A Novel Way of ‘Understanding’ Quantum Mechanics

Armin Nikkhah Shirazi*

University of Michigan
Ann Arbor, MI 48109

9/22/2011

Abstract

Written at a level appropriate for an educated lay audience, this paper attempts to give a primarily conceptual overview of a framework recently introduced in reference [3] by this author, which attempts to clarify what quantum mechanics tells us about reality. Physicists may find this paper useful because it focuses on the central ideas of the framework at a conceptual level, thereby lessening their unfamiliarity, an unavoidable feature of truly novel ideas. The author hopes that this article will motivate physicists to seriously evaluate the mathematical details of the framework given in the original reference.

..I think I can safely say that nobody understands Quantum Mechanics.”[1]

Richard Phillips Feynman

1 Introduction

In Feynman’s famous quote the term ‘understands’ is used in a very specific sense. The mathematics of quantum mechanics are relatively straightforward, and its predictions of testable outcomes are unambiguous. But if one dares to ask what it all means, then one descends a dark tunnel in which no interpretation seems to illuminate a definitive path toward ‘understanding’. Feynman himself evidently thought that our understanding of quantum theory is akin to the ancient Mayans’ understanding of the Venusian cycle: they could predict the pattern of its appearance in the sky with exquisite accuracy, but they did not ‘understand’ it because they had no knowledge of planetary dynamics[2].

To highlight the difficulties with ‘understanding’ quantum mechanics in Feynman’s sense, consider that according to standard quantum mechanics, a system exists in a superposition of mutually exclusive states until a process only defined
in the vaguest terms as a measurement ‘collapses’ the superposition to the one that ends up being measured. For quantum systems that change with time, each state in the superposition is also associated with a phase factor which imbues these systems with certain wave properties. Furthermore, the probability of locating a particle in a particular region can be found by integrating the absolute square of the state function when expressed in terms of position—often identified as the wave function of the particle—over that region. In an equivalent and arguably physically more intuitive formulation due to Feynman, one can think of a system as progressing in time in all possible configurations before it is “measured”. For a single particle, for instance, this means that the particle traverses all possible paths between two points, where each path is associated with a phase factor mathematically slightly different from the previously mentioned phase factors associated with states. The probability of finding the particle at the second point given that at a specified earlier moment it was at the first is given by integrating over all possible paths and their associated phase factors and taking the absolute square. This is why Feynman’s approach is often called the path integral formulation of quantum mechanics. Because each path can be associated with a history of the particle, the path integral approach is sometimes also called the sum over histories.

There are probably over a dozen interpretations of quantum mechanics, some with multiple variants, of which the most widely accepted is called the orthodox interpretation. It claims that a system really did not have any definite properties (i.e. really did exist in a superposition) until a particular one is measured, and that questions about what properties it had prior to measurement are not meaningful. A ‘measurement’, however, produces a real outcome, namely a reduction of the superposition state $|\Psi>,$ into what is called an eigenstate $|\psi_i>,,$ which is one of the states in the superposition that represent a possible measurement outcome. This is sometimes also referred to as the ‘wave-function collapse’ and can be represented as follows:

$$|\Psi> \xrightarrow{\text{Collapse}} |\psi_i>$$

(1)

Other interpretations, such as the 'Many Worlds' or Bohmian interpretations, tell stories about the meaning of the equations which are incompatible with one another and with the orthodox interpretation, but are nevertheless accepted by segments of the physics community because they are thought to be observationally indistinguishable.

Of all scientific disciplines, physics, being the closest to mathematics, is the one in which it is generally easiest to determine with conclusive definiteness whether a particular claim or interpretation of a theory is true or false. The current situation, in which several mutually incompatible interpretations of quantum mechanics co-exist, seems highly anomalous. This article will present a novel way of understanding quantum mechanics in Feynman’s sense, based on a recently developed framework called the Dimensional theory[3]. As unorthodox as it may seem at first exposure, it will turn out to basically confirm the orthodox interpretation.
2 Two Observations

Let’s first motivate this novel framework with two simple observations which ostensibly have nothing to do with quantum mechanics. The first is that, for objects of same shape, the ratio of the surface area to volume ($A/V$ ratio) increases as the object becomes smaller. For instance, compared to a ball of radius 1 meter, a ball of $10^{-11}$ meter (on the order of the Bohr radius) has an $A/V$ ratio that is hundred billion times larger. If we take the $A/V$ ratio as an indicator of relative dimensionality, this means that hypothetical objects in the regime in which quantum behavior dominates are unimaginably more two-dimensional than ordinary classical objects we observe in our everyday experience.

The second observation is that when one attempts to represent an object in a higher-dimensional space, then it takes on in the manner of an actualizable potentiality all possible values of the property it is intrinsically missing in the higher-dimensional space. For example, consider a point in a two-dimensional plane located at $(x_0, y_0)$

![Figure 1: An arbitrary point in an xy-plane](image)

its representation in three-dimensional space is as an infinitely long line, a superposition of an infinite number of points, as it were, parallel to the $z$-axis, all having the $x$ and $y$ coordinates in common.

![Figure 2: The same point in an xyz-space with the same xy coordinates as before now manifests itself as a superposition of an infinite number of points, one for each possible z-coordinate value](image)

If one attributes a $z$–coordinate to the point $(x_0, y_0)$ at $z_i$ (the subscript
denotes a particular value for $z$), the infinite superposition ‘collapses’ to the point $(x_0, y_0, z_i)$

![Diagram of a 3D coordinate system with a point at $(x_0, y_0, z_i)$]

Figure 3: The superposition of an infinite number of points "collapses" to the point $(x_0, y_1, z_i)$ as soon as the point $(x_0, y_0)$ attains an intrinsic coordinate specification along the $z$–axis.

Mathematically, this process that can be symbolized as follows:

$$\sum_{i=-\infty}^{\infty} z_i \bullet (x_0, y_0) \xrightarrow{\text{‘Collapse’}} (x_0, y_0, z_i)$$

(2)

Where $\bullet$ is meant to indicate the association of the point $(x_0, y_0)$ in the plane with the superposition of all possible $z$–values (i.e. $\sum_{i=-\infty}^{\infty} z_i$) in space.

We can make a subtle distinction to make this mathematical observation more pertinent to physics: In a strict mathematical sense, a point is identified with its associated set of coordinates, but if we interpret the point at $(x_0, y_0)$ as a quasi-physical object, in the same sense in which a point-like particle can occupy a spacetime event, which is, after all, nothing more than a point in the mathematical sense specified by four coordinates in a particular coordinate system, then there is a definite distinction between a quasi-physical point and the locus it occupies (what we would think of as a "point" in a mathematical sense). What we just called a superposition of points in space is then really a superposition of loci, of "empty slots", each with the capacity to be occupied by a quasi-physical point once the latter attains an intrinsic $z$–coordinate specification with its $z$–value. Once the quasi-physical point in the plane at $(x_0, y_0)$ attains an intrinsic $z$–coordinate, then the infinite superposition of loci in space “collapse” to the locus at $z_i \bullet (x_0, y_0)$ which is now occupied by the quasi-physical point $(x_0, y_0, z_i)$ in space.

Hopefully it is evident that there is an ontological distinction, i.e. a distinction based on the nature of existence, between the point before and after the collapse: Pre-collapse it is part of the manifestation in space of a quasi-physical point that exists in the plane, i.e. $z_i \bullet (x_0, y_0)$, post-collapse it manifests the quasi-physical point $(x_0, y_0, z_i)$ that exists in space.

The fact that quantum behavior generally dominates at very small scales and
the similarities between the point analogy and some basic aspects of quantum mechanics that present the conceptual difficulties mentioned in the introduction motivate the central concepts of the dimensional theory.

3 The Dimensional Theory

To begin, the theory uses the first observation to motivate an axiom that there exists a limit in which the $A/V$ ratio becomes infinite. In this limit space vanishes but area does not, and spacetime is postulated be replaced by a one-dimension lower analog. This analog is called areatime, a term chosen because it is less awkward than “2+1 spacetime” and because it connotes more strongly that its status in this theory is on the same footing as that of spacetime, as $2+1$ spacetimes often appear in toy theories that are explicitly not meant to be representative of reality because they are too simplified.

Of course, we cannot say much about hypothetical objects in this limit except for one thing: Since their existence is in areatime, in their frame they must have a history in areatime. But this means that they cannot have a spacetime history in their frame. This assertion can be justified by a geometric argument: According to special relativity, the proper time in spacetime, i.e. the time that passes in a rest frame, is proportional to the spacetime metric interval. The metric interval of spacetime and that of areatime must be distinct from each other because there are three length dimensions contributing to the former but only two contributing to the latter. Thus, the proper times in areatime and spacetime must also be distinct, in the sense that they must be two distinct time dimensions. Let us refer to the passage of proper time by the shorthand of “aging”. If a hypothetical object which ages in areatime did also age in spacetime, even at a different rate, we would simply interpret this as the passage along the same proper time dimension measured by two clocks at rest in that frame which are calibrated differently with respect to each other, not as the passage of time along two distinct proper time dimensions. So in order for the two time dimensions to remain distinct, an object in areatime cannot age in spacetime. Yet it must age in spacetime in order to have a spacetime history in its proper frame. Therefore, hypothetical objects existing in areatime cannot have a spacetime history in their frames.

The theory now takes the point analogy in the second observation and applies it to histories: It postulates that if an object does not have a spacetime history in its frame, it must manifest itself in spacetime as the superposition of all possible spacetime histories of the object into which it would transform if it were to emerge in spacetime. But just as in the point analogy, there is a distinction between the history in its frame and the histories that are part of the superposition: The latter are like “empty slots” because they are not created as a consequence of the passage of a proper time in the object’s frame. To distinguish these from the actual history experienced in a frame as a result of the passage of proper time we will refer to these as actualizable histories.

Since an object can only age along one proper time, how can the passage of
time for objects that exist in areatime be compared to the passage of time for spacetime observers? The theory postulates that a certain novel symmetry governs the comparison between the passage of time along the two proper time dimensions without leading to a net transformation. Avoiding a net transformation is critical because if one time dimension can be transformed such that it can be projected unto another, this immediately implies that they are no longer distinct time dimensions. This follows from the same argument as above with the two differently calibrated clocks.

A graphical representation of the symmetry is given below:

![Diagram of a postulated symmetry](image)

Figure 4: A graphical representation of a postulated symmetry which allows a comparison between the proper time $\tau_A$ of a system that exist in areatime and the proper time $\tau$ of an observer who exists in spacetime without transforming them into each other, which would destroy the perpendicularity of the proper times. The symmetry is represented in a way to suggest that it can be decomposed into two rotations of opposite sense, and these in turn can mathematically be transformed into complex conjugate exponential phases, which by substituting the postulated relation $\tau_A \Rightarrow \pm i\hbar/\mu c^2$ (one sign per rotation sense) can be identified with the familiar phases in the path integral formulation.

The idea is that a rotation permits a comparison of the passage of time, but is always cancelled by a rotation in the opposite sense, resulting in a zero net transformation. The symmetry can therefore be decomposed into two opposite rotations which, given the “smallness” of the system’s proper time $\tau_A$ compared to the observer’s proper time $\tau$ and under the substitution $\tau_A \Rightarrow \pm i\hbar/\mu c^2$ (given

---

1What is meant by the “smallness” of $\tau_A$ relative to $\tau$ is this: Because the system is postulated to exist in a lower size limit its associated proper time interval is necessarily finite, and when it is to be compared against any spacetime proper time interval larger in magnitude is postulated to repeat itself (i.e. appear periodic to us). But that means that a while a rotation over an angle $2\pi$ returns one to the initial position along $\tau_A$, it does not do so along $\tau$. This, in effect, turns the plane between the two proper times into a Riemann Sheet, and means that $\tau_A$ must be mathematically modeled by spacetime observers as an imaginary period. This is the physical explanation that this framework gives for why imaginary numbers appear explicitly in quantum theory.

6
by an axiom) can be transformed into two complex conjugate phases identical with those used in the path integral formulation. Each actualizable history is then associated with both phases, and the infinite superposition of histories associated with each phase then results in a mathematical expression that looks identical to the Feynman path integral, but is ontologically distinct because it is the integral of paths associated with actualizable, not actual histories.

Transforming from the path integral to the standard formulation is well understood[4, 5], and when performed here leads to the expression for the state of a system that is a superposition of what are called actualizable states. According to a final axiom, when a certain process (such as what we call an ‘interaction’) causes a system or parts of it to exceed the postulated limit in which spacetime vanishes, then constituent masses associated with those parts actualize, or emerge as actual objects in spacetime, which means that the superposition of actualizable states collapse to an actual state. When applied to the entire system, this leads to a state description that can be symbolized as follows:

\[
\begin{array}{c}
\Psi > \hspace{2cm} |\psi_i > \\
\end{array}
\]

This looks exactly like the first equation, except that a specific underlying physical cause is identified (‘actualization’ as a consequence of exceeding the limit in which spacetime vanishes) and an ontological distinction between pre- and immediate post-measurement states is formally introduced into the equation by underlining the immediate post-measurement state. The mathematical importance of the distinction is this: The states can be thought of as vectors in an abstract vector space, called the Hilbert space, and in standard quantum mechanics the pre- and immediate post-measurement states are on a completely equivalent footing. In particular, a rotation of the Hilbert space can turn the eigenstate on the right into a completely equivalent superposition state in what is called another basis. The ontological distinction breaks this equivalence: the underlined state “looks” like a vector in Hilbert Space, but isn’t because it is an actual, not an actualizable state, just as the quasi-physical point \((x_0, y_0, z_i)\) in our analogy “looks” like the empty slot \(z_i\) \(* (x_0, y_0)\) but isn’t. However, in the regime in which quantum behavior dominates, this transformation must be reversible because we already know empirically that shortly after such a collapse, the state once again spreads out in a superposition and hence must be considered actualizable. The details of the physical process underlying these ontological transformations (i.e. ‘actualization’ and ‘de-actualization’) are at present not known, but one might suspect that within the context of quantum field theory they may eventually be elucidated.

The original reference introducing this theory presents a mathematical derivation for the simplest possible case, a free particle, that starts with 5 axioms and leads to an expression identical to the Feynman path integral, except that it is the integral of paths associated with actualizable histories [3].
4 ‘Understanding’ Quantum Entanglement

This framework has obviously many implications, but here we will only conceptually address arguably the most mystifying aspect of quantum mechanics, something often called *quantum entanglement*. This refers to the prediction of standard quantum mechanics that for a quantum system consisting of multiple components described by a non-separable state, a measurement of the property of one component in the appropriate setup seems to instantaneously, as it were, determine the measurement outcomes of corresponding properties of other components, even if they are arbitrarily far away. The most famous example of this is called the *EPR paradox*, named after Einstein, Podolsky and Rosen, the authors of the paper which first introduced it[6]. It involves the measurement of the intrinsic angular momenta, called spins, of two entangled particles separated in space. Correlations between the measurement outcomes suggest that somehow the measurement outcomes of the two influence one another in an as yet undefined way. Einstein, who apparently never believed that this aspect of quantum mechanics truly describes reality, famously called this ‘spooky action at a distance’. Experiments performed in the last 40 years, however, have unequivocally confirmed that quantum mechanics does indeed provide the correct mathematical description of this aspect of nature [7, 8, 9], but how to ‘understand’ it is still an open question.

Given the above framework, there is a surprisingly simple way of interpreting this phenomenon. Once again, according to the Dimensional theory, when a system must be described in terms of a superposition, it actually exists in areatime. Therefore, a collapse event in one region of spacetime can affect other objects described by the same state in a manner completely free from constraints imposed by the spacetime metric because distance relations between them are only subject to the constraints of the areatime metric. More specifically, because by the fundamental assumption $\tau_A \Rightarrow \pm \hbar/mc^2$ (where $m$ must in this context be considered the mass of the combined system) the proper time associated with a system in areatime is incorporated into its description in terms of a state, the components of a system described by a non-separable common state must be associated with the same proper time and therefore the same metric interval in areatime, no matter how far apart they are from each other in spacetime.

This explanation may seem almost too facile and, particularly if one has not yet developed an intuition for how the components of a system can be independent of metric relations in a higher-dimensional space, it may be difficult to visualize. To help develop such an intuition, let us go back to our second observation and consider a similar analogy, but this time instead of a point in the plane consider a pattern, as in fig. 5 which can be decomposed into a square and a circle.
Prior to the attainment of an intrinsic \( z \)-coordinate specification, the representation of the overall pattern in space would once again be in terms of an infinitely long superposition parallel to the \( z \)-axis (i.e. two infinitely long cylinders with a circular and square cross section, respectively).

Let us now contrive the following situation: suppose either of the two components could attain an intrinsic specification at \(-z_0\) at time \( t_1\) and that the complementary component could attain a specification at \(+z_0\) at time \( t_2\), but that before \( t_1\) it was undetermined which component attains which specification. Then, the specification at \( t_1\) would immediately tell us that at time \( t_2\) the complementary component will attain an intrinsic specification at \(+z_0\). For instance, if the superposition of the circle component were to collapse onto the plane at \(-z_0\) at \( t_1\) then we would immediately know that an intrinsic specification of the rest of the overall pattern would lead to a collapse of the superposition of the square component onto the plane at \(+z_0\) at \( t_2\) (and vice versa), as depicted below:

Figure 5: A simple pattern in the xy-plane that can be decomposed into a square and a circle
Figure 6: A Collapse of the circle component onto the plane at \(-z_0\) at time \(t_1\) immediately tells us that the pattern that will be found in the plane at \(z_0\) at \(t_2\) is going to be a square, even though before the collapse of the circle component this was undetermined.

Notice that until either component has gained an intrinsic \(z\)–coordinate specification, the distance relationships between any point on the square and any point on the circle are governed by \(d\rho^2 \equiv dx^2 + dy^2\), not by \(dr^2 = dx^2 + dy^2 + dz^2\)!

If we did not appreciate this, then the immediate determination of one pattern based on the specification of the \(z\)–coordinate of the other might seem very strange to us since in space the two components after such specification are separated by a distance of \(2z_0\).

Analogously, in the EPR experiment, the determination of the spin of one particle seems to instantaneously determine the outcome of the measurement of the spin of another, no matter how far apart, because the intrinsic metric relations between what gives rise to the spins obey the areatime metric, not the spacetime metric. Actually, in a relativistic context we must express this more carefully: What the theory suggests is that the determination of the spin of one particle does not instantly determine the spin of the other (unless the measurements are simultaneous in that frame)-because, until the other measurement has been performed, there is no such thing as an actual second particle to have spin as a property—but once the second measurement is performed, the underlying process will in effect ‘force’ a particle to emerge out of areatime (i.e. actualize) with a spin that always complements the spin of the first. The reason for highlighting this subtlety is that if the two measurements are spacelike separated, then one can always find two inertial frames in which the order of measurements is reversed. Then, the assumption that a measurement of the spin of one in a particular inertial frame instantaneously collapses the state of the other either leads to a contradiction with the observation in another frame in which the order (i.e. the direction of the ‘influence’) is reversed, or requires that the first inertial frame be preferred, in gross violation of the principle of relativity. This has been previously called the EPR influence paradox and this framework, which assumes that not only do quantum objects not have well-defined properties prior
to being measured, but they do not even actually exist in spacetime until their properties are measured, easily resolves it [10]. If between the two measurements there is no actual second particle in spacetime with spin as a property, then this precludes the determination of a “direction of influence”: All one can determine are correlations after both measurements have been completed.

It is important to keep in mind that the mathematical details of the explanation given here in the non-relativistic limit are exactly those of standard non-relativistic quantum mechanics. What is supplied here is new meaning, not new mathematics. No change is necessary in the mathematical description of entangled states, save for the ontological distinction between actual and actualizable states. However, with this ontological distinction made explicit, and the underlying physical reason elucidated, there is a profound implication: Since what has been called a ‘measurement’ is really any process—whether observer-initiated or not—that causes a system that actually exists in areatime under certain circumstances to emerge beyond the postulated limit in spacetime, it must be ubiquitous. And this suggests that the collapse process is in fact the mechanism by which all of spacetime emerges from areatime!

5 A Testable Prediction

What sets a genuinely deeper theory apart from just another interpretation of quantum mechanics is that it makes new testable predictions that do not follow from standard quantum mechanics. This framework makes at least one definite, under current paradigms totally unexpected prediction for the outcome of an experiment, as yet unperformed, which is in principle extremely simple but in practice extremely difficult to perform with current technologies. The prediction is that electromagnetic radiation in transit does not produce gravitational fields. The reason according to the DT is that photons, having a zero proper time and therefore no actual history in spacetime, do not actually exist in it, and the energy associated with them can therefore not be localized in space (since localization in space presumes actual existence in spacetime). Unless absorbed by massive particles, they cannot set up gravitational fields.

Two clarifications are immediately in order: First, this prediction, which directly contradicts the prediction of General Relativity (GR) for the same situation [11], is not a claim that the latter is wrong, but only that this may represent an instance in which GR is applied beyond its domain of validity. GR has so far passed every experimental challenge with flying colors[12], which means however, that we really do not know where the boundaries of its domain of validity lie. Yet our experience with Newton’s and Maxwell’s theories teaches us that we should expect every theory—except perhaps for a hypothetical theory of everything—to have such a boundary, and given that GR on the face of it seems incompatible with quantum theory, we should expect such a boundary in relation to quantum objects to exist somewhere. This prediction makes that expectation precise in terms of a definite measurable outcome.

Second, this prediction does not really amount to a violation of conservation of
momentum, as one might be led to believe considering that we already know that gravity “bends” light, if one takes the ontological status of photons suggested by the DT seriously. According to the DT, the gravitational influence is due to the skewing of the path integral by the underlying spacetime curvature, not due to a gravitational interaction with the photons as objects “in and of themselves”, since they are presumed to actually exist in areatime. The backreaction occurs when the radiation is absorbed by massive objects which, now having a greater energy, set up slightly different gravitational fields than otherwise, and these eventually reach the original object that bent the light path (the “light path” being a classical notion that is physically incorrect) and act on it. When one frames this as an interaction between emitters and absorbers and considers that in terms of the spacetime interval the separation between the action-reaction pair is still null (assuming gravitational influences also travel at c), conservation of momentum does still apply.

Those who are still inclined to dismiss this as yet untested prediction out of hand are reminded that prior to the actual performance of the famous Michelson-Morley experiment, a prediction of their null result would in all likelihood have been immediately dismissed by any physicist of that age. If the current prediction is empirically confirmed, however, then both null results could be understood not as a coincidence but as consequences of the same currently unfamiliar idea, namely that photons do not actually exist in spacetime[13]. More generally, the prediction given here applies to any object describable by a superposition of positions, and in this sense, the object’s actualization according to the DT corresponds to the creation of a gravitational field. Photons, however, offer the most definitive way to falsify this prediction because they never actualize.

In the post-LHC age, one may well wonder what important new experiments could be performed in fundamental physics. Measuring the gravitational acceleration of a test particle due to a high-energy beam of radiation in its vicinity would be a highly interesting and worthwhile experiment. We should attempt to perform it regardless of the prediction given here.

6 Conclusion: Existence as a Physics Concept

This paper discussed at a conceptual level a recently introduced framework which allows one to ‘understand’ quantum mechanics in a novel way. It is hoped that the competent reader has been intrigued into examining the mathematical details in the original reference [3]. Whereas Feynman thought that our understanding of quantum mechanics was incomplete, it is well known that Einstein thought that it was quantum mechanics itself that was incomplete. If the framework introduced here is correct, it validates the orthodox interpretation and allows us to ‘understand’ it in Feynman’s sense, but it also suggests that standard quantum mechanics was incomplete after all, though probably not in a way that anyone anticipated. As presented here, the incompleteness is due to its failure to make an ontological distinction between pre- and immediate post-measurement states in its formalism. Com-
pleting the theory in this mathematically trivial manner, however, requires a radical shift in our conceptualization of physical existence. It requires us to move physical existence as a concept subject to investigation from the purview of philosophy to that of physics and seriously consider that physical objects can have more than one mode of existence. This may well be distasteful to some. But then, is it not possible that our fundamental physical knowledge has become so profound that we can no longer circumvent the explicit consideration of physical existence itself in order to ‘understand’ nature more deeply?

References


