-than-light influence, which contradicted relativity.

The crux of the EPR paradox lies in the phenomenon of quantum entanglement, where the quantum states of two or more objects are intertwined in such a way that the state of one object directly depends on the state of the other(s).

The EPR paper did not disprove quantum mechanics but rather sparked significant debate about its interpretation. The most profound aspect highlighted by EPR was entanglement (a term later popularized by Schrödinger). Entanglement challenges classical intuitions about separability and locality, suggesting that quantum particles can exhibit correlations that defy classical explanations.

In the 1960s, John Bell developed Bell's Theorem, which provided a testable criterion to distinguish between quantum mechanics and local hidden variable theories. Subsequent experiments, especially those by Alain Aspect and others, largely affirmed the non-local predictions of quantum mechanics, thus reinforcing the reality of entanglement and non-locality in the quantum realm. They also demonstrated that the correlations between entangled particles cannot be explained by any local hidden variable theory.

Infobox 0.9 The Einstein-Podolsky-Rosen (EPR) Arguments

In summary, the EPR arguments were pivotal in highlighting entanglement, a central feature of quantum mechanics, and in stimulating ongoing debates and experimental investigations into the nature of reality, locality, and the completeness of quantum theory. These arguments have led to significant advancements in quantum information theory, quantum cryptography, and quantum computing.

to predict the future. It also doesn't lead to contradictions because it only affects quantum reality. The distinction between classical and quantum reality is this: with classical reality, you can ask the system what its state is.

For instance, consider this small telescope. If I ask it what its shape is, I can examine it and note, "It's roughly cylindrical, with a circular top and bottom, and some surface irregularities." In this classical scenario, the telescope has clear and distinct properties that can be measured and stated without any uncertainty."

Home:
How about quantum reality?

Penrose:
The concept of quantum reality can be illustrated with the typical example of quantum mechanical spin. If I have a spin-half particle, I'm not allowed to ask what its state is (I can ask but won't get an answer). However, if I have a theory which tells me what I think its state should be, I can ask, "Is that your state?" Thus, you can ask to confirm the state. This is Einstein's criterion for an "element of reality." If, by not disturbing the system, you can confirm with probability 1 your suggestion of what the state is, that gives an element of reality to that state. The Einsteinian element of reality is the element of quantum reality. If this was a spin-half particle, and the spin happens to be, say, along that axis, then if my theory tells me the spin should be about that axis, I can measure the spin in that direction, and it will give me

a complete "yes" or "no" answer to that question.

If I don't have a specific theory and I try a different direction, I'll only get a probabilistic answer. In other words, I won't be able to determine the state with certainty. For quantum reality, you can't ask the system what its state is and expect a 100% answer. You can only confirm what you think the state is with 100% correctness. This is the criterion for quantum reality — you can confirm the state, but you can't directly ask what the state is and expect an unambiguous answer. Quantum reality seems to occur when a measurement is made. However, classical reality reverts back, and that's where space-time becomes involved.

This point is illustrated in a diagram in my latest article that was submitted for publication in the book called *Consciousness and Quantum Mechanics* edited by Shan Gao. The name of the article is *New Physics for the Orch-OR Consciousness Proposal* [26].

Sengupta:
You gave a series of lectures along with the late Stephen Hawking, which culminated in a debate about how universal and applicable the theories of quantum mechanics and relativity are (the lectures and debate were published in the 1996 book, *The Nature of Space and Time* [11]. In Chapter Seven (The Debate), Hawking remarked, "Reality is not a quantity you can test with litmus paper." While realism cannot be tested alone, we know that it can be tested in conjunction with other assumptions. For example, John Bell showed how to do this by combining it with the notion of locality. Leggett-Garg showed how macrorealism, combined with the notion of non-invasive measurements, can be tested. Hence, do you consider Hawking's statement strictly accurate?

Penrose:
I'm not sure what I said at that time. It was quite interesting because Stephen Hawking had difficulty speaking. He wanted

to know my questions in advance so he could prepare over the weekend. It wasn't as spontaneous as debates usually are.

Regarding your question, I have a slightly different answer for the "litmus paper" question. I believe that classical reality can be measured with litmus paper, but quantum reality cannot. I didn't have this distinction between the two realities at the time, so I'm giving you my current answer to that question.

Infobox 0.10 Macrorealism

Macrorealism (MR) is a classical notion in the philosophy of physics and quantum mechanics that encapsulates the idea of realism extended to macroscopic objects. It builds upon two key assumptions:

Macroscopic Realism (MR): A macroscopic system that is observable (large enough) will always be in a definite state irrespective of whether or not it is interacting with or observed by an external system. This means that macroscopic objects have properties that exist and are well-defined independently of measurements.

Noninvasive Measurability (NIM): It is possible, at least in principle, to determine the state of a macroscopic system without disturbing the system in any significant way. In other words, the measurement process can be noninvasive so that it does not alter the state of the system.

Combining these principles, the classical notion of macrorealism implies that:

- A macroscopic object possesses definite states at all times.
- These states can be measured without affecting them.

Macrorealism stands in contrast to quantum mechanics, which suggests that the properties of microscopic (and even macroscopic, under certain conditions) systems may not be well-defined until they are measured and that the act of measurement can influence the system's state. Concepts such as superposition and entanglement in quantum mechanics challenge the classical notion of macrorealism. One famous example highlighting this tension is Schrödinger's cat thought experiment, illustrating how quantum properties challenge our classical understanding of macrorealism.

It really depends on what kind of reality you're measuring. In classical reality, you can ask the system, "What is your state?"

and it will tell you whether it's acidic or alkaline, like litmus paper. However, in the case of a quantum state, you must know ahead of time what you think the state is. It's not exactly like litmus paper; it's the kind of reality that you can only confirm rather than request what it's like.

I didn't have that understanding at the time, so if I had to answer it now, I might have approached it slightly differently in terms of these two very different types of reality. It's more like the classical reality that you can use litmus paper for. However, this is not a perfect example because you might think that acidic and alkaline are analogous to spin up and spin down. Then, you could ask whether it's spin up or spin down because, in quantum mechanics, you have linear combinations of properties. So, you run into this difficulty even though it gives you a distinction between classical and quantum reality.

Home:
Is the difference between quantum reality and classical reality essentially due to the superposition principle?

Penrose:
This concept originates from that idea. Superpositions can have just as much reality. For example, you can have "this state plus that state" and "this state minus that state." You can have state A plus state B, and you can also have A minus B. These are orthogonal alternatives, so I can distinguish between those two. However, this is different from asking whether it's in a superposition of here or there. It would be interesting to consider that.

Yes. In order to understand why someone may have a different perspective than mine, which might seem extreme because it involves retro-causality,[13] it is crucial to distinguish between clas-

[13]Retrocausality is the idea that effects can precede their causes, challenging the traditional notion of causality. It is an intriguing but highly speculative notion that challenges our fundamental understanding of time, causation, and the natural order. Considered in

sical reality and quantum reality. While classical reality follows the laws of space-time, quantum reality has some retro-causal features that allow it to travel back in time. However, this doesn't necessarily mean that it's transmitting information in the traditional sense.

Home:
Could you give an example of how this distinction plays out?

Penrose:
Certainly. I wrote an article in a book titled *Consciousness and Quantum Mechanics* [3]. In the article, I discussed the differences between classical and quantum reality. I explained that in classical reality, you can ask the system what its state is, and it will tell you. However, in quantum reality, you cannot ask that kind of question. You may have an idea of what the state is, and you try and confirm it. The system will not tell you what its actual state is. If you could, you would violate the speed of light transmission. In other words, you would be able to send a signal back in time along the past light cone. However, this is not the transmission of a signal in the ordinary sense but rather the transmission of quantum reality back in time. It is important to have this understanding when discussing EPR situations (see Box # 0.7). Later, I regretted not having included this information in my article.

Home:
Interesting. So, how does this translate into the EPR scenarios?

Penrose:
In my article, I discussed the quantum and classical realities, but I do not talk about the EPR situation . However, the EPR example

---

various contexts, including quantum mechanics and philosophical discussions about time and fundamental laws, its examples include quantum interpretations like the transactional interpretation and time travel narratives in science fiction. Since the concept violates causal ordering, retrocausality can lead to paradoxes and challenge conventional interpretations of physical laws. The notion remains speculative and controversial, with no definitive proof or widespread acceptance in mainstream physics.

is a great way to explain the concept. Let's take the situation with spin $\frac{1}{2}$. We have a spin 0 state shared by Alice and Bob, and each of them has a spin $\frac{1}{2}$ state (this is possible because it initially came from a spin 0 state with Bob's state the opposite of Alice's state). When Alice measures her state, the results of her measurement are transmitted along the past light cone and change Bob's state to be the opposite of what Alice has measured.[14] It may seem like a message is being transmitted faster than light and going back in time, but it is actually just the transmission of quantum reality. Bob can only confirm by saying, "Is your state such and such?" If Alice gets it right, then the state will be a YES state. That is the most that Bob can see, and if he gets it wrong (NO state), it would only be a probabilistic answer. It is not an answer that will enable Bob to ascertain the state. If he could, then Alice would be able to send a signal back in time, but that's not possible because if that happened, we would rapidly run into a contradiction.

Home:
How does Einstein's notion of reality fit into this?

Penrose:
Einstein's notion of reality, commonly referred to as Einstein's dictum, comes into play in the example of the quantum realm. There have been many debates over whether the quantum state is indeed real in any sense, but Einstein argued that an experiment that provides an answer without disturbing the state gives an element of reality.

Thus, quantum reality can be characterized by Einstein's criterion. When you ask the system what it is, you can only confirm the state, which is quite curious. If Alice had not transmitted

[14]This phenomenon is known as quantum entanglement and has been observed in various experiments. It is a peculiar property of quantum mechanics that violates the classical notion of locality and suggests that particles can be connected in a way that transcends space and time.

her result to Bob, he wouldn't know what the quantum reality of his state is because it's only quantum reality. When he asks questions, he gets probabilistic answers.

Home:
So, there's a clear distinction?

Penrose:
Yes, there is an interesting distinction between quantum and classical reality. In classical reality, you can ask the state what it is, and it will tell you what its state is. Classical reality is what space-time determines. You could have a space-time in a superposition, but it will become one or the other after a certain period of time. Then, it will be as if that was the state all the time until it returned to the bifurcation point. I think it's entirely consistent. However, my article contains only half of my current arguments. I have written up the rest of my arguments in some notes, which I sent to quite a lot of people, but I am not sure if I sent them to you.

I recently gave a talk at the Oxford physics department. During the talk, I discussed classical and quantum reality. Someone in the audience mentioned John Bell, who had his own ideas about reality. However, I don't believe Bell's ideas align with mine. When I asked the person about Bell's notions, they couldn't recall the details.

Home:
As far as I know, Bell had the idea of physical correspondence, which means that what is observed should correspond to our perception. He discussed psycho-physical perception and famously stated, "What is interesting unless it is observed?" Even in classical physics, the observer plays a role. It's a beautiful quote. Let me read it to you – I have it here. I find it very intriguing, and it's one of my favorite quotes.

Penrose:
Has this been printed somewhere?

Home:
Yes, I am going to quote from Bell's book *Speakable and Unspeakable in Quantum Mechanics* [4, p. 125]. I thought it should be underscored more, especially in the context of quantum measurements. Bell says, "It could also be said that even in classical mechanics, the human observer is *implicit*, for what is interesting if not experienced? But even a human observer is no trouble (*in principle*) in classical theory—he can be included (in a schematic way) by postulating a "psycho-physical parallelism."—i.e., supposing his experience to be correlated with functions of the coordinates. This is *not* possible in quantum mechanics, where some kind of observer is not only essential but *essentially outside*." (italics ours)

Penrose:
That's a different viewpoint. I don't talk about observers at all. The distinction is not dependent on the observer. It's a different view altogether. Regarding an observer, is it a conscious being?

Home:
According to Bell, in quantum mechanics, the concept of an observer is not explicitly defined within the theory. However, it is necessary to have something within the theory that corresponds to observers and their conscious perception of outcomes. This is why the collapse mechanism, or reduction, is incorporated into the equation.

Penrose:
(18:20) I'd have to look into this. It doesn't sound very much the same.

Home:
I agree; it's not the same because you considered the Einstein-

Podolsky-Rosen (EPR) element of reality. To answer this question fully, it would be helpful to reference your article on consciousness and quantum mechanics and Stuart Hameroff's work.

Penrose:
I think it's very helpful to consider the EPR situation. Quantum reality does retroactively extend along the past light cone. It is intriguing because classical reality extends along the future light cone.

Home:
So, the principle of induction does not hold? Does causality operate backward in time?

Penrose:
Yes, but it's not strictly backward in time; it follows the path along the past light cone. This pathway affects reality, specifically quantum reality. If we were able to determine quantum reality clearly, a significant issue would arise—it would enable faster-than-light signaling, which defies the principles of special relativity.

What we typically consider a signal is classical information. However, in this context, we'd be dealing with a quantum information signal. Whether this equates to what others refer to as quantum information is another question I am uncertain about. It might indeed be quantum information, but I haven't confirmed if that terminology applies correctly here. It's certainly not classical information, though; I can confirm that.

Home:
However, in the derivation of Bell's inequality (Bell's theorem), the assumption involves causality progressing forward in time.

Penrose:
You could approach it that way. I need to review how the Bell inequalities operate in this context. Exploring this could indeed

prove to be intriguing.

Objections to Copenhagen Interpretation

Sengupta:
Let's shift our focus and discuss your main objections to the standard interpretation of quantum mechanics, known as the Copenhagen interpretation.

Penrose:
I can't recall the exact details of the Copenhagen interpretation, but I know it offers an explanation. Imagine you have a quantum state that becomes entangled with the environment. Initially, your understanding of reality is based on your quantum state (which always represents a quantum reality). However, as the environment is very complex, it becomes difficult to discern what's actually happening.

We choose to describe reality using the density matrix or density operator instead of the state vector for convenience.[15] This may seem like cheating because we are changing our ontological view. Originally, we believed reality is described by the wave function, but now we are shifting our perspective. When we adopt this new view, we can adjust our perspective until the density matrix becomes diagonal. We can then shift back to our original perspective and interpret these diagonal elements as probabilities of different alternatives. This allows us to view the density matrix as a probability mixture of these different alternatives.

Your ontological interpretation of the mathematical description

[15] Penrose explains this further in Ref. [14]: "According to the 'Copenhagen' viewpoint of Niels Bohr, the state vector is not considered to represent a quantum-level 'reality', but merely the 'state of mind' of the experimenter. On this view, the 'jumping' that occurs in the procedure of state-vector reduction is considered to be merely the result of a discontinuous change in the state of knowledge of the experimenter and not a physical change to which can be attributed a physical reality."

has changed twice, which makes it inconsistent. First, you said the wave function describes reality; then, you said the density matrix describes reality. This shift in interpretation, though clever, leads to inconsistency, as you then revert back to saying it's not a probability mixture of wave functions with probabilities given by the diagonal elements in the density matrix. This inconsistency doesn't provide a clear picture of how collapse takes place. An alternative explanation is needed for a more consistent description.

Home:
Can you provide further details regarding the alternative explanation?

Penrose:
The way I see it, the collapse of the state (reduction of state) is a physical process. You can find the probability by taking the squared modulus of the amplitudes associated with each eigenstate. However, this doesn't provide a consistent ontology. So, what does it actually mean to make a measurement?

You see, this is the problem: Making a measurement usually involves taking a measuring device out of the cupboard, noting the dial reading as A, B, or C, and then putting the device back into the cupboard. However, this approach is flawed because the measuring device is not being treated as a quantum device but rather as a classical one. Even though you get a good answer, you've essentially cheated.

I believe the idea is based on accepted physical principles from other areas of physics. However, it doesn't provide a complete theory with an equation that describes how the system behaves. It doesn't go that far.

Home:
You said that the "inconsistency doesn't provide a clear picture

of how collapse takes place." Could you explain this please.

Penrose:
Many people attempt to comprehend quantum mechanics without the collapse of the wave function. The argument, possibly from von Neumann or Landau, goes like this: In a quantum experiment involving the environment, the environmental degrees of freedom become entangled with the experiment, making the situation very complicated. To deal with this, a mathematical construct called the density matrix is used to describe the combined system of the experiment and the environment. The density matrix accounts for the degrees of freedom in the environment. By manipulating the density matrix, it can be transformed into a diagonal form, where the diagonal terms can be interpreted as probabilities. This process represents a shift in our understanding of reality, as we transition from considering the quantum state as real to regarding the density matrix as real due to the entanglement of the quantum state with the environment. I regard this as a double ontology shift.

Home:
Could you elaborate, please?

Penrose:
In other words, I have transitioned my understanding from the concept of state vector reality to density matrix reality and then back to a quantum probability mixture of quantum state realities, which leads to the design matrix and the associated probabilities for the collapse of the wave function. While this approach is practical for describing the collapse of the wave function, it's important to note that it does not provide a consistent view of the real world. It is a useful framework, and I don't oppose its use. It provides a way to account for the collapse of the wave function and discuss related concepts, but it's not a consistent representation of reality. You shift your perspective twice: first from state vector

reality to density matrix reality, and then back again from density matrix reality to a different probability mixture of raised states. Using the density matrix description allows you the freedom to switch to a different understanding of reality. The state vectors that were previously used to describe reality are no longer the states used to describe reality; a different set of states is used. It's a practical approach that provides a stable foundation, like finding a comfortable place to sit after climbing a mountain.

Alternative Interpretations: Bohmian Mechanics and Many-Worlds

Home:
You taught at Birkbeck College. Did you interact with David Bohm during your time there? We are interested in hearing your views on the Bohm model.

Penrose:
I remember being at Birkbeck College with Basil Hiley[16] and David Bohm. While I was in the maths department, they were in physics. We used to have regular discussions. I knew David Bohm reasonably well at that time. I heard about his views but wasn't too happy with them. He preferred determinism; he believed that the system somehow knew which way it was going to collapse based on something from its past that would determine its future direction. I heard a lot about these discussions but wasn't entirely happy with his views.

[16]Basil Hiley was David Bohm's longtime collaborator and co-author of his final book, "The Undivided Universe." Hiley is a theoretical physicist at Birkbeck College of the University of London. His interview with George Musser, a contributing editor at Scientific American, can be found here: [0].

Infobox 0.11 Hidden Variables

The American theoretical physicist David Bohm is known for his *hidden variables* interpretation of quantum mechanics (Bohmian mechanics, as it is called) that expanded upon the paradox famously introduced by Einstein, Boris Podolsky, and Nathan Rosen (EPR). Bohm had independently developed his model, being unaware of Louis de Broglie's 1926-1927 work. He became aware of it only after he sent the preprints of his 1952 [15], [16] paper to de Broglie and followed up by acknowledging it in the introduction of his paper. Bohmian mechanics is an alternative way to look at quantum phenomena, and his insights into hidden variables and non-locality in quantum mechanics led to John Bell's famous 1964 inequality.

Unlike the widely accepted Copenhagen interpretation, which suggests that a particle's position and momentum are defined only when measured, Bohmian mechanics argues that particles always have definite positions. It offers a deterministic and realist interpretation of quantum mechanics, contrasting with the dominant Copenhagen interpretation. It posits particles with definite trajectories guided by a pilot wave, thus reintroducing the classical concept of trajectory into the quantum realm. While statistically identical to standard quantum mechanics, Bohmian mechanics eliminates the measurement problem by eliminating the need for wave function collapse. Nonlocality is a central feature of Bohmian mechanics, arising from the influence of a quantum potential, Q.

When Bohm proposed his new quantum theory, most physicists were convinced of the completeness and finality of the standard interpretation. There was little motivation or value in researching an interpretation alternative to the standard one. So why did Bohm choose such a risky path? The answer was given by Bohm himself in the abstract of his classic 1952 paper—*A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. I*—in Physical Review [15]. With his customary clarity Bohm starts his paper with these elegant lines:

> The usual interpretation of the quantum theory is self-consistent, but it involves an assumption that cannot be tested experimentally, viz., that the most complete possible specification of an individual system is in terms of a wave function that determines only probable results of actual measurement processes. The only way of investigating the truth of this assumption is by trying to find some other interpretation of the quantum theory in terms of at present "hidden" variables, which in principle determine the precise behavior of an individual system, but which are in practice averaged over in measurements of the types that can now be carried out. In this paper [...] an interpretation of the quantum theory in terms of just such "hidden" variables is suggested.

Infobox 0.12 Hidden Variables Contd.

Bohmian mechanics derives its equations from the Schrödinger equation, introducing a quantum potential (Q) alongside the classical potential (V). The guiding equation:

$$\vec{v} = \frac{1}{m}\nabla S. \tag{1}$$

determines the particle's velocity (v) based on the gradient of the phase (S) of the wave function. The equation of motion:

$$\frac{d\vec{p}}{dt} = -\nabla(V + Q). \tag{2}$$

describes how the momentum (p) changes under the influence of both classical and quantum potentials.

The key features of Bohmian mechanics are:

- **Time evolution:** Governed by the continuity equation and the guidance condition, demonstrating that the probability density (P) evolves according to $|\psi|^2$.
- **No wave function collapse:** Measurement is a dynamic, many-body process, eliminating the need for an ad hoc collapse mechanism.
- **Non-locality:** Manifested through the quantum potential, demonstrating interconnectedness between particles even at large distances.

Bohm's theory did not gain the acceptance he desired. However, it was not entirely overlooked; prominent physicists like Einstein, Pauli, and Heisenberg authored articles explaining their objections to it. Despite initial criticism, the theory remains a provocative alternative interpretation of quantum mechanics, encouraging continuous research and debate about its implications for our comprehension of the quantum realm.

I recall a conversation with Bohm and was curious about his theory. Interestingly, his way of answering questions was similar to a wave function. When you asked him a question, he would provide a clear answer, but then he would start branching out into different topics, such as linguistics, without addressing the original question. It was as if his response would expand infinitely, and if it became too broad, you had to ask another question to bring it back to the point. He would then provide another clear

answer, only for the process to start over again.

Infobox 0.13 Many-Worlds Interpretation of Quantum Mechanics

The Many-Worlds Interpretation (MWI) is a cornerstone of quantum mechanics, proposing that the Universe branches into multiple parallel Universes whenever a measurement is made. This interpretation suggests that all possible outcomes of a quantum measurement occur but in different parallel branches of reality.[a]

The key tenets of MWI are:

1. **Unitary Evolution:** The MWI posits that the state vector of the Universe evolves strictly according to the Schrödinger equation. This implies that the Universe is inherently deterministic.
2. **Measurement and Branching**: When a measurement is made, the state vector of the system and the measuring apparatus become entangled. This entanglement leads to the branching of the Universe into multiple parallel realities, each corresponding to a different possible outcome of the measurement.
3. **Observer-Independent Reality**: The branching of the Universe is an objective process that occurs independently of any observer. This is in contrast to other interpretations of quantum mechanics that rely on the concept of a conscious observer to collapse the wave function.

The implications of MWI are:

1. **Parallel Universes:** A vast number of parallel Universes exist, each with its unique history and outcomes of quantum measurements.
2. **No Collapse:** Unlike other interpretations that invoke a wave function collapse, the MWI does not require such a process. All possible outcomes of measurement exist in different branches of the Universe.
3. **Determinism:** The MWI is a deterministic theory, as the evolution of the Universe is governed by the Schrödinger equation. However, branching the Universe into multiple realities can give the illusion of randomness.

[a] Hugh Everett wrote the original paper introducing the Many-Worlds Interpretation in 1957 [17]. An early paper discussing the MWI in relation to measurement problems was by H. D. Zeh [18]. A more recent book providing a comprehensive overview of the MWI can be found in the book by D. Wallace [5].

Infobox 0.14 Many-Worlds Interpretation of Quantum Mechanics. Contd.

Penrose's critique of the MWI often revolves around the role of consciousness and the nature of reality [14]. By suggesting that the various outcomes of quantum measurement are entangled with different states of an observer's consciousness, Penrose highlights a key implication of MWI — that parallel consciousness states coexist for every possible measurement outcome.

Penrose has consistently argued that consciousness might involve non-computable processes that conventional physics cannot fully explain, including the quantum mechanics underlying the MWI. If consciousness indeed involves processes beyond what current physics can describe, the entanglement of consciousness states across parallel Universes might not be a complete account of conscious experience.

The MWI suggests a proliferation of realities, each as "real" as the next. Penrose sees this as problematic because it challenges the notion of an objective reality that can be empirically observed and verified. If every measurement results in a different reality and numerous versions of an observer's consciousness, it complicates our understanding of a shared, empirical world.

The interpretation does not address the measurement problem in a manner that necessitates an intervention on the part of an observer. Instead, it sidesteps the issue by suggesting that all outcomes happen simultaneously in separate branches. However, it might not clearly explain why we experience one particular result as opposed to witnessing multiple at once.

Ultimately, Penrose advocates for interpretations of quantum mechanics that more fully account for the unexplained phenomena of consciousness and reality as observed by a singular consciousness, thus expressing skepticism about the stretch of MWI to adequately integrate these complex issues. His reservations are inclined towards theories that bridge the gap between quantum mechanics and general relativity, potentially via yet-to-be-understood phenomena such as quantum gravity or non-computable processes in consciousness.

Home:
Because of his diverse interests, I guess?

Penrose:
That's right.

Home:
So, after all these exercises and experimental tests, if quantum superposition continues to be found to be true and the collapse models are ruled out, how do we proceed? What are your thoughts on the Many Worlds idea?

Penrose:
I used to have a rather unconventional perspective on this. If you entertain the idea of the Many Worlds theory at any juncture in your life, the shorter the duration, the better. At one point, I was drawn to this idea, though my interest was brief, and I can't remember exactly how short. It's a natural outcome of resisting any changes to quantum mechanics, which works well as it is, but ultimately, it's not a very fulfilling solution. I used to worry about what would happen to all my friends. Why would my friends follow the same path if my consciousness follows a particular path through my decisions and choices? They would be veering off into a different Universe (a rather unsettling thought). It's not clear how this relates to one's conscious experiences. Do they follow a single path in these branching alternate Universes?

Home:
But once an outcome is realized, it is surely realized in one of the branches, isn't it?

Penrose:
I'm not sure whether they put you in one branch or whether you're spread out in all the branches. Consciousness doesn't seem to work like that; it doesn't spread out in that manner. I've never been convinced by that idea. However, it is a very popular concept among philosophers at Oxford.

Home:
In your debate with Hawking, you raised an interesting question. The wave function of a system representing a Schrödinger cat

system is a superposition of live + dead. One can also write it as (live + dead) + (live – dead), because this kind of superposition is possible in quantum mechanics. Then, if one could observe a live or a dead cat, and if there is also a possibility for distinguishing between a live + dead cat, the collapse would occur on that basis. Therefore, the choice then boils down to the preferred basis we are living in or making measurements. In that sense, your way of gravitation is that it puts a choice motivated by gravitational considerations.

Penrose:
Yes, so that would be the argument. It's the gravitationally well-defined space-times that give you the preferred basis. I'm not sure if I was thinking in terms of a preferred basis, but that would be the kind of argument.

Infobox 0.15 Space-time

The equations of relativity show that both space and time coordinates of any event must be combined mathematically to accurately describe an event. Space-time is a 4-dimensional entity since space consists of 3 dimensions and time is 1-dimensional. It is believed to be a 'continuum' because there are no missing points in space or instants in time, and both can be sub-divided without any apparent limit in size or duration. Physicists consider our world to be embedded in a 4-dimensional space-time continuum, and all events, places, moments in history, actions, and so on are described in terms of their location in space-time.

Home:
When I talk to supporters of the Many-Worlds idea, they say that the different branches coexist but are not allowed to communicate with one another. One observation can be made in one branch, and another observation can be made in another branch. Does that observation correspond to a live + dead state, as opposed to a live - dead state?

Penrose:
I'm not sure.

I attended philosophers' meetings occasionally and engaged in discussions with them.[17] However, I eventually stopped participating. I recall having debates with two philosophers and an American colleague, although I can't quite remember who that was. I expressed my frustration with the philosophers' seeming lack of effort in correcting physicists when they don't have a clear understanding of the world. I felt that they were not fulfilling their role.

Quantum Information and Quantum Computing

Sengupta:
We saw an article of yours in which you discussed whether the phenomenon of quantum state reduction could have a positive role in quantum computation. The paper came out in 1998 [14]. A lot has happened since then in the world of quantum computers. Have you refined your views even further?

Penrose:
I still don't know. Well, what I'm saying, in some sense, is this. There are many problems in mathematics, for example, which seem to go beyond [axiomatic steps]. You have a set of axioms, and then you can always see how to transcend those axioms. We use our understanding that allows us to transcend those axioms. Now, if you had a device which somehow could provide its own

[17] Penrose does not provide a direct answer to this question. But in [14], he does explain his viewpoint by saying the state vector in the 'many-worlds' interpretation, whose evolution is strictly unitary in accordance with the Schrodinger equation, 'is taken to represent reality.' Thus, in a measurement, all outcomes coexist, but they are each entangled with different states of the observer's consciousness, which all coexist, but each one experiences a different 'world' and encounters a different outcome of the measurement.

understanding of the system, then that could be very different. Now, is that a quantum computer in some sense? I don't know. Such a machine must be able to transcend computing in the ordinary sense. Transcending that must also involve the collapse of the wave function.

Home:
It goes back to understanding quantum computers, which is different from so-called AI computers, isn't it?

Penrose:
Well, I don't know. I always think of AI as artificial cleverness rather than intelligence.

Home:
Artificial intelligence, certainly. Is intelligence the broad understanding?

Penrose:
That would be my use of words, yes. In the idea of quantum computers as it currently exists, I don't know how they address the question of measurement. My understanding is that the measurement would always take place at the end.[18] After performing quantum computing, a measurement is taken of the system. However, if there was a way to incorporate a component within the machine that could harness the collapse of the wave function, it might be possible to achieve something extraordinary.

[18] In [19], Penrose notes that a quantum state reduction (or measurement) occurs at the end of a computational process when the result of the computation must be ascertained or be known. But the system must be probed occasionally (and therefore requires multiple measurements) since there is no good way of knowing whether the end of the computation has already been achieved or, in his words, whether the computation has 'halted.' Moreover, in quantum computation, procedures for 'purifying' the states used in the 'entanglement distillation', 'syndrome extraction', 'quantum error correction' etc., themselves make use of quantum state reduction.

Experiments and Future Directions

Penrose:
Ivette Fuentes is proposing an experiment to test the state reduction scheme that I'm proposing, using the time scale suggested initially by Diósi [20]. The timescale is the same, but the overall scheme isn't. The idea is to use this experiment to test the state reduction scheme.

Home:
What kind of experiment is she planning?

Penrose:
Ivette plans to utilize Bose-Einstein condensates and put them in a superposition of two locations at once. The trouble is that the current technicalities are too challenging, and you cannot perform the experiment directly under current conditions. It is impossible to place Bose-Einstein condensates in a superposition of two locations and then determine their specific location.

Infobox 0.16 Bridging Quantum Mechanics and General Relativity: Insights from Bose-Einstein Condensates. Contd.

If gravity's role in collapsing quantum states is confirmed, it could alter our understanding of quantum phenomena, possibly affecting quantum computing and our comprehension of concepts like entanglement. Philosophical implications suggest that we might be active participants in shaping the Universe through observation, making our role in the cosmos essential for studying these phenomena.

This confluence of ideas could lead to groundbreaking new physics, uniting quantum mechanics and relativity under certain conditions and posing new questions about the nature of observation and reality.

Overall, the paper presents an intriguing and potentially testable approach to addressing the long-standing challenge of unifying quantum mechanics and general relativity. While further research is needed to fully explore the implications of this proposal, it offers a promising avenue for understanding the fundamental nature of our Universe.

Home:
That's fascinating. Could you explain more about your argument from the late 1990s?

Penrose:
The argument aimed to show a conflict between two basic principles: the Principle of Equivalence (the Galileo-Einstein Principle of Equivalence) in general relativity and the principle of superposition in quantum mechanics. Imagine conducting a quantum experiment within Earth's gravitational field.

The Earth's field can be approached in two ways: First, you can think of it in a Newtonian way, where you consider the Earth's field as a force and include a term in the Hamiltonian for it. The other approach is to view it as an acceleration effect, using freely falling coordinates and stating there's no gravitational field.

Home:
What happens when you calculate using both approaches?

Penrose:
When you calculate using both approaches, you find that the answers are almost the same, but not quite, due to a difference in phase factor. This phase factor indicates two different quantum vacuums—one for each approach— yield different results. Initially, this may not seem important since your vacuum stays the same and you see no difference.

Then, I adjusted the experiment to involve a massive body in a superposition of two locations, adopting the Einsteinian-Galilean perspective. The problem is you can't do it consistently because your frame of reference gets superposed in two locations. Thus, you can't consistently talk about your coordinates. To proceed, I argue it is wrong because the Principle of Equivalence isn't being fully considered. However, we can estimate the error and integrate it over space, which results in energy uncertainty for the

system. By using the Heisenberg time-energy uncertainty principle, you then get a lifetime for that system. This suggests the quantum superposition of different gravitational states is unstable and has a definitive lifetime, which matches the Diósi lifetime.

Diósi worked this out independently about two years before I did. Based on his model, I guess collapse happens instantaneously, while in my view, the collapse relates to the bifurcation point.

If a relativistically invariant picture exists, with superposed locations transcending space-time where one ultimately disappears, and the other becomes a reality, you end up with a problem. It's not Lorentz invariant since different observer frames would see the collapse differently.

To resolve this, you must go back to the bifurcation point for consistency. In classical terms, this means reality goes back to that point, as though it had always collapsed, despite the collapse having a lifetime. This eliminates heating because there's no jump in the state of the system. I found trouble with the heating element anyway, as it seems to affect general relativity because a state jump disrupts Einstein's equations, potentially causing issues with objects like neutron stars due to the quantum collapse.

Home:
Because it goes back to the bifurcation point, there is no heating?

Penrose:
Correct. It's a strange concept indeed. I already had this viewpoint when the discussions came up. When people measured the temperature in the Gran Sasso mineshaft in Italy, I didn't expect any heating there.

I had already formed this point of view rather vaguely. I don't think I had emphasized my idea anywhere. I believe the article was already written by then (in this Shan Gao book). In this book, I mentioned my ideas about heating. What's not in the article is

the EPR argument. It was already becoming too long.

Let me come back to Ivette Fuentes's work. She has developed a method based on Bose-Einstein condensates to detect what we refer to as the "shaking of the building." This shaking is a result of the instability of the state, which is akin to an unstable particle with uncertainty in mass. The mass-energy uncertainty is linked to the particle's lifetime. Fuentes aims to conduct an experiment to observe the effects of this shaking (instability) prior to its collapse. She came across Oppenheim's paper [21] while working on her article, which initially concerned her as it seemed to anticipate her work. However, she eventually concluded that this was not the case. However, I cannot confirm this as I haven't studied the Oppenheim paper extensively.

Home:
The paper suggests that modifying and introducing some randomness in the coupling between classical and quantum space-times within classical general relativity can lead to decoherence in the quantum system. The authors explain that if the classical field interacts with the quantum system through a stochastic process, the classic space-time metric cannot be ruled out by the Feynman-Aharanov double-slit experiment. The author also mentions that "the quantum degrees of freedom inherit some classicality from space-time, which caus[es] the quantum system to localize. In a rough sense, this appears to be similar to the gravitationally induced decoherence of the wave function, conjectured by Karolyhazy, Diósi, and Penrose [9], [22], [23]."

Penrose:
This is quite possible because my argument was very specific. However, I haven't looked at the paper. I think I've only glanced at it but haven't studied it. It is pretty interesting.

Infobox 0.17 Is it time to rethink quantum gravity?

In a recent article, Jonathan Oppenheim explores the intriguing idea that, unlike other forces, gravity may not be quantized despite the prevailing contemporary belief [21]. The conceptual challenges in developing a quantum theory of gravity (that he dubs the *quantum-classical theory*) prompt us to consider alternative viewpoints. The article examines those conceptual challenges posed by quantizing gravity, particularly the issue of time and the Hamiltonian constraint; it suggests that a classical-quantum hybrid theory, where space-time is treated classically, could provide a more consistent framework.

Gravity creates a fabric of space-time where time flows at different rates, and matter causes space-time to bend, resulting in the force of gravity. The observer and their measurement tools are treated classically in traditional quantum theory. When this framework is applied to quantum gravity, which considers the entire Universe, including the observer, it presents significant conceptual challenges. Here, space-time is quantized, making it difficult to define meaningful measurements. Traditional quantum measurement relies on a classical concept of time and a clear separation between the observer and the observed. However, in quantum gravity, these distinctions become blurred, leading to what is known as the "problem of time."

Oppenheim examines the idea of a Universe in which space-time is classical, providing a fixed background for quantum matter where the interaction between the classical gravitational field and quantum fields could lead to new insights into the measurement problem and the transition from quantum systems to classicality. Furthermore, it could lead to decoherence, causing the quantum system to exhibit more classical behavior. He speculates that this interaction could offer a mechanism for quantum measurement without relying on the traditional measurement postulate or Born's rule. However, he acknowledges that this is a developing area of research with ongoing challenges.

Oppenheim used a clever thought experiment called the Feynman-Aharonov experiment to figure out if a regular gravitational field could mess with quantum superposition and interference. Let's dive into the details of this thought experiment now.

Infobox 0.18 Is it time to rethink quantum gravity? Contd.

The Feynman-Aharonov Gedanken Experiment and Gravity

In a standard double-slit experiment, a massive particle can pass through one of two slits. Depending on its position, the particle creates a unique gravitational field. In other words, if the particle goes through the left slit, it creates a slightly different gravitational field compared to when it goes through the right slit. If we assume the gravitational field is classical, we can measure it precisely (with arbitrary accuracy) to determine the particle's position without disturbing it. For instance, one could use a hypothetical, infinitely precise pendulum to detect the minute gravitational influence of the particle and determine its path. This information would reveal which slit the particle passed through.

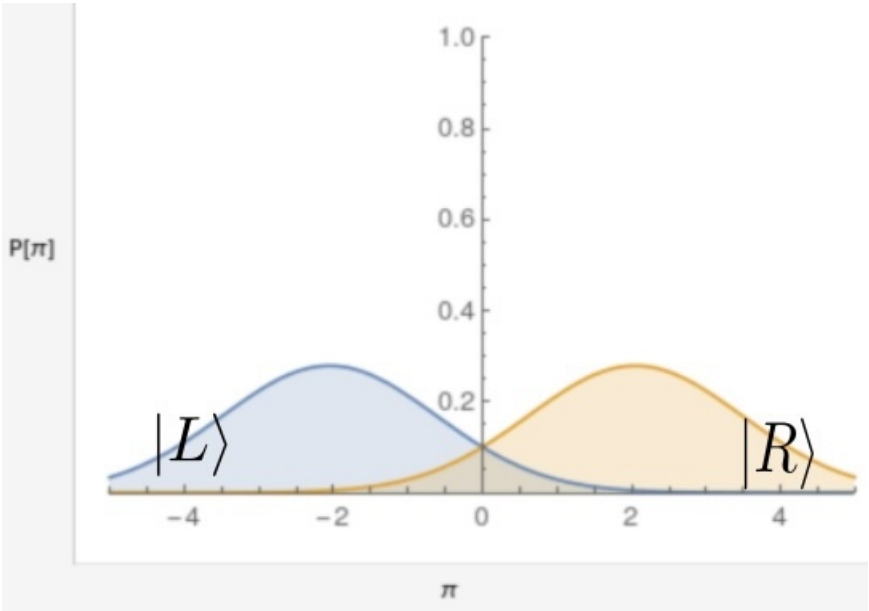

Figure 1: The Feynman-Aharonov thought experiment involves a particle in state $|L\rangle$ or $|R\rangle$ producing a probability distribution over states of the gravitational field, represented as a probability distribution P over a degree of freedom π. This means that the state of the gravitational field does not entirely determine whether the particle is in the $|L\rangle$ or $|R\rangle$ state, leading to the possibility of an interference pattern. A broader probability distribution results in a longer coherence time [22].

However, precisely measuring the gravitational field to determine the particle's path would contradict the fundamental principle of quantum superposition. Knowing which slit the particle went through would eliminate the possibility of it being in a superposition of states, which is necessary for the interference pattern to form on the screen behind the slits.

However, an interference pattern is an experimentally confirmed phenomenon in quantum mechanics.

Infobox 0.19 Is it time to rethink quantum gravity? Contd.

The apparent contradiction arises from the assumption that the classical gravitational field is deterministically produced by the quantum particle. Oppenheim proposes that the conflict could be resolved if the gravitational field responds to the quantum system through a stochastic (random) process. This randomness in the interaction is crucial for maintaining coherence and avoiding contradictions with experimental observations like the interference pattern in the Feynman-Aharonov thought experiment. The article suggests that this stochasticity might be the essential ingredient for understanding the emergence of classical probabilities from fundamentally quantum systems.

Instead of a single, definite gravitational field, imagine a probability distribution of possible field configurations for each slit the particle could go through. If there is some overlap in these probability distributions, measuring the gravitational field will not definitively reveal the particle's path. This overlap in gravitational field probabilities allows the quantum particle to remain in a superposition of states, preserving the possibility of interference. Since gravity is exceptionally weak, only a tiny amount of stochasticity would be needed to prevent the gravitational field from "giving away" the particle's position and destroying the interference pattern.

The Feynman-Aharonov thought experiment doesn't rule out a classical description of gravity when considering the possibility of a random response of the gravitational field. It highlights the challenges of reconciling quantum mechanics with a classical gravitational field and suggests that randomness may play a significant role in their interaction.

Oppenheim suggests that his conclusions are similar to the gravitationally induced decoherence of the wave function, as conjectured by Karolyhazy, Diosi, and Penrose.

Potential implications

Here we summarize the potential implications of a quantum-classical theory:

Infobox 0.20 Is it time to rethink quantum gravity? Contd.

Eliminating the Measurement Postulate: The traditional interpretation of quantum mechanics involves the measurement postulate, which is a separate process from the unitary evolution of the quantum system. In a classical-quantum theory of gravity, the classical nature of spacetime may induce decoherence in quantum systems. This induced decoherence could lead to the localization of quantum states, resembling classical behavior, without relying on the measurement postulate.

Emergence of Born's Rule: Born's rule gives the probability of different outcomes in quantum measurements, and it is often considered an additional postulate in standard quantum mechanics. In a classical-quantum gravity framework, the breakdown in predictability of the classical gravitational field, due to its stochastic interaction with quantum systems, might correspond to the probabilities associated with quantum measurements. This connection could provide a way to derive Born's rule naturally rather than assuming it as a separate axiom.

Role of Stochasticity: A key aspect of the proposed classical-quantum gravity models is the stochastic response of the gravitational field to quantum matter. This randomness in the interaction is crucial for maintaining coherence and avoiding contradictions with experimental observations like the interference pattern in the Feynman-Aharonov thought experiment. This stochasticity might be the essential ingredient for understanding the emergence of classical probabilities from fundamentally quantum systems.

Analogy to Gravitationally Induced Decoherence: A parallel can be drawn between the proposed measurement mechanism in classical-quantum gravity and the existing ideas of gravitationally induced decoherence. While these ideas are distinct, they both involve gravity playing a crucial role in driving quantum systems toward classical behavior.

The classical-quantum theory of gravity is still under development. Many challenges remain, such as proving the consistency and predictability of such theories. However, if successful, these theories could revolutionize our understanding of gravity and the foundations of quantum mechanics, particularly in how measurements lead to the emergence of the classical world from underlying quantum behavior.

Home:

Oppenheim is proposing a concept called gravitationally induced coherence, which involves the interaction between quantum mat-

ter and classical space-time. He suggests that there will inevitably be some randomness involved due to fluctuations in the intrinsic gravitational interaction. He proposes experiments to test how this may affect precision mass measurement. Additionally, he discusses the Penrose model and how it avoids heating. He also suggests that quantum measurement postulates are unnecessary because the gravitational effect causes decoherence.

Consciousness and Quantum Mechanics

Sengupta:
You have said, "Whatever consciousness is, it must be beyond computable physics." However, your views on consciousness have evolved with time, particularly about the role of quantum mechanics in this context. Can you please briefly summarize the present status of your views?

Home:
So, your view has evolved and become more detailed over time. It seems to be different from the early quantum mechanics belief that it is a conscious observer who causes the collapse of the wave function. Can you explain your perspective on this?

Penrose:
Yes, you're correct. I believe that the wave function's collapse is a physical process, not dependent on a conscious observer. In fact, consciousness itself is dependent on this physical process. My initial reasons for this belief stem from my attendance at lectures given by Stein in Cambridge on mathematical logic. Even though I was studying algebraic geometry at the time, these lectures introduced me to Turing machines and brought up Gödel's theorem in a way I had not encountered before. This led me to reconsider the nature of physical processes and consciousness.

Infobox 0.21 The Turing Machine

The Turing machine is a theoretical model of computation first proposed by Alan Turing, a British mathematician and computer scientist, in 1936 [24], [6]. This machine is considered one of the foundational concepts in computer science and provides a theoretical framework for understanding the limits and capabilities of computation.

The machine consists of an infinite tape with cells, a read/write head that can move back and forth on the tape, and a finite set of states that govern its behavior. The machine can read the symbol on the current cell, write a symbol, move the tape left or right, and change its state based on predefined rules.

Turing designed the machine as a thought experiment to investigate which problems could be solved algorithmically and explore the concept of "effective computability." He proved that a Turing machine could compute everything that is computable in principle, thus establishing the theoretical foundation for the limits of computation.

Today, the Turing machine is a fundamental concept that underpins the theory of computation. Although modern computers have evolved far beyond the simple design of a Turing machine, the concept remains relevant as a theoretical tool for understanding the fundamental principles of computation, algorithms, and computational complexity.

Modern computers are built on the principles of the von Neumann architecture, which includes components such as the central processing unit (CPU), memory, input/output devices, and control unit. While modern computers are much more complex and powerful than Turing machines, the fundamental concept of computation laid out by Turing still forms the basis for our understanding of computation and what can be achieved algorithmically.

Sengupta:
In this context, I just read in Max Tegmark's paper that Eugene Wigner proposed the connection between consciousness and the quantum measurement problem, although he changed his mind later and gave up the idea. I was wondering how your ideas of consciousness differ from Wigner's.

Penrose:
It's quite the opposite. The Wigner thinking doesn't explain what

consciousness is, that it's really something outside physics in a way that enacts the collapse of the wave function. I disagree with this idea. Interestingly, I had the opportunity to discuss this with Wigner in Princeton, perhaps over lunch. He was not as rigid in his views as I had anticipated, but he was certainly earnest in his consideration of the topic. Nonetheless, I maintain my argument against it.

The argument I would like to present is as follows: Imagine a distant planet, similar to Earth, with a comparable atmosphere. The only difference is that life has never emerged on this planet. Now, there is a well-known idea in chaos theory that suggests even the flapping of a butterfly's wings can have a profound effect on the weather. Now, suppose we send a space probe towards this planet. Since there are no conscious beings, the planet's atmosphere exists as a superposition of various states without any order or particularity. It becomes a chaotic and disorganized mixture of atmospheres.

The argument I am putting forward is based on the idea that quantum effects can lead to the existence of multiple simultaneous possibilities in the atmosphere of a planet. In this scenario, without any collapse of the wave function, the atmosphere becomes a chaotic combination of various possibilities. Now, let's consider a space probe that travels several light years away from Earth and captures a photograph of this superposed atmosphere. As the probe starts its return journey and sends the signal back to Earth, it takes several more light years for the signal to reach us. However, only when someone observes the picture of the superposed atmosphere does it retroactively determine the precise and distinct atmosphere that existed on the planet.

Home:
I'm having a hard time wrapping my head around this idea. The notion that a conscious observer looking at a photograph of the

atmosphere can determine whether it exists as a superposition or a concrete state seems far-fetched.

Penrose:
I agree; it's a difficult concept to grasp. Back then, I didn't have this argument, so I don't know what Wigner's reaction could have been to it. However, now I have to consider it from a different perspective. There's a possibility that the observation could have a retroactive effect. But personally, I find it highly unlikely that such a significant effect could manifest at a macroscopic scale.

Home:
If we consider human conscious perception/awareness of definite distinguishable outcomes to be causing wave function collapse, the collapse time has to be at least equal to or greater than the perception time. Studies have been conducted on this issue by examining the visual perception of a few photons in the eye and the dynamical model usually used for wave function collapse. Such studies have been reviewed and assessed by Adrian Kent [25].

Penrose:
If you were a frog, you would be able to see a single photon, whereas humans only need a few photons to form a visual image. It's possible to imagine experiments where quantum effects come into play. However, the challenge arises when the photons travel down the nerves, as by that time, the wave function has already collapsed. It might be more feasible to observe these effects within the microtubules.

Looking back, if I had been thinking clearly, I wouldn't have written my book *The Emperor's New Mind*. I'm not so sure now it was such a good idea. Serious scientists typically avoid writing about consciousness if they want to be taken seriously. I was disregarding the consequences. Anyway, what was particularly

curious was towards the end of the book when I describe some experiments conducted by Benjamin Libet.

These experiments involved patients undergoing open brain operations, often for the treatment of epilepsy, where the brain was exposed. Libet performed his experiments with the permission of the patients since their brain exposure during surgery allowed for this opportunity. Electrodes were placed both on the brain, specifically targeting areas associated with movement initiation, and sometimes on areas of the skin, such as the fingers. Patients were asked to report the timing of when they felt a stimulus or their intention to move by observing a fast-moving clock. They would indicate the moment they became conscious of the intention to act or the sensation. This setup revealed that brain activity initiating actions (readiness potential) occurred before the patients became aware of their intentions, challenging ideas about the nature of free will and conscious decision-making.

If the brain was stimulated in the same part, it would feel like the finger being stimulated, although a little bit different; it would be sort of fuzzy, not the same feeling, so the patient would know it was the brain stimulation and not the finger stimulation. That stimulation would not be felt instantaneously—it would be something like half a second later. When the fast-moving clock reached half a second later, the patient would say, "Yes, now I feel it." This would go on for a while.

Home:
But when the patient says, "Now I feel it," before that, he also felt the...

Penrose:
If it was the finger being stimulated, it would be almost instantaneous. If it was the brain, it would not be instantaneous; it would be something like half a second later. This is strange because if

you stimulate the finger first, then a quarter of a second later stimulate the brain, the patient does not feel the finger stimulation, even though it would have been felt otherwise. How could it be unfelt simply by doing something later? Then the brain stimulation is felt half a second later as before, but the finger stimulation is never felt. If the finger stimulation is done after the brain stimulation, then you feel it again. The big mystery is: how can something that is already being felt not be felt or somehow be unfelt?

Well, I don't know. This experiment was done, and I argued in my book *The Emperor's New Mind* that perhaps once perceptions of when something happens are delayed by something like half a second, you haven't really felt the stimulation.

Home:
Yet it depends on whether perception occurs in the brain or not. Because perception is built in the finger.

Penrose:
Have you really felt it instantaneously? The argument would be no. Because the clock is already delayed too and so all those things are delayed and you feel it's simultaneous with the clock at that point, because both things have been delayed. So, everything is delayed. That would be a consistent explanation.

However, I found this very implausible. Why?

I don't remember if I made these arguments in the book, but I certainly made them elsewhere. If you think of somebody, a tennis player or somebody playing ping pong. I hit the ball. The ball comes back to me. I decided to hit it that way, but not this way. I decided this almost at once, much less than half a second. I don't know... like a tenth of a second or less than that. Am I really deciding that, or is it an unconscious act? People who study perception and such subjects will say, "Oh no. You just

think you've done it consciously. It's really all unconscious. You didn't decide that. That decision was made unconsciously."

However, I don't believe that. Even more, I don't believe it because if you consider ordinary conversation—people talking to each other—it's much quicker than that. It's like ping pong. You don't have to wait half a second before you respond to a comment somebody has made. The conversation would be drawn out horribly. How could that be?

When you say that, I have no idea... except that I did suggest in *The Emperor's New Mind* that maybe one's experiences are just delayed—all experiences—by something like half a second. I didn't really talk about active experiences. If you want to do something with immediately recognized stimuli—playing a game like that or maybe playing a piano—and you want to decide whether to make your little finger hit this key a little bit more delicately than before, you must do this almost instantaneously. How could you do that?

Home:
What is your current view on this?

Penrose:
My current story comes basically from Stuart Hameroff. When I wrote *The Emperor's New Mind*, I had no idea how you could preserve quantum coherence in the brain. I learned about neurophysiology, nerve transmission, and the Hodgkin-Huxley theory. I thought there was no chance that could preserve quantum coherence. The signals get disturbed, and the electric fields in the brain get decoherent in no time. There's no way that could preserve the quantum conscious scope.

I came to believe that quantum reduction is somehow related to consciousness. This idea formed from my experience as a graduate student in Cambridge, attending lectures by Bondi on Godel's

theorem. I concluded that conscious understanding wasn't purely computational. I speculated that it must involve the collapse of the wave function, although I had no idea how it acted within neurophysiology.

When I wrote *The Emperor's New Mind*, I learned a fair amount about neurophysiology, but it didn't answer the question. Fortunately, this didn't stop me from writing the book. While I hoped to inspire young people, mainly retired individuals responded. However, one significant response was from Stuart Hameroff, who introduced me to microtubules.

Home:
What is the current status of your collaboration with Stuart Hameroff on quantum consciousness, and how does it relate to the differences between the cerebrum and the cerebellum?

Penrose:
Right, I thought microtubules might preserve coherence, giving a chance for quantum processes in the brain. We wrote papers together. Now, Hameroff thinks key ingredients of consciousness involve pyramidal cells, particularly in the cerebrum, not found in the cerebellum. This is promising because, although the cerebellum is organized more like a conventional computer and contains more neurons than the cerebrum, it lacks consciousness.

Home:
That's interesting. How do you align this with fast decision-making, like in sports?

Penrose:
The difficulty arises in accounting for the rapid, conscious decisions, such as those made by professional athletes. If the signals to pyramidal cells are delayed but crucial, how do these actions happen almost instantaneously? One hypothesis is that the different potential actions exist in a quantum superposition, ready

to be enacted based on a rapid, conscious choice.

Though the concept may sound crazy, multiple controlled muscle movements might be prepared in quantum states, chosen without conscious deliberation for each individual motion. So, while detailed movements might be unconscious, the overall decision could still be conscious and instantly enacted due to pre-existing quantum states.

Home:
It's certainly a provocative idea.

Penrose:
Indeed. It may seem bizarre, but such quantum theories might explain these rapid, controlled actions. We can further explore these possibilities through continued collaboration, particularly with Hameroff.

Books

[1] P. A. M. Dirac, *The Principles of Quantum Mechanics*. Oxford, Clarendon Press, 1930.

[2] R. Penrose, *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*. USA: Oxford University Press, Inc., 1989, ISBN: 0198519737.

[3] S. Gao, *Consciousness and Quantum Mechanics*. Oxford University Press, Jul. 2022, ISBN: 9780197501665. DOI: 10.1093/oso/9780197501665.001.0001. [Online]. Available: https://doi.org/10.1093/oso/9780197501665.001.0001.

[4] J. S. Bell and A. Aspect, *Speakable and unspeakable in quantum mechanics*. Cambridge University Press, 1987.

[5] D. Wallace, *The Emergent Multiverse: Quantum Theory according to the Everett Interpretation*. Oxford University Press, May 2012, ISBN: 9780199546961. DOI: 10.1093/acprof:oso/

9780199546961.001.0001. [Online]. Available: https://doi.org/10.1093/acprof:oso/9780199546961.001.0001.

[6] M. Sipser, *Introduction to the theory of Computation*. Course Technology, 2010.

Articles

[7] R. Penrose, "On the gravitization of quantum mechanics 1: Quantum state reduction," *Foundations of Physics*, vol. 44, pp. 557–575, 2014.

[8] S. Donadi, K. Piscicchia, C. Curceanu, L. Diósi, M. Laubenstein, and A. Bassi, "Underground test of gravity-related wave function collapse," *Nature Physics*, vol. 17, no. 1, pp. 74–78, Jan. 2021, ISSN: 1745-2481. DOI: 10.1038/s41567-020-1008-4. [Online]. Available: https://doi.org/10.1038/s41567-020-1008-4.

[9] L. Diósi, "Models for universal reduction of macroscopic quantum fluctuations," *Phys. Rev. A*, vol. 40, pp. 1165–1174, 3 Aug. 1989. DOI: 10.1103/PhysRevA.40.1165. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.40.1165.

[10] S. Carlip, "Quantum gravity: a progress report," *Reports on Progress in Physics*, vol. 64, no. 8, pp. 885–942, Jul. 2001. DOI: 10.1088/0034-4885/64/8/301. [Online]. Available: https://doi.org/10.1088/0034-4885/64/8/301.

[11] G. C. Ghirardi, A. Rimini, and T. Weber, "Unified dynamics for microscopic and macroscopic systems," *Phys. Rev. D*, vol. 34, pp. 470–491, 2 Jul. 1986. DOI: 10.1103/PhysRevD.34.470. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.34.470.

[12] P. Pearle, "Combining stochastic dynamical state-vector reduction with spontaneous localization," *Phys. Rev. A*, vol. 39, pp. 2277–2289, 5 Mar. 1989. DOI: 10.1103/PhysRevA.39.2277.

[Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.39.2277.

[13] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?" *Phys. Rev.*, vol. 47, pp. 777–780, 10 May 1935. DOI: 10.1103/PhysRev.47.777. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.47.777.

[14] R. Penrose, "Quantum Computation, entanglement and state reduction," *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 356, no. 1743, pp. 1927–1939, 1998. DOI: 10.1098/rsta.1998.0256.

[15] D. Bohm, "A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. I," *Phys. Rev.*, vol. 85, pp. 166–179, 2 Jul. 1951. DOI: 10.1103/PhysRev.85.166. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.85.166.

[16] D. Bohm, "A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. II," *Phys. Rev.*, vol. 85, pp. 180–193, 2 Jan. 1952. DOI: 10.1103/PhysRev.85.180. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.85.180.

[17] H. Everett, ""Relative State" Formulation of Quantum Mechanics," *Rev. Mod. Phys.*, vol. 29, pp. 454–462, 3 Jul. 1957. DOI: 10.1103/RevModPhys.29.454. [Online]. Available: https://link.aps.org/doi/10.1103/RevModPhys.29.454.

[18] H. D. Zeh, "On the Interpretation of Measurement in Quantum Theory," *Foundations of Physics*, vol. 1, no. 1, pp. 69–76, 1970. DOI: 10.1007/bf00708656.

[19] P. Pearle and E. Squires, "Gravity, Energy Conservation, and Parameter Values in Collapse Models," *Foundations of Physics*, vol. 26, no. 3, pp. 291–305, 1996. DOI: 10.1007/BF02069474.

[20] R. Howl, R. Penrose, and I. Fuentes, "Exploring the unification of quantum theory and general relativity with a bose–einstein condensate," *New Journal of Physics*, vol. 21, no. 4, p. 043 047, Apr. 2019. DOI: 10.1088/1367-2630/ab104a. [Online]. Available: https://dx.doi.org/10.1088/1367-2630/ab104a.

[21] J. Oppenheim, "Is it time to rethink quantum gravity?" *International Journal of Modern Physics D*, vol. 32, no. 14, 2342024-12, pp. 2 342 024–12, Oct. 2023. DOI: 10.1142/S0218271823420245. arXiv: 2310.12221 [gr-qc].

[22] F. Karolyhazy, "Gravitation and quantum mechanics of macroscopic objects," *Il Nuovo Cimento A (1965-1970)*, vol. 42, no. 2, pp. 390–402, Mar. 1966, ISSN: 1826-9869. DOI: 10.1007/BF02717926. [Online]. Available: https://doi.org/10.1007/BF02717926.

[23] R. Penrose, "On Gravity's role in Quantum State Reduction," *General Relativity and Gravitation*, vol. 28, no. 5, pp. 581–600, May 1996, ISSN: 1572-9532. DOI: 10.1007/BF02105068. [Online]. Available: https://doi.org/10.1007/BF02105068.

[24] A. M. Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem," *Proceedings of the London Mathematical Society*, vol. s2-42, no. 1, pp. 230–265, 1937. DOI: https://doi.org/10.1112/plms/s2-42.1.230. eprint: https://londmathsoc.onlinelibrary.wiley.com/doi/pdf/10.1112/plms/s2-42.1.230. [Online]. Available: https://londmathsoc.onlinelibrary.wiley.com/doi/abs/10.1112/plms/s2-42.1.230.

[25] A. Kent, "Perception Constraints on Mass-Dependent Spontaneous Localization," *arXiv e-prints*, arXiv:1806.10396, arXiv:1 Jun. 2018. DOI: 10.48550/arXiv.1806.10396. arXiv: 1806.10396 [quant-ph].

Incollection

[26] R. Penrose, "New Physics for the Orch-OR Consciousness Proposal," in *Consciousness and Quantum Mechanics*, Oxford University Press, Jul. 2022, ISBN: 9780197501665. DOI: 10.1093/oso/9780197501665.003.0014. eprint: https://academic.oup.com/book/0/chapter/376470269/chapter-pdf/49114883/oso-9780197501665-chapter-14.pdf. [Online]. Available: https://doi.org/10.1093/oso/9780197501665.003.0014.