ABSTRACT

The EPR-Bohm Paradox
and Kent's One-World Solution

Robert William Verrill, Ph.D.

Mentor: Alexander R. Pruss, Ph.D.

In this dissertation, I will consider in detail the EPR-Bohm paradox and the mysterious correlations that it exhibits, and I will explain why the EPR-Bohm paradox can lead one to think that our best physics suggests a many-worlds picture of reality. I will thus consider the many-worlds interpretation in some detail and explain why it is so attractive. But ultimately, my goal is not to endorse the many-worlds interpretation of quantum physics, but rather it is to argue that the reasons for endorsing the many- worlds interpretation are not compelling. My argument relies on a recent interpretation of quantum physics by the physicist Adrian Kent. Kent's interpretation shares several features with the many-worlds interpretation that make it seem attractive, yet in Kent's interpretation, there is only one world.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

To my parents who taught me the most valuable lessons of all

Introduction

In recent years, it has become increasingly common for popularizers of quantum physics to tell us that we need to let go of our naïve common sense understanding of reality. We're told we must replace this common sense understanding with something that at first seems very bizarre and counter-intuitive: a many-worlds interpretation of reality. This is the idea that whenever there is quantum indeterminacy among several possibilities, then all these possibilities are realized, and the actualization of these possibilities can be extrapolated up to the macroscopic level. Thus, many-worlds advocates, when reflecting on the famous Schrödinger's Cat thought experiment do not question the foundations of quantum mechanics on which the thought experiment is based, but rather they embrace the seemingly absurd conclusion of Schrödinger that the foundations of quantum mechanics imply that a cat could be both dead and alive. Many-worlds advocates thus speak of the cat being dead in one world and the cat being alive in another world that is just as real as the first.

In this dissertation, I will consider in detail the EPR-Bohm paradox and the mysterious correlations that it exhibits, and I will explain why the EPR-Bohm paradox can lead one to think that our best physics suggests a many-worlds picture of reality. I will thus consider the many-worlds interpretation in some detail and explain why it is so attractive. But ultimately, my goal is not to endorse the many-worlds interpretation of quantum physics, but rather it is to argue that the reasons for

endorsing the many-worlds interpretation are not compelling. My argument relies on a recent interpretation of quantum physics by the physicist Adrian Kent. Kent's interpretation shares several features with the many-worlds interpretation that make it seem attractive, yet in Kent's interpretation, there is only one world.

Like many interpretations of quantum physics, Kent's interpretation is highly speculative, so I am not making any claims about whether Kent's interpretation is actually true. Rather, I only aim to show that if one fully accepts the predictive power of quantum physics and accepts the truth of special relativity, then one is not obliged to accept a many-worlds interpretation of quantum physics, since Kent's interpretation offers us a viable alternative.

In this dissertation, I will do my best to avoid unnecessary mathematical jargon, but since so many of the ideas in quantum physics are expressed in mathematical terms, a certain amount of mathematics is unavoidable. I will be assuming a basic understanding of trigonometry and probability, but beyond that, I will endeavor to explain all the mathematical terminology I use as I go along. There will, however, be some sections which may be very challenging to readers who do not have much of a background in mathematics or physics. These sections will be marked with an asterisk *. There is also a lot of terminology from theoretical physics that I will need to invoke, so to aid the reader, I will use the convention of putting terminology in italics whenever the terminology is first defined, and there is also a list of notation at the end of this dissertation.

CHAPTER ONE

Confronting the EPR-Bohm Paradox

The central challenge that one-world interpretations of quantum physics must face is how to account for the mysterious correlations of the so-called EPR-Bohm paradox in a way that is consistent with Einstein's theory of special relativity. In this chapter, no prior knowledge of quantum theory will be assumed. I will therefore need to describe some key ideas of quantum theory, and I will do this in the context of the Stern-Gerlach experiment. I will then describe the EPR-Bohm paradox and the difficulty the traditional Copenhagen interpretation of quantum physics has in dealing with this paradox. A seemingly natural way to overcome this paradox is to supplement standard quantum theory with hidden variables. I will thus describe one way in which this can be done, and I will show how this leads to the remarkable inequality first derived by Bell. However, Bell's inequality is known to be experimentally violated. This means there must be something wrong with Bell's assumptions. I will therefore consider Shimony's analysis of the proof of Bell's inequality in which Shimony draws a distinction between parameter independence and outcome independence. Finally, following Butterfield, I will briefly explain why Shimony's analysis does not adequately resolve the EPR-Bohm Paradox.

Some of the key features of quantum physics are exhibited in the Stern-Gerlach experiment (see figure 1.1). In this experiment, silver atoms are heated in a furnace which randomly emerge from the furnace with various velocities. By aligning two plates with circular holes near the furnace, it is possible to select a subset of the emerging silver atoms having (approximately) the same momentum[1] to form a beam in one direction, the other silver atoms having been absorbed by the two plates. This beam of silver atoms is then directed between two magnets with the north pole of one magnet being aligned toward the south pole of the other magnet as shown in figure 1.1. Now silver atoms have a property somewhat analogous to the classical notion of
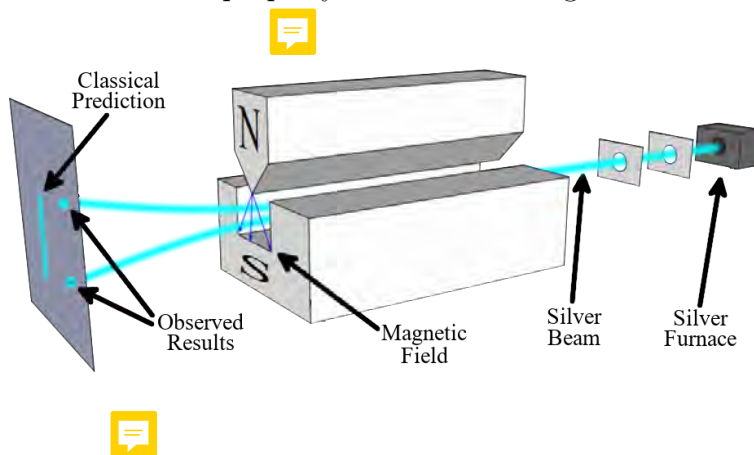
Figure 1.1. The Stern-Gerlach Experiment.[2]

---

1. The *momentum* of an atom is its mass multiplied by its velocity. See footnote 25 on page 21 for a definition of velocity.

2. Original diagram drawn by Theresa Knott. Labeling was modified for use in this dissertation. This image is licensed under the Creative Commons Attribution-Share Alike 4.0 International license. Source: https://commons.wikimedia.org/wiki/File:Stern-Gerlach_experiment_svg.svg.

angular momentum. For instance, a spinning top has angular momentum as shown in figure 1.2. Angular momentum is a *vector*, that is, it has direction and magnitude. In the case of a spinning top, the direction of the angular momentum would be parallel to the axis of rotation, pointing one way or the other depending on whether the rotation was clockwise or counterclockwise. The magnitude of the angular momentum would then be proportional to the *angular velocity* of the spinning top, so that the faster the top was spinning, the greater the angular velocity would be, and hence, the greater the angular momentum would be.
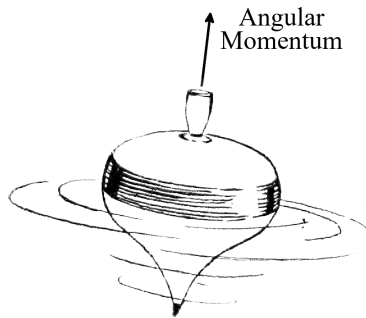
Angular
Momentum

Figure 1.2. Angular Momentum of a Spinning Top.[3]

Now if we tried to understand the angular momentum of a silver atom classically, we would expect the magnetic field of the two magnets to interact with the silver atom in a way that was determined by the relative direction of the silver atom's angular momentum compared to the direction of the magnetic field. Since we would expect the silver atom to have an entirely random angular momentum, we would expect it to be deflected by varying degrees either up or down in the direction of the magnetic field. Thus, if a detection screen were placed beyond the two magnets which the silver

atoms would hit, we would expect there to be a whole continuum of possible locations where the silver atoms would be detected. However, in reality, it is found that there are precisely two locations where the silver atoms hit the screen. It is as though the particles can spin either clockwise or anticlockwise, but that there is absolutely no variance in the angular speed at which they rotate. This is surprising. The angular momentum appears to be *quantized* in one of two directions, either parallel to the magnetic field or antiparallel to it.[4] Corresponding to this quantization of angular momentum, we say that the atom is either in the spin up state or the spin down state with respect to the direction of the magnetic field.
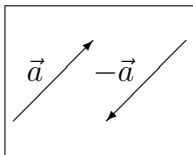
Figure 1.3. Meaning of antiparallel: the arrows in opposite directions are said to be antiparallel to one another.

If the direction of the magnetic field is implicitly understood, we write $|+\rangle$ and $|-\rangle$ for the spin up and spin down states of the atom respectively. We refer to the symbols $|+\rangle$ and $|-\rangle$ as *ket-vectors*, or simply as kets. We can think of the ket $|+\rangle$ for instance as shorthand for the proposition "the particle is in the spin up state." If we knew this

3. Drawing by Pearson Scott Foresman, Public domain, via Wikimedia Commons. Labeling was added for use in this dissertation. Original: https://commons.wikimedia.org/wiki/File:Top_(PSF).png.

4. See figure 1.3 for what is meant be antiparallel.

proposition to be true, we would know which of the two locations on the detection

screen the particle would end up if it were to travel between the two magnets of the

Stern-Gerlach apparatus. If we need to specify the spin with respect to a particular

direction of the magnetic field, say in the $\hat{\mathbf{a}}$-direction, we write the corresponding

spin up and down states as $|\hat{\mathbf{a}}+\rangle$ and $|\hat{\mathbf{a}}-\rangle$. For convenience, we write $\hat{\mathbf{a}}+$ and $\hat{\mathbf{a}}-$

respectively for the location that the particle would hit the detection screen.

The question then arises as to what happens when we rotate the magnetic field

around the axis of the particle beam in the Stern-Gerlach experimental setup. It turns

out that when we do this, the atoms are again detected in only one of two locations

(see figure 1.4).
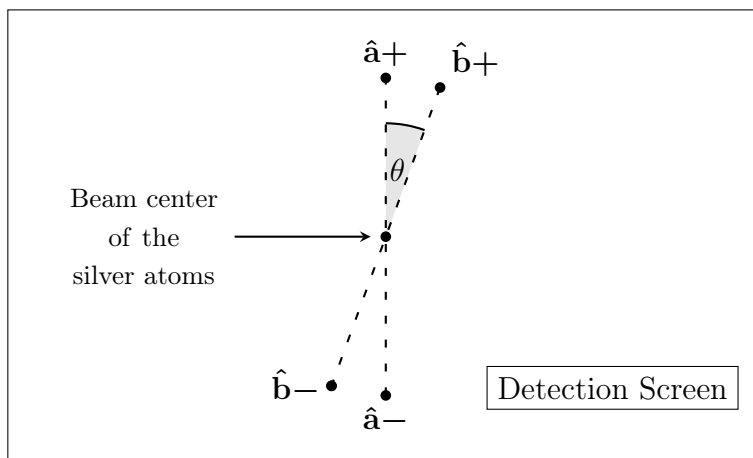
Figure 1.4. Locations of detections before and after rotating the magnetic field by an angle θ. Before rotation, the particles can be detected at either location $\hat{\mathbf{a}}+$ or location $\hat{\mathbf{a}}-$. After the rotation, particles can be detected at either location $\hat{\mathbf{b}}+$ or $\hat{\mathbf{b}}-$.

So suppose we knew the particle was in a spin state such that it was on course to

arrive at location $\hat{\mathbf{a}}+$ because we had previously directed it through another magnetic

field in the $\hat{\mathbf{a}}$-direction. For example, see figure 1.5 for how this might be done. In
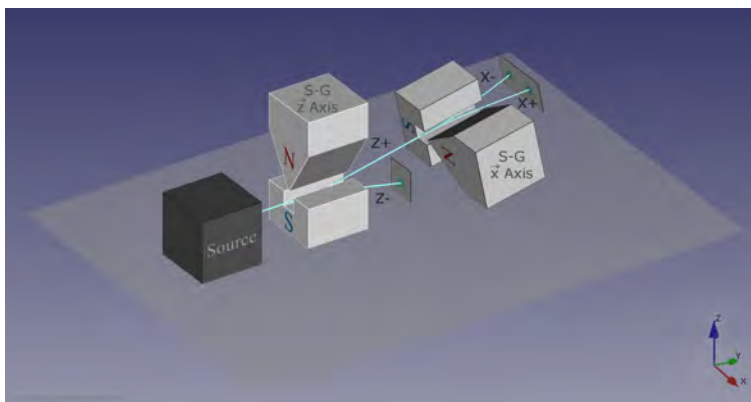
Figure 1.5. Two Stern-Gerlach experiments in sequence. By directing the beam of particles through one magnetic field first, the particles emerging in one of the two beams will be in a known spin state before they enter the second magnetic field.[5]

this experimental setup, the second magnetic field has been rotated by an angle of $90°$ with respect to the first magnetic field. But suppose we just rotated the second magnetic field by a very small angle θ with respect to the first magnetic field. Then we would expect the particle now to arrive at a location $\hat{\mathbf{b}}+$ close by to $\hat{\mathbf{a}}+$ as shown in figure 1.4. And this is what we notice experimentally for the most part. However, occasionally, the particle will arrive at location $\hat{\mathbf{b}}-$. The frequency of this occurrence becomes less and less the less and less the magnetic field is rotated (i.e. the smaller θ is), so that if the magnetic field is not rotated at all, i.e. $\theta = 0$, the particle will always arrive at location $\hat{\mathbf{a}}+$. To capture the probabilistic nature of these outcomes, we use the bra-ket notation. Thus, if $|\psi\rangle$ stands for either the $|\hat{\mathbf{a}}+\rangle$ or the $|\hat{\mathbf{a}}-\rangle$-state,

5. Diagram by MJasK. This file is licensed under the Creative Commons Attribution-Share Alike 4.0 International license. Source: https://commons.wikimedia.org/wiki/File:Stern-Gerlach_Analyzer_Sequential_Series_E2.png.

and $|\chi\rangle$ stands for either the $|\hat{\mathbf{b}}+\rangle$ or the $|\hat{\mathbf{b}}+\rangle$-state, then we define the bra-ket $\langle\psi|\chi\rangle \in \mathbb{C}$ to be a complex number[6] that satisfies the *Born Rule*, namely the rule that $|\langle\psi|\chi\rangle|^2$ is the probability that the particle will be found to be in state $|\psi\rangle$ given that we know that the particle is in state $|\chi\rangle$. For example, if $|\langle\psi|\chi\rangle|^2 = \frac{1}{4}$, and we performed the experiment a thousand times with the particle initially prepared in the $|\chi\rangle$-state, then we would expect the particle to be found in the $|\psi\rangle$-state in around 250 runs of the experiment. Given the Born Rule, we would thus expect $|\langle\hat{\mathbf{a}}-|\hat{\mathbf{a}}+\rangle|^2$ to be 0, from which it will follow that $\langle\hat{\mathbf{a}}-|\hat{\mathbf{a}}+\rangle$ has to be 0. The Born Rule also implies that $|\langle\hat{\mathbf{a}}+|\hat{\mathbf{a}}+\rangle|^2 = 1$. We will also insist that $\langle\psi|\chi\rangle = \overline{\langle\chi|\psi\rangle}$,[7] from which it will follow that $\langle\hat{\mathbf{a}}+|\hat{\mathbf{a}}+\rangle$ is a real number of modulus 1 (i.e. $+1$ or -1). By convention, we choose $\langle\psi|\chi\rangle$ such that $\langle\psi|\psi\rangle$ is a real and positive number, in which case we must have $\langle\hat{\mathbf{a}}+|\hat{\mathbf{a}}+\rangle = 1$. If we now rotate the magnetic field by an angle θ as indicated in figure 1.4, the particle will be detected either at location $\hat{\mathbf{b}}+$ or location $\hat{\mathbf{b}}-$. We can then ask the question "given that the particle is in state $|\hat{\mathbf{a}}+\rangle$, what is

---

6. Regarding the set of complex numbers $\mathbb{C}$, we will use the notation $i = \sqrt{-1}$. Complex conjugation will be denoted by an overline so that $\overline{x+iy} = x - iy$ for real numbers x and y. The modulus of a complex number $z = x+iy$ will then be given by $|z| = \sqrt{z\bar{z}} = \sqrt{x^2+y^2}$. Since the defining property of $\langle\psi|\chi\rangle$ is that $|\langle\psi|\chi\rangle|^2$ is the probability that the particle will be found to be in state $|\psi\rangle$ given that we know that the particle is in state $|\chi\rangle$, we have to choose an arbitrary phase to fully determine $\langle\psi|\chi\rangle$.

7. Note that this condition implies time symmetry: the probability a particle transitions from a state $|\chi\rangle$ to a state $|\psi\rangle$ will be the same as the probability a particle transitions from the state $|\psi\rangle$ to the state $|\chi\rangle$. This is in accord with the observation that closed quantum systems are symmetric on time-reversal. A more precise formulation of this statement is given by the CPT theorem e.g. see Steven Weinberg, *The quantum theory of fields. Volume 1, Foundations [electronic resource]*, Cambridge core (Cambridge, 1995), 244. This might at first seem surprising in the light of the fact that phenomena such as radioactive decay are not obviously time-symmetric. However, it turns out that this time asymmetry results from the quantum system not being closed. For more details, see Saverio Pascazio, "All you ever wanted to know about the quantum Zeno effect in 70 minutes," *44th Symposium on Mathematical Physics on New Developments in the Theory of Open Quantum Systems*, 2013, https://doi.org/10.1142/S1230161214400071, eprint: arXiv:1311.6645v1[quant-ph].

the probability that the particle will be found to be in state $|\hat{\mathbf{b}}+\rangle$?" According to the notation discussed above, this probability will be $|\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}+\rangle|^2$ where $\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}+\rangle$ is a complex number such that $\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}+\rangle = 1$ when $\theta = 0$ and $\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}+\rangle = 0$ when $\theta = 180°$. We would likewise expect $\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}-\rangle = 0$ when $\theta = 0$ and $\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}-\rangle = 1$ when $\theta = 180°$. Since $\cos 0° = \sin 90° = 1$ and $\cos 90° = \sin 0° = 0$, we might guess that in general $|\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}+\rangle| = |\cos(\theta/2)|$ and $|\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}-\rangle| = |\sin(\theta/2)|$.[8] Experimentation shows us that this guess is correct. This suggests that we can express the state $|\hat{\mathbf{b}}+\rangle$ in terms of the states $|\hat{\mathbf{a}}+\rangle$ and $|\hat{\mathbf{a}}-\rangle$. We thus suppose that

$$|\hat{\mathbf{b}}+\rangle = \alpha\,|\hat{\mathbf{a}}+\rangle + \beta\,|\hat{\mathbf{a}}-\rangle \tag{1.1.1a}$$

$$|\hat{\mathbf{b}}-\rangle = \bar{\alpha}\,|\hat{\mathbf{a}}-\rangle - \bar{\beta}\,|\hat{\mathbf{a}}+\rangle \tag{1.1.1b}$$

for complex numbers $\alpha, \beta \in \mathbb{C}$ such that $|\alpha|^2 + |\beta|^2 = 1$, and we suppose that the bra-ket has the *linearity* property so that $\langle\psi|\hat{\mathbf{b}}+\rangle = \alpha\,\langle\psi|\hat{\mathbf{a}}+\rangle + \beta\,\langle\psi|\hat{\mathbf{a}}-\rangle$ and $\langle\psi|\hat{\mathbf{b}}-\rangle = \bar{\alpha}\,\langle\psi|\hat{\mathbf{a}}-\rangle - \bar{\beta}\,\langle\psi|\hat{\mathbf{a}}+\rangle$ for any state $|\psi\rangle$. Then it will follow that $\langle\hat{\mathbf{b}}+|\hat{\mathbf{b}}-\rangle = 0$,[9] and

8. Besides the guess being correct for angles such as $0°$ and $180°$, a further reason for thinking this is a good guess follows from the fact that since a particle in the $|\hat{\mathbf{b}}+\rangle$-state must be measured to be in either the $|\hat{\mathbf{a}}+\rangle$-state or the $|\hat{\mathbf{a}}-\rangle$-state when measured along the $\hat{\mathbf{a}}$-axes, from the Born Rule we must have

$$|\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}+\rangle|^2 + |\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}-\rangle|^2 = 1.$$

But for any angle θ we also have

$$|\cos(\theta/2)|^2 + |\sin(\theta/2)|^2 = 1,$$

and so setting $|\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}+\rangle| = |\cos(\theta/2)|$ and $|\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}-\rangle| = |\sin(\theta/2)|$ is consistent with these two equations.

9. To see this, by putting $|\psi\rangle = |\hat{\mathbf{b}}+\rangle$, we will have $\langle\hat{\mathbf{b}}+|\hat{\mathbf{b}}-\rangle = \bar{\alpha}\,\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}-\rangle - \bar{\beta}\,\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}+\rangle$ by equation (1.1.1b). Since $\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}-\rangle = \overline{\langle\hat{\mathbf{a}}-|\hat{\mathbf{b}}+\rangle}$ we have $\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}-\rangle = \bar{\beta}$ by equation (1.1.1a), and likewise, since $\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}+\rangle = \overline{\langle\hat{\mathbf{a}}+|\hat{\mathbf{b}}+\rangle}$, we have $\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}+\rangle = \bar{\alpha}$. Therefore, $\langle\hat{\mathbf{b}}+|\hat{\mathbf{b}}-\rangle = \bar{\alpha}\bar{\beta} - \bar{\beta}\bar{\alpha} = 0$.

that $\langle \hat{\mathbf{b}}+|\hat{\mathbf{b}}+\rangle = \langle \hat{\mathbf{b}}-|\hat{\mathbf{b}}-\rangle = 1.$[10] If we then put $\alpha = \cos(\theta/2)$ and $\beta = \sin(\theta/2)$, it

will follow that $|\langle \hat{\mathbf{b}}+|\hat{\mathbf{a}}+\rangle| = |\cos(\theta/2)|$ and $|\langle \hat{\mathbf{b}}+|\hat{\mathbf{a}}-\rangle| = |\sin(\theta/2)|,$[11] and so with

these values for α and β we will have

$$|\hat{\mathbf{b}}+\rangle = \cos(\theta/2)\,|\hat{\mathbf{a}}+\rangle + \sin(\theta/2)\,|\hat{\mathbf{a}}-\rangle, \tag{1.1.2a}$$

$$|\hat{\mathbf{b}}-\rangle = \cos(\theta/2)\,|\hat{\mathbf{a}}-\rangle - \sin(\theta/2)\,|\hat{\mathbf{a}}+\rangle. \tag{1.1.2b}$$

1.2 The Copenhagen Interpretation and the EPR-Bohm Paradox

Given the equations (1.1.2a) and (1.1.2b) that relate the states $|\hat{\mathbf{a}}\pm\rangle$ and $|\hat{\mathbf{b}}\pm\rangle$ to

each other, we can calculate probabilities such as the probability a particle will be

measured to be in the $|\hat{\mathbf{a}}+\rangle$-state given that it is in the $|\hat{\mathbf{b}}+\rangle$-state. There however

arises the question of what the physical meaning of one of these states is. Clearly,

the $|\hat{\mathbf{b}}+\rangle$-state says something about the spin of a particle; but is this a complete

description of the particle's spin state? For the $|\hat{\mathbf{b}}+\rangle$-state only tells us what the

outcome of a spin measurement would be along one particular axis $\hat{\mathbf{b}}$. For a spin

measurement along another axis $\hat{\mathbf{a}} \neq \pm\hat{\mathbf{b}}$, $|\hat{\mathbf{b}}+\rangle$ only tells us the probabilities (via

equation (1.1.2a) and the Born Rule) that the measurement outcome would be $|\hat{\mathbf{a}}+\rangle$

---

10. To see this, by putting $|\psi\rangle = |\hat{\mathbf{b}}+\rangle$ and using equation (1.1.1a), we will have $\langle \hat{\mathbf{b}}+|\hat{\mathbf{b}}+\rangle = \alpha\langle \hat{\mathbf{b}}+|\hat{\mathbf{a}}+\rangle + \beta\langle \hat{\mathbf{b}}+|\hat{\mathbf{a}}-\rangle = \alpha\overline{\alpha} + \beta\overline{\beta} = |\alpha|^2 + |\beta|^2 = 1.$ By a similar calculation, we also see $\langle \hat{\mathbf{b}}+|\hat{\mathbf{b}}+\rangle = 1.$

11. To satisfy these criteria, α and β are only determined up to rotation by a complex number. Rotating a complex number $z \in \mathbb{C}$ just means multiplying it by a complex number λ of modulus 1 (i.e. $|\lambda| = 1$) to get λz. We would need to take into account this rotation factor if we considered the three-dimensional situation. Then, without loss of generality, $\alpha = \cos(\theta/2)$ and $\beta = e^{i\phi}\sin(\theta/2)$ where θ and ϕ are the polar and azimuthal angles respectively.

or $|\hat{\mathbf{a}}-\rangle$ – but the $|\hat{\mathbf{b}}+\rangle$-state doesn't determine either of these outcomes. So we want to know whether this indetermination of the measurement outcome along the $\hat{\mathbf{a}}$-axis is merely a reflection of our lack of knowledge of a more complete specification of the particle's spin state, or alternatively, whether the $|\hat{\mathbf{b}}+\rangle$-state is a complete description of the spin state of the particle so that there is no fact of the matter about what spin state the particle would be found to be in along the $\hat{\mathbf{a}}$-axis until a measurement of spin along the $\hat{\mathbf{a}}$-axis is made.

Now Bohr and Heisenberg believed the latter to be the case. This was because their mathematical formalism of quantum physics implied that there are physical quantities of particles that couldn't be simultaneously determined. For example, their mathematical formalism is incapable of representing a particle which has a definite spin in both the $\hat{\mathbf{a}}$-direction and the $\hat{\mathbf{b}}$-direction when $\hat{\mathbf{a}} \neq \pm\hat{\mathbf{b}}$. So when a particle that is in the $|\hat{\mathbf{b}}+\rangle$-state is measured along the $\hat{\mathbf{a}}$-axis and is found to be in the $|\hat{\mathbf{a}}+\rangle$-state, there is a so-called collapse of the $|\hat{\mathbf{b}}+\rangle$-state:

$$|\hat{\mathbf{b}}+\rangle = \cos(\theta/2)\,|\hat{\mathbf{a}}+\rangle + \sin(\theta/2)\,|\hat{\mathbf{a}}-\rangle \xrightarrow{\text{Collapse!!}} |\hat{\mathbf{a}}+\rangle$$

so that after the measurement, the particle is no longer in the $|\hat{\mathbf{b}}+\rangle$-state. This interpretation of the quantum state as a complete physical description in which the state collapses to another state upon measurement is known as the *Copenhagen Interpretation*.

Einstein, Podolsky, and Rosen, however, strongly objected to the Copenhagen Interpretation, and they introduced their EPR paradox to explain what troubled them.[12] The EPR paradox was originally expressed in terms of the position and momentum of a particle, but it was Bohm who translated the EPR paradox to the context of spin,[13] and this is the version we will consider here.

The EPR-Bohm paradox arises in the context of particle pairs known as spin singlets. A *spin singlet* describes the state of two particles which a single particle of zero spin has decayed into. For example, a high energy *photon*, that is, a particle of light, can decay into a negatively charged electron, and a positively charged positron (where a *positron* is a fundamental particle like an electron but of opposite charge). Since spin is a conserved physical quantity, the spin of the two particles q_A and q_B of a spin singlet state must be equal and opposite when measured along the same axis, no matter what direction this axis happens to point in. The existence of spin singlet states thus raises the question of what the physical mechanism or principle is that ensures two experimenters, Alice and Bob say, will always obtain opposite spin measurement results if Alice measures the spin of particle q_A, and Bob measures the spin of particle q_B along the same axis. There is of course no experiment that could prove that Alice and Bob will always obtain opposite spin measurements, and there are some interpretations of quantum physics such as the GRW spontaneous

---

12. See A. Einstein, B. Podolsky, and N. Rosen, "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?," *Physical review* 47, no. 10 (1935): 777–780.

13. e.g. see D. Bohm, *Quantum Theory* (Englewood Cliffs: Prentice-Hall, 1951), p. 29, Ch. 5 sec. 3, and Ch. 22 sec. 19.

collapse theory[14] which predict that very occasionally Alice and Bob would obtain the same spin measurement result. But in this dissertation, we will assume that all measurements are consistent with *standard quantum theory*. In other words, we assume that the physical world can be described by quantum states[15] that evolve over time according to the Schrödinger equation, and that the probability of a system being found to be in one state given that it is in another state will be given by the Born Rule. In particular, under the assumption of standard quantum theory, it will follow that Alice and Bob will always obtain opposite spin results when performing measurements along the same axis of two particles in the spin singlet state.

Now naively, one would expect that if the spin of q_A were to be measured, then this would have no effect on any spin-measurement of q_B. This assumption is a special case of *Einstein's locality principle*: For two spatially separated systems $\mathcal{S}_1$ and $\mathcal{S}_2$, the real factual situation of the system $\mathcal{S}_2$ should be independent of what is done to the system $\mathcal{S}_1$.[16] This principle seemed so intuitive to Einstein, he was convinced it must

---

14. See Giancarlo Ghirardi and Angelo Bassi, "Collapse Theories," in *The Stanford Encyclopedia of Philosophy*, Summer 2020, ed. Edward N. Zalta (Metaphysics Research Lab, Stanford University, 2020).

15. In standard quantum theory, we remain agnostic as to whether a quantum state provides a complete description of the physical world, or whether it needs to be supplemented by some additional information in order to obtain a complete description.

16. Einstein expressed this locality principle in his autobiographical notes: "But on one supposition we should, in my opinion, absolutely hold fast: the real factual situation of the system $\mathcal{S}_2$ is independent of what is done with the system $\mathcal{S}_1$, which is spatially separated from the former." Albert Einstein, *Albert Einstein, Philosopher Scientist*, ed. P. A. Schilp (Evanston, Illinois: Library of Living Philosophers, 1949), p. 85.

be true.[17] Writing to Max Born, Einstein famously said "Physics should represent a reality in time and space, free from spooky action at a distance."[18] Although Einstein offered no argument in support of his principle of locality, several arguments have been proposed in its favor. For instance, Van Laer makes the following argument:

> Suppose that body A is located in void space, and that body P is now brought into its vicinity. If P is acted on by A, then either A's activity "across" the intervening space pre-existed P's arrival, in which case an action existed without an object and without a correlative passion, both of which run counter to the *ratio* of action; or A's activity only occurred when P became present to it, in which case A would somehow have to have "known" P's becoming present, which is not consonant with the nature of mere body as such.[19]

Now if Einstein's locality principle holds, we would be able to attribute a state $|\alpha\rangle_A$ to particle q_A, and a state $|\beta\rangle_B$ to particle q_B, so that if Alice were to perform a Stern-Gerlach experiment on particle q_A in which one of the possible outcomes was a

17. Walter Isaacson, *Einstein : His Life and Universe* (London: Pocket, 2008), 449.

18. From a letter written March 3, 1947, Max Born, Hedwig Born, and Irene Born, *The Born-Einstein Letters : Friendship, Politics and Physics in Uncertain Times* (Basingstoke, 2005), 155. Quoted from Isaacson, *Einstein : His Life and Universe*, 450.

19. See P. H. van Laer, *Philosophico-scientific problems*, Duquesne studies. Philosophical series ; 3 (Pittsburgh: Duquesne University Press, 1953), 59–144. The quote here is from Stanley F. Grove, "Quantum Theory and Aquinas's Doctrine on Matter" (Doctoral dissertation, The Catholic University of America, 2008), 246, who in turn quotes a slightly modified version of Van Laer's argument found in James Tallarico, "Action at a Distance," *The Thomist* (Washington, etc) 25, no. 2 (1962): 257. Grove also appeals to the Aristotelian doctrine of hylomorphism in order to defend the principle of locality:

> What does it mean for one hylomorphic substance to act on another? It is for the one to effect accidental or (by way of the accidental) substantial change in the other. As the accidents involved are inseparable from spatiality or extension, so must be the operations which occur through them. That is to say, that the interactions of cosmic (nonintellectual) beings are in essence spatial (and therefore spatiotemporal): they involve that which characterizes the spatial as such, i.e., "parts outside parts," extendedness and adjacency. For one hylomorphic being to act on another (and for that other to "receive" the action of the one), there must be an ontological ground or commonality which is located precisely in their spatiality; they must "share space,"

spin state $|\alpha'\rangle_A$, then by the Born Rule, the probability Alice would find q_A to be in state $|\alpha'\rangle_A$ would be $|\langle\alpha'|\alpha\rangle_A|^2$. Likewise, if Bob were to perform a Stern-Gerlach experiment on particle q_B in which one of the possible outcomes was a spin state $|\beta'\rangle$, then the probability Bob would find q_B to be in state $|\beta'\rangle_B$ would be $|\langle\beta'|\beta\rangle_B|^2$.

Now in order to decide how to represent the joint state of the particles q_A and q_B, we recall that in probability theory, we say that two events X and Y are *statistically independent* if and only if

$$P(X, Y) = P(X)P(Y) \tag{1.2.1}$$

where $P(X)$ is the probability that X occurs, $P(Y)$ is the probability that Y occurs, and $P(X, Y)$ is the probability that both X and Y both occur. We define the *conditional probability* $P(X|Y)$ of X given Y to be

$$P(X|Y) = \frac{P(X, Y)}{P(Y)}. \tag{1.2.2}$$

From (1.2.1), it is easy to see that when X and Y are independent with $P(Y) \neq 0$, then $P(X|Y) = P(X)$. We also say that two events X and Y are *conditionally independent* given some third event Z if and only if

$$P(X, Y|Z) = P(X|Z)P(Y|Z) \tag{1.2.3}$$

---

not in violation of the principle of unilocality but along a common boundary, thereby forming a unity through relation, i.e., "existing with respect to one another."

where $P(X|Z)$ is the conditional probability that X occurs given Z, $P(Y|Z)$ is the conditional probability that Y occurs given Z, and $P(X,Y|Z)$ is the conditional probability that both X and Y occur given Z.

Now if q_A can be described by the $|\alpha\rangle_A$-state and q_B can be described by the $|\beta\rangle_B$-state, the Born Rule implies that the conditional probability (conditioned on q_A being in the $|\alpha\rangle_A$-state and q_B being in the $|\beta\rangle_B$-state) that Alice would measure her particle to be in the state $|\alpha'\rangle_A$ is not going to depend on $|\beta\rangle_B$, and so $P(\alpha'|\alpha,\beta) = P(\alpha'|\alpha)$.[20] Likewise, $P(\beta'|\alpha,\beta) = P(\beta'|\beta)$. Therefore, if Alice and Bob's measurement outcomes are conditionally independent given the prior states of their particles, then according to (1.2.3) and the Born Rule, we would obtain the conditional probability

$$P_{AB}(\alpha',\beta'|\alpha,\beta) = |\langle\alpha'|\alpha\rangle_A|^2 \times |\langle\beta'|\beta\rangle_B|^2 = |\langle\alpha'|\alpha\rangle_A \langle\beta'|\beta\rangle_B|^2 \qquad (1.2.4)$$

where $P_{AB}(\alpha',\beta'|\alpha,\beta)$ is the conditional probability Alice will measure q_A to be in the $|\alpha'\rangle_A$-state and Bob will measure q_B to be in the $|\beta'\rangle_B$-state given that q_A is initially in the $|\alpha\rangle_A$-state and q_B is initially in the state $|\beta\rangle$. This suggests that if we write $|\alpha\rangle_A |\beta\rangle_B$ for the state of the composite system of both particles, then the bra-ket of $|\alpha'\rangle_A |\beta'\rangle_B$ and $|\alpha\rangle_A |\beta\rangle_B$ would be

$$_B\langle\beta'|\,_A\langle\alpha'|\alpha\rangle_A |\beta\rangle_B = \langle\alpha'|\alpha\rangle_A \langle\beta'|\beta\rangle_B. \qquad (1.2.5)$$

---

20. $P(\alpha'|\alpha,\beta)$ denotes the conditional probability that q_A will be measured to be in the $|\alpha'\rangle_A$-state given that q_A is in the $|\alpha\rangle_A$-state and q_B is in the $|\beta\rangle_B$-state.

We extend this bracket to sums of states so that it satisfies the linearity property (see page 10). However, when the particles q_A and q_B form a spin singlet, it will not be possible to express their joint state as $|\alpha\rangle_A |\beta\rangle_B$ because otherwise, according to (1.2.4), we will always be able to find a direction $\hat{\mathbf{a}}$ such that $P_{AB}(\hat{\mathbf{a}}+, \hat{\mathbf{a}}+|\alpha, \beta) \neq 0$, whereas in reality, the state Z describing the singlet has to satisfy $P_{AB}(\hat{\mathbf{a}}+, \hat{\mathbf{a}}+|Z) = 0$ for all $\hat{\mathbf{a}}$. But it turns out that the summation of states:

$$|\Psi_{\mathrm{Bell}}\rangle = \frac{1}{\sqrt{2}}(|\hat{\mathbf{a}}+\rangle_A |\hat{\mathbf{a}}-\rangle_B - |\hat{\mathbf{a}}-\rangle_A |\hat{\mathbf{a}}+\rangle_B). \tag{1.2.6}$$

can describe the singlet state of the two particles. We refer to the state (1.2.6) as a *Bell state*.[21] If the composite system is in the Bell state $|\Psi_{\mathrm{Bell}}\rangle$, then according to the Born Rule, the probability that Alice measures her particle to be in state $|\alpha\rangle$ and Bob measures his particle to be in the state $|\beta\rangle$ will be:

$$\begin{aligned} P_{AB}(\alpha, \beta|\Psi_{\mathrm{Bell}}) &= |_B\langle\beta|_A\langle\alpha|\Psi_{\mathrm{Bell}}\rangle|^2 \\ &= \frac{1}{2}|\langle\alpha|\hat{\mathbf{a}}+\rangle_A \langle\beta|\hat{\mathbf{a}}-\rangle_B - \langle\alpha|\hat{\mathbf{a}}-\rangle_A \langle\beta|\hat{\mathbf{a}}+\rangle_B|^2 \end{aligned} \tag{1.2.7}$$

This means that whatever axis Bob decides to measure along, if Alice measures her particle along the $\hat{\mathbf{a}}$-axis, then the Born Rule predicts that she will measure the particle to be in either the $|\hat{\mathbf{a}}+\rangle_A$-state or the $|\hat{\mathbf{a}}-\rangle_A$-state, each with probability

---

21. By convention, the states $\frac{1}{\sqrt{2}}(|\hat{\mathbf{a}}+\rangle_A |\hat{\mathbf{a}}-\rangle_B + |\hat{\mathbf{a}}-\rangle_A |\hat{\mathbf{a}}+\rangle_B)$, $\frac{1}{\sqrt{2}}(|\hat{\mathbf{a}}+\rangle_A |\hat{\mathbf{a}}+\rangle_B - |\hat{\mathbf{a}}-\rangle_A |\hat{\mathbf{a}}-\rangle_B)$, and $\frac{1}{\sqrt{2}}(|\hat{\mathbf{a}}+\rangle_A |\hat{\mathbf{a}}+\rangle_B + |\hat{\mathbf{a}}-\rangle_A |\hat{\mathbf{a}}-\rangle_B)$ are also referred to as Bell states.

of $\frac{1}{2}$.[22] But also, the Born Rule implies that if both Alice and Bob measure their

respective particles along the same $\hat{\mathbf{a}}$-axis, then the probability Bob will measure his

particle to have the same spin as Alice's particle will be zero, and the probability that

Bob measures his particle to have the opposite spin from Alice's particle will be one.[23]

These probabilities predicted by the Born Rule using the Bell state $|\Psi_{\text{Bell}}\rangle$ correspond

to the probabilities observed experimentally. Also note that despite the appearance

of the formula (1.2.6), $|\Psi_{\text{Bell}}\rangle$ is independent of the axis $\hat{\mathbf{a}}$. That is, for any other

---

22. To see this, suppose Bob performs his measurement along an arbitrary axis $\hat{\mathbf{b}}$. Then the probability Alice measures her particle to be in the $|\hat{\mathbf{a}}+\rangle$-state will be

$$
\begin{aligned}
P_{AB}(\hat{\mathbf{a}}+, \hat{\mathbf{b}}+|\Psi_{\text{Bell}}) + P_{AB}(\hat{\mathbf{a}}+, \hat{\mathbf{b}}-|\Psi_{\text{Bell}}) = \\
= \frac{1}{2}\big| \langle\hat{\mathbf{a}}+|\hat{\mathbf{a}}+\rangle_A \ \langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}-\rangle_B - \langle\hat{\mathbf{a}}+|\hat{\mathbf{a}}-\rangle_A \ \langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}+\rangle_B \big|^2 \\
+ \frac{1}{2}\big| \langle\hat{\mathbf{a}}+|\hat{\mathbf{a}}+\rangle_A \ \langle\hat{\mathbf{b}}-|\hat{\mathbf{a}}-\rangle_B - \langle\hat{\mathbf{a}}+|\hat{\mathbf{a}}-\rangle_A \ \langle\hat{\mathbf{b}}-|\hat{\mathbf{a}}+\rangle_B \big|^2 \\
= \frac{1}{2}\big| \langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}-\rangle_B \big|^2 + \frac{1}{2}\big| \langle\hat{\mathbf{b}}-|\hat{\mathbf{a}}-\rangle_B \big|^2 = \frac{1}{2}
\end{aligned}
\tag{1.2.8}
$$

where on the last line we have used the fact that

$$
|\hat{\mathbf{a}}-\rangle_B = |\hat{\mathbf{b}}+\rangle_B \ \langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}-\rangle_B + |\hat{\mathbf{b}}-\rangle \ \langle\hat{\mathbf{b}}-|\hat{\mathbf{a}}-\rangle_B
$$

and

$$
\langle\hat{\mathbf{a}}-|\hat{\mathbf{a}}-\rangle_B = 1, \quad \langle\hat{\mathbf{b}}\pm|\hat{\mathbf{b}}\pm\rangle_B = 1, \quad \text{and} \quad \langle\hat{\mathbf{b}}\pm|\hat{\mathbf{b}}\mp\rangle_B = 0
$$

so that

$$
|\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}-\rangle_B|^2 + \langle\hat{\mathbf{b}}-|\hat{\mathbf{a}}-\rangle_B|^2 = 1.
$$

23. This is because the probability Alice and Bob will measure their particles to have the same spin will be

$$
\begin{aligned}
P_{AB}(\hat{\mathbf{a}}+, \hat{\mathbf{a}}+|\Psi_{\text{Bell}}) + P_{AB}(\hat{\mathbf{a}}-, \hat{\mathbf{a}}-|\Psi_{\text{Bell}}) = \\
= \frac{1}{2}\big| \langle\hat{\mathbf{a}}+|\hat{\mathbf{a}}+\rangle_A \ \langle\hat{\mathbf{a}}+|\hat{\mathbf{a}}-\rangle_B - \langle\hat{\mathbf{a}}+|\hat{\mathbf{a}}-\rangle_A \ \langle\hat{\mathbf{a}}+|\hat{\mathbf{a}}+\rangle_B \big|^2 \\
+ \frac{1}{2}\big| \langle\hat{\mathbf{a}}-|\hat{\mathbf{a}}+\rangle_A \ \langle\hat{\mathbf{a}}-|\hat{\mathbf{a}}-\rangle_B - \langle\hat{\mathbf{a}}-|\hat{\mathbf{a}}-\rangle_A \ \langle\hat{\mathbf{a}}-|\hat{\mathbf{a}}+\rangle_B \big|^2 = 0.
\end{aligned}
$$

It therefore follows that Alice and Bob will measure their particles to have different spins with probability 1, since for any proposition q and its negation $\neg q$ (e.g. q = Alice and Bob measure their particles to have the same spin, and $\neg q$ = Alice and Bob measure their particles to have different spins), the sum of the probabilities of q and of $\neg q$ conditioned on any other proposition Ψ that has

direction $\hat{\mathbf{b}}$, it can be shown that,

$$\frac{1}{\sqrt{2}}(|\hat{\mathbf{a}}+\rangle_A |\hat{\mathbf{a}}-\rangle_B - |\hat{\mathbf{a}}-\rangle_A |\hat{\mathbf{a}}+\rangle_B) = \frac{1}{\sqrt{2}}(|\hat{\mathbf{b}}+\rangle_A |\hat{\mathbf{b}}-\rangle_B - |\hat{\mathbf{b}}-\rangle_A |\hat{\mathbf{b}}+\rangle_B).^{24} \quad (1.2.9)$$

Therefore, if Alice had chosen to measure her particle along the $\hat{\mathbf{b}}$-axis rather than the $\hat{\mathbf{a}}$-axis, she would still obtain equal probabilities for finding her particle to be in either the state $|\hat{\mathbf{b}}+\rangle_A$ or $|\hat{\mathbf{b}}-\rangle_A$, and the same equal probabilities for Bob's measurement outcomes hold as well.

Now although attributing the state $|\Psi_{\text{Bell}}\rangle$ to a spin singlet successfully determines the experimentally observed probabilities for Alice and Bob's spin measurement outcomes, this fact spells trouble for the Copenhagen interpretation. For if we stipulate that $|\Psi_{\text{Bell}}\rangle$ is a complete description of the spin state of the spin-singlet, then assuming Alice makes her measurement first and obtains the measurement

---

non-zero probability (e.g. the particles q_A and q_B are in the $|\Psi_{\text{Bell}}\rangle$ state) must be equal to 1, i.e. we must have $P(q|\Psi) + P(\neg q|\Psi) = 1$.

24. To see this, using the transformation rules given in equation (1.1.2) we have

$$\frac{1}{\sqrt{2}}(|\hat{\mathbf{b}}+\rangle |\hat{\mathbf{b}}-\rangle - |\hat{\mathbf{b}}-\rangle |\hat{\mathbf{b}}+\rangle)$$

$$= \frac{1}{\sqrt{2}}((\cos(\theta/2)|\hat{\mathbf{a}}+\rangle + \sin(\theta/2)|\hat{\mathbf{a}}-\rangle)(\cos(\theta/2)|\hat{\mathbf{a}}-\rangle - \sin(\theta/2)|\hat{\mathbf{a}}+\rangle)$$

$$- (\cos(\theta/2)|\hat{\mathbf{a}}-\rangle - \sin(\theta/2)|\hat{\mathbf{a}}+\rangle)(\cos(\theta/2)|\hat{\mathbf{a}}+\rangle + \sin(\theta/2)|\hat{\mathbf{a}}-\rangle))$$

$$= \frac{1}{\sqrt{2}}(\cos(\theta/2)|\hat{\mathbf{a}}+\rangle \cos(\theta/2)|\hat{\mathbf{a}}-\rangle - \cos(\theta/2)|\hat{\mathbf{a}}+\rangle \sin(\theta/2)|\hat{\mathbf{a}}+\rangle$$

$$+ \sin(\theta/2)|\hat{\mathbf{a}}-\rangle \cos(\theta/2)|\hat{\mathbf{a}}-\rangle - \sin(\theta/2)|\hat{\mathbf{a}}-\rangle \sin(\theta/2)|\hat{\mathbf{a}}+\rangle$$

$$- \cos(\theta/2)|\hat{\mathbf{a}}-\rangle \cos(\theta/2)|\hat{\mathbf{a}}+\rangle - \cos(\theta/2)|\hat{\mathbf{a}}-\rangle \sin(\theta/2)|\hat{\mathbf{a}}-\rangle$$

$$+ \sin(\theta/2)|\hat{\mathbf{a}}+\rangle \cos(\theta/2)|\hat{\mathbf{a}}+\rangle + \sin(\theta/2)|\hat{\mathbf{a}}+\rangle \sin(\theta/2)|\hat{\mathbf{a}}-\rangle)$$

$$= \frac{1}{\sqrt{2}}((\cos^2(\theta/2) + \sin^2(\theta/2))|\hat{\mathbf{a}}+\rangle |\hat{\mathbf{a}}-\rangle - (\cos^2(\theta/2) + \sin^2(\theta/2))|\hat{\mathbf{a}}-\rangle |\hat{\mathbf{a}}+\rangle)$$

$$= \frac{1}{\sqrt{2}}(|\hat{\mathbf{a}}+\rangle |\hat{\mathbf{a}}-\rangle - |\hat{\mathbf{a}}-\rangle |\hat{\mathbf{a}}+\rangle).$$

outcome $|\hat{\mathbf{a}}+\rangle_A$, then as this measurement outcome is established there will be a collapse of the state

$$|\Psi_{\text{Bell}}\rangle = \frac{1}{\sqrt{2}} (|\hat{\mathbf{a}}+\rangle_A |\hat{\mathbf{a}}-\rangle_B - |\hat{\mathbf{a}}-\rangle_A |\hat{\mathbf{a}}+\rangle_B) \xrightarrow{\text{Collapse!!}} |\hat{\mathbf{a}}+\rangle_A |\hat{\mathbf{a}}-\rangle_B, \quad (1.2.10)$$

whereas if Alice makes her measurement first and obtains the measurement outcome $|\hat{\mathbf{a}}-\rangle_A$, there will be a state collapse

$$|\Psi_{\text{Bell}}\rangle = \frac{1}{\sqrt{2}} (|\hat{\mathbf{a}}+\rangle_A |\hat{\mathbf{a}}-\rangle_B - |\hat{\mathbf{a}}-\rangle_A |\hat{\mathbf{a}}+\rangle_B) \xrightarrow{\text{Collapse!!}} |\hat{\mathbf{a}}-\rangle_A |\hat{\mathbf{a}}+\rangle_B. \quad (1.2.11)$$

Thus, according to the Copenhagen interpretation, prior to Alice's measurement, the real factual situation of particle q_B will be indeterminate, but once Alice has made her measurement, if she obtains the outcome $|\hat{\mathbf{a}}+\rangle_A$, the real factual situation of particle q_B will be expressed by the determinate state $|\hat{\mathbf{a}}-\rangle_B$, and if she obtains the outcome $|\hat{\mathbf{a}}-\rangle_A$, the real factual situation of particle q_B will be expressed by the determinate state $|\hat{\mathbf{a}}+\rangle_B$. But in either case, in the state collapse that Alice brings about via her measurement, there is a violation of Einstein's locality principle.

But things get even worse when we take Einstein's theory of special relativity into consideration. For according to special relativity, the order in which Alice and Bob perform their measurements is going to depend on their respective velocities[25] and on

---

25. *Velocity* is a vector that specifies both the speed (the velocity's magnitude) of something and its direction. In three dimensions, a three-tuple $\mathbf{v} = (v_x, v_y, v_z)$ is needed to specify the velocity of something, where the v_x component specifies how fast the thing is moving along the x-axis, and

which inertial frame of reference one is in.[26] Thus, if we are moving at one velocity, it may appear that Alice makes her measurement first meaning that Alice causes the state collapse and thus affects the state of Bob's particle, whereas if we are moving at another velocity, it may appear that Bob makes his measurement first causing the state to collapse and thus affects the state of Alice's particle. Once the collapse has occurred, there is no further collapse when the second experimenter makes his or her measurement along the same axis as the first experimenter. Thus, the thesis of the Copenhagen interpretation that it is the act of measurement that causes the state to collapse is not compatible with Einstein's theory of special relativity, for if special relativity is correct, then one and the same measurement both will and will not cause the state to collapse depending upon which frame of reference one is in.

Therefore, if one is more convinced (as most physicists are) of the truth of special relativity than the truth of the Copenhagen interpretation, then one must reject the Copenhagen interpretation.[27] Thus, in the light of the EPR-Bohm paradox we need an alternative to the Copenhagen interpretation of quantum states that can account

---

similarly, the v_y and v_z components specify how fast the thing is moving along the y and z axes respectively.

26. In special relativity, an inertial frame of reference is a spacetime coordinate system (t, x, y, z) in which all objects which have no forces acting on them have trajectories that are straight lines. Thus, we can move to another inertial frame by moving to a reference frame with constant velocity $\mathbf{v}$ with respect to the first reference frame. In the case when $\mathbf{v} = (v, 0, 0)$, Einstein's theory of special relativity tells us that under such a "boost", spacetime coordinates will transform as $(t, \mathbf{x}) \to (t', x', y', z') = (\gamma(t - \frac{vx}{c^2}), \gamma(x - vt), y, z)$ where c is the speed of light and $\gamma = \left(\sqrt{1 - \frac{v^2}{c^2}}\right)^{-1}$. For details of the derivation of this transformation, see footnote 18 on page 192.

27. One might still try to imagine models of reality in which special relativity and the Copenhagen interpretation are compatible. For example, here's a scenario suggested by my advisor Alexander Pruss: suppose that whenever there are two spacelike separated measurements of an entangled system of two

for how the two experimenters of the paradox always obtain opposite results when they make their measurement along the same axis. The obvious approach to take is to suppose that the quantum state $|\Psi_{\text{Bell}}\rangle$ is an incomplete description of the spin state of the two particles. This is the hidden variables approach, but as we will see shortly, this approach is also problematic. But before we consider the hidden variables approach, we will turn to what is traditionally considered to be one of the most troubling aspect of the Copenhagen interpretation, namely the possibility of scenarios exemplified by Schrödinger's cat thought experiment.

1.3 Schrödinger's Cat

Schrödinger's famous thought experiment was first discussed in a 1935 paper in the context of the EPR paradox and the inherent problems it raised for the Copenhagen

---

particles, an angel flips a coin to decide which measurement collapses the system first in the causal order. Then if reality were like this, perhaps special relativity and the Copenhagen interpretation could be compatible. However, a problem with this model is that it entails the possibility of backwards in time causality. This is because according to the Copenhagen interpretation, a measurement causes the state to collapse if and only if the system being measured is in a superposition of different measurement outcomes immediately before the measurement is made. So if an angel flips a coin which determines that Alice is the person whose measurement makes the state collapse, then if we choose an inertial frame in which Bob makes his measurement before Alice, then since Bob's measurement doesn't cause the state to collapse, the state of the particle he is measuring must already be in one of the measurement outcome states. But since the angel determines that Alice makes the state collapse, this would mean that as the entangled state of both particles evolved via the Schrödinger equation towards the time when Alice makes her measurement, Bob's particle would have to be in a superposition of possible measurement outcomes until Alice made her measurement. So for Bob not to collapse the state, the act of Alice's measurement would have to change the history of Bob's particle so that it was in a definite outcome state immediately before he made his measurement. So if one is willing to grant the possibility of backwards in time causality so that some action could change the past, then maybe special relativity and the Copenhagen interpretation are compatible. Price takes seriously the possibility of backwards causality in Huw Price, "A Neglected Route to Realism about Quantum Mechanics," *Mind* (London, etc) 103, no. 411 (1994): 303–336. However, if one were to insist that the impossibility of backwards in time causality was an implicit assumption of special relativity and/or quantum mechanics, then at least in the model just considered here, it doesn't look like special relativity and the Copenhagen interpretation would be compatible.

interpretation.[28] According to the Copenhagen interpretation, given the Bell state

$$|\Psi_{\text{Bell}}\rangle = \frac{1}{\sqrt{2}}(\,|\hat{\mathbf{a}}+\rangle_A\,|\hat{\mathbf{a}}-\rangle_B - |\hat{\mathbf{a}}-\rangle_A\,|\hat{\mathbf{a}}+\rangle_B), \qquad\qquad (1.2.6 \text{ revisited})$$

there is no fact of the matter as to whether particle q_A is spin up and particle q_B is spin down (or vice versa) until an observation is made.

Now the strange goings-on at subatomic level might not initially give us much cause for concern about the nature of reality. For one might say that although it is an interesting curiosity that facts about subatomic particles are a bit fuzzy, our observation of definite facts in everyday life should convince us that such fuzziness can be brought into focus when we zoom out from the subatomic level to the macroscopic level. That is, we might suppose that the very many indefinite things on the small scale average out to give us something definite on the large scale. However, the Schrödinger's Cat thought experiment suggests that our confidence in there being definite facts at the macroscopic level is seriously undermined if the Copenhagen interpretation is correct. Schrödinger himself described the scenario in his thought experiment as ridiculous[29] indicating that he didn't think we should doubt the definiteness of facts

---

28. For Schrödinger's original reference to his thought experiment, see E. Schrödinger, "Die gegenwärtige Situation in der Quantenmechanik," *Die Naturwissenschaften* 23, no. 48 (November 1935): 807–812. For an English translation, see John D. Trimmer, "The Present Situation in Quantum Mechanics: A Translation of Schrödinger's "Cat Paradox" Paper," *Proceedings of the American Philosophical Society* 124, no. 5 (1980): 323–338.

29. See Trimmer, p. 328.

at the macroscopic level; rather, we should call into question the reasonableness of the Copenhagen interpretation.

The Schrödinger's cat thought experiment invites us to consider a scenario like the one depicted in (1.2.10) and (1.2.11), but instead of having two microscopic particles coupled together, we have a microscopic particle such as a radioactive atom coupled together with a macroscopic object such as a cat. Schrödinger suggests how this might be done. A cat is enclosed in a steel chamber in which there is a Geiger counter that is directed towards a small radioactive source, so that in the course of an hour, there is a probability of $\frac{1}{2}$ that the Geiger counter will click indicating that one of the radioactive atoms has decayed, and there is a probability of $\frac{1}{2}$ that the Geiger counter doesn't click because none of the radioactive atoms decay over the course of an hour. The Geiger counter itself is hooked up to a relay such that if the Geiger counter clicks, it releases a hammer which shatters a small flask of hydrocyanic acid causing the cat to die. The two possibilities are depicted in figure 1.6.
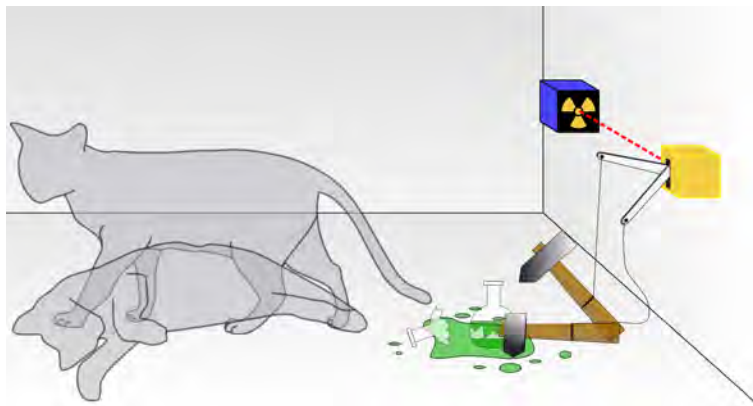
Figure 1.6. A depiction of Schrödinger's cat being both dead and alive.[30]

According to the Copenhagen interpretation, the cat will only enter into a determinate state once the box is opened at the end of the hour and an observation is made. Thus, there are two possibilities analogous to (1.2.10) and (1.2.11): either

$$\frac{1}{\sqrt{2}}\Big(|\text{No atoms decay}\rangle\, |\text{Cat Alive}\rangle - |\text{Atom decays}\rangle\, |\text{Cat Dead}\rangle \Big)$$

$$\xrightarrow{\text{Collapse!!}} |\text{No atoms decay}\rangle\, |\text{Cat Alive}\rangle\,,$$

or

$$\frac{1}{\sqrt{2}}\Big(|\text{No atoms decay}\rangle\, |\text{Cat Alive}\rangle - |\text{Atom decays}\rangle\, |\text{Cat Dead}\rangle \Big)$$

$$\xrightarrow{\text{Collapse!!}} |\text{Atom decays}\rangle\, |\text{Cat Dead}\rangle\,.$$

But before the box is opened and any measurement is made, the state of the cat and the atom will not be determinate, just as in the case of the spin-singlet where the spin states of the two particles are not determined to a definite value before any measurement is made. The Schrödinger's Cat thought experiment thus highlights that under the Copenhagen interpretation, there is no good reason to restrict indeterminacy to the microscopic level. If there is any indeterminacy at the microscopic level, then we should expect it at the macroscopic level as well, in which case it should be possible for a cat to be in an indeterminate state of being alive and dead. It is this possibility that Schrödinger found ridiculous.

---

30. Original by Dhatfield. This image is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license. Source: https://commons.wikimedia.org/wiki/File:Schrodingers_cat.svg

Given the problem with the Copenhagen interpretation that the EPR-Bohm paradox and Schrödinger's Cat highlight, it is tempting to suppose that states such as $|\hat{\mathbf{a}}+\rangle$ and $|\hat{\mathbf{a}}-\rangle$ merely represents our limited knowledge of a more complete physical state that would also include a specification of the particle's spin state along other axes besides the $\hat{\mathbf{a}}$-axis. If we were to make this supposition, there would be a fact of the matter, albeit unknown to us, concerning what spin state the particle would be found to be in were we to measure its spin along some $\hat{\mathbf{b}}$-axis for $\hat{\mathbf{b}} \neq \hat{\mathbf{a}}$. And even though we might decide not to measure the spin of the particle along the $\hat{\mathbf{b}}$-axis, there would still be this hidden fact about the particle's spin in this direction. Furthermore, since there would be no reason to suppose there was anything special about the $\hat{\mathbf{a}}$ or $\hat{\mathbf{b}}$-axes, it would then be reasonable to suppose that there were hidden facts about what spin direction the particle would be found to be in for every possible axis orientation. This would mean that a complete description of the particle's spin state would require an infinite list of outcomes for all the possible orientations we could configure the magnetic field of our Stern-Gerlach apparatus to be in. For example, the complete spin state of a particle which according to our limited knowledge was in the $|\hat{\mathbf{b}}+\rangle$-state could be depicted as $|\hat{\mathbf{a}}+; \hat{\mathbf{b}}+; \ldots\rangle$ or $|\hat{\mathbf{a}}-; \hat{\mathbf{b}}+; \ldots\rangle$, etc. where the ellipses would range over one of the two possible measurement outcomes for every other magnetic field orientation. However, because we would never in practice be able to perform all these experiments, and since only one such experiment would be needed to alter this

infinite list,[31] nearly all of the entries in this infinite list would remain forever hidden. Hence, this would be an example of a *hidden variables* interpretation of quantum theory. Moreover, if we're assuming Einstein's locality principle, it follows that any changes in these hidden spin outcomes for possible measurements of a specific particle can't affect the hidden variables of any other spatially separated localized particles. Therefore, given Einstein's locality principle, it is appropriate to refer to these hidden variables as *local* hidden variables.

Now although a local hidden variables theory seems rather intuitive, in 1964, John Bell derived an inequality based on the local hidden variables theory just described.[32] Moreover, it is now known that Bell's inequality can be violated experimentally.[33] This doesn't mean we must abandon all hidden variables theories. Rather it just means we need to abandon local hidden variables theories like the one Bell used in deriving his inequality. It is nevertheless somewhat paradoxical that local hidden variables cannot account for the correlations between the measurement outcomes of spin singlets in the EPR-Bohm paradox. Local hidden variables were meant to resolve the EPR-Bohm paradox, and so the violation of Bell's inequality really heightens this paradox. Therefore, any satisfactory interpretation of quantum physics has to face

---

31. In other words, it is assumed that directly measuring the particle will involve perturbing it so that its state will change.

32. In providing a derivation of Bell's inequality, we will follow Jim J. Napolitano and J. J. Sakurai, *Modern Quantum Mechanics* (Pearson Education, 2013), 241–249.

33. Aspect, Clauser and Zeilinger received the 2022 physics Nobel Prize for establishing the experimental violation of Bell's inequality. See "Press Release: The Nobel Prize in Physics 2022," *Nobel Prize Outreach*, October 2022, accessed January 2, 2023, https://www.nobelprize.org/prizes/physics/2022/summary/.

up to and resolve this paradox by offering a suitable alternative to a local hidden variables theory.

In order to describe Bell's inequality, we again consider two experimenters Alice and Bob making spin measurements on a spin singlet as described in section 1.2. So in each run of the experiment, a spin singlet consisting of two particles q_A and q_B will be generated with particle q_A being sent to Alice who measures q_A's spin in a direction of her choosing, and with particle q_B being sent to Bob who measures q_B's spin in a direction of his choosing. But now, instead of describing the two particles by the state $|\Psi_{\text{Bell}}\rangle$, we describe particles q_A and q_B in terms of all the spin outcomes one would obtain for every possible measurement axis in such a way that if q_A is spin up with respect to an axis $\hat{\mathbf{a}}$, then q_B would be spin down with respect to this axis. For example, if the complete spin state for q_A was given by $|\Psi_A\rangle = |\hat{\mathbf{a}}+; \hat{\mathbf{b}}+; \hat{\mathbf{c}}-; \ldots\rangle_A$, then the complete spin state for q_B would be given by $|\Psi_B\rangle = |\hat{\mathbf{a}}-; \hat{\mathbf{b}}-; \hat{\mathbf{c}}+; \ldots\rangle_B$.

We also assume that in each run of the experiment, Alice and Bob independently measure the spin of their particles along one of three possible directions $\hat{\mathbf{a}}$, $\hat{\mathbf{b}}$, and $\hat{\mathbf{c}}$, and that Einstein's locality principle holds. Furthermore, we assume that in each run of the experiment, the outcome of Alice's measurement will be statistically independent of any of the other measurement outcomes for different runs of the experiment, and for any of the three axes she measures along, she will get a spin up outcome or a spin down outcome with equal probability of $\frac{1}{2}$. Likewise, we assume Bob's measurement outcomes are also similarly independent between different runs of the experiment. We also assume that the $8 = 2^3$ states $|\hat{\mathbf{a}}\pm; \hat{\mathbf{b}}\pm; \hat{\mathbf{c}}\pm\rangle_A$ exhaust all the possible states for

Alice's particles that can be distinguished from one another by making one of the three possible measurement choices available. Thus, Alice can distinguish between the $|\hat{\mathbf{a}}+; \hat{\mathbf{b}}+; \hat{\mathbf{c}}+\rangle_A$-state and the $|\hat{\mathbf{a}}+; \hat{\mathbf{b}}+; \hat{\mathbf{c}}-\rangle_A$-state by making a measurement along the $\hat{\mathbf{c}}$-axis, though if she happened to make her measurement along the $\hat{\mathbf{a}}$ or $\hat{\mathbf{b}}$-axis, she wouldn't be able to distinguish between these two states. But in principle, she can distinguish between these two states if she happens to make her measurement along the right axis, in this case the $\hat{\mathbf{c}}$-axis. We similarly assume the states $|\hat{\mathbf{a}}\pm; \hat{\mathbf{b}}\pm; \hat{\mathbf{c}}\pm\rangle_B$ exhaust all the possible states for Bob's particles that he can distinguish between, and we assume that if Alice and Bob measure the particle along the same axis, they will always obtain opposite results from one another. For instance, if Alice's particle is in state $|\hat{\mathbf{a}}+; \hat{\mathbf{b}}+; \hat{\mathbf{c}}+\rangle_A$, then Bob's particle must be in state $|\hat{\mathbf{a}}-; \hat{\mathbf{b}}-; \hat{\mathbf{c}}-\rangle_B$. Now suppose the experiment is run N times for large N,[34] and let N_i be the number of times particle q_A is in the ith state so that[35] $N = \sum_{i=1}^{8} N_i$ as shown in table 1.1.

We define $P_{AB}(\hat{\mathbf{a}}+, \hat{\mathbf{b}}+)$ to be the probability that Alice measures particle q_A to be at location $\hat{\mathbf{a}}+$ on her detection screen and Bob measures particle q_B to be at location $\hat{\mathbf{b}}+$ on his detection screen where this probability is conditioned only on the assumption that the two particles are prepared in such a way as to guarantee Alice and Bob obtain opposite spin measurement outcomes if they perform their measurements along the same axis. We similarly define the probabilities for all other combinations of detection locations. It is relatively easy to calculate all these probabilities in terms

34. N has to be large since a frequentist definition of probability is being assumed.

35. The notation $\sum_{i=1}^{8} N_i$ is shorthand for $N_1 + N_2 + N_3 + N_4 + N_5 + N_6 + N_7 + N_8$.

Table 1.1. Spin-components of particles q_A and q_B in the hidden-variable theory

Population	Particle $\boldsymbol{q_A}$	Particle $\boldsymbol{q_B}$
N_1	$|\hat{\mathbf{a}}+;\hat{\mathbf{b}}+;\hat{\mathbf{c}}+\rangle_A$	$|\hat{\mathbf{a}}-;\hat{\mathbf{b}}-;\hat{\mathbf{c}}-\rangle_B$
N_2	$|\hat{\mathbf{a}}+;\hat{\mathbf{b}}+;\hat{\mathbf{c}}-\rangle_A$	$|\hat{\mathbf{a}}-;\hat{\mathbf{b}}-;\hat{\mathbf{c}}+\rangle_B$
N_3	$|\hat{\mathbf{a}}+;\hat{\mathbf{b}}-;\hat{\mathbf{c}}+\rangle_A$	$|\hat{\mathbf{a}}-;\hat{\mathbf{b}}+;\hat{\mathbf{c}}-\rangle_B$
N_4	$|\hat{\mathbf{a}}+;\hat{\mathbf{b}}-;\hat{\mathbf{c}}-\rangle_A$	$|\hat{\mathbf{a}}-;\hat{\mathbf{b}}+;\hat{\mathbf{c}}+\rangle_B$
N_5	$|\hat{\mathbf{a}}-;\hat{\mathbf{b}}+;\hat{\mathbf{c}}+\rangle_A$	$|\hat{\mathbf{a}}+;\hat{\mathbf{b}}-;\hat{\mathbf{c}}-\rangle_B$
N_6	$|\hat{\mathbf{a}}-;\hat{\mathbf{b}}+;\hat{\mathbf{c}}-\rangle_A$	$|\hat{\mathbf{a}}+;\hat{\mathbf{b}}-;\hat{\mathbf{c}}+\rangle_B$
N_7	$|\hat{\mathbf{a}}-;\hat{\mathbf{b}}-;\hat{\mathbf{c}}+\rangle_A$	$|\hat{\mathbf{a}}+;\hat{\mathbf{b}}+;\hat{\mathbf{c}}-\rangle_B$
N_8	$|\hat{\mathbf{a}}-;\hat{\mathbf{b}}-;\hat{\mathbf{c}}-\rangle_A$	$|\hat{\mathbf{a}}+;\hat{\mathbf{b}}+;\hat{\mathbf{c}}+\rangle_B$

of the values N_i from table 1.1,[36] or alternatively by simply measuring the frequency of these different outcomes for where Alice and Bob detect their particles. Note that the values of N_i will be unknown, but on the assumption that there is a fact of the matter of which states in table 1.1 obtain, and on the assumption that the states to which the N_i correspond exhaust all the possible states for Alice's and Bob's particles, we can show that[37]

$$P_{AB}(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+) \leq P_{AB}(\hat{\mathbf{a}}+,\hat{\mathbf{c}}+) + P_{AB}(\hat{\mathbf{c}}+,\hat{\mathbf{b}}+). \qquad (1.4.1)$$

This inequality is known as *Bell's inequality*, and it follows from Einstein's locality principle. However, as already mentioned, when this experiment is actually performed, we can choose the three axes so that Bell's inequality is violated. Nevertheless, it also

---

36. e.g. $P_{AB}(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+) = \frac{N_3+N_4}{N}$, $P_{AB}(\hat{\mathbf{a}}+,\hat{\mathbf{c}}+) = \frac{N_2+N_4}{N}$, $P_{AB}(\hat{\mathbf{c}}+,\hat{\mathbf{b}}+) = \frac{N_3+N_7}{N}$.

37. This inequality follows since

$$P_{AB}(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+) = \frac{N_3+N_4}{N} \leq \frac{N_2+N_4+N_3+N_7}{N} = P_{AB}(\hat{\mathbf{a}}+,\hat{\mathbf{c}}+) + P_{AB}(\hat{\mathbf{c}}+,\hat{\mathbf{b}}+).$$

turns out that this violation of Bell's inequality is entirely predictable if we assume that the state of the spin singlet consisting of the two particles q_A and q_B is given by the Bell state:

$$|\Psi_{\text{Bell}}\rangle = \frac{1}{\sqrt{2}}(\,|\hat{\mathbf{a}}+\rangle_A\,|\hat{\mathbf{a}}-\rangle_B - |\hat{\mathbf{a}}-\rangle_A\,|\hat{\mathbf{a}}+\rangle_B). \tag{1.4.2}$$

When the spin singlet is in the $|\Psi_{\text{Bell}}\rangle$-state, it can be shown that

$$P_{AB}(\hat{\mathbf{a}}+, \hat{\mathbf{b}}+) = \frac{1}{2}\sin^2(\theta/2) \tag{1.4.3}$$

where θ is the angle between the $\hat{\mathbf{a}}$-axis and $\hat{\mathbf{b}}$-axis.[38] Then taking the angle between the $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$-axes to be 90°, and the $\hat{\mathbf{c}}$-axis to be at 45° to both the $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$-axes, we would find that $P_{AB}(\hat{\mathbf{a}}+, \hat{\mathbf{b}}+) = \frac{1}{4}$ and $P_{AB}(\hat{\mathbf{a}}+, \hat{\mathbf{c}}+) + P_{AB}(\hat{\mathbf{c}}+, \hat{\mathbf{b}}+) = 0.1464\ldots$, and so Bell's inequality would be violated if we assumed that the probability of each outcome is determined by the Bell state (1.2.6).

1.5 Isolating the Culprit

Given the experimental violation of Bell's inequality, a strategy some philosophers of physics take is to reexamine the assumptions that lead to Bell's Inequality. Because of the observed violation of Bell's inequality, one of these assumptions will have to be

---

38. To see why this is, let $P_A(\hat{\mathbf{a}}+)$ be the probability that Alice would detect her particle at location $\hat{\mathbf{a}}+$ given that she is making a measurement along the $\hat{\mathbf{a}}$-axis, and let $P_{BA}(\hat{\mathbf{b}}+|\hat{\mathbf{a}}+)$ be the probability that Bob will detect his particle at location $\hat{\mathbf{b}}+$ given that he is making a measurement

discarded. The false assumption that is used to prove Bell's Inequality is sometimes referred to as *the culprit*.[39] We therefore need to isolate the culprit, that is, we need to decide which assumption we should discard while keeping in mind that we wish to maintain a theory that is compatible with the experimental findings of quantum physics and special relativity.

Shimony noticed that there are two key assumptions in the proof of Bell's Inequality that might be identified as the culprit. He refers to one assumption as Outcome Independence (OI), and to the other assumption as Parameter Independence (PI).[40] Shimony argued that if we only denied OI, then the proof of Bell's Inequality would fail to go through. Yet by continuing to assume PI, there is a sense in which special relativity is not obviously violated. Shimony therefore thought that denying OI and assuming PI was sufficient to ensure peaceful coexistence between quantum theory and special relativity. In other words, Shimony thought OI was the culprit.

---

along the $\hat{\mathbf{b}}$-axis and Alice has detected her particle at location $\hat{\mathbf{a}}+$. Given that the joint state of the particles is given by equation (1.2.6), $P_A(\hat{\mathbf{a}}+) = \frac{1}{2}$. But also note that if Alice has detected her particle at location $\hat{\mathbf{a}}+$, then Bob's particle must be in state $|\hat{\mathbf{a}}-\rangle$. From the Born Rule (see page 9) and equation (1.1.2a) it follows that

$$P_{BA}(\hat{\mathbf{b}}+|\hat{\mathbf{a}}+) = |\langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}-\rangle_B|^2 = \sin^2(\theta/2).$$

Therefore,

$$P_{AB}(\hat{\mathbf{a}}+, \hat{\mathbf{b}}+) = P_A(\hat{\mathbf{a}}+)P_{BA}(\hat{\mathbf{b}}+|\hat{\mathbf{a}}+) = \frac{1}{2}\sin^2(\theta/2).$$

39. This is the terminology Butterfield uses following Abner Shimony. e.g. see Jeremy Butterfield, "Peaceful Coexistence: Examining Kent's Relativistic Solution to the Quantum Measurement Problem," 2017, 1, eprint: arXiv:1710.07844.

40. See A. Shimony, "Events and processes in the quantum world," in *Search for a Naturalistic World View: Volume II: natural science and metaphysics* (1986; Cambridge: Cambridge University Press, 1993), 146–147.

1.6 Parameter Independence

To explain Shimony's[41] notion of Parameter Independence, we suppose we have an experimental setup similar to the experimental setup described in the previous section. Thus, we suppose there are two particles labeled q_A and q_B, and that a measurement can be made on particle q_A at one location (e.g. Alice's laboratory), and a measurement can be made on particle q_B at some other location (e.g. Bob's laboratory). We will assume that Alice can make a choice of one of n measurements to be made. These are labeled $a_1, \ldots, a_n$. For example, a_1 might be a measurement of q_A's spin along the z-axis, whereas a_2 might be the measurement of q_A's spin along an axis that is at a $45°$ angle to the z-axis etc. We use the variable x to denote Alice's choice so that $x = a_i$ for some $i \in \{1, \ldots, n\}$. If Alice chooses to make measurement a_i (i.e. $x = a_i$), the measurement outcome is labeled A_i, and this outcome can take values $+1$ or -1. For example, Alice could use the convention in which $+1$ corresponds to a spin up outcome, and -1 corresponds to a spin down outcome. We will use the variable X to denote the measurement outcome Alice obtains, so for example, if Alice chooses to make the a_1 measurement so that $x = a_1$ and obtains the outcome $A_1 = 1$, then $X = 1$. Similarly, we use the notation b_i, y, and B_i, Y to correspond to the measurement choices and measurement outcomes for Bob respectively.

We now suppose that there is a complete state $\lambda \in \Lambda$ describing both q_A and q_B that is independent of Alice and Bob's measurement choices, but that encodes all

---

41. See Shimony, "Events and processes in the quantum world," 146–147 and Butterfield, "Peaceful Coexistence: Examining Kent's Relativistic Solution to the Quantum Measurement Problem," 7–9.

other features that would influence the corresponding measurement outcomes. Here, the domain Λ of all such complete states will depend on how the two particles are prepared and the model we are assuming. We also assume that q_A and q_B are initially coupled together in such a way that Alice and Bob would always get opposite results when they made their measurements in the same direction. For instance, for $n = 3$, we might assume a model in which

$$\Lambda = \big\{(A_1, A_2, A_3, B_1, B_2, B_3) : A_1,\ A_2,\ A_3 = \pm 1,\ B_i = -A_i\big\}. \qquad (1.6.1)$$

In this case, $\lambda \in \Lambda$ would fully determine Alice and Bob's measurement outcomes along the three axes. This would be like the model described in the proof of Bell's Inequality with all the states of Λ being described in table 1.1 of section 1.2.

However, in general, we don't insist on such determinism. Rather, we suppose that given a complete state $\lambda \in \Lambda$, and given that Alice makes a measurement choice x and Bob makes a measurement choice y, then there will be a probability $P_{\lambda,x,y}(X,Y)$ representing the probability Alice gets outcome X and Bob gets outcome Y. In deterministic models in which there are outcomes,[42] $P_{\lambda,x,y}(X,Y)$ will have values restricted to either 0 or 1. In non-deterministic models, there will have to be some situations when $P_{\lambda,x,y}(X,Y)$ has a value strictly between 0 and 1. For example, if we assume the Copenhagen interpretation model, we could take λ to be the Bell state (1.2.6). Then it follows from equation (1.2.9) that as long as Alice and Bob's

---

42. In chapter 2, we will consider the many-worlds interpretation which is a deterministic model in which there aren't outcomes.

measurement choices x and y are in the same direction, then $P_{\lambda,x,y}(1,-1) = 1/2$.

Incidentally, we also note that equation (1.2.9) implies the domain Λ consists of a single state:

$$\Lambda = \left\{ \frac{1}{\sqrt{2}} \left(|\hat{\mathbf{a}}+\rangle_A \, |\hat{\mathbf{a}}-\rangle_B - |\hat{\mathbf{a}}-\rangle_A \, |\hat{\mathbf{a}}+\rangle_B \right) \right\}. \tag{1.6.2}$$

In both models (1.6.1) and (1.6.2), we see that if we define

$$P_{A,\lambda,x,y}(X) = P_{\lambda,x,y}(X,1) + P_{\lambda,x,y}(X,-1), \tag{1.6.3}$$

$$P_{B,\lambda,x,y}(Y) = P_{\lambda,x,y}(1,Y) + P_{\lambda,x,y}(-1,Y), \tag{1.6.4}$$

then $P_{A,\lambda,x,y}(X)$ (which is the probability Alice gets the output X regardless of what Bob's output is) is independent of Bob's choice of measurement y. And likewise, $P_{B,\lambda,x,y}(Y)$ (which is the probability Bob gets the output Y regardless of what Alice's output is) is independent of Alice's choice of measurement x.[43] In other models, however, it's possible that such independence does not hold. So to distinguish between such possibilities, we say a model satisfies *Parameter Independence* (PI) if and only

---

43. To see that this is true for model (1.6.1), it is obvious that $P_{A,\lambda,x,y}(X) = 1$ or 0 regardless of what y is. As for model (1.6.2), it is straightforward to show that $P_{A,\lambda,x,y}(X) = 1/2$ and $P_{B,\lambda,x,y}(Y) = 1/2$ for any X, Y. E.g. for $x = \hat{\mathbf{a}}$ and $y = \hat{\mathbf{b}}$, by (1.2.9), we can assume the two particles are in the state

$$|\zeta\rangle = \frac{1}{\sqrt{2}} \left(|\hat{\mathbf{b}}+\rangle_A \, |\hat{\mathbf{b}}-\rangle_B - |\hat{\mathbf{b}}-\rangle_A \, |\hat{\mathbf{b}}+\rangle_B \right).$$

Since the inner product on the composite system is given by $\langle\xi'|\xi\rangle = \langle\psi'|\psi\rangle_A \langle\chi'|\chi\rangle_B$ for $|\xi\rangle = |\psi\rangle_A \, |\chi\rangle_B$ and $|\xi'\rangle = |\psi'\rangle_A \, |\chi'\rangle_B$, it follows that

$$_A\langle\hat{\mathbf{a}}+| \, _B\langle\hat{\mathbf{b}}\pm|\zeta\rangle = \mp\frac{1}{\sqrt{2}} \langle\hat{\mathbf{a}}+|\hat{\mathbf{b}}\mp\rangle_A.$$

if $P_{A,\lambda,x,y}(X)$ is independent of y, and $P_{B,\lambda,x,y}(Y)$ is independent of x. In particular, PI holds in the model (1.6.1) and in the Copenhagen interpretation model (1.6.2). In other words, PI holds if and only if (1.6.3) and (1.6.4) hold for all λ, x, y, X, and Y. If PI fails to hold in a model, we say that the model satisfies *Parameter Dependence* (PD).

1.7 The Problem with Parameter Dependence

As is well known, special relativity implies that it is impossible to send signals faster than the speed of light. To see why this is, first recall that in order for light to travel at a constant speed regardless of how fast the source of light is traveling, it is necessary for clocks moving relative to an observer to run more slowly. For a proof of this fact, suppose Alice is an experimenter on earth and Bob is an experimenter on a spaceship which is travelling close to the speed of light away from the earth. We can imagine each of them having a simple clock in their possession which consists of a photon travelling back and forth between two mirrors one meter apart with a tick of the clock happening each time the photon hits one of the mirrors. If both their clocks are parallel to each other (i.e. the line joining Alice's two mirrors is parallel to

---

Therefore, by the Born Rule (see page 9)

$$P_{\lambda,\hat{\mathbf{a}},\hat{\mathbf{b}}}(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+) + P_{\lambda,\hat{\mathbf{a}},\hat{\mathbf{b}}}(\hat{\mathbf{a}}+,\hat{\mathbf{b}}-) = \frac{1}{2}|\langle\hat{\mathbf{a}}+|\hat{\mathbf{b}}-\rangle_A|^2 + \frac{1}{2}|\langle\hat{\mathbf{a}}+|\hat{\mathbf{b}}+\rangle_A|^2.$$

But since

$$|\hat{\mathbf{a}}+\rangle_A = \langle\hat{\mathbf{b}}+|\hat{\mathbf{a}}+\rangle_A \, |\hat{\mathbf{b}}+\rangle_A + \langle\hat{\mathbf{b}}-|\hat{\mathbf{a}}+\rangle_A \, |\hat{\mathbf{b}}-\rangle_A$$

it follows that

$$|\langle\hat{\mathbf{a}}+|\hat{\mathbf{b}}+\rangle_A|^2 + |\langle\hat{\mathbf{a}}+|\hat{\mathbf{b}}-\rangle_A|^2 = 1.$$

Therefore,

$$P_{\lambda,\hat{\mathbf{a}},\hat{\mathbf{b}}}(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+) + P_{\lambda,x,y}(\hat{\mathbf{a}}+,\hat{\mathbf{b}}-) = \frac{1}{2}.$$

the line joining Bob's two mirrors) and perpendicular to Bob's motion with respect to Alice, then Alice would observe the distance the photon travels between each tick of Bob's clock to be longer than the distance the photon travels between each tick of her clock. Since Alice would observe both the photon of her own clock and the photon of Bob's clock to travel at the same speed, she would therefore observe Bob's clock to tick more slowly than her own clock. This means that if their clocks were originally synced when Bob left the earth, and Alice determined Bob's clock to have ticked n_B times since his departure, then she would observe her own clock to have ticked $n_A = \gamma n_B$ times, where $\gamma > 1$ depends on how fast Bob is moving relative to her.

Now if Alice could send messages to Bob much faster than the speed of light, we could suppose she sends out a message to Bob when her clock has just ticked n_A times. If the message was transmitted almost instantaneously, Bob would receive the message at approximately the time when his clock had just ticked n_A/γ times. But if Bob then immediately replied to Alice's message with his message also transmitting almost simultaneously, then since Bob will also notice Alice's clock ticking more slowly with respect to his own clock, he will determine that Alice receives his reply message at approximately the time when her clock reads n_A/γ^2 ticks. But since $\gamma > 1$, this means that Alice will receive a reply to her message to Bob before she has sent her original message to him. This is clearly absurd.

Using a slightly more technical argument[44] one can also see that an absurdity follows even when the signal transmission isn't almost instantaneous, just so long as the signal propagates faster than the speed of light. Special relativity therefore implies that it is impossible to send signals faster than the speed of light.

Now a serious problem with PD is that it suggests the possibility of superluminal signalling. In particular, if Alice happened to know what λ was for each run of the experiment, and if Bob made the same measurement, then because the distribution of Alice's outcomes will depend on Bob's choice of measurement, with enough runs of the experiment, Alice should be able to work out what measurement Bob is making. And this should be possible even if Alice and Bob are separated by many light years. So it seems that faster than light communication could be possible. The only thing preventing such communication would be Alice's lack of knowledge of λ.

---

44. The more technical argument is as follows: we work in one spatial dimension and suppose that Bob is travelling away from Alice with velocity v on a trajectory (t, vt) in Alice's spacetime coordinates. At time t_A, she sends a message to Bob which travels at a velocity $v_m > v$ so that her message can catch up with Bob. Her message to Bob therefore has a trajectory $(t_A + T, v_m T)$ where T is the interval of time after she has sent her message. At the time t_{AB} in Alice's time coordinates when her message reaches Bob, her coordinates for Bob and her coordinates for her message location must coincide. Therefore,

$$(t_{AB}, vt_{AB}) = (t_A + T_{AB}, v_m T_{AB}),$$

where T_{AB} is the time it takes for her message to reach Bob as measured by Alice. Solving for t_{AB}, we therefore find that Alice's clock will show the time

$$t_{AB} = \frac{t_A}{1 - v/v_m}$$

at the moment Bob receives her message. But since Alice determines Bob's clock to be running more slowly than her own clock by a factor of γ, it follows that Bob's clock will show the time

$$t_B = \frac{t_A}{\gamma(1 - v/v_m)}$$

the moment he receives Alice's message.

Now suppose Bob immediately replies to Alice's message with his message propagating back to Alice with velocity $-v_m$. In Bob's coordinates, Alice will have a trajectory $(t, -vt)$, and Bob's message to her will have a trajectory $(t_B + T, -v_m T)$. Since these two trajectories must coincide at the moment

39

One might still respond to this argument against PD by saying that in reality, Alice does not know anything about λ and hence it won't be possible for Bob to send signals faster than the speed of light to Alice. However, it would nevertheless be very strange if the validity of special relativity hung on what human beings were capable of knowing rather than on the laws that actually governed physical reality itself. But even if it was metaphysically impossible for human beings to know λ, this would not allay the fears that adherents of special relativity would have against PD. For the real issue is not so much the possibility that Alice could translate a message that was transmitted from Bob at superluminal speed. Rather, the issue is that Bob could do something whose effect was propagated at superluminal speed. If PI fails, then superluminal propagation of effects will be possible. It's just that if Alice is to know the cause of this effect, she will need to know λ. But even if she doesn't know λ, it still seems as though some unknown effect has been transmitted to her superluminally,

---

Bob's reply reaches Alice, it follows that Bob's clock will show a time t_{BA} which satisfies the equation

$$(t_{BA}, -vt_{BA}) = (t_B + T_{BA}, -v_m T_{BA})$$

where T_{BA} is the time it takes for Bob's message to Alice as measured by Bob. Solving for t_{BA}, we therefore find that Bob's clock will show the time

$$t_{BA} = \frac{t_B}{1 - v/v_m}$$

when Alice receives Bob's reply. But since Bob determines Alice's clock to be running more slowly than his own clock by a factor of γ, it follows that Alice's clock will show the time

$$t_{AA} = \frac{t_B}{\gamma(1 - v/v_m)} = \frac{t_A}{\gamma^2(1 - v/v_m)^2}$$

the moment she receives Bob's reply. Now an absurdity will only arise if $t_{AA} < t_A$ since then Alice would receive a reply from Bob before she had sent her message. Thus, an absurdity will arise if both $\gamma^2(1 - v/v_m)^2 > 1$ and $v_m > v$, where as already mentioned, the second inequality must hold in order for Alice's message to reach Bob and vice versa. Now as we will see in footnote 18 on page 192,

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}.$$

and with the superluminal propagation of an effect comes the possibility that an effect might occur before its cause. Such a possibility would be analogous to how Alice could receive a reply to a message from Bob before she had sent the original message to him.

1.8 Parameter Dependence in the Bohmian Interpretation

One model which exhibits PD is the *Bohmian interpretation* of standard quantum theory. In this interpretation, it is assumed that at any instant of time t, the particles q_A and q_B will have definite positions $\mathbf{x}_A$ and $\mathbf{x}_B$ and definite momenta $\mathbf{p}_A$ and $\mathbf{p}_B$ respectively. But in addition to the positions and momenta of the particles, it is also assumed that there is a so-called *pilot wave*

$$\psi(\mathbf{x}_A, \mathbf{x}_B, t) = r(\mathbf{x}_A, \mathbf{x}_B, t)e^{iS(\mathbf{x}_A, \mathbf{x}_B, t)} \tag{1.8.1}$$

A simple calculation therefore shows that the inequality $t_{AA} < t_A$ implies

$$v > \frac{2v_m}{1 + v_m^2/c^2}.$$

Combining this with the inequality $v < v_m$, we see that an absurdity will be possible if

$$v_m > \frac{2v_m}{1 + v_m^2/c^2}$$

or equivalently, if $v_m > c$.

where $r(\mathbf{x}_A, \mathbf{x}_B, t) > 0$ is the modulus of $\psi(\mathbf{x}_A, \mathbf{x}_B, t)$, and the real-valued function[45]

$S(\mathbf{x}_A, \mathbf{x}_B, t)$ is the complex phase[46] of $\psi(\mathbf{x}_A, \mathbf{x}_B, t)$. The time evolution of the pilot wave

is deterministically governed by the Schrödinger equation, and the phase $S(\mathbf{x}_A, \mathbf{x}_B, t)$

relates the positions $\mathbf{x}_A$ and $\mathbf{x}_B$ to the momenta $\mathbf{p}_A$ and $\mathbf{p}_B$ via the gradient[47] of S:

$$\mathbf{p}_A = \boldsymbol{\nabla}_A S(\mathbf{x}_A, \mathbf{x}_B), \qquad \mathbf{p}_B = \boldsymbol{\nabla}_B S(\mathbf{x}_A, \mathbf{x}_B). \tag{1.8.2}$$

In other words, if we fix $\mathbf{x}_B$ and consider S to be just a function of $\mathbf{x}_A$, then the

momentum $\mathbf{p}_A$ is in the direction and has the magnitude of the steepest ascent of S

considered as a function of $\mathbf{x}_A$. The momentum $\mathbf{p}_B$ is determined similarly.

In reality, we don't know the exact positions of all the particles, but based on

what we know about an experimental setup, we can average over our uncertainty and

45. A *function* is a mathematical object that takes an element from a set called the function's *domain* as its input, and returns an output from a set called the function's *range*. A *real-valued function* is a function whose range is contained within the set of real numbers. If x is in the domain of a function f, we denote the corresponding output of the function f as $f(x)$. The sin function is an example of a real-valued function whose domain is the set of real numbers and whose range is the set of real numbers between -1 and 1. Some functions need to take several mathematical objects as input to produce an output. For example, the function S in equation (1.8.1) takes two vectors $\mathbf{x}_A$ and $\mathbf{x}_B$, and a real number t as input from which it produces an output denoted by $S(\mathbf{x}_A, \mathbf{x}_B, t)$.

46. The *phase* of a complex number z is the angle θ (in radians) such that $z = r(\cos\theta + i\sin\theta)$ where $r > 0$ is a real number called the *modulus*. Since $\cos\theta + i\sin\theta = e^{i\theta}$, it follows from (1.8.1) that $r(\mathbf{x}_A, \mathbf{x}_B, t) > 0$ is the modulus of $\psi(\mathbf{x}_A, \mathbf{x}_B, t)$, and $S(\mathbf{x}_A, \mathbf{x}_B, t)$ is the phase of $\psi(\mathbf{x}_A, \mathbf{x}_B, t)$.

47. Given a smooth function F that maps a vector $\mathbf{x}$ to a real number $F(\mathbf{x})$, the *gradient* of F at $\mathbf{x}$, written $\boldsymbol{\nabla}F(\mathbf{x})$ is the vector which points in the direction of the steepest ascent of F at $\mathbf{x}$ and whose magnitude is determined by the rate of the steepest ascent at $\mathbf{x}$. When F is a smooth function mapping two vectors $\mathbf{x}_A$ and $\mathbf{x}_B$ to a real number $F(\mathbf{x}_A, \mathbf{x}_B)$, we write $\boldsymbol{\nabla}_A F(\mathbf{x}_A, \mathbf{x}_B)$ for the gradient of F at $\mathbf{x}_A$ when F is just considered as a function of $\mathbf{x}_A$ with $\mathbf{x}_B$ fixed, and we write $\boldsymbol{\nabla}_B F(\mathbf{x}_A, \mathbf{x}_B)$ for the gradient of F at $\mathbf{x}_B$ when F is just considered as a function of $\mathbf{x}_B$ with $\mathbf{x}_A$ fixed.

recover exactly the same predictions that standard quantum theory would make.[48] So for instance, our knowledge of the experimental setup above should enable us to know that q_A and q_B are contained within a bounded region V, and the experimental setup should also enable us to work out the probability $p(V_i, V_j)$ that particle q_A will be in a region V_i, and particle q_B will be in a region V_j, where the V_i are small non-overlapping regions such that $V = \bigcup_i V_i$. If we are interested in some physical quantity $O(\mathbf{x}_A, \mathbf{x}_B)$ that depends on the positions $\mathbf{x}_A$ and $\mathbf{x}_B$ of the two particles, then when the regions V_i are sufficiently small so that almost everywhere,[49] $O(\mathbf{x}_i, \mathbf{x}_j)$ varies negligibly for any $\mathbf{x}_i \in V_i$ and $\mathbf{x}_j \in V_j$, the average value (also known as the *expectation value*)

$$\langle O \rangle = \sum_{i,j} p(V_i, V_j) O(\mathbf{x}_i, \mathbf{x}_j) \tag{1.8.3}$$

---

48. See David Bohm, "A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. I," *Physical review* 85, no. 2 (1952): 166–179 and David Bohm, "A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. II," *Physical review* 85, no. 2 (1952): 180–193.

49. I am using the expression "almost everywhere" in a measure theoretic sense to mean that the set of all points where discontinuities occur has measure 0. The motivation for this qualification is that whilst it might seem reasonable to suppose that there could be physical quantities that are discontinuous in some places, it doesn't seem reasonable to suppose that there are physical quantities that could be pathologically discontinuous, e.g. we wouldn't expect there to be a physical quantity that took the value of 1 unit when all the coordinates of a location were rational and a value of 2 units when one of the coordinates was irrational. Given this qualification about $O(\mathbf{x}_A, \mathbf{x}_B)$ being continuous almost everywhere, then even though $O(\mathbf{x}_A, \mathbf{x}_B)$ could be discontinuous at some points so that there could be some $V_i \times V_j$ cells in which $O(\mathbf{x}_A, \mathbf{x}_B)$ varied non-negligibly for $(\mathbf{x}_A, \mathbf{x}_B) \in V_i \times V_j$, it would still be the case that the sum of the $p(V_i, V_j) \max_{(\mathbf{x}_i, \mathbf{x}_j) \in V_i \times V_j} |O(\mathbf{x}_i, \mathbf{x}_j)|$ terms for such cells would be negligible. This follows from the Lebesgue-Vitali theorem that states that any bounded function on a compact interval is Riemann integrable if and only if it is continuous almost everywhere. See Wikipedia contributors, *Riemann integral — Wikipedia, The Free Encyclopedia*, [Online; accessed 07-July-2023], 2023, https://en.wikipedia.org/wiki/Riemann_integral.

calculated in the Bohmian interpretation turns out to be the same as the expectation value[50] for O predicted by standard quantum theory.[51]

To see why PI fails to hold in the Bohmian interpretation, we first note that since the Bohmian interpretation makes the same predictions as standard quantum theory when averaged over all the hidden variables, the violation of Bell's inequality (1.4.1) implies there must be some hidden variable λ and choices of measurement directions $\hat{\boldsymbol{a}}$, $\hat{\boldsymbol{b}}$, and $\hat{\boldsymbol{c}}$ such that

$$P_{\lambda,\hat{\boldsymbol{a}},\hat{\boldsymbol{b}}}(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+) > P_{\lambda,\hat{\boldsymbol{a}},\hat{\boldsymbol{c}}}(\hat{\mathbf{a}}+,\hat{\mathbf{c}}+) + P_{\lambda,\hat{\boldsymbol{c}},\hat{\boldsymbol{b}}}(\hat{\mathbf{c}}+,\hat{\mathbf{b}}+).^{52} \tag{1.8.4}$$

Since physics in the Bohmian interpretation is deterministic, probabilities must be either 0 or 1. Therefore, the only way (1.8.4) can be satisfied is for

$$P_{\lambda,\hat{\boldsymbol{a}},\hat{\boldsymbol{b}}}(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+) = 1, \tag{1.8.8}$$

---

50. See (2.2.1) for the definition of the expectation value of an observable in standard quantum theory.

51. In this explanation, I've refrained from using measure theory, but basically this explanation is saying that when we construct a measure μ on $V \times V$ based on our knowledge of the experimental setup, $\int_{V \times V} O(\mathbf{x}_i, \mathbf{x}_j) \mathrm{d}\mu$ will be the same as the expectation value for O predicted by standard quantum theory.

52. To see why this is, we consider the observable $O_{\hat{\mathbf{a}}+,\hat{\mathbf{b}}+}(\mathbf{x}_A, \mathbf{x}_B)$ which returns 1 if and only if the particle q_A will end up at location $\hat{\mathbf{a}}+$ and particle q_B will end up at location $\hat{\mathbf{b}}+$ as determined by the pilot wave, and otherwise $O_{\hat{\mathbf{a}}+,\hat{\mathbf{b}}+}(\mathbf{x}_A, \mathbf{x}_B)$ returns 0. Then by (1.8.3),

$$P_{AB}(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+) = \sum_{\lambda} p_{\lambda} P_{\lambda,\hat{\boldsymbol{a}},\hat{\boldsymbol{b}}}(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+) \tag{1.8.5}$$

where $P_{AB}(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+)$ is the probability described in section 1.4, λ ranges over the pairs of indices (i,j) corresponding to the cell $V_i \times V_j$, $p_{\lambda} = p(V_i, V_j)$ for $\lambda = (i,j)$, and we have used the fact that $P_{\lambda,\hat{\boldsymbol{a}},\hat{\boldsymbol{b}}}(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+) = O_{\hat{\mathbf{a}}+,\hat{\mathbf{b}}+}(\mathbf{x}_i, \mathbf{x}_j)$ for almost all $(\mathbf{x}_i, \mathbf{x}_j) \in V_i \times V_j$. Now if (1.8.4) is false, then for

$$P_{\lambda,\hat{a},\hat{c}}(\hat{a}+,\hat{c}+) = 0, \tag{1.8.9}$$

$$P_{\lambda,\hat{c},\hat{b}}(\hat{c}+,\hat{b}+) = 0. \tag{1.8.10}$$

We suppose that PI holds, and we will try to arrive at a contradiction.[53] If both Alice

and Bob make their measurement in the $\hat{c}$-direction, there are two possibilities: either

Alice measures q_A to be in the state $\hat{c}-$ and Bob measures q_B to be in the state $\hat{c}+$,

or Alice measures q_A to be in the state $\hat{c}+$ and Bob measures q_B to be in the state

$\hat{c}-$. So expressed in terms of probabilities, these two possibilities are equivalent to

either

$$P_{\lambda,\hat{c},\hat{c}}(\hat{c}-,\hat{c}+) = 1 \qquad \text{and} \qquad P_{\lambda,\hat{c},\hat{c}}(\hat{c}+,\hat{c}-) = 0, \tag{1.8.11}$$

or

$$P_{\lambda,\hat{c},\hat{c}}(\hat{c}+,\hat{c}-) = 1 \qquad \text{and} \qquad P_{\lambda,\hat{c},\hat{c}}(\hat{c}-,\hat{c}+) = 0. \tag{1.8.12}$$

---

all λ, $\hat{a}$, and $\hat{b}$,

$$P_{\lambda,\hat{a},\hat{b}}(\hat{a}+,\hat{b}+) \leq P_{\lambda,\hat{a},\hat{c}}(\hat{a}+,\hat{c}+) + P_{\lambda,\hat{c},\hat{b}}(\hat{c}+,\hat{b}+) \tag{1.8.6}$$

Therefore, since $p_\lambda \geq 0$, it follows that

$$\sum_\lambda p_\lambda P_{\lambda,\hat{a},\hat{b}}(\hat{a}+,\hat{b}+) \leq \sum_\lambda p_\lambda P_{\lambda,\hat{a},\hat{c}}(\hat{a}+,\hat{c}+) + \sum_\lambda p_\lambda P_{\lambda,\hat{c},\hat{b}}(\hat{c}+,\hat{b}+). \tag{1.8.7}$$

It would therefore follow from (1.8.7) and (1.8.5) that Bell's inequality (1.4.1) would hold. But since (1.4.1) is experimentally violated, it is not the case that (1.8.6) holds for all λ, $\hat{a}$, and $\hat{b}$. Hence, there must be some λ, $\hat{a}$, and $\hat{b}$, for which (1.8.4) holds.

53. For another proof of this result, see J. S. Bell, "On the Einstein Podolsky Rosen paradox," *Physics (New York. 1964)* 1, no. 3 (1964): 195–200.

It will be sufficient to show that (1.8.11) leads to a contradiction, since then we can just swap Alice and Bob to show that (1.8.12) also leads to a contradiction. So note that

$$P_{\lambda,\hat{c},\hat{c}}(\hat{\mathbf{c}}+,\hat{\mathbf{c}}-) + P_{\lambda,\hat{c},\hat{c}}(\hat{\mathbf{c}}-,\hat{\mathbf{c}}-) = 0. \tag{1.8.13}$$

Therefore, since we are assuming PI,

$$P_{\lambda,\hat{a},\hat{c}}(\hat{\mathbf{a}}+,\hat{\mathbf{c}}-) + P_{\lambda,\hat{a},\hat{c}}(\hat{\mathbf{a}}-,\hat{\mathbf{c}}-) = 0. \tag{1.8.14}$$

In particular,

$$P_{\lambda,\hat{a},\hat{c}}(\hat{\mathbf{a}}+,\hat{\mathbf{c}}-) = 0. \tag{1.8.15}$$

But by (1.8.8), we know that

$$P_{\lambda,\hat{a},\hat{b}}(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+) + P_{\lambda,\hat{a},\hat{b}}(\hat{\mathbf{a}}+,\hat{\mathbf{b}}-) = 1, \tag{1.8.16}$$

so using this together with PI, we must have

$$P_{\lambda,\hat{a},\hat{c}}(\hat{\mathbf{a}}+,\hat{\mathbf{c}}+) + P_{\lambda,\hat{a},\hat{c}}(\hat{\mathbf{a}}+,\hat{\mathbf{c}}-) = 1. \tag{1.8.17}$$

But by (1.8.9) and (1.8.15)

$$P_{\lambda,\hat{a},\hat{c}}(\hat{\mathbf{a}}+,\hat{\mathbf{c}}+) + P_{\lambda,\hat{a},\hat{c}}(\hat{\mathbf{a}}+,\hat{\mathbf{c}}-) = 0. \qquad (1.8.18)$$

Since (1.8.17) contradicts (1.8.18), the assumption (1.8.11) must be false if PI is to hold. Similarly, we will find that (1.8.12) must be false if PI is to hold. So we can only conclude that PI fails to hold in the Bohmian interpretation. But we can conclude even more than that: any deterministic hidden variable model that gives the same predictions as standard quantum theory when averaged over the hidden variables must violate PI.

1.9 Outcome Independence

Although a PI violation can account for the violation of Bell's Inequality, this is not the only possible culprit to consider. Another assumption of Bell's Inequality that might be violated is *Outcome Independence* (OI). Outcome independence is the assumption

$$P_{\lambda,x,y}(X,Y) = P_{A,\lambda,x,y}(X) \cdot P_{B,\lambda,x,y}(Y), \qquad (1.9.1)$$

where $P_{A,\lambda,x,y}(X)$ and $P_{B,\lambda,x,y}(Y)$ are defined in equations (1.6.3) and (1.6.4) respectively. So the difference between OI and PI is the following: with OI, given Alice and Bob's choice of measurements x and y, and the hidden variable λ, Alice and Bob's measurement outcomes will be statistically independent from one another, whereas

with PI, given Alice's choice of measurement x and the hidden variable λ, whatever measurement choice Bob makes, this will have absolutely no effect on the probabilities of Alice's measurement outcomes, and similarly, Alice's choice of measurement will have absolutely no effect on the probabilities of Bob's measurement outcomes.

Now we can see that if OI holds in any model which gives the same predictions as standard quantum theory when averaged over the hidden variables, then PI must be violated in such a model. For if both PI and OI hold, then for any measurement choices $\hat{\mathbf{a}}$, $\hat{\mathbf{b}}$, and $\hat{\mathbf{c}}$, and hidden variable λ, we have

$$
\begin{aligned}
P_{\lambda,\hat{\mathbf{a}},\hat{\mathbf{c}}}(\hat{\mathbf{a}}+,\hat{\mathbf{c}}+) &= P_{A,\lambda,\hat{\mathbf{a}},\hat{\mathbf{c}}}(\hat{\mathbf{a}}+) \cdot P_{B,\lambda,\hat{\mathbf{a}},\hat{\mathbf{c}}}(\hat{\mathbf{c}}+) \\
&= \Big(P_{\lambda,\hat{\mathbf{a}},\hat{\mathbf{c}}}(\hat{\mathbf{a}}+,\hat{\mathbf{c}}+) + P_{\lambda,\hat{\mathbf{a}},\hat{\mathbf{c}}}(\hat{\mathbf{a}}+,\hat{\mathbf{c}}-) \Big) \cdot \Big(P_{\lambda,\hat{\mathbf{a}},\hat{\mathbf{c}}}(\hat{\mathbf{a}}+,\hat{\mathbf{c}}+) + P_{\lambda,\hat{\mathbf{a}},\hat{\mathbf{c}}}(\hat{\mathbf{a}}-,\hat{\mathbf{c}}+) \Big) \\
&= \Big(P_{\lambda,\hat{\mathbf{a}},\hat{\mathbf{c}}}(\hat{\mathbf{a}}+,\hat{\mathbf{c}}+) + P_{\lambda,\hat{\mathbf{a}},\hat{\mathbf{c}}}(\hat{\mathbf{a}}+,\hat{\mathbf{c}}-) \Big) \cdot \Big(P_{\lambda,\hat{\mathbf{c}},\hat{\mathbf{c}}}(\hat{\mathbf{c}}+,\hat{\mathbf{c}}+)^{\nearrow 0} + P_{\lambda,\hat{\mathbf{c}},\hat{\mathbf{c}}}(\hat{\mathbf{c}}-,\hat{\mathbf{c}}+) \Big) \\
&= \Big(P_{\lambda,\hat{\mathbf{a}},\hat{\mathbf{b}}}(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+) + P_{\lambda,\hat{\mathbf{a}},\hat{\mathbf{b}}}(\hat{\mathbf{a}}+,\hat{\mathbf{b}}-) \Big) \cdot P_{\lambda,\hat{\mathbf{c}},\hat{\mathbf{c}}}(\hat{\mathbf{c}}-,\hat{\mathbf{c}}+) \\
&\geq P_{\lambda,\hat{\mathbf{a}},\hat{\mathbf{b}}}(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+) \cdot P_{\lambda,\hat{\mathbf{c}},\hat{\mathbf{c}}}(\hat{\mathbf{c}}-,\hat{\mathbf{c}}+)
\end{aligned}
$$

(1.9.2)

Similarly, we have

$$
\begin{aligned}
P_{\lambda,\hat{c},\hat{b}}(\hat{c}+,\hat{b}+) &= P_{A,\lambda,\hat{c},\hat{b}}(\hat{c}+) \cdot P_{B,\lambda,\hat{c},\hat{b}}(\hat{b}+) \\
&= \left(P_{\lambda,\hat{c},\hat{b}}(\hat{c}+,\hat{b}+) + P_{\lambda,\hat{c},\hat{b}}(\hat{c}+,\hat{b}-) \right) \cdot \left(P_{\lambda,\hat{c},\hat{b}}(\hat{c}+,\hat{b}+) + P_{\lambda,\hat{c},\hat{b}}(\hat{c}-,\hat{b}+) \right) \\
&= \left(\overset{0}{\cancel{P_{\lambda,\hat{c},\hat{c}}(\hat{c}+,\hat{c}+)}} + P_{\lambda,\hat{c},\hat{c}}(\hat{c}+,\hat{c}-) \right) \cdot \left(P_{\lambda,\hat{c},\hat{b}}(\hat{c}+,\hat{b}+) + P_{\lambda,\hat{c},\hat{b}}(\hat{c}-,\hat{b}+) \right) \\
&= P_{\lambda,\hat{c},\hat{c}}(\hat{c}+,\hat{c}-) \cdot \left(P_{\lambda,\hat{a},\hat{b}}(\hat{a}+,\hat{b}+) + P_{\lambda,\hat{a},\hat{b}}(\hat{a}-,\hat{b}+) \right) \\
&\geq P_{\lambda,\hat{c},\hat{c}}(\hat{c}+,\hat{c}-) \cdot P_{\lambda,\hat{a},\hat{b}}(\hat{a}+,\hat{b}+).
\end{aligned}
$$

$$(1.9.3)$$

But since the hidden variable λ is assumed to be independent of Alice and Bob's measurement, and since Alice and Bob will always get opposite results when they make the same choice of measurement, it follows that

$$
P_{\lambda,\hat{c},\hat{c}}(\hat{c}+,\hat{c}-) + P_{\lambda,\hat{c},\hat{c}}(\hat{c}-,\hat{c}+) = 1. \tag{1.9.4}
$$

Therefore, putting (1.9.2), (1.9.3), and (1.9.4) together, we have

$$
P_{\lambda,\hat{a},\hat{c}}(\hat{a}+,\hat{c}+) + P_{\lambda,\hat{c},\hat{b}}(\hat{c}+,\hat{b}+) \geq P_{\lambda,\hat{a},\hat{b}}(\hat{a}+,\hat{b}+). \tag{1.9.5}
$$

We have thus proved that OI and PI together imply Bell's Inequality (1.8.4). But since Bell's Inequality does not hold in reality, it follows that if OI is always true, then PI must be violated.

49

In the case of deterministic models in which there are outcomes,[54] OI necessarily holds. To see why, we note that since the model is deterministic, either (1) $P_{A,\lambda,x,y}(X) = P_{B,\lambda,x,y}(Y) = 1$, or (2) $P_{A,\lambda,x,y}(X) = 0$ or $P_{B,\lambda,x,y}(Y) = 0$. In case (1), then by definition of the probability $P_{\lambda,x,y}(X,Y)$, we clearly have $P_{\lambda,x,y}(X,Y) = 1$, and so

$$P_{\lambda,x,y}(X,Y) = 1 = P_{A,\lambda,x,y}(X)P_{B,\lambda,x,y}(Y).$$

As for case (2), it clearly follows from the definition of the probability $P_{\lambda,x,y}(X,Y)$ that $P_{\lambda,x,y}(X,Y) = 0$, and so

$$P_{\lambda,x,y}(X,Y) = 0 = P_{A,\lambda,x,y}(X)P_{B,\lambda,x,y}(Y).$$

Therefore, (1.9.1) holds in either case. It therefore follows that OI holds in any deterministic model. In particular, OI holds under the Bohmian interpretation.

When it comes to the Copenhagen interpretation of quantum physics, however, OI fails to hold. For instance, if $x = y = \hat{\mathbf{a}}$, then $P_{\lambda,\hat{\mathbf{a}},\hat{\mathbf{a}}}(\hat{\mathbf{a}}+, \hat{\mathbf{a}}+) = 0$, but $P_{A,\lambda,\hat{\mathbf{a}},\hat{\mathbf{a}}}(\hat{\mathbf{a}}+) = P_{B,\lambda,\hat{\mathbf{a}},\hat{\mathbf{a}}}(\hat{\mathbf{a}}+) = 1/2$,[55] and so in this case, (1.9.1) is false.

---

54. For deterministic models in which there are outcomes, the probability of a particular outcome given a complete description of a system will be either 0 or 1. For models in which there are no outcomes, such as the many-worlds interpretation which will be discussed in the next chapter, we can't speak of outcome independence.

55. See footnote 43 on page 36.

1.10 Peaceful Coexistence of Special Relativity and Quantum Physics

So far we've seen that PI holds in the Copenhagen interpretation (section 1.6), that OI holds in the Bohmian interpretation (section 1.9), and that due to the violation of Bell's inequality, it is not possible for both OI and PI to hold in a model that is consistent with physical reality (section 1.9). Nevertheless, as long as PI holds, the failure of OI does not enable Bob to send messages to Alice faster than light because Bob only has control over the measurement choice he makes rather than the outcome he observes. Assuming Bob's mental states have no effect on the measurement outcome, there is nothing he can do to influence his outcome, so although Alice will be able to work out or at least get some information about Bob's measurement outcome if she already happens to know which choice of measurement he has made, she will not be able to work out which measurement Bob makes (or even whether Bob has made a measurement at all) by measuring the outcome of her particle. For Shimony,[56] this inability to send superluminal messages between Alice and Bob when PI holds and OI is violated was deemed sufficient for standard quantum theory and special relativity to peacefully coexist.

---

56. See Shimony, "Events and processes in the quantum world," 146–147.

However, Butterfield is not satisfied with Shimony's solution to peaceful coexistence.[57] Firstly, he notes that proofs of non-superluminal signaling[58] make no assumptions about spacetime locations.[59] One would have thought that any proof that superluminal signalling between two points is impossible would have to show that a signal cannot be transmitted from one point to the other in less time than the time it takes light to travel between the two points. But if nothing is said about the location of these two points or what is so special about the speed of light compared to the speed of any other particle, then there does not seem to be enough information in the premises to draw the desired conclusion that superluminal signaling is impossible in quantum physics.

Secondly, Butterfield notes that Shimony thinks peaceful coexistence of quantum physics and special relativity is guaranteed by the denial of OI and the acceptance of PI. However, OI itself depends on the (often) rather vague notion of what an outcome really is. For instance, in the many-worlds interpretation, it is not clear that there are any outcomes at all. Rather, there is just a universal quantum state that tells us the probability of certain outcomes, if there were such things as outcomes – it doesn't tell us that there really are any outcomes. In the next chapter, we will discuss

57. See Butterfield, "Peaceful Coexistence: Examining Kent's Relativistic Solution to the Quantum Measurement Problem," p. 12.

58. e.g. see Michael Redhead, *Incompleteness, nonlocality, and realism : a prolegomenon to the philosophy of quantum mechanics* (Oxford : New York: Clarendon Press ; Oxford University Press, 1987), p. 113–116; David Bohm and B. J Hiley, *The undivided universe : an ontological interpretation of quantum theory* (London: Routledge, 1993), p. 139–140

59. For a definition of a spacetime location, see page 123.

the many-worlds interpretation and why denying the reality of outcomes is such an unsatisfactory way of guarantying peaceful coexistence of quantum physics and special relativity. But as is clear from other interpretations of quantum physics, the notion of what an outcome is doesn't have to be vague. In the Bohmian interpretation of quantum physics, it is very clear what an outcome of an experiment is since all the particles have definite positions and momenta. Because of this, the pointers and displays of measuring devices which are made up of particles will have definite readouts which will correspond to the definite positions of particles being measured (assuming the measurement device is working properly). So unlike the many-worlds interpretation, measurements in the Bohmian interpretation have definite outcomes, and hence there is only a single world in the Bohmian interpretation of quantum physics. But as we've just seen, the problem with the Bohmian interpretation is the violation of PI.

Thus, a satisfactory account of the peaceful coexistence of quantum physics and special relativity requires an interpretation of quantum physics in which not only PI holds, but also an interpretation of quantum physics that has special relativity built into it (thus satisfying Butterfield's first objection), and in which we can make sense of what it means to be an outcome (thus satisfying Butterfield's second objection).[60]

---

60. If PI holds, then OI can't hold in a theory that makes the same predictions as standard quantum theory. But one might still worry that a violation of OI will require some sort of superluminal causation. After all, if there is a correlation between results, there needs to be some sort of explanation of the correlation. Putting it anthropomorphically, how is Bob's system to know that Alice's system collapsed, or vice versa, if there is no superluminal causality? One response to this worry is to suggest that maybe there is no collapse in the correct theory. There can still be outcomes in theories that don't have collapses, just as in the Bohmian interpretation. Furthermore, one might suppose that there could be some cause which atemporally determines Alice and Bob's measurements to have different outcomes. For example, one could imagine a supernatural being surveying a four dimensional

To fully address Butterfield's first objection would require quantum field theory, and this would be beyond the scope of this dissertation. But a more modest aspiration that would go some way to address Butterfield's first objection would be to insist on an interpretation of quantum physics which has a clear notion of outcome and which also has a property known as Lorentz invariance. This provides a motivation for the consideration of Kent's theory of quantum physics that has this property of Lorentz invariance. But before we consider Kent's theory, in the next chapter we will examine the many-worlds interpretation of quantum physics discussing both its appeal and its drawbacks.

1.11 Summary

In this chapter, we have considered the EPR-Bohm paradox and the problem it raises of how to account for the mysterious correlations between the measurement outcomes of two observers measuring the spin properties of spin singlets. The question the EPR-Bohm paradox raises is how one can account for these correlations in a way that is consistent with special relativity and the predictions of standard quantum theory.

---

spacetime landscape, and that this being had the power to make different events correlate with each other. So for instance, if the supernatural being saw that Alice and Bob were going to measure the spin of their respective particles along the same axis, the supernatural being could make it so that Alice got spin up and Bob got spin down with respect to the axis of measurement. Although physicists may find this scenario rather far-fetched, it doesn't entail that there is any superluminal causation by which Alice's particle affects Bob's particle or vice versa. Theologians might also be troubled by this scenario if it meant that the supernatural being could be causally affected by the measurement choices that Alice and Bob make. There are, however, ways of overcoming this theological objection. e.g. if God is the supernatural being in question, then some theologians would be happy with the idea of God both causing Alice and Bob to will the measurements settings they use, and causing them to get opposite results (possibly via some secondary cause) should God cause them to make their measurements along the same axis.

The Copenhagen interpretation in which the act of observation causes the quantum state to collapse does not seem to be consistent with special relativity. Theories such as the GRW interpretation which posit the spontaneous collapse of quantum states make predictions that violate the predictions of standard quantum theory, and to date, there is no experimental evidence for such violations.[61]

If the proposed theory posits the existence of hidden variables in addition to the traditional quantum state of standard quantum theory, then it is not possible for the theory to satisfy both Parameter Independence (PI) and Outcome Independence (OI) because otherwise, Bell's inequality would hold in this theory, and this is not consistent with the violation of Bell's inequality in physical reality. A denial of PI would allow for the superluminal propagation of signals if one knew the hidden variables, and so without a compelling reason for why these hidden variables must be unknown,

---

61. To see why the GRW interpretation implies a violation of standard quantum theory, we recall that the GRW interpretation supposes that there is a random spontaneous collapse of entangled states of particles which is frequent enough so that we don't have superpositions of macroscopic objects such as live and dead cats, but such that this collapse is infrequent enough to account for the correlations of that have been experimentally observed in EPR-type experiments. Nevertheless, GRW still violates the predictions of standard quantum theory, for according to standard quantum theory, if one particle of a spin singlet is sent to Alice and the other particle of a spin singlet is sent to Bob, then so long as there are no other particle interactions with the two particles of the spin singlet prior to Alice and Bob's measurement, then they will always obtain opposite results. This follows from how the state evolves according to the Schrödinger equation. But now suppose GRW is true so that after a very long time, the spin singlet collapses to a state in which Alice's particle points up in the spin z-direction and Bob's particle points down in the spin z-direction, and furthermore suppose Alice and Bob perform their measurements along the x-axis. Then since the two particles are no longer entangled, there will be a non-zero probability that Alice and Bob will measure their particles to have the same spin along the x-axis. This is not a possibility that standard quantum theory predicts. Granted, the probability of a collapse for any given entangled state is going to be extremely small, so it is very unlikely that Alice and Bob will obtain the same spin measurements, but if such collapses do occur as GRW predicts, then Alice and Bob can get the same spin measurements, and this will constitute a violation of the predictions of standard quantum theory which predicts Alice and Bob should always obtain different spin measurements. Perhaps GRW is right which would entail that standard quantum theory is slightly wrong, but in this dissertation, I will not be considering this possibility, and will instead assume that the predictions that standard quantum theory makes are always right.

denying PI does not sit easily with special relativity in which superluminal signalling is impossible. The denial of OI seems more promising, but this approach by itself is not sufficient to provide an adequate account of the mysterious correlations of the EPR-Bohm paradox since it does not address the thorny issue of what we mean by outcome or whether there is in fact any physical reality to outcomes at all.

One theory that denies the reality of experimental outcomes is the many-worlds interpretation. In the next chapter, we will discuss the many-worlds interpretation in some detail, and we will see why it does not provide a satisfactory account of the correlations of the EPR-Bohm paradox.

CHAPTER TWO

The Measurement Problem and the Many-Worlds Interpretation

In the previous chapter we discussed the EPR-Bohm paradox which exhibits the mysterious correlations of measurement outcomes for measurements made on spin singlets by two different observers. We saw that the standard Copenhagen interpretation of quantum physics (which posits that upon measurement an instantaneous state collapse to a definite measurement outcome occurs) must be rejected if one is to accept Einstein's theory of special relativity. Furthermore, we saw that the experimental violation of Bell's inequality implies that a local hidden variables theory must also be rejected. We also discussed Shimony's suggestion that it would be acceptable to have non-local hidden variables so long as they satisfied Parameter Independence (PI) but failed to satisfy Outcome Independence (OI), since Bell's inequality wouldn't then follow from these two assumptions, and the denial of OI is a sufficiently weak form of non-locality that it doesn't imply superluminal signaling. Thus, Shimony argued that accepting PI and denying OI was sufficient for the peaceful coexistence of quantum physics and special relativity. However, Shimony's solution rests on there being a clear notion of what an outcome is. But the question of what an outcome is and even whether there are such things as measurement outcomes is very controversial and forms an important part of what is known as the measurement problem.

The measurement problem actually consists of three parts: (1) the preferred basis problem, (2) the nonobservability of interference at the macroscopic level, and (3) the problem of outcomes.[1] These three parts of the measurement problem will be discussed in some detail later on in this chapter. Decoherence theory is able to resolve parts (1) and (2) of the measurement problem, and following Schlosshauer, we will explain how this is done, but as we will see (also following Schlosshauer), decoherence theory is not able to resolve part (3) of the measurement problem, the problem of outcomes.

The problem of outcomes arises when we suppose that the quantum state gives a complete description of the system it is describing.[2] Throughout this chapter we will make this assumption, that is, we will assume there are no hidden variables, whether local or otherwise; rather, a quantum state whose evolution is described by the Schrödinger equation provides the most complete description possible of a physical system.

In this chapter, we will be discussing the measurement problem in some detail, and in particular, we will be examining the many-worlds interpretation of quantum physics which attempts to sidestep the problem of outcomes by refusing to acknowledge

---

1. See Maximilian Schlosshauer, *Decoherence and the Quantum-to-Classical Transition* (Berlin: Springer-Verlag, 2007), 50

2. We need to make a small caveat here – the example of the quantum states we have described so far such as $|\hat{\mathbf{a}}\rangle$ and $|\Psi_{\text{Bell}}\rangle$ only give a complete description of the spin state of a system. These states don't say anything about the position of the system's particles. Nevertheless, we could in principle write down an expression for the state that did include this positional information. We will therefore think of the expressions such as $|\hat{\mathbf{a}}\rangle$ and $|\Psi_{\text{Bell}}\rangle$ as abbreviations for quantum states that also include a complete specification of the particle positions which in the context of quantum physics will be a wave function that determines the probability the particles can be detected in a particular region.

the reality of outcomes. Although the many-worlds interpretation is mathematically appealing, we'll see that from a philosophical point of view, it is woefully inadequate. Hence, any account of the mysterious correlations of the EPR-Bohm paradox (as described in chapter 1) that depended on the many-worlds interpretation would be unsatisfactory. Nevertheless, an understanding of the many-worlds interpretation and how it relates to the problem of outcomes will prove helpful when we come to evaluate Kent's theory of quantum physics.

2.1 A Preliminary Consideration of the Many-Worlds Interpretation

In the previous chapter, we considered experiments for which there are only two possible measurement outcomes (given the choice of measurement made), and we saw that in order to calculate the probabilities of these two outcomes, we can do this by positing that the state describing the item being measured is in a superposition of two states each of which corresponds to a definite measurement outcome. For example, in the case of a spin singlet, the quantum state of the two particles can be expressed as the Bell state

$$|\Psi_{\text{Bell}}\rangle = \frac{1}{\sqrt{2}}(\,|\hat{\mathbf{a}}+\rangle_A\,|\hat{\mathbf{a}}-\rangle_B - \,|\hat{\mathbf{a}}-\rangle_A\,|\hat{\mathbf{a}}+\rangle_B) \qquad \text{(1.2.6 revisited)}$$

which is a superposition of the two definite measurement outcome states $|\hat{\mathbf{a}}+\rangle_A\,|\hat{\mathbf{a}}-\rangle_B$ and $|\hat{\mathbf{a}}-\rangle_A\,|\hat{\mathbf{a}}+\rangle_B$. But at this stage in our line of reasoning, it is too early to resort to a many-worlds interpretation of the Bell state where the first component corresponds to a world in which Alice detects her particle at location $\hat{\mathbf{a}}+$ on her detection screen and

Bob detects his particle at location $\hat{\mathbf{a}}-$ on his detection screen, and where the second component corresponds to a world in which Alice detects her particle at location $\hat{\mathbf{a}}-$ on her detection screen and Bob detects his particle at location $\hat{\mathbf{a}}+$ on his detection screen. Such an interpretation would be premature because as mentioned on page 20, for any other axis $\hat{\mathbf{b}}$, the transformation rules in equation (1.1.2) imply that

$$
\begin{aligned}
\frac{1}{\sqrt{2}}(\,|\hat{\mathbf{a}}+\rangle_A\,|\hat{\mathbf{a}}-\rangle_B &- |\hat{\mathbf{a}}-\rangle_A\,|\hat{\mathbf{a}}+\rangle_B) \\
&= \frac{1}{\sqrt{2}}(\,|\hat{\mathbf{b}}+\rangle_A\,|\hat{\mathbf{b}}-\rangle_B - |\hat{\mathbf{b}}-\rangle_A\,|\hat{\mathbf{b}}+\rangle_B).
\end{aligned}
\tag{1.2.9 revisited}
$$

Similarly, given the transformation rules in equation (1.1.2), we should resist the temptation to interpret a state of the form $\frac{1}{\sqrt{2}}(\,|\hat{\mathbf{a}}+\rangle + |\hat{\mathbf{a}}-\rangle)$ as representing two worlds, one in which the particle is in the state $|\hat{\mathbf{a}}+\rangle$, and another in which the particle is in the state $|\hat{\mathbf{a}}-\rangle$. For according to equation (1.1.2a), the much more obvious interpretation is that this state just describes one world in which the particle is in the state $|\hat{\mathbf{b}}+\rangle$ where the angle between the $\hat{\mathbf{a}}$ and the $\hat{\mathbf{b}}$ axis is 90°.[3]

In order to make a case for a many-worlds interpretation, we need to discuss decoherence theory. Decoherence theory considers how a system interacts with its environment, and it allows us to understand what kinds of measurements can be made on the system. In order to discuss decoherence theory and its relevance to the many-worlds interpretation, we first need to introduce the mathematical formalism of standard quantum theory.

----

3. This is because when $\theta = 90°$, $\sin(\theta/2) = \cos(\theta/2) = \frac{1}{\sqrt{2}}$, so $|\hat{\mathbf{b}}+\rangle = \frac{1}{\sqrt{2}}(\,|\hat{\mathbf{a}}+\rangle + |\hat{\mathbf{a}}-\rangle)$ in equation (1.1.2a) with $\theta = 90°$.

Given a possible kind of measurement (e.g. measuring the spin of a particle along a particular axis), there will be a mathematical object called an *observable* which encodes all the possible measurement outcomes for this particular kind of measurement. The precise mathematical definition of an observable is as follows: an observable of a physical system is a Hermitian operator that acts on the Hilbert space of states describing the physical system. In order to understand what this definition means, there are a number of things we need to explain: what a Hilbert space is, what a Hermitian operator is, and how a Hermitian operator relates to a particular kind of measurement. In order to explain all this, it will be helpful to keep in mind the simple example of an experimenter Alice making a spin measurement on a particle along an axis $\hat{\mathbf{a}}$. In performing this measurement, we suppose she has a physical measurement device which we denote as $O_{\hat{\mathbf{a}}+}$ and which outputs 1 if the particle is in the spin $|\hat{\mathbf{a}}+\rangle$-state and 0 if the particle is in the spin $|\hat{\mathbf{a}}-\rangle$-state. This measurement device will have a corresponding observable which we will denote by $\hat{O}_{\hat{\mathbf{a}}+}$ and which we will describe shortly once we have defined what a Hilbert space is.[4]

To motivate the definition of a Hilbert space, recall that the states $|\hat{\mathbf{a}}+\rangle$ and $|\hat{\mathbf{a}}-\rangle$ representing the spin of a particle can be added to give the spin state $|\hat{\mathbf{b}}+\rangle$ and $|\hat{\mathbf{b}}-\rangle$ as seen in equation (1.1.2). Also, recall that if we have two states $|\psi\rangle$ and $|\chi\rangle$, we can define their bra-ket $\langle\chi|\psi\rangle$ to be a complex number satisfying the Born Rule. In

---

4. Note that in contrast to the physical object $O_{\hat{\mathbf{a}}+}$, we're taking $\hat{O}_{\hat{\mathbf{a}}+}$ to be a mathematical object, a Hermitian operator on a Hilbert space.

other words, $|\langle \chi | \psi \rangle|^2$ is the probability $P(\chi | \psi)$ that the particle will be found to be in state $|\chi\rangle$ given that we know that the particle is in state $|\psi\rangle$. We thus imposed the assumption that $\langle \psi | \psi \rangle = 1$ for any state $|\psi\rangle$.

Now in order to arrive at a definition of a Hilbert space, we first need to relax this normalization condition $\langle \psi | \psi \rangle = 1$. Thus, if $|\psi\rangle$ is a state and $\lambda \in \mathbb{C}$ is any non-zero complex number, then we allow $|\psi'\rangle = \lambda |\psi\rangle$ also to be a state with the caveat that $|\psi'\rangle$ represents exactly the same physical state as $|\psi\rangle$, and that $\langle \psi' | \psi' \rangle = |\lambda|^2 \langle \psi | \psi \rangle$. We define $\|\psi\| = \sqrt{\langle \psi | \psi \rangle}$ and we say that $|\psi\rangle$ has been *normalized* when $\|\psi\| = 1$. Now when calculating probabilities, we need to remember to include a normalization factor. Thus, the probability that the particle will be found to be in state $|\chi\rangle$ given that we know that the particle is in state $|\psi\rangle$ will now be $P(\chi | \psi) = \frac{|\langle \chi | \psi \rangle|^2}{\|\psi\| \|\chi\|}$.[5] It is dropping the assumption $\|\psi\| = 1$ on the states of a physical system that gives rise to the mathematical structure known as a Hilbert Space.

A *Hilbert space* is a set H in which

1. any two members of H can be added to obtain another member of H,

2. any member of H can be multiplied by any complex number to obtain another member of H,

3. one can take the bra-ket of any two members of H to obtain a complex number

subject to some natural axioms.[6]

---

5. If there is no such normalization factor because $\|\psi\| = 0$, then $|\psi\rangle$ does not represent a physical state, so the probability the system is ever in this state will be zero, and so in this case we will set $P(\chi | \psi) = P(\psi | \chi) = 0$.

A very simple example of a Hilbert space would be the set of states

$$\{\alpha \,|\hat{\mathbf{a}}+\rangle + \beta \,|\hat{\mathbf{a}}-\rangle : \alpha, \beta \in \mathbb{C}\}.$$

As we will soon see, the observable corresponding to the measurement device $O_{\hat{\mathbf{a}}+}$ will be the operator $\hat{O}_{\hat{\mathbf{a}}+}$ that sends the state $\alpha \,|\hat{\mathbf{a}}+\rangle + \beta \,|\hat{\mathbf{a}}-\rangle$ to the state $\alpha \,|\hat{\mathbf{a}}+\rangle$.

More generally, suppose we have an experimental setup (for example the Stern-Gerlach experiment) where a physical system can be in one of several measurable states $|\psi_1\rangle, \ldots, |\psi_N\rangle \in H$. The physical system could also be in a state described by a sum of some of the $|\psi_1\rangle, \ldots, |\psi_N\rangle$, but by saying the system is in one of these $|\psi_1\rangle, \ldots, |\psi_N\rangle$ measurable states, we mean that there is a measuring device that

6. More formally, a complex Hilbert space H is a complex vector space possessing a bra-ket. By a *complex vector space*, we mean a set V such that the following axioms are satisfied
  - $\psi + (\chi + \zeta) = (\psi + \chi) + \zeta, \,\forall \lambda, \chi, \zeta \in V$
  - $\psi + \chi = \chi + \psi, \,\forall \psi, \chi \in V$
  - there exists an element $\mathbf{0} \in V$ such that $\psi + \mathbf{0} = \psi, \,\forall \psi \in V$.
  - $\forall \psi \in V$ there exists an element $-\psi \in V$ such that $\psi + (-\psi) = \mathbf{0}$.
  - $\forall \lambda, \mu \in \mathbb{C}$ (i.e. in the set of complex numbers – this is why it is called a *complex* vector space), and $\psi \in V$, $\lambda(\mu\psi) = (\lambda\mu)\psi$,
  - for the scalar $1 \in \mathbb{C}$, $1\psi = \psi, \,\forall \psi \in V$,
  - $\lambda(\psi + \chi) = \lambda\psi + \lambda\chi, \forall \psi, \chi \in V$ and $\lambda \in \mathbb{C}$
  - $(\lambda + \mu)\psi = \lambda\psi + \mu\psi, \,\forall \lambda, \mu \in \mathbb{C}$ and $\psi \in V$.

A *Hilbert space H* is a complex vector space possessing a bra-ket. Strictly speaking, a Hilbert space also has a property called completeness, but this property need not concern us here. In quantum theory, elements of H are expressed in terms of kets, $|\cdot\rangle$. Kets behave like vectors, so for $|\psi\rangle, |\chi\rangle \in H$ and $\lambda, \mu \in \mathbb{C}$, we have $\lambda |\psi\rangle + \mu |\chi\rangle = |\lambda\psi + \mu\chi\rangle$. The bra-ket of $|\psi\rangle$ and $|\chi\rangle$ is then written as $\langle\psi|\chi\rangle$, and it satisfies the following axioms:
  - $\langle\psi|\chi\rangle \in \mathbb{C}, \,\forall \psi, \chi \in H$,
  - $\langle\psi|\chi\rangle = \overline{\langle\chi|\psi\rangle}, \,\forall \psi, \chi \in H$,
  - $\langle\psi|\psi\rangle \geq 0, \,\forall\, |\psi\rangle \in H$ and $\langle\psi|\psi\rangle = 0$ if and only if $|\psi\rangle = \mathbf{0}$,
  - $\langle\zeta|\lambda\psi + \mu\chi\rangle = \lambda \langle\zeta|\psi\rangle + \mu \langle\zeta|\chi\rangle, \,\forall\, |\psi\rangle, |\chi\rangle, |\zeta\rangle \in H$ and $\lambda, \mu \in \mathbb{C}$.

will always give the same measurement outcome whenever the system is in the same measurable state.[7] We also assume *orthonormality*, that is we assume $\langle \psi_i | \psi_i \rangle = 1$ and $\langle \psi_i | \psi_j \rangle = 0$ for $i \neq j$ so that if the system is measured to be in the $|\psi_j\rangle$-state, then there would be zero probability that it could then be measured to be in the $|\psi_i\rangle$-state for $i \neq j$.

Now suppose that for each measurable state $|\psi_i\rangle$, we assign a real number o_i. There might be a very natural way of doing this, such as assigning o_i to be the angle by which a pointer of a measurement device is deflected when the system is in the state $|\psi_i\rangle$, but the assignment could be as ad hoc as we wished – we can just think of it as the measurement value an experimenter records when he or she observes a particular measurement outcome. Given such an assignment of measurement values, the corresponding observable $\hat{O}$ would be a mapping of states to states satisfying the following rules:

1. $\hat{O} |\psi_i\rangle = o_i |\psi_i\rangle$

2. $\hat{O}(\lambda |\psi\rangle + \mu |\chi\rangle) = \lambda \hat{O} |\psi\rangle + \mu \hat{O} |\chi\rangle$ for all states $|\psi\rangle, |\chi\rangle \in H$ and complex numbers $\lambda, \mu \in \mathbb{C}$.

When a mapping $\hat{O}$ satisfies rule 2., we refer to $\hat{O}$ as an *operator* on H. Since the measurement device $O_{\hat{a}+}$ outputs a value of 1 when the particle is in the $|\hat{a}+\rangle$-state and 0 when the particle is in the $|\hat{a}-\rangle$-state, it is now clear from rule 1 and 2 why the

---

7. When the state is described as a non-trivial sum of the measurable states, we no longer have such certainty, and instead we can only speak of probabilities based on the coefficients on the measurable states.

64

corresponding observable $\hat{O}_{\hat{\mathbf{a}}+}$ will be the operator that sends the state $\alpha \, |\hat{\mathbf{a}}+\rangle + \beta \, |\hat{\mathbf{a}}-\rangle$ to $\alpha \, |\hat{\mathbf{a}}+\rangle$.

For a given physical state $|\psi\rangle \in H$, we define the expected measurement value (usually referred to as the *expectation value*) of $\hat{O}$ to be

$$\langle \hat{O} \rangle_\psi \stackrel{\text{def}}{=} \sum_{i=1}^{N} p_i o_i, \tag{2.2.1}$$

where p_i for $i = 1, \ldots, N$ is the probability the particle will be found to be in the state $|\psi_i\rangle$ given that it is in the state $|\psi\rangle$ so that $\sum_{i=1}^{N} p_i = 1$. If we were to perform the same measurement corresponding to the observable $\hat{O}$ on many systems that were in the state $|\psi\rangle$ and calculated the average measurement outcome from all these measurements, then this average should be approximately equal to the value given in (2.2.1) if standard quantum theory is correct. Because this expectation value $\langle \hat{O} \rangle_\psi$ will depend on the state $|\psi\rangle$ the system is in, we include the subscript ψ. For instance, if the system is in the $|\psi_i\rangle$-state, then the expectation value of the measurement will be o_i since the probability that the state is in the state $|\psi_i\rangle$ will be 1, so that $p_i = 1$ and $p_j = 0$ for all $j \neq i$. But more generally, if the system was in an arbitrary state $|\psi\rangle = \sum_{i=1}^{N} \alpha_i \, |\psi_i\rangle$ with $\|\psi\| = 1$, then it turns out that

$$\langle \hat{O} \rangle_\psi = \langle \psi | \hat{O} | \psi \rangle .^8 \tag{2.2.2}$$

We can see that this formula is correct in the simple example when the system is in

the $|\psi_i\rangle$-state, for in this case, the expectation value should be o_i, and we clearly have

$\langle \psi_i | \hat{O} | \psi_i \rangle = o_i$ since $\hat{O} |\psi_i\rangle = o_i |\psi_i\rangle$ and $\langle \psi_i | \psi_i \rangle = 1$. Hence, $\langle \hat{O} \rangle_{\psi_i} = \langle \psi_i | \hat{O} | \psi_i \rangle$ as

expected.

To say that $\hat{O}$ is *Hermitian* is to say that $\langle \psi | \hat{O} | \psi \rangle$ is a real number for any arbitrary

state $|\psi\rangle$. Thus, the observable $\hat{O}$ defined by the two criteria above is a Hermitian

operator acting on the Hilbert space of states H since we are assuming the o_i are all

real numbers. Roughly speaking, we can assume[9] that given a Hermitian operator

$\hat{O}$ on a Hilbert space of states H, any state $|\psi\rangle \in H$ can be expressed as a (possibly

infinite) sum

$$|\psi\rangle = \sum_{i=1}^{N} \alpha_i |\psi_i\rangle \tag{2.2.3}$$

where $\hat{O} |\psi_i\rangle = o_i |\psi_i\rangle$ for some set of states $|\psi_i\rangle$ referred to as *eigenstates* of $\hat{O}$, and

real numbers o_i referred to as *eigenvalues* of $\hat{O}$. We will typically assume that the

$|\psi_i\rangle$ are orthonormal. Orthonormality of the $|\psi_i\rangle$ will entail that the coefficients α_i

---

8. To see this, note that $\alpha_i = \langle \psi_i | \psi \rangle$ from which it follows that if we define the mapping $I = \sum_{i=1}^{N} |\psi_i\rangle\langle\psi_i|$ then $I |\psi\rangle = \sum_{i=1}^{N} |\psi_i\rangle \langle \psi_i | \psi \rangle = \sum_{i=1}^{N} \langle \psi_i | \psi \rangle |\psi_i\rangle = |\psi\rangle$. Therefore,

$$\langle \psi | \hat{O} | \psi \rangle = \langle \psi | \hat{O} I | \psi \rangle = \sum_{i=1}^{N} \langle \psi | \hat{O} | \psi_i \rangle \langle \psi_i | \psi \rangle = \sum_{i=1}^{N} o_i \langle \psi | \psi_i \rangle \langle \psi_i | \psi \rangle = \sum_{i=1}^{N} o_i |\langle \psi_i | \psi \rangle|^2$$

$$= \sum_{i=1}^{N} o_i p_i = \langle \hat{O} \rangle_{\psi}.$$

9. Strictly speaking, we require a Hermitian operator to have a property known as compactness for this assumption to hold.

will be uniquely determined by the formula $\alpha_i = \langle \psi_i | \psi \rangle$. Thus, this set of eigenstates $\{ |\psi_1\rangle , \ldots |\psi_N\rangle \}$ satisfies the criterion for being a *basis* of the Hilbert space of states H, namely, every state $|\psi\rangle \in H$ can be uniquely expressed by a summation of the form given in equation (2.2.3).[10] We refer to an expression of the form (2.2.3) as a *linear combination* of the basis $\{ |\psi_1\rangle , \ldots |\psi_N\rangle \}$.

2.3 The Preferred Basis Problem[11]

Now just because we can have an observable $\hat{O}$, there is no guarantee that there is a measuring device that could determine whether the system was in one of the eigenstates of $\hat{O}$. For instance, if $|\text{Cat Alive}\rangle$ is the physical state in which a cat is alive, and $|\text{Cat Dead}\rangle$ is the physical state in which the same cat is dead, then although there are measuring devices that can distinguish between the $|\text{Cat Alive}\rangle$-state and the $|\text{Cat Dead}\rangle$-state,[12] there are no known measuring devices that can distinguish between the $\frac{1}{\sqrt{2}}(|\text{Cat Alive}\rangle + |\text{Cat Dead}\rangle)$-state and the $\frac{1}{\sqrt{2}}(|\text{Cat Alive}\rangle - |\text{Cat Dead}\rangle)$-state. On the other hand, there are measuring devices that can distinguish between the $\frac{1}{\sqrt{2}}(|\hat{\mathbf{a}}+\rangle + |\hat{\mathbf{a}}-\rangle)$-state and the $\frac{1}{\sqrt{2}}(|\hat{\mathbf{a}}+\rangle - |\hat{\mathbf{a}}-\rangle)$-state in a Stern-Gerlach experiment. Why the difference?

---

10. Note that although for a given basis, equation (2.2.3) will be unique, there will be many different bases, and the α_i coefficients will depend on which basis is chosen.

11. See Schlosshauer, *Decoherence and the Quantum-to-Classical Transition*, 53–55.

12. For example, we assume that human beings can be thought of as such measuring devices.

This question is at the heart of the *preferred basis problem*. As mentioned already, a basis is just a set of states via which all other states of the system can be uniquely expressed. For instance, we can express the state $\frac{1}{\sqrt{2}}(\,|\text{Cat Alive}\rangle + |\text{Cat Dead}\rangle)$ uniquely as a sum of elements from the basis $\{\,|\text{Cat Alive}\rangle\,, |\text{Cat Dead}\rangle\}$, and thus we think of $\frac{1}{\sqrt{2}}(\,|\text{Cat Alive}\rangle + |\text{Cat Dead}\rangle)$ as being a superposition of the $|\text{Cat Alive}\rangle$ and $|\text{Cat Dead}\rangle$ basis states. However, we can also uniquely express $|\text{Cat Alive}\rangle$ in terms of the basis $\{\frac{1}{\sqrt{2}}(\,|\text{Cat Alive}\rangle + |\text{Cat Dead}\rangle), \frac{1}{\sqrt{2}}(\,|\text{Cat Alive}\rangle - |\text{Cat Dead}\rangle)\}$.[13] Nevertheless, we would not tend to think of $|\text{Cat Alive}\rangle$ as being in a superposition of the $\frac{1}{\sqrt{2}}(\,|\text{Cat Alive}\rangle + |\text{Cat Dead}\rangle)$ and $\frac{1}{\sqrt{2}}(\,|\text{Cat Alive}\rangle - |\text{Cat Dead}\rangle)$ basis states. That is, we have a preference for the basis $\{\,|\text{Cat Alive}\rangle\,, |\text{Cat Dead}\rangle\}$ over the basis $\{\frac{1}{\sqrt{2}}(\,|\text{Cat Alive}\rangle + |\text{Cat Dead}\rangle), \frac{1}{\sqrt{2}}(\,|\text{Cat Alive}\rangle - |\text{Cat Dead}\rangle)\}$. We refer to a basis whose basis states our measuring device can distinguish between as a *preferred basis*. As will be shown in section 2.5, decoherence theory offers a very elegant solution to the preferred basis problem.

---

13. i.e. $|\text{Cat Alive}\rangle = \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}(\,|\text{Cat Alive}\rangle + |\text{Cat Dead}\rangle)\right) + \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}(\,|\text{Cat Alive}\rangle - |\text{Cat Dead}\rangle)\right).$

2.4 Decoherence Theory[*][14]

Before we can show how decoherence theory solves the preferred basis problem, we will first need to look at decoherence theory in general. To understand what's going on in decoherence theory, there are a number of things we need to discuss, namely

1. composite systems

2. entanglement

3. density matrices and traces

4. coherence

5. partial traces and reduced density matrices

6. the von Neumann measurement scheme

7. decoherence

2.4.1 Composite Systems

First we need to consider *composite systems*. We thus assume there is a distinction between what is being measured and the rest of physical reality. We denote the system that is being measured by $\mathcal{S}$ and the rest of physical reality by $\mathcal{E}$. We will refer to $\mathcal{E}$ as the environment, and we will denote the composite system of $\mathcal{S}$ and $\mathcal{E}$ by $\mathcal{U}$. We will often indicate that $\mathcal{U}$ is a composite of systems $\mathcal{S}$ and $\mathcal{E}$ by writing $\mathcal{U} = \mathcal{S} + \mathcal{E}$.

∗. As mentioned in the introduction on page 2, sections marked with an asterisk may be challenging to readers who do not have a mathematics or physics background.

14. For more details see Schlosshauer, *Decoherence and the Quantum-to-Classical Transition*, ch. 2.

The system $\mathcal{S}$ could be something microscopic like a silver atom, or something much bigger such as a cat or even a planet or galaxy.

Now suppose we have an observable (i.e. any Hermitian operator) $\hat{O}_\mathcal{S}$ that acts on the Hilbert space $H_\mathcal{S}$ of states of $\mathcal{S}$. As already mentioned, this means that we can find orthonormal eigenstates $|\psi_i\rangle_\mathcal{S}$ of $\hat{O}_\mathcal{S}$ and corresponding eigenvalues o_i such that any state $|\psi\rangle_\mathcal{S} \in H_\mathcal{S}$ can be uniquely expressed as a sum $|\psi\rangle_\mathcal{S} = \sum_{i=1}^{M} \alpha_i |\psi_i\rangle_\mathcal{S}$. We will often include the subscript $\mathcal{S}$ on the ket-vectors in order to make it clear that these ket-vectors belong to the Hilbert space $H_\mathcal{S}$. At other times we will omit these subscripts when it is clear what system we are talking about, but for the time being, we will keep these subscripts in place.

Now let us suppose we have a basis of normalized (but not necessarily orthonormal) states $\{|\chi_i\rangle_\mathcal{E} : i\}$ for the state space $H_\mathcal{E}$ of $\mathcal{E}$. In other words, every state $|\chi\rangle_\mathcal{E} \in H_\mathcal{E}$ can be uniquely expressed as a linear combination $|\chi\rangle_\mathcal{E} = \sum_{i=1}^{N} \beta_i |\chi_i\rangle_\mathcal{E}$. It is then assumed we will be able to express every state $|\xi\rangle_\mathcal{U} \in H_\mathcal{U}$ of the composite system $\mathcal{U}$ as a linear combination

$$|\xi\rangle_\mathcal{U} = \sum_{i=1}^{M}\sum_{j=1}^{N} \gamma_{i,j} |\psi_i\rangle_\mathcal{S} \, |\chi_j\rangle_\mathcal{E} \, . \tag{2.4.1}$$

Thus, we assume there are no emergent physical properties describing the composite system $\mathcal{U}$ that couldn't be expressed in terms of the component subsystems $\mathcal{S}$ and $\mathcal{E}$.

The Hilbert space $H_{\mathcal{U}}$ is endowed with the bra-ket $\langle \xi' | \xi \rangle_{\mathcal{U}}$ such that if $|\xi\rangle_{\mathcal{U}} = |\psi\rangle_{\mathcal{S}} |\chi\rangle_{\mathcal{E}}$ and $|\xi'\rangle_{\mathcal{U}} = |\psi'\rangle_{\mathcal{S}} |\chi'\rangle_{\mathcal{E}}$, then $\langle \xi' | \xi \rangle_{\mathcal{U}} = \langle \psi' | \psi \rangle_{\mathcal{S}} \langle \chi' | \chi \rangle_{\mathcal{E}}$ where we have again used subscripts to indicate which Hilbert space the bra-ket corresponds to.

2.4.2 Entanglement

By defining the bra-ket on the Hilbert space $H_{\mathcal{U}}$ as we have done, we are making the assumption that if we define $P(\psi', \chi' | \psi, \chi)$ to be the probability the composite system would be measured to be in the $|\psi'\rangle_{\mathcal{S}} |\chi'\rangle_{\mathcal{E}}$-state given that the composite system was known to be in the $|\psi\rangle_{\mathcal{S}} |\chi\rangle_{\mathcal{E}}$-state, then $P(\psi', \chi' | \psi, \chi) = P(\psi' | \psi) P(\chi' | \chi)$. A consequence of this assumption is that if the composite system is in the $|\psi\rangle_{\mathcal{S}} |\chi\rangle_{\mathcal{E}}$-state, the probability of finding system $\mathcal{S}$ to be in any particular state in $H_{\mathcal{S}}$ will be independent[15] of the state in $H_{\mathcal{E}}$ describing $\mathcal{E}$. For this reason, when the state $|\xi\rangle_{\mathcal{U}} \in H_{\mathcal{U}}$ describing the composite system $\mathcal{U}$ is expressible as a product state $|\xi\rangle_{\mathcal{U}} = |\psi\rangle_{\mathcal{S}} |\chi\rangle_{\mathcal{E}}$, we say that $\mathcal{S}$ and $\mathcal{E}$ are *not entangled* with one another. On the other hand, when the state $|\xi\rangle_{\mathcal{U}} \in H_{\mathcal{U}}$ of the composite system $\mathcal{U}$ cannot be expressed as a product state, we say that $\mathcal{S}$ and $\mathcal{E}$ are *entangled*. For example, if $|\xi\rangle_{\mathcal{U}} = \frac{1}{\sqrt{2}} (|\psi_1\rangle_{\mathcal{S}} |\chi_1\rangle_{\mathcal{E}} + |\psi_2\rangle_{\mathcal{S}} |\chi_2\rangle_{\mathcal{E}})$ with $|\psi_1\rangle_{\mathcal{S}} \not\propto |\psi_2\rangle_{\mathcal{S}}$ and $|\chi_1\rangle_{\mathcal{E}} \not\propto |\chi_2\rangle_{\mathcal{E}}$,[16] then $\mathcal{S}$ and $\mathcal{E}$ would be entangled with one another.

---

15. Here we are using the standard probabilistic definition of independence as given by (1.2.1) on page 16: two events X and Y are independent if and only if $P(X \text{ and } Y \text{ occur}) = P(X \text{ occurs}) P(Y \text{ occurs})$

16. Here the notation $|\psi_1\rangle_{\mathcal{S}} \propto |\psi_2\rangle_{\mathcal{S}}$ means there exists some α such that $|\psi_1\rangle_{\mathcal{S}} = \alpha |\psi_2\rangle_{\mathcal{S}}$, in which case $|\xi\rangle_{\mathcal{U}} = \frac{1}{\sqrt{2}} |\psi_2\rangle_{\mathcal{S}} (\alpha |\chi_1\rangle_{\mathcal{E}} + |\chi_2\rangle_{\mathcal{E}})$. Thus, if $|\psi_1\rangle_{\mathcal{S}} \propto |\psi_2\rangle_{\mathcal{S}}$, then $\mathcal{S}$ and $\mathcal{E}$ would not be entangled. This is why in the above example, we assume $|\psi_1\rangle_{\mathcal{S}} \not\propto |\psi_2\rangle_{\mathcal{S}}$, that is, we assume there is no such α such that $|\psi_1\rangle_{\mathcal{S}} = \alpha |\psi_2\rangle_{\mathcal{S}}$, and for the same reason we assume $|\chi_1\rangle_{\mathcal{S}} \not\propto |\chi_2\rangle_{\mathcal{E}}$.

Now given the observable $\hat{O}_{\mathcal{S}}$ acting on $H_{\mathcal{S}}$, we can naturally define the observable $\hat{O}_{\mathcal{U}}$ acting on $H_{\mathcal{U}}$ so that

$$\hat{O}_{\mathcal{U}} \left| \xi \right\rangle_{\mathcal{U}} = \sum_{i=1}^{M} \sum_{j=1}^{N} \gamma_{i,j} \hat{O}_{\mathcal{S}} \left| \psi_i \right\rangle_{\mathcal{S}} \left| \chi_j \right\rangle_{\mathcal{E}} = \sum_{i=1}^{M} \sum_{j=1}^{N} \gamma_{i,j} o_i \left| \psi_i \right\rangle_{\mathcal{S}} \left| \chi_j \right\rangle_{\mathcal{E}}. \qquad (2.4.2)$$

Just as in equation (2.2.2), for a given normalized state $\left| \xi \right\rangle_{\mathcal{U}} \in H_{\mathcal{U}}$, the expectation value of the observable $\hat{O}_{\mathcal{U}}$ will be $\langle \hat{O}_{\mathcal{U}} \rangle_{\xi} = \langle \xi | \hat{O}_{\mathcal{U}} | \xi \rangle_{\mathcal{U}}$. It is easy to see that if $\left| \xi \right\rangle_{\mathcal{U}} = \left| \psi \right\rangle_{\mathcal{S}} \left| \chi \right\rangle_{\mathcal{E}}$ (i.e. $\mathcal{S}$ and $\mathcal{E}$ are not entangled), then $\langle \hat{O}_{\mathcal{U}} \rangle_{\xi} = \langle \hat{O}_{\mathcal{S}} \rangle_{\psi}$.[17] Thus, when $\mathcal{S}$ and $\mathcal{E}$ are not entangled with one another, it is possible to say things about $\mathcal{S}$ independently of the current state of the environment $\mathcal{E}$. In this case we need have no knowledge of the information about $\mathcal{E}$ encapsulated in the state $\left| \chi \right\rangle_{\mathcal{E}}$ to determine the expectation value $\langle \hat{O}_{\mathcal{U}} \rangle_{\xi}$.

However, for a general entangled state $\left| \xi \right\rangle_{\mathcal{U}} = \sum_{i=1}^{M} \sum_{j=1}^{N} \gamma_{i,j} \left| \psi_i \right\rangle_{\mathcal{S}} \left| \chi_j \right\rangle_{\mathcal{E}}$, $\langle \hat{O}_{\mathcal{U}} \rangle_{\xi}$ will typically depend on the $\left| \chi_j \right\rangle_{\mathcal{E}}$-states and the coefficients $\gamma_{i,j}$. Nevertheless, despite there being a huge amount of information contained within these $\left| \chi_j \right\rangle_{\mathcal{E}}$-states and the $\gamma_{i,j}$, if we are only interested in making measurements on the system $\mathcal{S}$, nearly all this information can be discarded. In order to see how this is done, we need to generalize the notion of a state to that of a density matrix.

_____

17. This is because by definition, if $\left| \xi \right\rangle_{\mathcal{U}} = \left| \psi \right\rangle_{\mathcal{S}} \left| \chi \right\rangle_{\mathcal{E}}$ and $\left| \xi' \right\rangle_{\mathcal{U}} = \left| \psi' \right\rangle_{\mathcal{S}} \left| \chi' \right\rangle_{\mathcal{E}}$, then $\langle \xi' | \xi \rangle_{\mathcal{U}} = \langle \psi' | \psi \rangle_{\mathcal{S}} \langle \chi' | \chi \rangle_{\mathcal{E}}$. We will also have $\hat{O}_{\mathcal{U}} \left| \xi \right\rangle = \hat{O}_{\mathcal{S}} \left| \psi \right\rangle_{\mathcal{S}} \left| \chi \right\rangle_{\mathcal{E}}$. Thus, assuming both $\left| \psi \right\rangle_{\mathcal{S}}$ and $\left| \chi \right\rangle_{\mathcal{E}}$ are normalized, we have $\langle \hat{O}_{\mathcal{U}} \rangle_{\xi} = \langle \xi | \hat{O}_{\mathcal{U}} | \xi \rangle = \langle \psi | \hat{O}_{\mathcal{S}} | \psi \rangle_{\mathcal{S}} \langle \chi | \chi \rangle_{\mathcal{E}} = \langle \hat{O}_{\mathcal{S}} \rangle_{\psi}$.

2.4.3 Density Matrices and Traces

Given a normalized state $|\psi\rangle$ in any Hilbert space H, its density matrix will be the operator $\hat{\rho}_\psi \overset{\text{def}}{=} |\psi\rangle\langle\psi|$ which acts on H by sending an arbitrary state $|\psi'\rangle$ to $\langle\psi|\psi'\rangle\,|\psi\rangle$. Note that $\hat{\rho}_\psi$ is a Hermitian operator.[18] Also note that if we had a measuring device that returned the output 1 if a system was in the state $|\psi\rangle$ and 0 if the system was in a state $|\chi\rangle$ with $\langle\psi|\chi\rangle = 0$, the density matrix $\hat{\rho}_\psi$ would be the observable corresponding to this measurement. The expectation value of this measurement for an initial normalized state $|\psi'\rangle$ would then be $\langle\psi'|\hat{\rho}_\psi|\psi'\rangle = |\langle\psi|\psi'\rangle|^2 = P(\psi|\psi')$. In particular, if the system was initially in the state $|\psi\rangle$, the expectation value of this measurement would be 1.

Now it turns out that if we have an arbitrary orthonormal basis $\{\,|\phi_i\rangle : i\}$ of H and any observable $\hat{O}$ on H, then

$$\langle\hat{O}\rangle_\psi = \sum_i \langle\phi_i|\hat{\rho}_\psi\hat{O}|\phi_i\rangle\,.^{19} \tag{2.4.3}$$

Since this expression can be shown to be independent of which basis we choose,[20] we have a well-defined function called the *trace*, written as $\mathrm{Tr}(\cdot)$, which maps any

---

18. This is because for any arbitrary state $|\psi'\rangle$, $\langle\psi'|\hat{\rho}_\psi|\psi'\rangle = \langle\psi'|\psi\rangle\,\langle\psi|\psi'\rangle = \overline{\langle\psi|\psi'\rangle}\,\langle\psi|\psi'\rangle = |\langle\psi|\psi'\rangle|^2$, and so $\langle\psi'|\hat{\rho}_\psi|\psi'\rangle$ is real, and from this it follows that $\hat{\rho}_\psi$ is Hermitian.

19. To see this, as we saw in footnote 8 on page 66, if we define the mapping $I = \sum_{i=1}^{N} |\phi_i\rangle\langle\phi_i|$ then $I\,|\psi\rangle = |\psi\rangle$. Therefore, $\langle\hat{O}\rangle_\psi = \langle\psi|\hat{O}|\psi\rangle = \langle\psi|\hat{O}I|\psi\rangle = \sum_i \langle\psi|\hat{O}|\phi_i\rangle\,\langle\phi_i|\psi\rangle = \sum_i \langle\phi_i|\psi\rangle\,\langle\psi|\hat{O}|\phi_i\rangle = \sum_i \langle\phi_i|\hat{\rho}_\psi\hat{O}|\phi_i\rangle$.

20. To see this, we first note that for any orthonormal basis $\{\,|\phi_i\rangle : i\}$ of H, and any two operators $\hat{A}$ and $\hat{B}$ acting on H, using the fact that $I = \sum_i |\phi_i\rangle\langle\phi_i|$ is the identity operator on H, we have the

operator $\hat{A}$ acting on H to a value in $\mathbb{C}$ according to the formula

$$\mathrm{Tr}(\hat{A}) = \sum_i \langle \phi_i | \hat{A} | \phi_i \rangle. \tag{2.4.4}$$

Thus, it follows from equations (2.4.3) and (2.4.4) that

$$\langle \hat{O} \rangle_\psi = \mathrm{Tr}(\hat{\rho}_\psi \hat{O}). \tag{2.4.5}$$

So far we have defined the density matrix $\hat{\rho}_\psi$ corresponding to a normalized state $|\psi\rangle$. More generally, a *density matrix* $\hat{\rho}$ is defined to be a Hermitian operator with positive eigenvalues such that $\mathrm{Tr}(\hat{\rho}) = 1$. We will write $M(H)$ for the set of all density matrices on H. Since we are assuming[21] that for any Hermitian operator, there is an orthonormal basis of the Hilbert space consisting of eigenstates of the Hermitian operator, we can find an orthonormal basis $\{ |\psi_i\rangle : i \}$ of eigenstates of $\hat{\rho}$ considered as

commutativity property

$$\sum_i \langle \phi_i | \hat{A}\hat{B} | \phi_i \rangle = \sum_i \langle \phi_i | \hat{A} \sum_j |\phi_j\rangle\langle\phi_j| \hat{B} | \phi_i \rangle = \sum_{ij} \langle \phi_j | \hat{B} | \phi_i \rangle \langle \phi_i | \hat{A} | \phi_j \rangle = \sum_j \langle \phi_j | \hat{B}\hat{A} | \phi_j \rangle.$$

Now suppose that $\{ |\phi_i'\rangle : i \}$ is another orthonormal basis of H. Then we can define the operator $\hat{U}$ such that $\hat{U} |\phi_i\rangle = |\phi_i'\rangle$. We can also define the operator $\hat{U}^*$ (known as the *adjoint* of $\hat{U}$) such that $\langle \phi_i' | \psi \rangle = \langle \phi_i | \hat{U}^* | \psi \rangle$ for any state $|\psi\rangle \in H$. Since $\langle \phi_i' | \phi_j' \rangle = \langle \phi_i | \hat{U}^* | \phi_j' \rangle$ for all i, j, it will follow that $\hat{U}^* |\phi_j'\rangle = |\phi_j\rangle$. Therefore, $\hat{U}\hat{U}^* = I$. Using this fact together with the commutativity property, we have
$$\sum_i \langle \phi_i' | \hat{O} | \phi_i' \rangle = \sum_i \langle \phi_i | \hat{U}^* \hat{O} \hat{U} | \phi_i \rangle = \sum_i \langle \phi_i | \hat{O} \hat{U} \hat{U}^* | \phi_i \rangle = \sum_i \langle \phi_i | \hat{O} | \phi_i \rangle.$$

21. As mentioned earlier, we are making the assumption that Hermitian operators are compact.

a Hermitian operator with corresponding eigenvalues p_i such that

$$\hat{\rho} = \sum_i p_i \, |\psi_i\rangle\langle\psi_i| \, . \tag{2.4.6}$$

The condition $\text{Tr}(\hat{\rho}) = 1$ will then imply that $\sum_i p_i = 1$. Now we could think of the operator $\hat{\rho}$ as corresponding to a measurement which gave the output p_i when the system was in the state $|\psi_i\rangle$. However, we can alternatively think of $\hat{\rho}$ as describing a system which is known to be in one of the $|\psi_i\rangle$-states, but that we only know it is in the $|\psi_i\rangle$-state with probability p_i. Then given that $\hat{\rho}$ describes all we know about the system, the expectation value $\langle\hat{O}\rangle_\rho$ for an observable $\hat{O}$ on the system can be shown to be

$$\langle\hat{O}\rangle_\rho = \text{Tr}(\hat{\rho}\hat{O}), \, ^{22} \tag{2.4.7}$$

which is a generalization of (2.4.5). Thus, we can think of a density matrix $\hat{\rho} \in M(H)$ as a generalization of a state ket-vector $|\psi\rangle \in H$, since for every $|\psi\rangle \in H$ there corresponds a density matrix $\hat{\rho} = |\psi\rangle\langle\psi| \in M(H)$. Because of this identification, $\hat{\rho} = |\psi\rangle\langle\psi|$ is referred to as a *pure state*. On the other hand, the converse does not hold: if $\hat{\rho} = \sum_i p_i \, |\psi_i\rangle\langle\psi_i| \in M(H)$ with more than one of the $p_i > 0$, then there will not be a corresponding $|\psi\rangle \in H$ such that $\hat{\rho} = |\psi\rangle\langle\psi|$. In this case, when $\hat{\rho}$ is

---

22. This follows since $\text{Tr}(\hat{\rho}\hat{O}) = \sum_i p_i \, \text{Tr}(|\psi_i\rangle\langle\psi_i| \, \hat{O}) = \sum_i p_i \, \langle\hat{O}\rangle_{\psi_i}$ which will be the expectation value of $\hat{O}$ given that $\hat{\rho}$ encapsulates our knowledge of the system.

interpreted as describing a system that is definitely in one of the $|\psi_i\rangle$-states with probability p_i, then we will refer to $\hat{\rho}$ as a *mixed state*.

2.4.4 Coherence

Now suppose that the system $\mathcal{S}$ is initially in a superposition state $|\psi\rangle = \sum_i c_i |s_i\rangle$ with $\sum_i |c_i|^2 = 1$. Then the corresponding density matrix on $\mathcal{S}$ will be

$$|\psi\rangle\langle\psi| = \sum_{ij} c_i \overline{c_j} |s_i\rangle\langle s_j| .$$

When a density matrix has non-zero $|s_i\rangle\langle s_j|$-components for $i \neq j$, we say that there is *coherence* between the $|s_i\rangle$ and $|s_j\rangle$-states.[23] Thus, for the density matrix $|\psi\rangle\langle\psi|$ there will be coherence between the $|s_i\rangle$ and $|s_j\rangle$-states so long as both c_i and c_j are non-zero. Decoherence is a process (to be described shortly) by which the $|s_i\rangle\langle s_j|$-components (for $i \neq j$) of a density matrix restricted to a subsystem of a composite system appear to vanish.

2.4.5 Partial Traces and Reduced Density Matrices

As already mentioned, if we have a general entangled state on a composite system $\mathcal{U} = \mathcal{S} + \mathcal{E}$ of the form $|\xi\rangle_{\mathcal{U}} = \sum_{i,j} \gamma_{i,j} |\psi_i\rangle_{\mathcal{S}} |\chi_j\rangle_{\mathcal{E}}$, there is a huge amount of information in all the $\gamma_{i,j}$. However, most of this information can be discarded if we are only interested in making measurements on the system $\mathcal{S}$. We can't typically encapsulate this information in the form of a state $|\psi\rangle \in H_{\mathcal{S}}$, but we can encapsulate this

---

23. The fact that a density matrix can be written out in terms of $|s_i\rangle\langle s_j|$-components explains why we refer to a density matrix as a density *matrix*. For example, if our state space has a basis of just two states $\{|s_1\rangle, |s_2\rangle\}$, and if $\hat{\rho} = a |s_1\rangle\langle s_1| + b |s_1\rangle\langle s_2| + c |s_2\rangle\langle s_1| + d |s_2\rangle\langle s_2|$, then we can identify $\hat{\rho}$

information in the form of a density matrix $\hat{\rho}_{\mathcal{S}} \in M(H_{\mathcal{S}})$ which as mentioned on page 75 can be thought of as a generalization of a state ket-vector $|\psi\rangle \in H_{\mathcal{S}}$. In this subsection, we will show how the density matrix $\hat{\rho} = |\xi\rangle\langle\xi| \in M(H_{\mathcal{U}})$ can be reduced to a density matrix $\hat{\rho}_{\mathcal{S}} \in M(H_{\mathcal{S}})$ which encapsulates all the information needed to calculate expectation values of observables on $\mathcal{S}$. The reduced density matrix $\hat{\rho}_{\mathcal{S}}$ is derived from $\hat{\rho}$ via an operation called the partial trace.

In the context of a composite system $\mathcal{U} = \mathcal{S} + \mathcal{E}$, when taking traces, we will need to be more specific over which basis we are taking the trace over. If $\{|\psi_i\rangle : i\}$ is an orthonormal basis of $H_{\mathcal{S}}$ and $\{|\chi_j\rangle : j\}$ is an orthonormal basis of $H_{\mathcal{E}}$, then $\{|\xi_{ij}\rangle \overset{\text{def}}{=} |\psi_i\rangle |\chi_j\rangle : i, j\}$ will be an orthonormal basis of $H_{\mathcal{U}}$. For an operator $\hat{A}_{\mathcal{S}}$ of $H_{\mathcal{S}}$, we define

$$\text{Tr}_{\mathcal{S}}(\hat{A}_{\mathcal{S}}) = \sum_i \langle\psi_i|\hat{A}_{\mathcal{S}}|\psi_i\rangle \, ,$$

and for an operator $\hat{A}_{\mathcal{U}}$ of $H_{\mathcal{U}}$, we define

$$Tr_{\mathcal{U}}(\hat{A}_{\mathcal{U}}) = \sum_{ij} \langle\xi_{ij}|\hat{A}_{\mathcal{U}}|\xi_{ij}\rangle \, . \tag{2.4.8}$$

This is just what we would expect the traces to be for operators on $H_{\mathcal{S}}$ and on $H_{\mathcal{U}}$ respectively. But we also need the notion of a *partial trace* for an operator $\hat{A}_{\mathcal{U}}$ on $H_{\mathcal{U}}$:

$$\text{Tr}_{\mathcal{E}}(\hat{A}_{\mathcal{U}}) = \sum_j \langle\chi_j|\hat{A}_{\mathcal{U}}|\chi_j\rangle \, .^{24} \tag{2.4.9}$$

Note that whereas $\text{Tr}_{\mathcal{U}}(\hat{A}_{\mathcal{U}})$ is just a number, the partial trace $Tr_{\mathcal{E}}(\hat{A}_{\mathcal{U}})$ is an operator that acts on $H_{\mathcal{S}}$. To see why this is, consider the simple example of when $\hat{A}_{\mathcal{U}} = |\xi_{ij}\rangle\langle\xi_{lk}|$.

The operator $\hat{A}_{\mathcal{U}}$ would send the state $|\xi_{lk}\rangle$ to $|\xi_{ij}\rangle$ and all the other $|\xi_{l'k'}\rangle$-states of $H_{\mathcal{U}}$ to 0. But in order to define the partial trace as given in equation (2.4.9), we need to know what $\hat{A}_{\mathcal{U}}|\chi\rangle$ is and then what $\langle\chi|\hat{A}_{\mathcal{U}}|\chi\rangle$ is. In the case when $\hat{A}_{\mathcal{U}} = |\xi_{ij}\rangle\langle\xi_{lk}|$, we stipulate that $\hat{A}_{\mathcal{U}}|\chi\rangle$ is the operator that sends the state $|\psi\rangle \in H_{\mathcal{S}}$ to the state $\langle\psi_l|\psi\rangle \langle\chi_k|\chi\rangle |\xi_{ij}\rangle \in H_{\mathcal{U}}$. Furthermore, if we stipulate that $\langle\chi|\xi_{ij}\rangle = \langle\chi|\chi_j\rangle |\psi_i\rangle \in H_{\mathcal{S}}$, it follows that $\langle\chi|\hat{A}_{\mathcal{U}}|\chi\rangle$ will be the operator $\langle\chi|\chi_j\rangle \langle\chi_k|\chi\rangle |\psi_i\rangle\langle\psi_l|$ that sends the state $|\psi\rangle \in H_{\mathcal{S}}$ to the state $\langle\chi|\chi_j\rangle \langle\chi_k|\chi\rangle \langle\psi_l|\psi\rangle |\psi_i\rangle \in H_{\mathcal{S}}$. By (2.4.9), we therefore find that

$$\text{Tr}_{\mathcal{E}}(|\xi_{ij}\rangle\langle\xi_{lk}|) = \begin{cases} |\psi_i\rangle\langle\psi_l| & \text{if } j = k, \\ \\ 0 & \text{if } j \neq k. \end{cases} \tag{2.4.11}$$

And since any arbitrary operator $\hat{A}_{\mathcal{U}}$ on $H_{\mathcal{U}}$ can be expressed as a sum

$$\hat{A}_{\mathcal{U}} = \sum_{ijkl} \mu_{ijkl} |\xi_{ij}\rangle\langle\xi_{lk}|, \tag{2.4.12}$$

---

with the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. If we then identify the state $|\psi\rangle = x|s_1\rangle + y|s_2\rangle$ with the column vector $\begin{pmatrix} x \\ y \end{pmatrix}$, then the state $\hat{\rho}|\psi\rangle$ would be identified with the column vector under matrix multiplication $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$. The trace of a density matrix is then just the sum of the diagonal elements (top left to bottom right) of the matrix. Decoherence with respect to a particular basis occurs when the off-diagonal elements of the density matrix vanish.

24. The similarity between the notation of (2.4.8) and (2.4.9) could be confusing, for whereas $\langle\xi_{ij}|\hat{A}_{\mathcal{U}}|\xi_{ij}\rangle$ is just the inner product of two states $|\xi_{ij}\rangle$ and $\hat{A}_{\mathcal{U}}|\xi_{ij}\rangle$ in $H_{\mathcal{U}}$, and hence a complex number, the expression $\langle\chi_j|\hat{A}_{\mathcal{U}}|\chi_j\rangle$ in (2.4.9) is not an inner product of two states in $H_{\mathcal{U}}$. The

we can use equation (2.4.11) to find that

$$\mathrm{Tr}_{\mathcal{E}}(\hat{A}_{\mathcal{U}}) = \sum_{ijl} \mu_{ijjl} |\psi_i\rangle\langle\psi_l| . \qquad (2.4.13)$$

Now it turns out that given a density matrix $\hat{\rho}$ on $H_{\mathcal{U}}$ and an observable $\hat{O}_{\mathcal{S}}$ of $H_{\mathcal{S}}$ (which induces an observable $\hat{O}_{\mathcal{U}}$ on $H_{\mathcal{U}}$ in the obvious way, e.g. $\hat{O}_{\mathcal{U}} |\xi_{ij}\rangle = (\hat{O}_{\mathcal{S}} |\psi_i\rangle) |\chi_j\rangle$), we have the important formula

$$\langle \hat{O}_{\mathcal{U}} \rangle_{\rho} = \mathrm{Tr}_{\mathcal{S}}(\hat{\rho}_{\mathcal{S}} \hat{O}_{\mathcal{S}}) \qquad (2.4.14)$$

where

$$\hat{\rho}_{\mathcal{S}} = \mathrm{Tr}_{\mathcal{E}}(\hat{\rho}).^{25} \qquad (2.4.15)$$

---

natural domain of the operator $\hat{A}_{\mathcal{U}}$ is $H_{\mathcal{U}}$, the space of states of the composite system $\mathcal{U}$, whereas $|\xi_j\rangle \in H_{\mathcal{E}}$ is only a state of the environment $\mathcal{E}$. However, there is a natural way to define a linear operator $F(\hat{A}_{\mathcal{U}})$ corresponding to $\hat{A}_{\mathcal{U}}$ that sends states in $H_{\mathcal{E}}$ to operators from $H_{\mathcal{S}}$ to $H_{\mathcal{U}}$, i.e. for $|\chi\rangle \in H_{\mathcal{E}}$, $F(\hat{A}_{\mathcal{U}})(|\chi\rangle)$ is an operator that sends a state $|\psi\rangle \in H_{\mathcal{S}}$ to a state $F(\hat{A}_{\mathcal{U}})(|\chi\rangle)|\psi\rangle \in H_{\mathcal{U}}$. For example, if $\hat{A}_{\mathcal{U}} = |\xi_{ij}\rangle\langle\xi_{lk}|$, then $F(\hat{A}_{\mathcal{U}})(|\chi\rangle)|\psi\rangle = \langle\psi_l|\psi\rangle \langle\chi_k|\chi\rangle |\xi_{ij}\rangle$. Furthermore, there is a natural way to define a function G that maps a pair of variables $(\langle\chi|, |\xi\rangle)$ to $G(\langle\chi|, |\xi\rangle) \in H_{\mathcal{S}}$ where $|\chi\rangle \in H_{\mathcal{E}}$ and $|\xi\rangle \in H_{\mathcal{U}}$. For example, on the basis states $|\xi_{ij}\rangle$ we have $G(\langle\chi|, |\xi_{ij}\rangle) = \langle\chi|\chi_j\rangle |\psi_i\rangle$. Using this notation, we can then define the operator $\langle\chi|\hat{A}_{\mathcal{U}}|\chi\rangle$ that maps $H_{\mathcal{S}}$ to itself so that

$$\langle\chi|\hat{A}_{\mathcal{U}}|\chi\rangle (|\psi\rangle) = G\big(\langle\chi|, F(\hat{A}_{\mathcal{U}})(|\chi\rangle)|\psi\rangle\big).$$

For example, if $\hat{A}_{\mathcal{U}} = |\xi_{ij}\rangle\langle\xi_{lk}|$, then

$$\langle\chi|\hat{A}_{\mathcal{U}}|\chi\rangle (|\psi\rangle) = \langle\chi|\chi_j\rangle \langle\chi_k|\chi\rangle \langle\psi_l|\psi\rangle |\psi_i\rangle. \qquad (2.4.10)$$

By inserting (2.4.10) into (2.4.9), we obtain (2.4.11), and we can then extend (2.4.11) by linearity to obtain the partial trace for any $\hat{A}_{\mathcal{U}}$ as in equation (2.4.13).

We refer to $\hat{\rho}_\mathcal{S}$ as the *reduced density matrix* of $\hat{\rho}$.

Note that if $\hat{\rho} = |\xi\rangle\langle\xi|$ with $|\xi\rangle = |\psi\rangle\,|\chi\rangle$ so that $\mathcal{S}$ and $\mathcal{E}$ are not entangled, then $\hat{\rho}_\mathcal{S} = |\psi\rangle\langle\psi|$.[26] This is what we should expect, since if $\mathcal{S}$ and $\mathcal{E}$ are not entangled, then the expectation values of observables defined on $\mathcal{S}$ should be independent of the state of $\mathcal{E}$, and by equation (2.4.14), this independence is seen to hold when $\hat{\rho}_\mathcal{S}$ is independent of any states on $\mathcal{E}$.

More generally, for an entangled state $|\xi\rangle = \sum_{i,j} \gamma_{i,j} |\psi_i\rangle\,|\chi_j\rangle$, from equations (2.4.7) and (2.4.14), we have $\langle\hat{O}_\mathcal{U}\rangle_\rho = \langle\hat{O}_\mathcal{S}\rangle_{\rho_\mathcal{S}}$. This means that when it comes to taking expectation values of measurements on a subsystem $\mathcal{S}$ that is part of a composite system $\mathcal{U} = \mathcal{S} + \mathcal{E}$ which is in the state $|\xi\rangle \in H_\mathcal{U}$, the subsystem $\mathcal{S}$ behaves as though it was described by the density matrix $\hat{\rho}_\mathcal{S}$. However, there is a rather subtle point one needs to be aware of here.[27] For in general, as we saw in equation (2.4.6), any density matrix $\hat{\rho}_\mathcal{S} \in M(H_\mathcal{S})$ can be expressed as a sum $\hat{\rho}_\mathcal{S} = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, and this can be *thought of* as corresponding to the system $\mathcal{S}$ being in one of the $|\psi_i\rangle$-states,

---

25. To see this, following Schlosshauer, *Decoherence and the Quantum-to-Classical Transition*, 46, we have

$$\langle\hat{O}_\mathcal{U}\rangle_\rho = \mathrm{Tr}_\mathcal{U}(\hat{\rho}\hat{O}_\mathcal{U}) = \sum_{ij} \langle\xi_{ij}|\hat{\rho}\hat{O}_\mathcal{U}|\xi_{ij}\rangle = \sum_i \langle\psi_i|\left(\sum_j \langle\chi_j|\hat{\rho}|\chi_j\rangle\right)\hat{O}_\mathcal{S}\,|\psi_i\rangle$$
$$= \sum_i \langle\psi_i|\,\hat{\rho}_\mathcal{S}\hat{O}_\mathcal{S}\,|\psi_i\rangle = \mathrm{Tr}_\mathcal{S}(\hat{\rho}_\mathcal{S}\hat{O}_\mathcal{S}).$$

26. To see this, we recall that the partial trace $\mathrm{Tr}_\mathcal{E}$ is independent of which orthonormal basis $\{|\chi_j\rangle : j\}$ we choose for $\mathcal{E}$. Therefore, if $\hat{\rho} = |\xi\rangle\langle\xi|$ with $|\xi\rangle = |\psi\rangle\,|\chi\rangle$, we can choose $|\chi_1\rangle = |\chi\rangle$ and all other $|\chi_i\rangle$ such that $\langle\chi_i|\chi\rangle = 0$. Then $\mathrm{Tr}_\mathcal{E}(\hat{\rho}) = \sum_j \langle\chi_j|\hat{\rho}|\chi_j\rangle = |\psi\rangle\langle\psi|$.

27. It is unfortunate that many physicists fail to pick up on this subtlety with the result that they form the erroneous belief that decoherence theory can by itself solve the measurement problem (of

but that we only know it is in the $|\psi_i\rangle$-state with probability p_i. If this was the correct interpretation of $\hat{\rho}_S$, then as explained on page 76, we would refer to $\hat{\rho}_S$ as a *mixed state*. But just because we can think of $\hat{\rho}_S$ in this way, it doesn't follow that S really is in one of these $|\psi_i\rangle$-states and that we are only ignorant of which state it is. When S is entangled with $\mathcal{E}$ there is no fact of the matter regarding which state S is in. Rather, there are only facts of the matter for the composite system $\mathcal{U}$, e.g. the fact of the matter is that $\mathcal{U}$ is in the state $|\xi\rangle$ rather than some other state of $H_\mathcal{U}$. Therefore, $\mathcal{U}$ is really in a pure state with density matrix $|\xi\rangle\langle\xi|$. Because we cannot give an ignorance interpretation to $\hat{\rho}_S$, d'Espagnat referred to density matrices of this sort as being *improper mixtures*.[28] But despite this subtle distinction between mixed states and improper mixtures, we have nevertheless succeeded in showing how a density matrix $\hat{\rho} = |\xi\rangle\langle\xi| \in M(H_\mathcal{U})$ can be reduced to a density matrix $\hat{\rho}_S \in M(H_S)$ which encapsulates all the information needed to calculate expectation values of observables on S.[29]

---

outcomes) when in fact it can't. For a further discussion of the problem of outcomes, see Schlosshauer, *Decoherence and the Quantum-to-Classical Transition*, 57–60.

28. See Bernard d' Espagnat, *Conceptual foundations of quantum mechanics*, 2nd ed., completely rev., enl., reset., Mathematical physics monograph series ; 20 (Reading, Mass.; London: W. A. Benjamin, 1976), ch. 6.2 – cited in Butterfield, "Peaceful Coexistence: Examining Kent's Relativistic Solution to the Quantum Measurement Problem," p. 19.

29. For a more philosophical discussion of the points raised in this subsection, see Guido Bacciagaluppi, "The Role of Decoherence in Quantum Mechanics," in *The Stanford Encyclopedia of Philosophy*, Fall 2020, ed. Edward N. Zalta (Metaphysics Research Lab, Stanford University, 2020).

2.4.6 The von Neumann Measurement Scheme

We are now in a position to consider the *von Neumann measurement scheme.*[30] Instead of considering the whole of physical reality, for the time being, we just consider a physical system $\mathcal{S}$ and a measuring device $\mathcal{A}$. This division reflects the fact that a scientist doesn't measure the system $\mathcal{S}$ directly, but rather observes a measuring device $\mathcal{A}$ that is affected by $\mathcal{S}$. The measuring device $\mathcal{A}$ has the characteristic that it has a normalized ready state $|a_r(t_0)\rangle$ at initial time t_0 and that there is an orthonormal basis $\{|s_i\rangle : i\}$ of $H_{\mathcal{S}}$, and normalized states $|a_i(t)\rangle$ of $\mathcal{A}$ such that

1. for any $t \geq t_0$ we have the evolution of the states $\quad |s_i\rangle \, |a_r(t_0)\rangle \xrightarrow{\text{time evolution}}$ $|s_i\rangle \, |a_i(t)\rangle$ so that $\mathcal{S}$ and $\mathcal{A}$ do not become entangled when $\mathcal{S}$ is initially in state $|s_i\rangle$ and $\mathcal{A}$ is initially in state $|a_i(t_0)\rangle$, and

2. there exists $\delta > 0$ such that if $t > t_0 + \delta$, then $\langle a_i(t)|a_j(t)\rangle \approx 0$ for $i \neq j$.[31]

These two criteria characterize the von Neumann measurement scheme. The orthonormal basis $\{|s_i\rangle : i\}$ of $H_{\mathcal{S}}$ for which these two criteria hold are called *pointer states* of $\mathcal{A}$. These pointer states will be determined by the dynamics of the composite system $\mathcal{S} + \mathcal{A}$ as well as the relative configuration of $\mathcal{S}$ with respect to $\mathcal{A}$. For instance, if $\mathcal{S}$ is a silver atom and $\mathcal{A}$ is a Stern-Gerlach apparatus, then the configuration and dynamics of the system will determine a fixed axis $\hat{\mathbf{a}}$ relative to the Stern-Gerlach configuration $\mathcal{A}$ such that the states $|\hat{\mathbf{a}}+\rangle$ and $|\hat{\mathbf{a}}-\rangle$ of $\mathcal{S}$ don't get entangled with

---

30. See Schlosshauer, *Decoherence and the Quantum-to-Classical Transition*, 50–53 for more details.

31. More precisely, we should say that for all $\epsilon > 0$ there exists $\delta > 0$ such that if $t > t_0 + \delta$, then $|\langle a_i(t)|a_j(t)\rangle| < \epsilon$ for $i \neq j$.

$\mathcal{A}$, that is, there exists $\delta > 0$ such that $|\hat{\mathbf{a}}\pm\rangle\, |a_r(t_0)\rangle \xrightarrow{\text{time evolution}} |\hat{\mathbf{a}}\pm\rangle\, |a_\pm(t)\rangle$ with

$\langle a_+(t)|a_-(t)\rangle \approx 0$ for $t > t_0 + \delta$.[32, 33] Since no entanglement occurs with the silver

atom and the Stern-Gerlach apparatus when the silver atom is in the $|\hat{\mathbf{a}}\pm\rangle$-state, then

in this situation, we can interact with the apparatus to find out whether the particle

is in the $|\hat{\mathbf{a}}+\rangle$-state or the $|\hat{\mathbf{a}}-\rangle$-state without changing the spin state of the silver

atom. Indeed, we should expect an experimental apparatus to have this property of

non-entanglement with the measurement outcomes it reports, for otherwise, every

scientist who looked at the measurement device couldn't be sure that the spin state

of the silver atom being measured remained unchanged whenever the apparatus was

observed, and so the scientists couldn't expect there to be any agreement among

themselves regarding which spin-state the silver atom was in. Thus, the basis of $H_\mathcal{S}$ for

which entanglement doesn't occur is a preferred basis. However, if we were to consider

a different basis, say $\{\frac{1}{\sqrt{2}}(|\hat{\mathbf{a}}+\rangle + |\hat{\mathbf{a}}-\rangle), \frac{1}{\sqrt{2}}(|\hat{\mathbf{a}}+\rangle - |\hat{\mathbf{a}}-\rangle)\}$,[34] then assuming that the

configuration of $\mathcal{A}$ remained unchanged, entanglement between $\mathcal{S}$ and $\mathcal{A}$ would occur

---

32. Strictly speaking, we would need more information to describe states in $H_\mathcal{S}$ besides the spin, so we should really express this scenario in terms of $\{\,|s_{i,+}\rangle \stackrel{\text{def}}{=} |\hat{\mathbf{a}}+, i\rangle : i\,\} \cup \{\,|s_{i,-}\rangle \stackrel{\text{def}}{=} |\hat{\mathbf{a}}-, i\rangle : i\,\}$ and $\{\,|a_{i,+}(t)\rangle : i\,\} \cup \{\,|a_{i,-}(t)\rangle : i\,\}$ where the i-indices encode all the additional information beyond spin.

33. Although we only require that $\langle a_+(t)|a_-(t)\rangle \approx 0$ for $t > t_0 + \delta$ rather than demanding $\langle a_+(t)|a_-(t)\rangle = 0$, we can think of the scientist who observes the apparatus as determining whether the apparatus is either in one of two normalized state $|a'_+(t)\rangle$ or $|a'_-(t)\rangle$ where $\langle a'_+(t)|a'_-(t)\rangle = 0$ and $\langle a'_\pm(t)|a'_\pm(t)\rangle \approx 1$, so that the scientist can confidently assert that the particle is in the state $|\hat{\mathbf{a}}+\rangle$ if for instance the measurement device is found to be in the state $|a'_+(t)\rangle$. Because $\langle a'_+(t)|a_-(t)\rangle$ is only very small, but not identically zero, in theory, the particle could be in the $|\hat{\mathbf{a}}-\rangle$-state, but we're assuming that such a possibility would be as likely as a significant violation of the Second Law of Thermodynamics, say.

34. According to equation (1.1.2), this basis would correspond to measuring the spin in an axis at right angles to $\hat{\mathbf{a}}$.

since then $\frac{1}{\sqrt{2}}(|\hat{\mathbf{a}}+\rangle \pm |\hat{\mathbf{a}}-\rangle) |a_r(t_0)\rangle \xrightarrow{\text{time evolution}} \frac{1}{\sqrt{2}}(|\hat{\mathbf{a}}+\rangle |a_+(t)\rangle \pm |\hat{\mathbf{a}}-\rangle |a_-(t)\rangle)$.

Thus, $\{\frac{1}{\sqrt{2}}(|\hat{\mathbf{a}}+\rangle + |\hat{\mathbf{a}}-\rangle), \frac{1}{\sqrt{2}}(|\hat{\mathbf{a}}+\rangle - |\hat{\mathbf{a}}-\rangle)\}$ would not be a preferred basis. In this

case, if $\mathcal{S}$ was in the $\frac{1}{\sqrt{2}}(|\hat{\mathbf{a}}+\rangle + |\hat{\mathbf{a}}-\rangle)$-state, a scientist would measure $\mathcal{A}$ to be in the

$|a_+(t)\rangle$-state with probability $\frac{1}{2}$. But having measured $\mathcal{A}$ to be in the $|a_+(t)\rangle$-state,

the scientist would continue to observe $\mathcal{A}$ to be in the $|a_+(t)\rangle$-state because of the

subsequent non-entanglement of $\mathcal{S}$ with $\mathcal{A}$ when $\mathcal{S}$ is in the $|\hat{\mathbf{a}}+\rangle$-state and $\mathcal{A}$ is in the

$|a_+(t)\rangle$-state. Note that this situation is somewhat analogous to when we have the

Bell-state (1.2.6), so that when Bob measures his particle to be in the $|\hat{\mathbf{a}}-\rangle$-state, he

knows that Alice's particle is in the $|\hat{\mathbf{a}}+\rangle$-state. Likewise, in the von Neumann mea-

surement scheme, if the scientist measures $\mathcal{A}$ to be in the $|a_+(t)\rangle$-state for $t > t_0 + \delta$,

he will then (almost certainly) know[35] that the system $\mathcal{S}$ will be in the $|\hat{\mathbf{a}}+\rangle$-state.

In the case where $\mathcal{S}$ has more than two states, we can write a generic normalized

state of the composite system $\mathcal{U} = \mathcal{S} + \mathcal{A}$ as $|\Psi(t)\rangle = \sum_i c_i |\xi_i(t)\rangle$ where $|\xi_i(t)\rangle =$

$|s_i\rangle |a_i(t)\rangle$. There will then be coherence between $|\xi_i(t)\rangle$ and $|\xi_j(t)\rangle$ for the density

matrix $\hat{\rho}(t) \overset{\text{def}}{=} |\Psi(t)\rangle\langle\Psi(t)|$ so long as both c_i and c_j are non-zero. However, if we are

only interested in observables $\hat{O}_\mathcal{S}$ on $H_\mathcal{S}$, then we only need to consider the reduced

density matrix $\hat{\rho}_\mathcal{S}(t) = \text{Tr}_\mathcal{A}(\hat{\rho}(t))$. Initially, at time t_0 we have $|a_i(t_0)\rangle = |a_r(t_0)\rangle$

so $|\Psi(t_0)\rangle = |\psi\rangle |a_r(t_0)\rangle$ where $|\psi\rangle = \sum_i c_i |s_i\rangle$ which we assume to be normalized.

Thus, initially, $\mathcal{S}$ would not be entangled with $\mathcal{A}$, and therefore the density matrix

---

35. We say that Bob is *almost* certain rather than *completely* certain because $\langle a_+(t)|a_-(t)\rangle$ will be very nearly zero rather than identically zero as discussed in footnote 33 on page 83.

describing $\mathcal{S}$ would be $\hat{\rho}_\mathcal{S}(t_0) = |\psi\rangle\langle\psi|$.[36] Hence, if we consider $\hat{O}_\mathcal{S} = |\psi\rangle\langle\psi|$ as an observable on $\mathcal{S}$ corresponding to a measurement[37] that records the value 1 if the system is in the $|\psi\rangle$ and 0 if the system is in a state $|\psi'\rangle$ with $\langle\psi'|\psi\rangle = 0$, then both intuitively[38] and by equation (2.4.14),[39] we would have $\langle\hat{O}_\mathcal{U}\rangle_{\rho(t_0)} = 1$. But if the scientist is to measure $\mathcal{S}$ to be in the $|\psi\rangle$-state, the expectation value $\langle\hat{O}_\mathcal{U}\rangle_{\rho(t)}$ would have to be 1 for times t discernibly greater than t_0.

However, if more than one of the c_i are non-zero, then the scientist will not be able to measure the system $\mathcal{S}$ to be in the $|\psi\rangle$-state for any discernible length of time. To see why this is, we first note that

$$\hat{\rho}_\mathcal{S}(t) = \sum_i |c_i|^2 |s_i\rangle\langle s_i| + \sum_{i \neq j} c_i \overline{c_j} \langle a_j(t)|a_i(t)\rangle |s_i\rangle\langle s_j| .^{40} \qquad (2.4.16)$$

---

36. Recall if $|\xi\rangle_\mathcal{U} = |\psi\rangle_\mathcal{S} |\chi\rangle_\mathcal{E}$ (i.e. $\mathcal{S}$ and $\mathcal{E}$ are not entangled), then $\langle\hat{O}_\mathcal{U}\rangle_\xi = \langle\hat{O}_\mathcal{S}\rangle_\psi$ as explained in footnote 26 on page 80.

37. This is a measurement we conduct by some means other than looking at the apparatus $\mathcal{A}$.

38. I.e. we would expect the expectation value of $\hat{O}_\mathcal{U}$ to be 1 if we knew that $\mathcal{S}$ was in the state $|\psi\rangle$ with probability 1.

39. I.e. given that $\hat{\rho}_\mathcal{S}(t_0) = |\psi\rangle\langle\psi| = \hat{O}_\mathcal{S}$, and that $\hat{O}_\mathcal{S}^2 = \hat{O}_\mathcal{S}$, and $\text{Tr}_\mathcal{S}(\hat{O}_\mathcal{S}) = 1$, it follows that $\langle\hat{O}_\mathcal{U}\rangle_{\rho(t_0)} = \text{Tr}_\mathcal{S}(\hat{\rho}_\mathcal{S}(t_0)\hat{O}_\mathcal{S}) = \text{Tr}_\mathcal{S}(\hat{O}_\mathcal{S}) = 1$.

Now because $\langle a_j(t)|a_i(t)\rangle \approx 0$ for $t > t_0 + \delta$, it follows that $\hat{\rho}_\mathcal{S} \approx \sum_i |c_i|^2 \, |s_i\rangle\langle s_i|$ for $t > t_0 + \delta$. It will then follow that $\langle \hat{O}_\mathcal{U}\rangle_{\rho(t)} = \sum_i |c_i|^4$,[41] and this will only be 1 if only one of the c_i is 1 and all the other c_i are 0. Hence, if more than one of the c_i are non-zero, the scientist will not be able to measure the system $\mathcal{S}$ to be in the $|\psi\rangle$-state for any discernible length of time.

2.4.7 Decoherence

Note that although for the original density matrix $|\psi\rangle\langle\psi|$ there is coherence between the states $|s_i\rangle$ and $|s_j\rangle$, this coherence effectively disappears when the system $\mathcal{S}$ interacts with the measuring device $\mathcal{A}$ (i.e. the $|s_i\rangle\langle s_j|$-coefficients of $\hat{\rho}_\mathcal{S}$ are approximately zero for $t > t_0 + \delta$). This is what we mean by *decoherence*: the coherence has effectively disappeared. The *decoherence time* δ which is the time it takes for $\langle a_i(t)|a_j(t)\rangle$ to go from 1 when $t = t_0$ to approximately zero when $t = t_0 + \delta$ will depend on what situation we are considering, but very often this time will be extremely small. For instance if we were measuring neurons firing in the brain, the decoherence time will

---

40. To see this, it is sufficient to show that $\mathrm{Tr}_\mathcal{A}(|\xi_i(t)\rangle\langle\xi_j(t)|) = \langle a_j(t)|a_i(t)\rangle\,|s_i\rangle\langle s_j|$ for then we will obtain the first summand of $\hat{\rho}_\mathcal{S}$ from the fact that $|a_i(t)\rangle$ are normalized, and we will obtain the second summand by linearity of $\mathrm{Tr}_\mathcal{A}(\cdot)$. Well, taking $\{|\phi_k\rangle : k\}$ to be an orthonormal basis of $H_\mathcal{A}$, we have

$$\mathrm{Tr}_\mathcal{A}(|\xi_i(t)\rangle\langle\xi_j(t)|) = \sum_k \langle\phi_k|\left(|\xi_i(t)\rangle\langle\xi_j(t)|\right)|\phi_k\rangle = \sum_k \langle\phi_k|a_i(t)\rangle\,\langle a_j(t)|\phi_k\rangle\,|s_i\rangle\langle s_j|$$

$$= \langle a_j(t)|\left(\sum_k |\phi_k\rangle\langle\phi_k|\right)|a_i(t)\rangle\,|s_i\rangle\langle s_j| = \langle a_j(t)|a_i(t)\rangle\,|s_i\rangle\langle s_j|,$$

where we have used the fact that $I = \sum_k |\phi_k\rangle\langle\phi_k|$ is the identity operator on $H_\mathcal{A}$.

41. This is because

$$\mathrm{Tr}_\mathcal{S}(\hat{\rho}_\mathcal{S}\,|\psi\rangle\langle\psi|) \approx \mathrm{Tr}_\mathcal{S}(\sum_i |c_i|^2\,|s_i\rangle\langle s_i|\sum_{jk} c_j\overline{c_k}\,|s_j\rangle\langle s_k|)$$

$$= \mathrm{Tr}_\mathcal{S}(\sum_{ik} |c_i|^2 c_i\overline{c_k}\,|s_i\rangle\langle s_k|) = \sum_l \sum_{ik} \langle s_l|\,|c_i|^2 c_i\overline{c_k}\,|s_i\rangle\,\langle s_k|s_l\rangle = \sum_i |c_i|^4.$$

typically be of the order $\delta = 10^{-19}\,\text{s}$.[42] It is because of decoherence that we can't expect the system $\mathcal{S}$ to remain in the state $|\psi\rangle = \sum_i c_i |s_i\rangle$ for any discernible length of time, unless $|\psi\rangle$ is proportional to one of the $|s_i\rangle$-states.

Also note that when decoherence occurs, we say the coherence *effectively* disappears, insofar as the coherence will not be measurable if we only consider observables just acting on $H_{\mathcal{S}}$. Thus, after decoherence has taken place, if we restrict our attention to the system $\mathcal{S}$ alone, it will be experimentally indistinguishable[43] from the situation where $\mathcal{S}$ is known to be in one of the $|s_i\rangle$-states, but where we only know that it is in the $|s_i\rangle$-state with probability $|c_i|^2$. Nevertheless, the coherence is still there, since if we chose to consider more general observables on the composite $\mathcal{S} + \mathcal{A}$, the $|\xi_i(t)\rangle\langle\xi_j(t)|$-coefficients of $\hat{\rho}$ will continue to be $c_i \overline{c_j}$ which will in general will be non-zero, and as a whole, at time t the composite system will be in the state $|\Psi(t)\rangle = \sum_i c_i |\xi_i(t)\rangle$ with probability 1.

2.5 A Solution to the Preferred Basis Problem[44]

We can now see how decoherence theory solves the preferred basis problem. Although up to this point we have been focusing on how a system $\mathcal{S}$ interacts with

---

42. For details of this estimate see Schlosshauer, *Decoherence and the Quantum-to-Classical Transition*, 370.

43. Recall the discussion following equation (2.4.6) on page 75 as well as the discussion on page 80.

44. For more details, see Schlosshauer, *Decoherence and the Quantum-to-Classical Transition*, 71–84

a measuring apparatus $\mathcal{A}$, we can generalize to the situation in which a system $\mathcal{S}$ interacts with its environment $\mathcal{E}$. We can still define pointer states in the same way as we did on page 82. These pointer states will then make up the preferred basis. The two defining criteria of pointer states entail that pointer states will remain stable and immune to decoherence effects.

Since physicists have a good understanding of how different systems interact, they are able to explain what it is about a basis that makes it a preferred basis. The details of their analysis need not concern us here, but it's possible to show that for macroscopic and mesoscopic objects, states specified in terms of position decohere with one another very rapidly.[45] This explains why we don't detect $\frac{1}{\sqrt{2}}(\,|\text{Cat Alive}\rangle +$ $|\text{Cat Dead}\rangle)$ and $\frac{1}{\sqrt{2}}(\,|\text{Cat Alive}\rangle - |\text{Cat Dead}\rangle)$-states, but we do detect $|\text{Cat Alive}\rangle$ and $|\text{Cat Dead}\rangle$-states. Also note that $|\text{Cat Alive}\rangle$ does indeed have the property that it is immune to decoherence effects, for if we were to express $|\text{Cat Alive}\rangle$ in terms of the basis $\{\,|\psi_+\rangle\,,\,|\psi_-\rangle\}$ where $|\psi_\pm\rangle = \frac{1}{\sqrt{2}}(\,|\text{Cat Alive}\rangle \pm |\text{Cat Dead}\rangle)$, then $|\text{Cat Alive}\rangle = \frac{1}{\sqrt{2}}(\,|\psi_+\rangle + |\psi_-\rangle)$. The corresponding density matrix would then be

$$|\text{Cat Alive}\rangle\langle\text{Cat Alive}| = \frac{1}{\sqrt{2}}(|\psi_+\rangle\langle\psi_+| + |\psi_-\rangle\langle\psi_-| + |\psi_+\rangle\langle\psi_-| + |\psi_-\rangle\langle\psi_+|). \quad (2.5.1)$$

Since in normal situations, the left-hand side of equation (2.5.1) will remain unperturbed by the environment, the coefficients of the off-diagonal terms $|\psi_\pm\rangle\langle\psi_\mp|$ will also remain as they are; that is, $|\psi_\pm\rangle$ and $|\psi_\mp\rangle$ will not decohere with one another. It

---

45. e.g. See the discussion in Schlosshauer, *Decoherence and the Quantum-to-Classical Transition*, 94.

is only in very contrived situations such as when the cat's environment is a poison releasing device coupled to a radioactive atom that |Cat Alive⟩ will no longer be a pointer state with respect to this environment.

2.6 The Problem of Nonobservability of Interference and its Solution[46]

Before we consider the many-worlds interpretation in detail, it will be helpful to consider the role that decoherence plays in the removal of quantum interference at the macroscopic scale, as it is this lack of quantum interference between mutually exclusive states that justifies our belief that a system is in a definite state rather than in a superposition of alternative realities. The question of why quantum interference typically disappears at macroscopic scales is referred to as the problem of the nonobservability of interference.

We can explain this problem in the context of the double slit experiment: As figure 2.1 indicates, when a beam of particles is incident on a double slit, the particles that are detected on the detection screen are distributed according to a distribution pattern which either exhibits quantum interference as shown on the left in the figure, or does not exhibit such interference as shown on the right. Small particles like electrons and photons will tend to exhibit quantum interference, whereas mesoscopic particles will not typically exhibit quantum interference.[48]

---

46. See Schlosshauer, *Decoherence and the Quantum-to-Classical Transition*, 55–57, 63–65.

48. There are exceptions to this rule. For example, a *Superconducting Quantum Interference Devices (SQUID)* can demonstrate quantum interference even at macroscopic scales where a large superconductor can enter into a superposition of two states with the current flowing in opposite
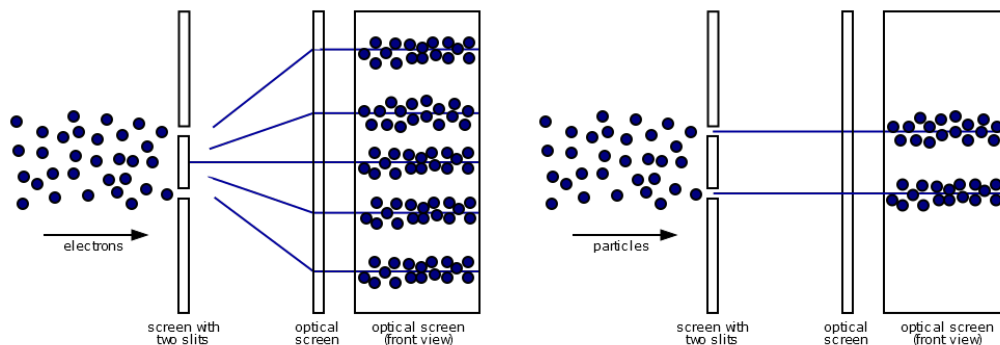
(A) Particles exhibiting interference.   (B) Particles not exhibiting interference.

Figure 2.1. The Double-Slit Experiment. Particles are incident on a double slit. In diagram (A), the particles are exhibiting an interference pattern, whereas in diagram (B), the particles are not exhibiting an interference pattern. Whether or not there is interference will depend on factors such as the size of the particles and whether it can be ascertained which slit the particle went through. The larger the particles are, or the more information available as to which slit the particle went through, the less likely the particles will exhibit interference.[47]

To explain what is going on, we suppose that when just the top slit is open, the normalized state of the particle is $|\psi_1\rangle$, whereas if just the bottom slit is open, we suppose that the normalized state of the particle is $|\psi_2\rangle$, and when both slits are open, we suppose that the state of the particle will be $\frac{1}{\sqrt{2}}(|\psi_1\rangle + |\psi_2\rangle)$. Now let the variable x describe the position on the detection screen. For instance, we might take $x = 0$ to be the center of the detection screen, and take positive values of x as corresponding to positions on the upper part of the screen, and negative values of x as corresponding to positions on the lower part of the screen, but the precise convention we adopt won't

directions in each state. See Schlosshauer, *Decoherence and the Quantum-to-Classical Transition*, ch. 6.

matter. Then we define the $|x\rangle$-state[49] as the physical state describing the particle to be exactly located at position x on the screen. Note that the state $|x\rangle$ is indexed by a continuous parameter, x. This is in contrast to the basis of states $|s_i\rangle$ which we have been considering up until now which are indexed by discrete values of i such as $i = 1, 2, \ldots$. Because of this difference, we need to use calculus to deal with $|x\rangle$-states in a rigorous manner, but such details will not concern us here. In reality, because of the Heisenberg uncertainty principle, a particle is never in just one $|x\rangle$-state, but rather the particle will be in a superposition of many $|x\rangle$-states, which may or may not be concentrated around a particular location, x_0 say. The more concentrated these $|x\rangle$-states of this superposition are concentrated around a particular location x_0, the more the particle will have the particle-like characteristic of being localized in one place. But if the $|x\rangle$-states of this superposition are more spread out, the particle will have more wave-like characteristics. So when physicists speak of particles, often they are not thinking of physical entities that are very localized in position, as non-physicists would think. Nevertheless, at the moment the particle is detected on the detection screen, it does seem to be highly localized.

Given a state $|\psi\rangle$ for a so-called particle, we define the function $\psi(x) = \langle x|\psi\rangle$. Because of the continuous nature of the variable x (in contrast to the discrete nature

---

48. Diagrams (A) and (B) are by inductiveload, and are Public domain, via Wikimedia Commons. Sources: https://commons.wikimedia.org/wiki/File:Two-Slit_Experiment_Electrons.svg and https://commons.wikimedia.org/wiki/File:Two-Slit_Experiment_Particles.svg.

49. $|x\rangle$ is not really a state in the proper sense. With the states we've seen so far, when $|\phi\rangle$ and $|\psi\rangle$ have been normalized, then $|\langle\phi|\psi\rangle|^2$ will be a conditional probability, and hence at most 1. However, $|x\rangle$ cannot be normalized. This is because the bracket $\langle x|y\rangle$ is defined to be $\langle x|y\rangle = \delta(x - y)$ where

of i in a basis $\{\,|s_i\rangle : i\})$, the function

$$\rho(x) = |\psi(x)|^2 \tag{2.6.2}$$

determines a probability density for a range of outcomes rather than a probability for a specific outcome. Here, we do not need to go into the details of probability densities,[50] but roughly speaking, the greater the value of $\rho(x)$, the greater will be the relative probability of detecting the particle in the vicinity of location x. Thus, if $\rho(x') = 0$ for all x' in the vicinity of x, then the particle would not be detected in the vicinity of location x.

Now if $|\psi\rangle = \frac{1}{\sqrt{2}}(\,|\psi_1\rangle + |\psi_2\rangle)$, then $\psi(x) = \frac{1}{\sqrt{2}}(\psi_1(x) + \psi_2(x))$. Therefore, the corresponding probability density will be

$$|\psi(x)|^2 = \frac{1}{2}(|\psi_1(x)|^2 + |\psi_2(x)|^2 + 2\,\mathrm{Re}(\overline{\psi_1(x)}\psi_2(x)).^{51}$$

$\delta(x)$ is the *Dirac delta* function such that

$$\delta(x) = \begin{cases} \infty & \text{if } x = 0, \\ 0 & \text{if } x \neq 0, \end{cases} \tag{2.6.1}$$

and has the property that $\int \mathrm{d}x\,\delta(x)\,f(x) = f(0)$ for any continuous function $f(x)$. The theory of distributions allows one to deal rigorously with Dirac delta functions. E.g. see Walter Rudin, *Functional Analysis*, Second Edition (McGraw-Hill, 1991), ch. 6.

50. But if you are interested, a probability density $\rho(x)$ for a random variable X that has real values is a function such that $\rho(x) \geq 0 \; \forall x \in \mathbb{R}$, and that $\int_{\mathbb{R}} \rho(x)\mathrm{d}x = 1$ and the probability that X has a value in the subset $U \subset \mathbb{R}$ is $\int_U \rho(x)\mathrm{d}x$.

Now when the detection screen is far away from the double slits, we will have

$|\psi_1(x)|^2 \approx |\psi_2(x)|^2$ for x near the center point on the screen. However, depending on

slight changes in the value of x from the center point on the screen, sometimes $\psi_1(x)$

and $\psi_2(x)$ will be in phase so that $\psi_1(x) \approx \psi_2(x)$, in which case $|\psi(x)|^2 \approx 2|\psi_1(x)|^2$.

But sometimes $\psi_1(x)$ and $\psi_2(x)$ will be out of phase so that $\psi_1(x) \approx -\psi_2(x)$, in which

case $|\psi(x)|^2 \approx 0$. Hence, we get the interference pattern as shown in figure 2.1 (A).

Now in order to consider how decoherence affects interference, we let

$$|\Psi(t)\rangle = \frac{1}{\sqrt{2}}\left(|\psi_1\rangle\, |E_1(t)\rangle + |\psi_2\rangle\, |E_2(t)\rangle \right)$$

be the state of the composite system $\mathcal{U} = \mathcal{S} + \mathcal{E}$ where $\mathcal{S}$ is a particle that has gone

through the double slit and will be detected on the detection screen, and $\mathcal{E}$ is the local

environment of the experimental set up. The expression for $|\Psi(t)\rangle$ indicates that we

are assuming $\mathcal{S}$ doesn't become entangled with $\mathcal{E}$ when $\mathcal{S}$ is in the state $|\psi_1\rangle$ or $|\psi_2\rangle$.

Corresponding to $|\Psi(t)\rangle$, we can define the density matrix $\hat{\rho}(t) = |\Psi(t)\rangle\langle\Psi(t)|$. We can

also define the observable[52] $|x\rangle\langle x|_\mathcal{S}$ for the system $\mathcal{S}$ so that $|x\rangle\langle x|_\mathcal{S}\, |\psi\rangle_\mathcal{S} = \psi(x)\, |x\rangle_\mathcal{S}$.

As we saw in equation (2.4.2) on page 72, we can naturally extend the action of $|x\rangle\langle x|_\mathcal{S}$

---

51. Here Re means the real part of a complex number. Thus, if the complex number $z = \alpha + i\beta$ for real numbers α and β, then $\mathrm{Re}(z) = \alpha$. To see why the above equation holds, we recall that $|z|^2 = z\bar{z}$ and that $\mathrm{Re}(z) = \frac{1}{2}(z + \bar{z})$. Therefore, if $z = \frac{1}{\sqrt{2}}(v + w)$ for complex number v and w, then $|z|^2 = \frac{1}{\sqrt{2}}(v+w)\frac{1}{\sqrt{2}}\overline{(v+w)} = \frac{1}{2}(v+w)\overline{(v+w)} = \frac{1}{2}(v\bar{v}+w\bar{w}+w\bar{v}+v\bar{w}) = \frac{1}{2}(|v|^2+|w|^2+w\bar{v}+\overline{w\bar{v}}) = \frac{1}{2}(|v|^2 + |w|^2 + 2\,\mathrm{Re}(\bar{v}w))$.

52. Note that we only call $|x\rangle\langle x|_\mathcal{S}$ an observable in an analogical sense since it is not a compact Hermitian operator acting on the Hilbert space of states $H_\mathcal{S}$. If we were being more rigorous, we would need to consider a Hermitian operator of the form $\int \sigma(x)\, |x\rangle\langle x|_\mathcal{S}\, dx$ for an appropriate test function $\sigma(x)$.

to $H_\mathcal{U}$.[53] This allows us to define

$$\rho_\mathcal{U}(x,t) \stackrel{\text{def}}{=} \text{Tr}_\mathcal{U}(\hat{\rho}(t)_\mathcal{U} \, |x\rangle\langle x|_\mathcal{U}).$$

In the specific case when $\mathcal{S}$ and $\mathcal{E}$ are not entangled so that $\hat{\rho}_\mathcal{U}(t) = |\xi(t)\rangle\langle\xi(t)|_\mathcal{U}$ with $|\xi(t)\rangle_\mathcal{U} = |\psi\rangle_\mathcal{S} \, |E(t)\rangle_\mathcal{E}$ for normalized states $|\psi\rangle_\mathcal{S}$ and $|E(t)\rangle_\mathcal{E}$, we have $\rho_\mathcal{U}(x,t) = |\psi(x)|^2$ which is equal to the probability density function $\rho(x)$ we saw in equation (2.6.2).[54] However, if $|E_1(t)\rangle_\mathcal{E}$ is not proportional to $|E_2(t)\rangle_\mathcal{E}$, then $|\Psi(t)\rangle$ will be an entangled state of $\mathcal{S}$ and $\mathcal{E}$. But whether or not $\mathcal{S}$ and $\mathcal{E}$ are entangled, we can still use equation (2.4.16) to calculate the partial trace:

$$\hat{\rho}_\mathcal{S}(t) = \frac{1}{2}(|\psi_1\rangle\langle\psi_1|_\mathcal{S} + |\psi_2\rangle\langle\psi_2|_\mathcal{S} + \langle E_2(t)|E_1(t)\rangle_\mathcal{E} \, |\psi_1\rangle\langle\psi_2|_\mathcal{S} + \langle E_1(t)|E_2(t)\rangle_\mathcal{E} \, |\psi_2\rangle\langle\psi_1|_\mathcal{S}).$$

By equations (2.4.7) and (2.4.14), we therefore have

$$\rho_\mathcal{U}(x,t) = \frac{1}{2}\left(|\psi_1(x)|^2 + |\psi_2(x)|^2 + 2\,\text{Re}\left(\langle E_2(t)|E_1(t)\rangle_\mathcal{E} \, \overline{\psi_2(x)}\psi_1(x)\right)\right).^{[55]} \qquad (2.6.3)$$

---

53. Strictly speaking, it is not $|x\rangle\langle x|_\mathcal{S}$ that is extended to act on $H_\mathcal{U}$, but rather a Hermitian operator of the form $\int \sigma(x) \, |x\rangle\langle x|_\mathcal{S} \, dx$ for an appropriate test function $\sigma(x)$ that is extended to $H_\mathcal{U}$. For a state $|\xi\rangle_\mathcal{U} = |\psi\rangle_\mathcal{S} \, |E\rangle_\mathcal{E}$, the action of $|x\rangle\langle x|_\mathcal{U}$ on $|\xi\rangle_\mathcal{U}$ gives the 'state' $|x\rangle\langle x|_\mathcal{U} \, |\xi\rangle_\mathcal{U} \stackrel{\text{def}}{=} \psi(x) \, |x\rangle_\mathcal{S} \, |E\rangle_\mathcal{E}$, but since this is not normalizable, we have to 'smear' it by integrating it with respect to the test function $\sigma(x)$.

54. To see this, note that we can ignore $\mathcal{E}$ in calculating $\langle|x\rangle\langle x|_\mathcal{U}\rangle_\xi$ since when $\mathcal{S}$ and $\mathcal{E}$ are not entangled, $\langle|x\rangle\langle x|_\mathcal{U}\rangle_\xi = \langle|x\rangle\langle x|_\mathcal{S}\rangle_\psi$ as explained in footnote 17 on page 72. We can therefore just consider $\mathcal{S}$ and drop the subscripts. Furthermore, as we saw in equation (2.4.7), $\langle\hat{O}\rangle_\psi = \text{Tr}(\hat{\rho}\hat{O})$ where $\hat{\rho} = |\psi\rangle\langle\psi|$. We can thus take an orthonormal basis $\{|\psi_1\rangle, |\psi_2\rangle, \ldots\}$ of $H_\mathcal{S}$ with $|\psi_1\rangle = |\psi\rangle$. Then $\rho(x) = \text{Tr}(\hat{\rho} \, |x\rangle\langle x|) = \text{Tr}(|\psi\rangle\langle\psi| \, |x\rangle\langle x|) = \sum_i \langle\psi_i|\psi\rangle \, \langle\psi|x\rangle \, \langle x|\psi_i\rangle = \langle\psi_1|\psi\rangle \, \langle\psi|x\rangle \, \langle x|\psi_1\rangle = \langle\psi|x\rangle \, \langle x|\psi\rangle = \overline{\langle x|\psi\rangle} \, \langle x|\psi\rangle = |\langle x|\psi\rangle|^2 = |\psi(x)|^2$.

Thus, if $\langle E_1(t)|E_2(t)\rangle_{\mathcal{E}} \approx 0$ then $\rho_{\mathcal{U}}(x,t) \approx \frac{1}{2}\left(|\psi_1(x)|^2 + |\psi_2(x)|^2\right)$ and so we would

observe a distribution pattern not exhibiting interference as shown in figure 2.1 (B).

On the other hand, if $\langle E_1(t)|E_2(t)\rangle_{\mathcal{E}} \not\approx 0$ we would get a distribution pattern exhibiting

interference as shown in figure 2.1 (B). Therefore, since it is often possible to determine

the asymptotic behavior of $\langle E_1(t)|E_2(t)\rangle_{\mathcal{E}}$,[56] decoherence theory gives us a means of

determining whether or not quantum interference will be exhibited.

2.7 The Problem of Outcomes

In the last two sections we have seen how decoherence theory solves the preferred

basis problem and the problem of the nonobservability of interference. However, there

is a third fundamental problem in quantum physics which decoherence theory is unable

to solve. This is the problem of outcomes. As discussed in subsection 2.4.6, in the

---

55. The calculation is as follows:

$$\rho_{\mathcal{U}}(x,t) = \mathrm{Tr}_{\mathcal{U}}(\hat{\rho}(t)_{\mathcal{U}} |x\rangle\langle x|_{\mathcal{U}}) = \langle |x\rangle\langle x|_{\mathcal{U}}\rangle_{\rho(t)} = \mathrm{Tr}_{\mathcal{S}}(\hat{\rho}_{\mathcal{S}}(t) |x\rangle\langle x|_{\mathcal{S}})$$

$$= \mathrm{Tr}_{\mathcal{S}}\left(\frac{1}{2}(|\psi_1\rangle\langle\psi_1|_{\mathcal{S}} + |\psi_2\rangle\langle\psi_2|_{\mathcal{S}}\right.$$
$$\left. + \langle E_2(t)|E_1(t)\rangle_{\mathcal{E}} |\psi_1\rangle\langle\psi_2|_{\mathcal{S}} + \langle E_1(t)|E_2(t)\rangle_{\mathcal{E}} |\psi_2\rangle\langle\psi_1|_{\mathcal{S}}) |x\rangle\langle x|_{\mathcal{S}}\right)$$

$$= \mathrm{Tr}_{\mathcal{S}}\left(\frac{1}{2}(\langle\psi_1|x\rangle_{\mathcal{S}} |\psi_1\rangle\langle x|_{\mathcal{S}} + \langle\psi_2|x\rangle_{\mathcal{S}} |\psi_2\rangle\langle x|_{\mathcal{S}}\right.$$
$$\left. + \langle E_2(t)|E_1(t)\rangle_{\mathcal{E}} \langle\psi_2|x\rangle_{\mathcal{S}} |\psi_1\rangle\langle x|_{\mathcal{S}} + \langle E_1(t)|E_2(t)\rangle_{\mathcal{E}} \langle\psi_1|x\rangle_{\mathcal{S}} |\psi_2\rangle\langle x|_{\mathcal{S}})\right)$$

$$= \frac{1}{2}\left(\langle\psi_1|x\rangle_{\mathcal{S}} \langle x|\psi_1\rangle_{\mathcal{S}} + \langle\psi_2|x\rangle_{\mathcal{S}} \langle x|\psi_2\rangle_{\mathcal{S}}\right.$$
$$\left. + \langle E_2(t)|E_1(t)\rangle_{\mathcal{E}} \langle\psi_2|x\rangle_{\mathcal{S}} \langle x|\psi_1\rangle_{\mathcal{S}} + \langle E_1(t)|E_2(t)\rangle_{\mathcal{E}} \langle\psi_1|x\rangle_{\mathcal{S}} \langle x|\psi_2\rangle_{\mathcal{S}}\right)$$

$$= \frac{1}{2}\left(\overline{\langle x|\psi_1\rangle_{\mathcal{S}}} \langle x|\psi_1\rangle_{\mathcal{S}} + \overline{\langle x|\psi_2\rangle_{\mathcal{S}}} \langle x|\psi_2\rangle_{\mathcal{S}}\right.$$
$$\left. + \langle E_2(t)|E_1(t)\rangle_{\mathcal{E}} \overline{\langle x|\psi_2\rangle_{\mathcal{S}}} \langle x|\psi_1\rangle_{\mathcal{S}} + \overline{\langle E_2(t)|E_1(t)\rangle_{\mathcal{E}}} \overline{\langle x|\psi_2\rangle_{\mathcal{S}}} \langle x|\psi_1\rangle_{\mathcal{S}}\right)$$

$$= \frac{1}{2}\left(|\psi_1(x)|^2 + |\psi_2(x)|^2 + 2\,\mathrm{Re}\left(\langle E_2(t)|E_1(t)\rangle_{\mathcal{E}} \overline{\psi_2(x)}\psi_1(x)\right)\right).$$

56. i.e. whether or not $\langle E_1(t)|E_2(t)\rangle_{\mathcal{E}} \to 0$ as $t \to \infty$ and how fast this convergence might take place.

von Neumann measurement scheme, it is supposed that for the measurement of a physical system $\mathcal{S}$ to take place, it must interact with a measuring device $\mathcal{A}$ which together satisfy the conditions 1. and 2. on page 82. If $\mathcal{S}$ is initially in a superposition of states $|\psi\rangle = \sum_i c_i |s_i\rangle$ then for $\mathcal{A}$ to measure $\mathcal{S}$, it is necessary for the combined system $\mathcal{S} + \mathcal{A}$ to enter into a superposition

$$|\psi\rangle |a_r(t_0)\rangle \xrightarrow{\text{time evolution}} \sum_i c_i |s_i\rangle |a_i(t)\rangle . \tag{2.7.1}$$

However, although the evolution described in (2.7.1) must take place if $\mathcal{A}$ is to measure $\mathcal{S}$, it is not sufficient. When one takes the partial trace of $|\psi\rangle\langle\psi|$ over $\mathcal{A}$, then according to (2.4.16),

$$\text{tr}_{\mathcal{A}}(|\psi\rangle\langle\psi|) \xrightarrow{\text{time evolution}} \sum_i |c_i|^2 |s_i\rangle\langle s_i| . \tag{2.7.2}$$

But as noted on page 81, we cannot give an ignorance interpretation to $\sum_i |c_i|^2 |s_i\rangle\langle s_i|$ for as d'Espagnat puts it, this is an improper mixture. When considered together, the system $\mathcal{S}$ and the apparatus $\mathcal{A}$ remain in the superposition described by (2.7.1), and so none of the measurement outcomes from the set of possible outcomes $\{ |s_i\rangle : i \}$ have actually occurred. The problem of explaining how the composite system $\mathcal{S} + \mathcal{A}$ goes from being in the state $\sum_i c_i |s_i\rangle |a_i(t)\rangle$ to a state $|s_i\rangle |a_i(t)\rangle$ is known as the *problem of outcomes*.

2.8 The Many-Worlds Interpretation

Not everyone is convinced that the problem of outcomes is a genuine problem. In particular, physicists such as Everett, De Witt, Deutsch, Zeh, and Zurek[57] who endorse the many-worlds interpretation of quantum physics effectively argue that there are no outcomes in the traditional sense. In this section, we give an account of the many-worlds interpretation of quantum physics and why physicists find it attractive. To this end, let us consider a physical universe $\mathcal{U} = \mathcal{S} + \mathcal{A} + \mathcal{P}_A + \mathcal{P}_B + \mathcal{E}$ consisting of subsystems $\mathcal{S}, \mathcal{A}, \mathcal{P}_A, \mathcal{P}_B$ and $\mathcal{E}$. $\mathcal{S}$ is the physical system under investigation; $\mathcal{A}$ is some measuring apparatus that interacts with $\mathcal{S}$; $\mathcal{P}_A$ and $\mathcal{P}_B$ are the physical systems corresponding to two scientists, Alice and Bob who observed the apparatus $\mathcal{A}$; and $\mathcal{E}$ is the remainder of the physical universe $\mathcal{U}$. For convenience, we define the composite subsystem $\mathcal{V} = \mathcal{S} + \mathcal{A} + \mathcal{P}_A + \mathcal{P}_B$ so that $\mathcal{U} = \mathcal{V} + \mathcal{E}$. As above on page 82, we assume that there is an orthonormal basis $\{ |s_i\rangle : i \}$ of $H_\mathcal{S}$ which we again refer to as pointer states, but now we assume that there are ready states $|a_r(t)\rangle \in H_\mathcal{A}$, $|p_{r,A}(t)\rangle \in H_{\mathcal{P}_A}$, $|p_{r,B}(t)\rangle \in H_{\mathcal{P}_B}$, and $|E_r(t)\rangle \in H_\mathcal{E}$ and that for each i, there are normalized states $|a_i(t)\rangle \in H_\mathcal{A}$, $|p_{i,A}(t)\rangle \in H_{\mathcal{P}_A}$, $|p_{i,B}(t)\rangle \in H_{\mathcal{P}_B}$, and $|E_i(t)\rangle \in H_\mathcal{E}$ such that

---

57. For a short history of the development of the many-worlds interpretation, see Schlosshauer, *Decoherence and the Quantum-to-Classical Transition*, 336–337. The central idea underlying the many-worlds was first described in Hugh Everett, ""Relative State" Formulation of Quantum Mechanics," *Reviews of modern physics* 29, no. 3 (1957): 454–462. Everett's idea and the idea of branching worlds was popularized in Bryce S. DeWitt, "Quantum mechanics and reality," *Physics today* 23, no. 9 (1970): 30–35 and D Deutsch, "Quantum theory as a universal physical theory," *International journal of theoretical physics* (New York, NY) 24, no. 1 (1985): 1–41. Zeh and Zurek used Everett's ideas in their development of decoherence theory, e.g. see H.D. Zeh, "There are no quantum jumps, nor are there particles," *Physics letters. A* (Amsterdam) 172, no. 4 (1993): 189–192 and Wojciech H. Zurek, "Decoherence, einselection and the existential interpretation (the rough guide)," *Philosophical*

1. for any $t \geq 0$ we have the evolution of the states

$$|s_i\rangle \, |a_r(t)\rangle \, |p_{r,A}(t)\rangle \, |p_{r,B}(t)\rangle \, |E_r(t)\rangle$$

$$\xrightarrow{\text{time evolution}} |s_i\rangle \, |a_i(t)\rangle \, |p_{i,A}(t)\rangle \, |p_{i,B}(t)\rangle \, |E_i(t)\rangle \, ,$$

2. there exists $\delta > 0$ such that if $t > t_0 + \delta$, then for $i \neq j$, $\langle a_i(t)|a_j(t)\rangle \approx 0$, $\langle p_{i,A}(t)|p_{j,A}(t)\rangle \approx 0$, $\langle p_{i,B}(t)|p_{j,B}(t)\rangle \approx 0$ and $\langle E_i(t)|E_j(t)\rangle \approx 0$.[58]

We also suppose that the $|p_{i,A}(t)\rangle$-state would describe actions of Alice consistent with her observing the apparatus being in the $|a_i(t)\rangle$-state, for example, her writing down in her log book that the apparatus is in the $|a_i(t)\rangle$-state or telling her colleague that this is the case. Likewise, we assume the $|p_{i,B}(t)\rangle$-state is consistent with Bob also observing the apparatus to be in the $|a_i(t)\rangle$-state.

Now suppose the initial (normalized) state of $\mathcal{S}$ is $|\psi\rangle = \sum_i c_i \, |s_i\rangle$, so that the state for the composite system $\mathcal{U} = \mathcal{V} + \mathcal{E}$ is $|\Psi(t)\rangle = \sum_i c_i \, |\xi_i(t)\rangle \, |E_i(t)\rangle \in H_{\mathcal{U}}$ where $|\xi_i(t)\rangle = |s_i\rangle \, |a_i(t)\rangle \, |p_{i,A}(t)\rangle \, |p_{i,B}(t)\rangle \in H_{\mathcal{V}}$. We also define the corresponding density matrix for the composite system $\hat{\rho} = |\Psi\rangle\langle\Psi|$. If we were unable to make any observations on $\mathcal{E}$, then the partial trace $\hat{\rho}_{\mathcal{V}}(t) = \text{Tr}_{\mathcal{E}}(\hat{\rho}(t))$ will contain all the information we need to work out the expectation values for any observables of $\mathcal{V}$. So

58. Again, recall footnote 31 on page 82.

just as with equation (2.4.16), we will have

$$
\begin{aligned}
\hat{\rho}_{\mathcal{V}}(t) &= \sum_i |c_i|^2 \, |\xi_i(t)\rangle\langle\xi_i(t)| + \sum_{i \neq j} c_i \overline{c_j} \, \langle E_j(t)|E_i(t)\rangle \, |\xi_i(t)\rangle\langle\xi_j(t)| \\
&\approx \sum_i |c_i|^2 \, |\xi_i(t)\rangle\langle\xi_i(t)|
\end{aligned}
\tag{2.8.1}
$$

for $t > t_0 + \delta$. Then the expectation values of any observables on $\mathcal{V}$ will be indistinguishable from the scenario in which $\mathcal{V}$ is actually in one of the $|\xi_i(t)\rangle$-states with probability $|c_i|^2$.[59] It would nevertheless be incorrect for us to conclude on the basis of decoherence theory alone that $\mathcal{V}$ actually was in one of those $|\xi_i(t)\rangle$-states, since equation (2.8.1) is based on a subjective distinction between $\mathcal{V}$ and $\mathcal{E}$ in the decomposition $\mathcal{U} = \mathcal{V} + \mathcal{E}$. Human scientists make this distinction to reflect the fact that they can only perform measurements on $\mathcal{V}$ and can't measure $\mathcal{E}$. But if a super-intelligent being could measure everything in $\mathcal{U}$, then such a being would not say that $\mathcal{V}$ was in one of the $|\xi_i(t)\rangle$-states, but rather that $\mathcal{U}$ was in the state $|\Psi(t)\rangle$. As we have already discussed on pages 80–81, the density matrix $\hat{\rho}_{\mathcal{V}}(t)$ is not a mixed state, but is an improper mixture.

Now if we define the observables $\hat{O}_{i,\mathcal{A}}(t) = |p_{i,A}(t)\rangle\langle p_{i,A}(t)|$ that would measure the behavior of Alice, and the observables $\hat{O}_{i,\mathcal{B}}(t) = |p_{i,B}(t)\rangle\langle p_{i,B}(t)|$ that would measure the behavior of Bob, then for $t > t_0 + \delta$, we see that $\hat{O}_{i,\mathcal{A}}(t)\hat{O}_{j,\mathcal{B}}(t)\,|\Psi(t)\rangle \approx 0$, when $i \neq j$. This means that when we consider $\hat{O}_{i,\mathcal{A}}(t)\hat{O}_{j,\mathcal{B}}(t)$ as an observable acting on $\mathcal{V}$, the expectation value $\langle\hat{O}_{i,\mathcal{A}}(t)\hat{O}_{j,\mathcal{B}}(t)\rangle_{\rho_{\mathcal{V}}(t)}$ will be approximately zero for $i \neq j$.

---

59. Again recall the discussion following equation (2.4.6) on page 75. There is the question of uniqueness of $\hat{\rho}_{\mathcal{V}} = \sum_i |c_i|^2 \, |\xi_i(t)\rangle\langle\xi_i(t)|$. If all the $|c_i|^2$ are unique, then if we have another

What this means is that if we consider ourselves as observing Alice and Bob observing the apparatus, then after time $t_0 + \delta$, the probability we would see Alice and Bob disagreeing with each other concerning their observations of the apparatus would be approximately 0. On the other hand, since $\hat{O}_{i,\mathcal{A}}(t)\hat{O}_{i,\mathcal{B}}(t)\,|\Psi(t)\rangle \approx c_i\,|\xi_i(t)\rangle\,|E_i(t)\rangle$ for $t > t_0 + \delta$, it follows that $\langle \hat{O}_{i,\mathcal{A}}(t)\hat{O}_{i,\mathcal{B}}(t)\rangle_{\rho_\mathcal{V}(t)} = |c_i|^2$. We would thus observe Alice and Bob observing the apparatus to be in the $|a_i(t)\rangle$-state with probability $|c_i|^2$.

But note that on the assumption that there are no hidden variables, if we did actually make such an observation and this observation corresponded to reality, then the quantum state $|\Psi(t)\rangle$ would have had to have changed to $|\xi_i(t)\rangle\,|E_i(t)\rangle$, since before our observation when $|\Psi(t)\rangle$ was a complete description of $\mathcal{U}$, we would say Alice and Bob will measure the $|a_i(t)\rangle$-state with probability $|c_i|^2$, but when we are actually seeing them measuring the $|a_i(t)\rangle$-state, we would have to say that now the probability they are measuring the $|a_i(t)\rangle$-state is 1, and hence we would say that the system was in the $|\xi_i(t)\rangle\,|E_i(t)\rangle$-state. Whether or not the process of the state going from $|\Psi(t)\rangle$ to $|\xi_i(t)\rangle\,|E_i(t)\rangle$ was instantaneous or took a non-infinitesimal amount of time, this interpretation would be susceptible to the problems already discussed with the Copenhagen interpretation in section 1.2.

But in the *many-worlds interpretation*, rather than assuming that $|\Psi(t)\rangle = \sum_i c_i\,|\xi_i(t)\rangle\,|E_i(t)\rangle$ is the complete description of $\mathcal{U}$ that enables us to work the probability of certain outcomes, we simply say that $|\Psi(t)\rangle$ is a complete description

decomposition $\hat{\rho}_{\mathcal{S}+\mathcal{A}+\mathcal{P}_A+\mathcal{P}_B} = \sum_i |c_i|^2\,|\xi_i'(t)\rangle\langle\xi_i'(t)|$ it follows that $|\xi_i(t)\rangle \propto |\xi_i'(t)\rangle$. But even if some of the $|c_i|^2$ are the same, criteria 1 and 2 above will ensure that states with the same value of $|c_i|^2$ will be determined up to permutation.

of the state of $\mathcal{U}$, and we drop the assumption that we need to interpret this state as describing probabilities of outcomes. Thus, a many-worlds adherent would say we can understand what the state of $\mathcal{U}$ is on its own terms without the need to appeal to any other extrinsic principle such as measurement. Just as we don't puzzle over how to interpret what a sphere is in terms of an extrinsic principle, we don't need to puzzle over how to interpret the space of states of $\mathcal{U}$. We can think of the mathematical formalism $|\Psi(t)\rangle = \sum_i c_i |\xi_i(t)\rangle |E_i(t)\rangle$ describing the state of $\mathcal{U}$ as being somewhat akin to a point lying on a sphere given by the equation $x^2 + y^2 + z^2 = 1$.[60] Although we might be tempted to interpret $|\Psi(t)\rangle$ as describing the probability of outcomes, we are not obliged to do so, since these probabilities can instead be understood to be grounded in the symmetries the system possesses rather than in terms of the frequency of how many measurement outcomes are likely to occur. For instance, when we see a coin and judge that it will come up heads with probability $\frac{1}{2}$ and tails with probability $\frac{1}{2}$, we intuit this by looking at the symmetry of the coin rather than tossing the coin millions of times and counting how often it comes up heads and how often it comes up tails. Thus, we might suppose $|\Psi(t)\rangle$ has analogous symmetries that allow us to determine its c_i coefficients without the need to posit any of the $|\xi_i(t)\rangle |E_i(t)\rangle$ measurement outcomes being realized.

As for the decomposition $|\Psi(t)\rangle = \sum_i c_i |\xi_i(t)\rangle |E_i(t)\rangle$ in terms of the $|\xi_i(t)\rangle |E_i(t)\rangle$ basis states, decoherence theory gives us a natural account of why we should choose this

---

60. On the assumption that $|\Psi(t)\rangle$ is normalized, we could think of $|\Psi(t)\rangle$ as specifying a point on the (possibly infinite-dimensional) hypersphere $\{(c_1, c_2 \ldots) : \sum_i |c_i|^2 = 1\}$.

basis rather than any other. When $\mathcal{U}$ is in the state $|\Psi(t_0)\rangle = \left(\sum_i c_i |\xi_i(t_0)\rangle \right) |E_r(t)\rangle$,
we can think of this state as describing one world, W say. But once $t > t_0 + \delta$ so
that $\langle E_i(t)|E_j(t)\rangle \approx 0$ for $i \neq j$, we can think of each $|\xi_i(t)\rangle \, |E_i(t)\rangle$-component as a
different world W_i. Thus, for $t > t_0 + \delta$, we say the world W has *branched* into as
many-worlds W_i for which the c_i are non-zero.

But why should we think that there are literally many worlds? Well, from an
ontological point of view, one might very well think that there is really only one world
and that this world is described by $|\Psi(t)\rangle$; it would be a rather weird world since the
entanglement between $\mathcal{V}$ and $\mathcal{E}$ would mean there wouldn't be any absolute matters
of fact describing $\mathcal{V}$. But it might not be a bad thing to say that the "many" in the
many-worlds interpretation is really just a figure of speech that we shouldn't take too
literally. After all, a common objection to the many-worlds interpretation is that it is
ontologically extravagant and that we should appeal to Occam's Razor. But if we just
say that there is actually only one world described by $|\Psi(t)\rangle$ then this "many"-worlds
interpretation is actually rather parsimonious from an ontological point of view.

But if by literal, we mean descriptive rather than ontological, it does seem rather
natural to say that there are literally many worlds. For even though we might
initially suspect that the worlds W_i and W_j are not clearly delineated given the fact
that $\langle E_i(t)|E_j(t)\rangle$ is very small but not zero for $i \neq j$, we can nevertheless expect
$\langle \xi_i(t)|\xi_j(t)\rangle$ to be identically zero for $i \neq j$, just as we can expect $\langle \hat{\mathbf{a}}+|\hat{\mathbf{a}}-\rangle$ to be

identically zero.[61] Thus, if we define $|W_i(t)\rangle = |\xi_i(t)\rangle \, |E_i(t)\rangle$, then the $\langle W_i(t)|W_j(t)\rangle$ will be identically zero for $i \neq j$, and so we would in fact be able to clearly delineate these worlds.

Still, the supposition that $|\Psi(t)\rangle = \sum_i c_i \, |W_i(t)\rangle$ with $\langle W_i|W_j\rangle = 0$ for $i \neq j$ is not of itself sufficient justification for describing the state $|\Psi(t)\rangle$ as a composition of mutually exclusive world descriptions given by the $|W_i(t)\rangle$. After all, the fact that

$$
\begin{aligned}
|\text{Cat Alive}\rangle = &\frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{2}} (|\text{Cat Alive}\rangle + |\text{Cat Dead}\rangle) \right) \\
&+ \frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{2}} (|\text{Cat Alive}\rangle - |\text{Cat Dead}\rangle) \right)
\end{aligned}
$$

does not incline us to think of the state $|\text{Cat Alive}\rangle$ as being composed of the mutually exclusive cat states $\frac{1}{\sqrt{2}}(|\text{Cat Alive}\rangle + |\text{Cat Dead}\rangle)$ and $\frac{1}{\sqrt{2}}(|\text{Cat Alive}\rangle - |\text{Cat Dead}\rangle)$.

But the key justification for describing the state $|\Psi(t)\rangle$ as a composition of the mutually exclusive $|W_i(t)\rangle$-states is the fact that the states $|\xi_i(t)\rangle$ and $|\xi_j(t)\rangle$ decohere for $i \neq j$, that is, the off-diagonal entries $|\xi_i(t)\rangle\langle\xi_j(t)|$ of the reduced density matrix $\hat{\rho}_\mathcal{V}(t)$ will tend to zero, and as we saw in section 2.6, it will follow that quantum interference effects between $|\xi_i(t)\rangle$ and $|\xi_j(t)\rangle$ will then tend to zero. Thus, when it comes to observables defined on $\mathcal{V}$, using equations (2.4.14) and (2.8.1), we can calculate the expectation value of an observable $\hat{O}_\mathcal{V}$ as a weighted sum of expectation

61. It is also reasonable to suppose that in situations such as the double-slit experiment described on page 90 that $\langle \psi_1(t)|\psi_2(t)\rangle$ is identically zero. This is because $\langle \psi_1(t)|\psi_2(t)\rangle = 0$ when t is the time at which the particle is going through the slit, and this will remain zero because of a property of the time evolution operator known as unitarity.

values for each of the states $|\xi_i(t)\rangle$:

$$\langle \hat{O}_\mathcal{U} \rangle_{\Psi(t)} \approx \sum_i |c_i|^2 \langle \hat{O}_\mathcal{V} \rangle_{\xi_i(t)} \tag{2.8.2}$$

where $\hat{O}_\mathcal{U}$ is the natural extension of $\hat{O}_\mathcal{V}$ to $H_\mathcal{U}$ analogous to the extension given in equation (2.4.2) with $\mathcal{S}$ replaced by $\mathcal{V}$. The fact that (2.8.2) is only an approximation suggests that the time at which branching occurs is not well-defined. All that we can do is choose a sufficiently large time interval δ such that for $t > t_0 + \delta$, the approximation (2.8.2) meets our desired level of accuracy.

Despite this vagueness on when branching occurs, we can still form a natural and well-defined notion of worlds according to the following definition: a set $\{W_i : i\}$ is the set of worlds for a universe $\mathcal{U} = \mathcal{V} + \mathcal{E}$ when

1. W_i is a description of $\mathcal{U}$ given by $|W_i(t)\rangle = |\xi_i(t)\rangle |E_i(t)\rangle$ where $|\xi_i(t)\rangle$ is a state of $\mathcal{V}$ and $|E_i(t)\rangle$ is a state of $\mathcal{E}$,

2. $\mathcal{U}$ is in the state $|\Psi(t)\rangle = \sum_i c_i |W_i(t)\rangle$ with all $c_i \neq 0$.

3. $\langle \xi_i(t)|\xi_j(t)\rangle = 0$ for $i \neq j$,

4. $\langle E_i(t)|E_j(t)\rangle \to 0$ as $t \to \infty$ for $i \neq j$, and the convergence is such that for any observable $\hat{O}_\mathcal{V}$ defined on $\mathcal{V}$, $\langle \hat{O}_\mathcal{U} \rangle_{\Psi(t)} \to \sum_i |c_i|^2 \langle \hat{O}_\mathcal{V} \rangle_{\xi_i(t)}$.

The partition of the universe $\mathcal{U}$ into a system $\mathcal{V}$ and environment $\mathcal{E}$ is somewhat arbitrary, but in many situations, decoherence theory suggest that the convergence $\langle E_i(t)|E_j(t)\rangle \to 0$ for $i \neq j$ is extremely rapid. For example, we could take a single neuron in the brain to be the system $\mathcal{V}$, and the rest of the universe to be the environment

104

$\mathcal{E}$. Corresponding to whether or not the neuron is firing, we would expect there to be environmental states $|E_{\mathcal{V} \text{ firing}}(t)\rangle$ and $|E_{\mathcal{V} \text{ not firing}}(t)\rangle$ respectively. Tegmark models this situation and estimates that within 10^{-19} seconds, $\langle E_{\mathcal{V} \text{ firing}}(t)|E_{\mathcal{V} \text{ not firing}}(t)\rangle$ will be negligible.[62]

Note that according to the above definition of a set of worlds, the description $|\Psi(t)\rangle$ is rather trivially a world – we just take the environment $\mathcal{E}$ to be empty so that there would be only one $|E_i(t)\rangle$ which would be the vacuum state. So there is at least one world according to this definition. There is a question of whether there could be more than one world, and this would depend on whether we could really have a non-trivial decomposition $\mathcal{U} = \mathcal{V} + \mathcal{E}$, for the supposition that there is such a decomposition requires that it is possible to distinguish $\mathcal{V}$ and $\mathcal{E}$. But this might not in fact be possible. For instance, if the ultimate fate of the universe was that it would collapse into a singularity, then there would come a point at which it wouldn't be possible to make a distinction between $\mathcal{V}$ and $\mathcal{E}$. But despite this possible concern, the above definition makes it seem plausible that there could be many well-defined worlds W_i.[63]

When we look at a particular $|\xi_i(t)\rangle$, it will look like it is describing a fairly classical world with scientists performing their measurements and agreeing about what

---

62. See M Tegmark, "Importance of quantum decoherence in brain processes," *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics* (United States) 61, no. 4 Pt B (2000): 4194–4206. For a summary of Tegmark's paper, see Schlosshauer, *Decoherence and the Quantum-to-Classical Transition*, 368–371.

63. For the purposes of this chapter, plausibility is enough. I am only trying to show why physicists might find the many-worlds interpretation of quantum physics attractive. I am certainly not trying

they measure. And as long as the $|\xi_i(t)\rangle$-states remain pointer states with respect to $|E_i(t)\rangle$, no branching will occur. But typically, a $|\xi_i(t)\rangle$-state will not indefinitely remain a pointer state with respect to $|E_i(t)\rangle$. We can think of how this happens with the Stern-Gerlach experiment. For if one Stern-Gerlach apparatus has its magnetic field oriented in the $\hat{\mathbf{a}}$-direction, then $|\hat{\mathbf{a}}+\rangle$ and $|\hat{\mathbf{a}}-\rangle$ will be pointer states for a silver atom in the vicinity of this apparatus. But if the same silver atom then travels onward to another Stern-Gerlach apparatus with its magnetic field now oriented in the $\hat{\mathbf{b}}$-direction with $\hat{\mathbf{b}} \neq \hat{\mathbf{a}}$, $|\hat{\mathbf{a}}+\rangle$ and $|\hat{\mathbf{a}}-\rangle$ will no longer be pointer states with respect to their environment, and so branching will occur. But this is not necessarily a problem for the definition of many-worlds given above on page 104, for when a $|\xi_i(t)\rangle$-state does not indefinitely remain a pointer state with respect to $|E_i(t)\rangle$, we can just rewrite $|\xi_i(t)\rangle$ as a sum of pointer states $|\xi_{ij}(t)\rangle$ and $|E_i(t)\rangle$ as a sum of their respective environments $|E_{ij}(t)\rangle$, and then $|E_i(t)\rangle$ will be like a ready state for the $|E_{ij}(t)\rangle$. Assuming we can do this so that the $|\xi_{ij}(t)\rangle$ are orthogonal to the $|\xi_{i'j'}(t)\rangle$ when $i' \neq i$ or $j' \neq j$, then we would still be able to have well-defined worlds according to the definition given above.

2.9 The Many-Worlds Solution to the EPR-Bohm Paradox

We are now in a position to consider how proponents of the many-worlds interpretation attempt to resolve the EPR-Bohm paradox. We thus suppose there is a spin-singlet as described in section 1.2 consisting of two particles q_A and q_B that are

---

to show that the many-worlds interpretation is the most convincing and satisfactory interpretation of quantum physics.

in the entangled Bell state

$$|\Psi_{\text{Bell}}\rangle = \frac{1}{\sqrt{2}}(\,|\hat{\mathbf{a}}+\rangle_A\,|\hat{\mathbf{a}}-\rangle_B - |\hat{\mathbf{a}}-\rangle_A\,|\hat{\mathbf{a}}+\rangle_B). \qquad (1.2.6\text{ revisited})$$

If we assume that the two experimenters Alice and Bob (themselves constituting physical systems $\mathcal{P}_A$ and $\mathcal{P}_B$ respectively) set their Stern-Gerlach apparatuses $\mathcal{A}_A$ and $\mathcal{A}_B$ to measure their respective particles along the same axis, then by (1.2.9), we can assume they both perform their measurements along the $\hat{\mathbf{a}}$-axis. This means that $|\hat{\mathbf{a}}+\rangle_A\,|\hat{\mathbf{a}}-\rangle_B$ and $|\hat{\mathbf{a}}-\rangle_A\,|\hat{\mathbf{a}}+\rangle_B$ will be pointer states of the composite system $\mathcal{A}_A + \mathcal{A}_B + \mathcal{P}_A + \mathcal{P}_B$. There will thus be ready states $|a_{r,A}(t)\rangle \in H_{\mathcal{A}_A}$, $|a_{r,B}(t)\rangle \in H_{\mathcal{A}_B}$, $|p_{r,A}(t)\rangle \in H_{\mathcal{P}_A}$ and $|p_{r,B}(t)\rangle \in H_{\mathcal{P}_B}$, and normalized states $|a_{\pm,A}(t)\rangle \in H_{\mathcal{A}_A}$, $|a_{\pm,B}(t)\rangle \in H_{\mathcal{A}_B}$, $|p_{\pm,A}(t)\rangle \in H_{\mathcal{P}_A}$ and $|p_{\pm,B}(t)\rangle \in H_{\mathcal{P}_B}$ such that

$$|\hat{\mathbf{a}}+\rangle_A\,|\hat{\mathbf{a}}-\rangle_B\,|a_{r,A}(t)\rangle\,|a_{r,B}(t)\rangle\,|p_{r,A}(t)\rangle\,|p_{r,B}(t)\rangle$$

$$\xrightarrow{\text{time evolution}}\,|\hat{\mathbf{a}}+\rangle_A\,|\hat{\mathbf{a}}-\rangle_B\,|a_{+,A}(t)\rangle\,|a_{-,B}(t)\rangle\,|p_{+,A}(t)\rangle\,|p_{-,B}(t)\rangle\,,$$

and similarly

$$|\hat{\mathbf{a}}-\rangle_A\,|\hat{\mathbf{a}}+\rangle_B\,|a_{r,A}(t)\rangle\,|a_{r,B}(t)\rangle\,|p_{r,A}(t)\rangle\,|p_{r,B}(t)\rangle$$

$$\xrightarrow{\text{time evolution}}\,|\hat{\mathbf{a}}-\rangle_A\,|\hat{\mathbf{a}}+\rangle_B\,|a_{-,A}(t)\rangle\,|a_{+,B}(t)\rangle\,|p_{-,A}(t)\rangle\,|p_{+,B}(t)\rangle\,.$$

It will therefore follow that

$$|\Psi_{\text{Bell}}\rangle \, |a_{r,A}(t)\rangle \, |a_{r,B}(t)\rangle \, |p_{r,A}(t)\rangle \, |p_{r,B}(t)\rangle$$

$$\xrightarrow{\text{time evolution}} \frac{1}{\sqrt{2}} \, |\hat{\mathbf{a}}+\rangle_A \, |\hat{\mathbf{a}}-\rangle_B \, |a_{+,A}(t)\rangle \, |a_{-,B}(t)\rangle \, |p_{+,A}(t)\rangle \, |p_{-,B}(t)\rangle \qquad (2.9.1)$$

$$+ \frac{1}{\sqrt{2}} \, |\hat{\mathbf{a}}-\rangle_A \, |\hat{\mathbf{a}}+\rangle_B \, |a_{-,A}(t)\rangle \, |a_{+,B}(t)\rangle \, |p_{-,A}(t)\rangle \, |p_{+,B}(t)\rangle \,.$$

Thus, in the language of the many-worlds interpretation, the first summand of (2.9.1) corresponds to a world in which Alice observes her particle to be spin up, and Bob observes his particle to be spin down, and the second summand of (2.9.1) corresponds to a world in which Alice observes her particle to be spin down, and Bob observes his particle to be spin up. So in each world, Alice and Bob obtain opposite results. But if on the other hand, Bob chooses to make his measurement along a different axis, then $|\hat{\mathbf{a}}+\rangle_B$ and $|\hat{\mathbf{a}}-\rangle_B$ won't be pointer states for the composite system $\mathcal{A}_B + \mathcal{P}_B$, and so if $|a'_{r,B}(t)\rangle \, |p'_{r,B}(t)\rangle$ is the ready state for Bob's measurement choice, we must assume that

$$|\hat{\mathbf{a}}\pm\rangle_B \, |a'_{r,B}(t)\rangle \, |p'_{r,B}(t)\rangle \xrightarrow{\text{time evolution}} |E_{\pm,B}(t)\rangle$$

for some entangled state $|E_{\pm,B}(t)\rangle$ of the composite system $q_B + \mathcal{A}_B + \mathcal{P}_B$ with $\langle E_{\pm,B}(t)|E_{\pm,B}(t)\rangle = 1$ and $\langle E_{\pm,B}(t)|E_{\mp,B}(t)\rangle = 0$. It will then follow that

$$|\Psi_{\text{Bell}}\rangle \, |a_{r,A}(t)\rangle \, |a'_{r,B}(t)\rangle \, |p_{r,A}(t)\rangle \, |p'_{r,B}(t)\rangle$$

$$\xrightarrow{\text{time evolution}} \frac{1}{\sqrt{2}} \, |\hat{\mathbf{a}}+\rangle_A \, |a_{+,A}(t)\rangle \, |p_{+,A}(t)\rangle \, |E_{-,B}(t)\rangle \qquad (2.9.2)$$

$$+ \frac{1}{\sqrt{2}} \, |\hat{\mathbf{a}}-\rangle_A \, |a_{-,A}(t)\rangle \, |p_{-,A}(t)\rangle \, |E_{+,B}(t)\rangle \, .$$

Thus, treating $q_B + \mathcal{A}_B + \mathcal{P}_B$ as though it were an environment of $q_A + \mathcal{A}_A + \mathcal{P}_A$, then according to the definition of a world on page 104, the first summand of (2.9.2) will correspond to a world in which Alice observes her particle to be spin up, and the second summand of (2.9.2) will correspond to a world in which Alice observes her particle to be spin down, and this will be the case regardless of what axis Bob chooses to make his measurement along. Moreover, since there is no state collapse when Bob makes his measurement, and since Bob's choice has no effect on the state

$$|\xi_{\pm,A}(t)\rangle = |\hat{\mathbf{a}}\pm\rangle_A \, |a_{\pm,A}(t)\rangle \, |p_{\pm,A}(t)\rangle$$

describing Alice's observation, the many-worlds interpretation gives us no reason to worry about there being a violation of special relativity. So that is how proponents of the many-worlds interpretation attempt to resolve the EPR-Bohm paradox.

2.10 Evaluating the Many-Worlds Interpretation

Given the account in the previous two sections of the many-worlds interpretation of quantum physics and how it attempts to resolve the EPR-Bohm paradox, it is understandable why physicists would find it so attractive. Although we can't specify an exact moment at which branching occurs, the idea of branching and of there being many worlds itself is not particularly mysterious. This can all be explained in terms of the dynamics of the system and the environment, and decoherence theory allows us to understand why the interference effects that are the hallmark of quantum physics generally disappear on the macroscopic level.

There are other advantages of the many-worlds interpretation besides these which we need not discuss here.[64] But for all the advantages of the many-worlds hypothesis, there is one fundamental problem, and that is its patent absurdity. It seems that we should be able to say whether a cat is alive or dead without having to say what state the rest of the universe is in. However, the many-worlds interpretation suggests that for any subsystem of the universe, we will in general only be able to say what state it is in with respect to the state of the rest of the universe. For example, if the state $\mathcal{S}$ is the system constituting a cat-wise configuration of particles and $\mathcal{E}$ is the rest of the

---

64. More details can be found in Schlosshauer, *Decoherence and the Quantum-to-Classical Transition* and E. Joos et al., *Decoherence and the Appearance of a Classical World in Quantum Theory* (Springer Berlin Heidelberg, 2013).

universe, then given that the composite system $\mathcal{U} = \mathcal{S} + \mathcal{E}$ is described by the state

$$|\Psi(t)\rangle = \frac{1}{\sqrt{2}}\big(\,|\text{Cat Alive}\rangle_{\mathcal{S}}\,|E_{\text{Cat Alive}}\rangle_{\mathcal{E}} + |\text{Cat Dead}\rangle_{\mathcal{S}}\,|E_{\text{Cat Dead}}\rangle_{\mathcal{E}}\,\big),$$

then we are in no position to make an absolute matter of fact claim about the system $\mathcal{S}$ and say the cat is dead or the cat is alive. Rather we have to say with respect to the environment described by $|E_{\text{Cat Alive}}\rangle_{\mathcal{E}}$, the cat is alive, and with respect to the environment $|E_{\text{Cat Dead}}\rangle_{\mathcal{E}}$, the cat is dead. Moreover, according to the many-worlds hypothesis, the branching into multiple worlds doesn't just occur in rare instances, such as in Schrödinger's cat-type experiments. On the contrary, branching is supposed to be happening all the time.

In order to convey how ubiquitous branching is in the many-worlds interpretation, one just has to consider the behavior of an electron. If we suppose that a free electron is initially described by a wave packet whose width is around 10^{-10} m (which is of the order of the width of and atom), then according to the Schrödinger equation which governs how wave packets evolve over time, after one second the width of the wave packet will have spread to a width of around 1000 km.[65] The only reason electrons remain localized rather than spreading out to such vast distances is because the electron is continually interacting with its environment. But according to the many-worlds interpretation, these continual interactions of the electron with its environment will result in a continual branching out of worlds corresponding to the possible locations

---

65. See Schlosshauer, *Decoherence and the Quantum-to-Classical Transition*, 117.

the environment localizes the particle to. Therefore, because the electron rapidly gets entangled with its environment, we cannot establish matter of fact claims about where the electron is localized to – we can only establish matter of fact claims about the composite system of the electron and its environment which expands with astonishing rapidity.

Because of the ubiquity of branching in the many-worlds interpretation, this interpretation appears to undermine the conditions for the possibility of doing science, for surely one of the conditions for the possibility of doing science is that measuring devices exist, but it doesn't look like there are such things as measuring devices in the many-worlds interpretation. To see why, consider the properties a measuring device should possess. Firstly, it must be capable of interacting either directly or indirectly[66] with another entity which is the thing to be measured. Secondly, there must be some kind of correspondence between the range of states the measuring device can be in and the range of states the thing being measured can be in. Thirdly, when the measuring device interacts with the thing being measured, the measuring device should enter into the state that corresponds to the state of the thing being measured.[67] But according to the many-worlds interpretation, what is taken to be a measuring device will in general become entangled with the thing that is being measured, and so

66. It is acceptable for a measuring device to interact with an environment that has interacted with the thing that is being measured.

67. Although the act of measuring may change the state of the thing being measured, a measuring device should still be able to tell us what the state of the thing being measured is in immediately after the measuring device has interacted with it.

there will be no fact of the matter regarding what state the measuring device is in. Rather there will at the very most only be a fact of the matter regarding the state of the composite system that includes both the measuring device and the thing being measured.[68] Thus, in the many-worlds interpretation, there are no measuring devices satisfying the criteria one would expect a measuring device to satisfy. And so without such measuring devices, it does not appear to be possible to do science according to what we normally mean by science.

In the many-worlds interpretation, it is also very difficult to make any sense of the Hilbert space formalism on which the many-worlds interpretation depends. Using the Born Rule to interpret the Hilbert space of states is not an option since the Born Rule is a rule that assigns probabilities to outcome states given some initial state, and hence the Born Rule presupposes that there are outcomes, but the many-worlds interpretation denies this presupposition.[69] And even if we can give some meaning to the states of the Hilbert space for the universe, there will never be the kinds of states for subsystems of the universe that we would expect there to be. For instance, there would never be states such as $|\text{living cat}\rangle$ describing a physical system constituting a

---

68. In fact, it is questionable whether we can even make a matter of fact claim about there being a measuring device at all – rather, in the many-worlds interpretation, we can only say there is a superposition of states in which there is a measuring device in existence with respect to some environments, and of states in which there is not a measuring device with respect to other environments.

69. Albert and Loewer discuss the difficulty of making sense of Born Rule probabilities in the many-worlds interpretation, and so they propose a many-minds interpretation as a way of addressing this problem – see David Albert and Barry Loewer, "Interpreting the Many Worlds Interpretation," *Synthese (Dordrecht)* (Dordrecht) 77, no. 2 (1988): 195–213. However, Pruss argues that the many-minds interpretation leads to the problem of skepticism, and it also leads to some ethical problems – see Alexander R. Pruss, "A Traveling Forms Interpretation of Quantum Mechanics," in *Neo-Aristotelian Perspectives on Contemporary Science*, 1st ed. (Routledge, 2018), 106–109.

cat. Rather, there would only be states of the form

$$\alpha \left|\text{living cat}\right\rangle \left|\chi\right\rangle + \beta \left|\text{not a living cat}\right\rangle \left|\chi'\right\rangle \qquad (2.10.1)$$

for non-zero α and β, and (2.10.1) is not the state of a cat. Thus, the many-worlds interpretation implies that no cats exist.

Another reason for rejecting the many-worlds interpretation is that intuitively, it seems obvious that I can know I am alive without needing to know the state of the rest of the universe, but the many-worlds interpretation does not allow me to make this absolute matter of fact claim. So from both a common sense point of view and a scientific point of view, the many-worlds interpretation really is absurd.

Of course some hypotheses may initially seem absurd, but once the hypothesis has been fully explained, it can appear far more plausible. For instance, time dilation in special relativity might initially sound absurd to some people, but once one has a better grasp of special relativity and is open to the possibility that systems moving close to the speed of light with respect to ourselves might have properties rather different to systems that move with much slower speeds, then special relativity doesn't seem absurd at all. However, the many-worlds interpretation as presented here is different in this regard since it is not hypothesizing about some extreme situation. It is hypothesizing about ordinary situations. In order to accept the many-worlds interpretation, the arguments in its favor would have to be at least as convincing as the common sense beliefs it is calling into question such as the belief that we

can do science and my own personal belief that I am alive. But arguments for the many-worlds interpretation clearly fail to meet this criterion. Some people may choose to embrace the absurdity of the many-worlds interpretation and reject the most basic notions of common sense. But when a hypothesis entails an absurd conclusion, a reasonable person would surely think it better to reject the hypothesis rather than embrace the absurdity.

But in rejecting a hypothesis because of its absurd consequences, it doesn't mean that absolutely everything in the hypothesis needs to be rejected, for a hypothesis might be formulated in terms of sub-hypotheses, some of which might be very plausible and which don't of themselves entail absurdities, in which case something of the original hypothesis might be salvageable. In the case of the many-worlds hypothesis, I believe it does have something that is salvageable, namely decoherence theory. In the next chapter I will consider Adrian Kent's one-world interpretation of quantum physics in which the basic ideas of decoherence theory remain intact.

2.11 Summary

In this chapter I have given an account of the many-worlds interpretation and discussed why anyone would find this interpretation attractive. In order to provide this account, it was necessary to explain some of the mathematical formalism of standard quantum theory. In this formalism, states of a system are thought of as belonging to a space of states called a Hilbert space. A Hilbert space has an inner product that

enables us to calculate how likely one state is to be found in another state via the Born Rule.

By associating experiments with operators (which are referred to as observables) that act on the Hilbert space, we can calculate the average outcome of an experiment if the experiment is performed many times.

However, this formalism raises some problems. For the space of observables in the mathematical formalism of quantum theory is much larger than the kinds of measurements we can perform. For instance, we can perform measurements by which we can distinguish between a cat being alive and a cat being dead, but we can't perform measurements that distinguish between different superposition states of cats. This is the preferred basis problem. Also, there is the question of why very small objects exhibit interference, whereas larger objects like cats don't exhibit interference. This is the problem of the nonobservability of interference.

Decoherence theory enables us to answer these two problems by making a distinction between a system and its environment. When this distinction is made, we can work out how the system's interaction with its environment affects the observations made on the system. By averaging over the unknown states of the environment as the system interacts with it, we can form a generalized notion of a state for the system via an operation called the partial trace. This enables us to obtain a reduced density matrix describing the system. The reduced density matrix contains all the information necessary to calculate expectation values of observables. In many situations, via a process called decoherence, the reduced density matrix essentially becomes diagonal,

so that it is as though the system is in an unknown state corresponding to one of the diagonal entries of the reduced density matrix. The criterion for decoherence to occur enables us to solve the preferred basis problem and the problem of the nonobservability of interference.

However, decoherence theory doesn't enable us to solve what is known as the problem of outcomes. For decoherence theory relies on a subjective distinction between a system and its environment, but when one considers the composite of the system and its environment together, the composite system is in a superposition of states rather than a definite state. Many-worlds adherents conclude that the different states of this superposition are just as real as each other.

Despite some attractive features of the many-worlds interpretation, it calls for such a radical departure from common sense that it undermines our most basic belief that there are objective facts about reality. The proposed solution to the EPR-Bohm paradox offered by the many-world interpretation is therefore deeply unsatisfactory. Thus, we have good reason to look elsewhere in our search for a solution to the EPR-Bohm paradox. We will therefore continue our search in the next chapter by examining Kent's interpretation of quantum physics.

A description of Kent's Theory of Quantum Physics

In this chapter, I will describe Kent's theory of quantum physics, but before doing this, it is worth briefly reminding ourselves of the problem in quantum physics that we wish to address.

In chapter 1, we discussed the EPR-Bohm paradox and the problem of trying to account for the mysterious correlation of spin measurements on two spatially separated particles. We saw that the Copenhagen interpretation of quantum physics is unable to satisfactorily resolve this paradox because it posits that there is an instantaneous collapse of the state upon measurement, but the idea of an instantaneous collapse does not make sense in special relativity where there is no such thing as an instant of time.[1]

We also saw that any hidden variables theory that makes the same predictions as quantum mechanics when averaged over the hidden variables (and hence violates

---

1. Since instantaneous spatially extended collapse makes no sense in special relativity, some philosophers have entertained the possibility of there being some additional spacetime structure so that instantaneous collapse does make sense. This involves defining a foliation, that is a partition of spacetime into a continuous series of three-dimensional hypersurfaces, where each hypersurface defines what it means to be an instant. The state of the universe at any one instant is a state of one of these hypersurfaces. Sometimes the state will change from one hypersurface to another via Schrödinger evolution, but at other times, when there are state collapses, there will be a discontinuity in the transition between the hypersurfaces. Maudlin discusses some of the issues associated with adding a foliated structure to spacetime – see T Maudlin, "Space-time in the quantum world," in *Bohmian mechanics and quantum theory : an appraisal [electronic resource]*, by James T. Cushing, Arthur Fine, and Sheldon Goldstein, Boston studies in the philosophy of science ; Volume 184 (Dordrecht, 1996), 285–307.

Bell's inequality) cannot satisfy both parameter independence (PI) and outcome independence (OI). Shimony proposed that quantum theory and special relativity could peacefully coexist if we accepted PI and rejected OI, but as Butterfield[2] points out, Shimony's proposal does not address the problem of what an outcome is despite his proposal relying on there being outcomes.

As discussed in chapter 2, the problem of outcomes remains an unresolved part of the measurement problem, and the many-worlds interpretation that attempts to sidestep the problem of outcomes is deeply unsatisfactory. But as well as critiquing Shimony's proposal, Butterfield thinks that a suitable interpretation of quantum physics could provide what is missing in Shimony's account. It is for this reason that Butterfield highlights Kent's theory of quantum physics.

In this chapter, we will just focus on describing Kent's theory, and we will postpone our evaluation of whether Kent's theory can adequately resolve the EPR-Bohm paradox until chapter 4. In describing Kent's theory of quantum physics, we will focus on the ideas Kent presents in his 2014 paper.[3]

Kent's theory of quantum physics has some similarities in common with the Bohmian interpretation. Firstly, there is no quantum state collapse in Kent's theory. Secondly, some additional variables beyond standard quantum theory (i.e. in addition

---

2. See Butterfield, "Peaceful Coexistence: Examining Kent's Relativistic Solution to the Quantum Measurement Problem"

3. Adrian Kent, "Lorentzian Quantum Reality: Postulates and Toy Models," 2014, https://doi.org/10.1098/rsta.2014.0241, eprint: arXiv:1411.2957.

to the quantum state) are included in Kent's theory. And thirdly, Kent's theory is a one-world interpretation of quantum physics. I'll consider these three features of Kent's theory in some detail as I describe his theory. I'll also explain how to perform the expectation calculations that are central to Kent's theory. And finally, I'll present an account of Kent's toy model that provides a simple example of how the ideas of his theory fit together.

I don't claim any significant degree of originality in this chapter, but I do provide some details regarding my own understanding of Kent's interpretation. For example, in section 3.2.2, I discuss the additional variables of Kent's interpretation in far more detail that Kent does, and in section 3.4, I state explicitly the equations (3.4.4), (3.4.5), and (3.4.7) that are used in the calculation of Kent's beables.

3.1 The No-collapse Feature of Kent's Theory

We first consider the no-collapse feature of Kent's theory. This is a feature that belongs both to the many-worlds interpretation and to the Bohmian interpretation. In all three interpretations, the quantum state deterministically evolves according to the Schrödinger equation. The Schrödinger equation itself describes how a quantum state evolves over time when there are no outside influences. The precise formula for the Schrödinger equation need not concern us here, but all we need to know is that the Schrödinger equation determines a so-called *unitary operator* $U(t', t)$. What this means is that if a system is in a state $|\psi\rangle$ at time t, then it will be in the state $|\psi'\rangle = U(t', t)|\psi\rangle$ at time t'. A unitary operator U has the property that if

$|\psi'\rangle = U\,|\psi\rangle$ and $|\chi'\rangle = U\,|\chi\rangle$, then

$$\langle\chi'|\psi'\rangle = \langle\chi|\psi\rangle.^4 \tag{3.1.1}$$

Under the Copenhagen interpretation, a system will evolve unitarily for the most part, but there will typically be a non-unitary change in the state describing the system whenever there is a measurement.[5] However, in non-collapse models such as

---

4. A unitary operator U must also be linear so that for any two states $|\psi\rangle$ and $|\phi\rangle$ and complex numbers α and β, we have

$$U(\alpha\,|\psi\rangle + \beta\,|\phi\rangle) = \alpha U\,|\psi\rangle + \beta U\,|\phi\rangle\,,$$

and furthermore, a unitary operator must have the property that it is invertible: there is a linear operator U^{-1} such that UU^{-1} and $U^{-1}U$ are the identity operator I, i.e. $U^{-1}U\,|\psi\rangle = UU^{-1}\,|\psi\rangle = |\psi\rangle$ for any state $|\psi\rangle$.

5. Note that to say that the change in a state is non-unitary when a measurement is made is not to say that there is a non-unitary collapse operator that maps the quantum state to an eigenstate of some observable. Such a mapping would not make sense, since the collapse is not deterministic given the initial state. However, one could have a well-defined mapping from a time value t to the quantum state of the system $|\psi(t)\rangle$ at time t. We then say that a system changes unitarily if and only if there is a unitary operator $U(t_1, t_0)$ for any two times t_0 and t_1 such that whenever the state of the system at time t_0 is given by $|\psi(t_0)\rangle$, then the state of the system at time t_1 is given by $|\psi(t_1)\rangle = U(t_1, t_0)\,|\psi(t_0)\rangle$, and that for an intermediate time t, $U(t_1, t_0) = U(t_1, t)U(t, t_0)$. So to say that the change in a state is non-unitary when a measurement is made is to say that the state $|\psi(t)\rangle$ describing the system does not change unitarily in the process of making a measurement. Now to see why this is the case under the Copenhagen interpretation, we suppose that at time t_0 a system is in the state $|\psi(t_0)\rangle$ and that as long as no measurements are made up until a time $t \geq t_0$, the state evolves to a state $|\psi^{(U)}(t)\rangle = U(t, t_0)\,|\psi(t_0)\rangle$ where $U(t, t_0)$ is a unitary operator determined by Schrödinger's equation. Furthermore, we suppose that there is a measurable quantity with which we associate an observable $\hat{O}$ so that whenever the state of the system is an eigenstate of $\hat{O}$, the value of the measurable quantity for the system will be a determinate value and equal to the corresponding eigenvalue of $\hat{O}$. At time t_0, we can express $|\psi(t_0)\rangle$ as a linear combination

$$|\psi(t_0)\rangle = \sum_i c_i\,|s_i(t_0)\rangle$$

where the $|s_i(t_0)\rangle$ are eigenstates of $\hat{O}$ with distinct eigenvalues. As long as no measurement is made, this will evolve as

$$|\psi^{(U)}(t)\rangle = \sum_i c_i U(t, t_0)\,|s_i(t_0)\rangle\,.$$

We assume that as the state $|s_i(t_0)\rangle$ evolves to the state $|s_i(t_1)\rangle$ from time t_0 to t_1, it remains an eigenstate of $\hat{O}$ with approximately the same eigenvalue. This assumption is based on the principle that in practice, performing a measurement is not instantaneous, but rather must take place over a

121

the Bohmian interpretation, the many-worlds interpretation, and Kent's theory, the quantum state always evolves unitarily.

3.2 The Additional Variables of Kent's Theory

The second similarity Kent's theory has with the Bohmian interpretation is that it posits the reality of some additional variables beyond standard quantum theory (i.e. in addition to the quantum state[6]). Some details in the section are rather technical, so I've split this section into two subsections: the first section gives an overview, and the second section is more technical and can be skipped if desired.

time interval, and so the eigenstate and eigenvalue must be stable enough over this time interval so as to specify a definite outcome. We also assume that when the system is already in an eigenstate $|s_i(t_0)\rangle$ of the observable $\hat{O}$, it will evolve unitarily as $|s_i(t)\rangle = U(t, t_0) |s_i(t_0)\rangle$ for t between t_0 and t_1, and that performing the measurement corresponding to $\hat{O}$ will have no effect on the system when it is an eigenstate $|s_i(t)\rangle$ of $\hat{O}$ – otherwise we couldn't be sure that whenever we looked at the measurement readout that we weren't changing the value of the quantity we were trying to measure.

Now according to the Copenhagen interpretation, when the measurement corresponding to $\hat{O}$ is made, the system must enter into one of the eigenstates of the observable $\hat{O}$, and at time t_1 shortly after the measurement has been made, the probability the system will be in the $|s_i(t_1)\rangle$-state given that it was in the $|\psi(t_0)\rangle$-state at time t_0 will be $|\langle s_i(t_1)|\psi^{(U)}(t_1)\rangle|^2$ in accordance with the Born rule. So taking $|\psi(t_1)\rangle$ to be proportional to $|s_i(t_1)\rangle$ for some i, we see that for $j \neq i$, $\langle s_j(t_1)|\psi(t_1)\rangle = 0$. This is because eigenstates of a Hermitian operator that have different eigenvalues must be orthogonal. However, since $U(t_1, t_0)$ is unitary,

$$\langle s_j(t_1)|\psi^{(U)}(t_1)\rangle = \langle s_j(t_0)|\psi^{(U)}(t_0)\rangle = c_j.$$

So we see that $|\psi^{(U)}(t_1)\rangle \neq |\psi(t_1)\rangle$ if $\psi(t_0)$ is not initially in an eigenstate of $\hat{O}$, and hence $|\psi(t)\rangle$ doesn't evolve unitarily up to time t_1 as $|\psi^{(U)}(t)\rangle$ does.

6. We may wish to think of these additional variables as hidden variables, but we are not obliged to since we don't speculate on whether these additional variables are necessarily unknowable. Rather, we just see them as supplementing the quantum state so as to provide a complete description of the system.

3.2.1 Overview

In the Bohmian interpretation, the additional variables are the positions and momenta of all the particles, whereas in Kent's theory, the additional variables specify the mass-energy density on a three-dimensional distant future spacelike hypersurface S in spacetime.

To describe what is meant by a three-dimensional hyperspace S, we will need some terminology and notation that is used in special relativity. A *spacetime location* is a point (x^1, x^2, x^3) in three-dimensional space at a particular instant of time t, and hence described by four numbers (x^0, x^1, x^2, x^3) where $x^0 = ct$ and where c is the speed of light.[7] We will use the convention of boldface type to depict spatial locations, e.g. $\boldsymbol{x} = (x^1, x^2, x^3)$, and non-boldface type to depict a spacetime location, e.g. $x = (x^0, x^1, x^2, x^3)$.

Now a key insight of special relativity is that there is no such thing as absolute time. So for instance, two spacetime locations might seem to be simultaneous from one frame of reference, but another person travelling at a different velocity would judge with equal propriety the same two spacetime locations to be non-simultaneous. But it is not the case that for any two spacetime locations we can always find a frame of reference in which the two spacetime locations are simultaneous – sometimes this is not possible. We refer to distinct spacetime locations that could be simultaneous in

7. Multiplying time by the speed of light means that x^0 is a distance like x^1, x^2, and x^3.

some frame of references as being *spacelike-separated*. For example, the two spacetime

locations o and a in figure 3.1 are spacelike-separated.
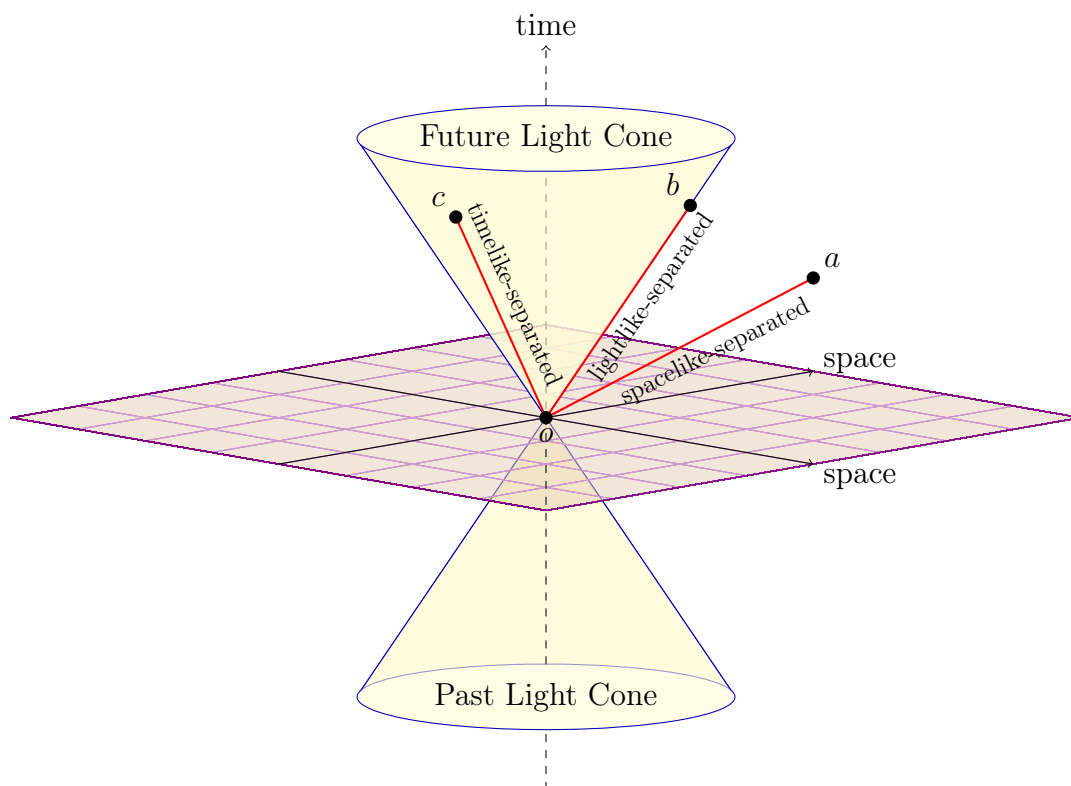
Figure 3.1. The meaning of spacelike, timelike and lightlike-separation when there are two space dimensions and one time dimension.

There are also distinct spacetime locations in spacetime that could be connected

by a beam of light such as the two spacetime locations o and b in figure 3.1. Such

spacetime locations are referred to as being *lightlike-separated*. For any given spacetime

location, the spacetime locations that are lightlike-separated from it form two cones[8]

called the future light cone and the past light cone as shown in figure 3.1. Because

---

8. Strictly speaking, the set of spacetime locations that are lightlike-separated from a given spacetime location form the surface of a cone rather than a cone (which is a convex object). But among physicists, the terminology light cone has stuck.

light appears to travel at the same speed no matter what frame of reference one uses, the light cone of a spacetime location remains invariant when one changes from one reference frame to another. In other words, if another spacetime location lies on the light cone of a spacetime location in one frame of reference, then it must lie on the light cone of this spacetime location in every frame of reference.

Figure 3.1 also depicts two spacetime locations o and c that are *timelike-separated*. Such spacetime locations lie within the light cones of each other, and when two spacetime locations are timelike-separated, it is always possible to choose a frame of reference in which the two spacetime locations are located at the same point in space, but with one spacetime location occuring after the other depending on which spacetime location is in the future light cone of the other.

Now a three-dimensional spacelike hypersurface S in spacetime is a maximal[9] three-dimensional surface in which any two spacetime locations of S are spacelike-separated. In this dissertation, we will assume that all hypersurfaces are spacelike and maximal. Kent assumes that the hypersurface S is in the distant future of an expanding universe so that nearly all the particles that can decay have already done so, and that all the particles that are not bound together are very far from each other so that the probability of any particle collisions is very small. In other words, all the interesting physics in the universe has played its course before S.

---

9. That is, it cannot be extended any further along any of its three dimensions, so it is not a small local surface contained within a boundary.

At every spacetime location $x \in S$, there is a quantity $T_S(x)$ called the *mass-energy density*.[10] The important thing to note about $T_S(x)$ is that it does not depend on which frame of reference one is in.[11] This property is in contrast to many physical properties that do depend on which frame of reference one is in. For example, the kinetic energy of an object will depend on the calculated velocity of the object, and this velocity will in turn depend on the frame of reference in which this calculation is done.

Now Kent supposes that a notional measurement is made over the whole of S to determine a measurement outcome for $T_S(x)$ which we denote as $\tau_S(x)$ for every $x \in S$. It is these $\tau_S(x)$ that are the additional variables of his theory. How this determination of $T_S(x)$ comes about is up to one's philosophical preferences. For example, one could suppose that it was simply by divine fiat that this determination of $T_S(x)$ comes about.[12]

Clearly, Kent is not assuming that such a measurement could be performed on S in reality, hence his reason for referring to the measurement as fictitious/notional.[13] However, Kent does seem to be saying that there would be a fact of the matter

---

10. The definition of $T_S(x)$ will be discussed in section 3.2.1.

11. The reason for why this is will be discussed in section 3.2.1.

12. I will discuss my philosophical preference in the final chapter.

13. Kent uses the word fictitious when referring to these measurements, e.g. Kent, "Lorentzian Quantum Reality: Postulates and Toy Models," 3. Butterfield refers to these measurements as notional e.g. Butterfield, "Peaceful Coexistence: Examining Kent's Relativistic Solution to the Quantum Measurement Problem," 17.

about what value $T_S(x)$ would have approximately had at each $x \in S$ if it had been measured.[14] Given that there is some fact of the matter about the value $\tau_S(x)$ that $T_S(x)$ takes for every $x \in S$ were $T_S(x)$ to be measured, Kent constructs a picture of physical reality that evolves in a way that is consistent with $T_S(x)$ being $\tau_S(x)$ for $x \in S$.

Kent's proposal might initially sound like he is suggesting that physical reality evolves unitarily from an initial state of the universe $|\Psi_0\rangle$ describing a hypersurface S_0 to a state $|\Psi_S\rangle$ describing the hypersurface S, and then a measurement of $T_S(x)$ is made yielding a value $\tau_S(x)$ for all $x \in S$ which in turn changes the past by a process of backwards causality so that the history of the universe evolving from S_0 is now consistent with $T_S(x)$ having the determinate value $\tau_S(x)$. However, it is not necessary to think that Kent's interpretation requires backwards causality. Rather, we can think of the values $\tau_S(x)$ for $T_S(x)$ as being determined primordially before the evolution of physical reality begins. In other words, we first suppose that the universe is initially in a determinate state $|\Psi_0\rangle$ that describes a hypersurface S_0. Given this state, there are many possibilities for the measurement outcome $T_S(x)$ on the distant future hypersurface S that can be worked out from $|\Psi_S\rangle$, but we don't consider $|\Psi_S\rangle$ as describing the state of actuality of the hypersurface S – all $|\Psi_S\rangle$ does is describe

---

14. Kent talks about taking the asymptotic limit as S tends to the infinite future of S_0. In saying this, he supposes that the physics we are interested in happens between two hypersurfaces S_0 and S_1. Then given some measurement outcome τ_S on a hypersurface S after S_1, for every hypersurface S' after S, there is a range of possible measurements $\tau_{S'}$ on S' that together occur with approximately the same probability as the measurement outcome of τ_S on S and such that the beables between S_0 and S_1 determined by any one of the τ'_S outcomes on S' in this range would be approximately the same as the beables determined by τ_S on S. See Kent, "Lorentzian Quantum Reality: Postulates and Toy Models," 3.

possibilities. One of the possible outcomes $\tau_S(x)$ is then selected,[15] and then physical reality evolves in a way that is consistent with $T_S(x)$ being $\tau_S(x)$ if a measurement were to be made on S. One can therefore avoid positing the need for backwards causality by insisting that the states that unitarily evolve from S_0 don't describe the actual state of the universe. Rather, there is just one history,[16] and this history is consistent with $T_S(x)$ being $\tau_S(x)$ for $x \in S$.

Now we saw on page 66, that for any measurement on a physical system, we can associate an observable whose eigenstates describe a possible state of the physical system in which the measurement is given by the observable's corresponding eigenvalue. Now in the case of the notional measurement of $T_S(x)$, there will be lots of observables, denoted by $\hat{T}_S(x)$, corresponding to all the different places $x \in S$ where the measurement could be made. Therefore, if a measurement were made on S for all $x \in S$, the corresponding state of S would not just be an eigenstate of one observable $\hat{T}_S(x)$, but rather, it would have to be an eigenstate of every observable $\hat{T}_S(x)$ for every $x \in S$. We will therefore need to speak of simultaneous $\hat{T}_S$-eigenstates and simultaneous $\hat{T}_S$-eigenvalues in order to specify the quantum state of S on which $T_S(x)$ has a determinate value of $\tau_S(x)$ for every $x \in S$. This terminology will be described in more detail on page 136. Given that there is such a simultaneous $\hat{T}_S$-eigenstate,

15. Kent doesn't say anything about the selection of $\tau_S(x)$ beyond its selection being consistent with the Born Rule.

16. The role that $\tau_S(x)$ plays in determining this one history will be described in sections 3.3 and 3.4.

we can then work out via the Born Rule the probability this state would be selected given the initial state of the universe.

3.2.2 Technical Details[*]

In order to specify more precisely the additional variables that Kent's theory requires, we need to discuss the Tomonaga-Schwinger picture of relativistic quantum physics.[17] In order to explain this formulation, it is helpful to first consider the distinction between the Heisenberg picture and the Schrödinger picture of standard quantum theory.

In the *Heisenberg picture*, the states describing a system do not change over time. Rather, the observables change over time. So for instance, if there is a time-independent state $|\Phi\rangle$ describing a system and there is some measurable quantity whose expectation value we wish to know at time t given the state $|\Phi\rangle$, then we will need a time dependent observable $\hat{\boldsymbol{O}}(t)$,[18] say, corresponding to the measurable quantity at time t from which we can calculate the expectation value $\langle\Phi|\hat{\boldsymbol{O}}(t)|\Phi\rangle$ at time t given the system is in state $|\Phi\rangle$. In the context of quantum field theory, any observable $\hat{\boldsymbol{O}}(t)$ in the Heisenberg picture will be expressible as a sum (or integral) of observables of the form $\hat{\boldsymbol{O}}(t, \boldsymbol{x})$, where $\hat{\boldsymbol{O}}(t, \boldsymbol{x})$ is an observable of some quantity at a

---

17. See Julian Schwinger, "Quantum Electrodynamics. I. A Covariant Formulation," *Physical review* 74, no. 10 (1948): 1439–1461; S. Tomonaga, "On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields," *Progress of theoretical physics* (Tokyo) 1, no. 2 (1946): 27–42

18. See footnote 22 for an explanation of the convention of using a boldface font for this observable.

particular time t and spatial location $\boldsymbol{x}$.[19]

The Heisenberg picture is contrasted with the *Schrödinger picture* in which the observables do not change over time, but rather the states change over time. So for instance, if there is a time-dependent state $|\Phi(t)\rangle$ describing a system at a specific time t and there is some measurable quantity whose expectation value we wish to know at time t given the state $|\Phi(t)\rangle$, then we will only require a time-independent observable $\hat{\boldsymbol{O}}$, say, corresponding to the measurable quantity from which we can calculate the expectation value $\langle\Phi(t)|\hat{\boldsymbol{O}}|\Phi(t)\rangle$. As in the Heisenberg picture, we can introduce a spatial dependence into the observables so that any observable $\hat{\boldsymbol{O}}$ is expressible as a sum (or integral) of observables of the form $\hat{\boldsymbol{O}}(\boldsymbol{x})$ where $\hat{\boldsymbol{O}}(\boldsymbol{x})$ is an observable of some quantity at a particular spatial location $\boldsymbol{x}$.

Now despite the Schrödinger and Heisenberg pictures taking these different perspectives, they are nevertheless physically equivalent. This is because in both pictures, there is a unitary operator $U(\Delta t)$ for any time interval Δt such that $U(\Delta t)|\Phi(t)\rangle = |\Phi(t+\Delta t)\rangle$, and $U(\Delta t)\hat{\boldsymbol{O}}(t,\boldsymbol{x})U(\Delta t)^{-1} = \hat{\boldsymbol{O}}(t+\Delta t,\boldsymbol{x})$. Thus, given the Schrödinger picture, to get the Heisenberg picture, all we need to do is the following: firstly, we fix a time t_0 and let all the states of the form $|\Phi(t_0)\rangle$ at time t_0 in the Schrödinger picture be the state space for the Heisenberg picture; then for any Schrödinger picture observable

---

19. For example, in quantum electrodynamics (which is one kind of quantum field theory), the observables will be expressible in terms of fields such as the *four-vector potential* $A^\mu(x)$ and the bispinor field $\psi(x)$ which are defined at all spacetime locations $(t,\boldsymbol{x}) = (t,x^1,x^2,x^3)$. The four-vector potential $A^\mu(x)$ can be used to determine the electromagnetic field, and the bispinor field $\psi(x)$ can

$\hat{\boldsymbol{O}}(\boldsymbol{x})$, we define the corresponding Heisenberg picture time-dependent observable

$$\hat{\boldsymbol{O}}(t, \boldsymbol{x}) = U(t - t_0)\hat{\boldsymbol{O}}(\boldsymbol{x})U(t - t_0)^{-1}.$$

Conversely, to move from the Heisenberg picture to the Schrödinger, we first fix a reference time t_0. Then for any state $|\Phi\rangle$ and observable $\hat{\boldsymbol{O}}(\boldsymbol{x}) \overset{\text{def}}{=} \hat{\boldsymbol{O}}(t, \boldsymbol{x})$ in the Heisenberg picture, the corresponding Schrödinger picture time-dependent state at time t will be $U(t-t_0)|\Phi\rangle$, and the corresponding Schrödinger picture time-independent observable will be $\hat{\boldsymbol{O}}(t_0, \boldsymbol{x})$.[20]

Now if there is a quantity we wish to measure at time t_0 with corresponding observable $\hat{\boldsymbol{O}}(\boldsymbol{x}) \overset{\text{def}}{=} \hat{\boldsymbol{O}}(t_0, \boldsymbol{x})$, then in both pictures, the expectation value of this measurable quantity given $|\Phi\rangle \overset{\text{def}}{=} |\Phi(t_0)\rangle$ will be

$$\langle\Phi(t_0)|\hat{\boldsymbol{O}}(\boldsymbol{x})|\Phi(t_0)\rangle = \langle\Phi(t_0)|\hat{\boldsymbol{O}}(t_0, \boldsymbol{x})|\Phi(t_0)\rangle = \langle\Phi|\hat{\boldsymbol{O}}(t_0, \boldsymbol{x})|\Phi\rangle \qquad (3.2.1)$$

Since the left-hand side of (3.2.1) is the Schrödinger picture expectation value of $\hat{\boldsymbol{O}}(\boldsymbol{x})$, and the right-hand side of (3.2.1) is the Heisenberg picture expectation value of $\hat{\boldsymbol{O}}(t_0, \boldsymbol{x})$, it follows that whatever picture we choose, it will make no difference to

---

be used to determine the electric current density. In the Heisenberg picture, these fields will have corresponding Hilbert space operators at each spacetime location x from which expectation values can be calculated at the spacetime location x for a given time-independent state.

20. Note that if there is a non-unitary aspect to the evolution of the Schrödinger picture time-dependent state $|\Phi(t)\rangle$ (such as would be the case if there was a collapse), then we can't expect the Schrödinger picture and the Heisenberg picture to be equivalent. This is because there could be many different states that could evolve to $|\Phi(t)\rangle$ if collapses are permitted, and so we wouldn't

the calculated expectation values of observables – in other words, the two pictures are physically equivalent.

Now although it is easy to move between both the Schrödinger and Heisenberg pictures, they both give a privileged status to hypersurfaces of the form $t = $ const. However, according to special relativity, there are no privileged hypersurfaces. One of the advantages of the Tomonaga-Schwinger picture is that it gives no privileged status to any class of hypersurfaces, but rather all hypersurfaces are placed on the same footing. We are going to see that the expectation value $\langle \Phi(t_0)|\hat{\boldsymbol{O}}(t_0, \boldsymbol{x})|\Phi(t_0)\rangle$ of equation (3.2.1) is a special case of what Tomonaga and Schwinger consider more generally.

We first note that if we can calculate $\langle \Phi(t_0)|\hat{\boldsymbol{O}}(t_0, \boldsymbol{x})|\Phi(t_0)\rangle$ for any t_0 and any $\boldsymbol{x}$, then we can calculate all the expectation values that might interest us. But we also note that in the expectation value $\langle \Phi(t_0)|\hat{\boldsymbol{O}}(t_0, \boldsymbol{x})|\Phi(t_0)\rangle$, the $|\Phi(t_0)\rangle$-state is the state of a hypersurface $t = t_0$, and $(t_0, \boldsymbol{x})$ is a spacetime location on this hypersurface. Now if we are to place all hypersurfaces on the same footing, then in specifying expectation values, we should be just as content in specifying expectation values of the form $\langle \Psi[S]|\hat{O}(x)|\Psi[S]\rangle$, where S is any hypersurface, $|\Psi[S]\rangle$ is any state of this hypersurface,[21] x is any spacetime location on the hypersurface S, and where $\hat{O}(x)$

---

expect there to be an invertible operator $U(\Delta t)$ in terms of which we could define the Heisenberg picture observable $\hat{\boldsymbol{O}}(t, \boldsymbol{x})$.

21. The convention of using square brackets such as in $|\Psi[S]\rangle$ indicates that the thing in question is a *functional*. Functions and functionals are closely related. A function f is a mapping from one set (the domain) to another set (the codomain), such that each input yields a single output. The typical convention is to use round brackets to denote the output, e.g. $f(x)$ where x is the input. A

is any observable of S.[22] The Tomonaga-Schwinger picture thus works with states of the form $|\Psi[S]\rangle$ for any hypersurface S, and observables of the form $\hat{O}(x)$ with $x \in S$ acting on the state space of the hypersurface S from which one can calculate the expectation value $\langle\Psi[S]|\hat{O}(x)|\Psi[S]\rangle$.

In order to construct $|\Psi[S]\rangle$ and $\hat{O}(x)$, Schwinger introduces a unitary operator[23] $U[S]$ that maps the $|\Phi\rangle$-state of the Heisenberg picture to the corresponding $|\Psi[S]\rangle$-state that describes the state of the hypersurface S, i.e. $|\Psi[S]\rangle = U[S]|\Phi\rangle$. Schwinger then defines the observable

$$\hat{O}(x) = U[S]\hat{\boldsymbol{O}}(x)U[S]^{-1} \tag{3.2.2}$$

on S where x is any spacetime location on S, and where $\hat{\boldsymbol{O}}(x)$ is any Heisenberg picture observable. Ostensibly, $\hat{O}(x)$ depends on the surface S, but Schwinger shows that under conditions that are readily satisfied, $\hat{O}(x)$ is independent of the hypersurface

---

functional g, on the other hand, is a function that maps a space of functions or other mathematical objects (such as surfaces or volumes) to some value. The typical convention is to use square brackets to denote the output, e.g. $g[y]$ where y is the input function or other mathematical object. So in the present case, $|\Phi[\cdot]\rangle$ is a functional that takes a surface S as input to produce a state $|\Phi[S]\rangle$ as output.

22. Here I am following the convention of Schwinger of always using non-boldface type to indicate Tomonaga-Schwinger picture observables, and boldface type to indicate Heisenberg picture and Schrödinger picture observables. See Schwinger, "Quantum Electrodynamics. I. A Covariant Formulation," p. 1448.

23. See Schwinger, p. 1448.

S.[24] Also, since

$$\langle \Psi[S] | \hat{O}(x) | \Psi[S] \rangle = \langle \Phi | \hat{\boldsymbol{O}}(x) | \Phi \rangle , \qquad (3.2.4)$$

the Tomonaga-Schwinger picture will give the same physics as the Heisenberg and Schrödinger picture. In order to avoid our notation becoming cluttered, we will sometimes drop the $[S]$ and just write $|\Psi\rangle$ instead of $|\Psi[S]\rangle$, and say that $|\Psi\rangle$ is a state of the hypersurface S, and we will speak of the Hilbert space H_S of all such states of the hypersurface S so that we can write $|\Psi\rangle \in H_S$.[25] The Hilbert space H_S is not to be confused with the Hilbert space $H_{\mathcal{S}}$ of the previous chapter which is the Hilbert space describing a system $\mathcal{S}$.

---

24. The required condition is that

$$i\hbar \frac{\delta U[S]}{\delta S(x)} = \mathcal{H}(x) U[S]$$

where $\mathcal{H}(x)$ is a Hermitian operator that is a Lorentz invariant function of the field quantities at the spacetime location x and has the dimensions of an energy density, and where the *functional derivative* $U[S]$ is given by

$$\frac{\delta U[S]}{\delta S(x)} = \lim_{\delta\omega \to 0} \frac{U[S'] - U[S]}{\delta\omega} \qquad (3.2.3)$$

where S' is a surface that only differs from S in the vicinity of x, and where $\delta\omega$ is the volume enclosed by S and S'. The Hermitian operator

$$\mathcal{H}(x) = -(1/c) j^{\mu}(x) A_{\mu}(x)$$

has the desired property where $j^{\mu}(x)$ is the *current density* and where $A^{\mu}(x)$ is the four-vector potential of the electromagnetic field. With this choice for $\mathcal{H}(x)$, Schwinger shows that $\Box A^{\mu}(x) = 0$ and $\partial_{\mu} A^{\mu}(x) |\Psi[S]\rangle = 0$, where $\Box = \partial_{\mu}\partial^{\mu}$ is the *d'Alembert operator* – see Schwinger, "Quantum Electrodynamics. I. A Covariant Formulation," p. 1449-1450.

25. Though to be clear, the H_S are really identical for all hypersurfaces S since each H_S is the image of the unitary operator $U[S]$ acting on the Heisenberg-picture Hilbert space, and the image of a unitary operator is always equal to the Hilbert space it is acting on.

We are now in a position to come back to the question of what the additional variables of Kent's theory are. As mentioned on page 126, for a given hypersurface S, there will be a mass-energy density $T_S(x)$. Corresponding to this, there will be a Heisenberg picture observable $\hat{\boldsymbol{T}}_S(x)$, and from this we can construct the Tomonaga-Schwinger observable $\hat{T}_S(x) = U[S]\hat{\boldsymbol{T}}_S(x)U[S]^{-1}$.[26] These mass-energy density observables have the property that if x and y are any two spacetime locations of S, then $\hat{T}_S(x)$ and $\hat{T}_S(y)$ will commute. In other words,

$$\hat{T}_S(x)\hat{T}_S(y) = \hat{T}_S(y)\hat{T}_S(x).$$

The commutativity of all the $\hat{T}_S(x)$ for $x \in S$ means that if $|\Gamma\rangle \in H_S$ is an eigenstate of $\hat{T}_S(x)$, then for any $y \in S$, $\hat{T}_S(y)|\Gamma\rangle$ is also an eigenstate of $\hat{T}_S(x)$ with the same eigenvalue as $|\Gamma\rangle$.[27] The invariance of any $\hat{T}_S(x)$-*eigenspace*[28] under the action of $\hat{T}_S(y)$ means that we can create an orthonormal basis of H_S consisting of simultaneous

---

26. Note that $\hat{T}_S(x)$ will depend on S. The remark above about $\hat{O}(x)$ not depending on S does not apply here since the independence of $\hat{O}(x)$ from S assumes that the Heisenberg picture observable $\hat{\boldsymbol{O}}(x)$ is independent of S, but this is not the case for $\hat{\boldsymbol{T}}_S(x)$. However, $\hat{T}_S(x)$ will only depend on S in the vicinity of x, so if S' only differs from S outside a neighborhood of x, then $\hat{T}_S(x) = \hat{T}_{S'}(x)$.

27. To see this, suppose that τ is the eigenvalue of $\hat{T}_S(x)$ corresponding to the eigenstate $|\Gamma\rangle$. Then by commutativity of $\hat{T}_S(x)$ and $\hat{T}_S(y)$ we have

$$\hat{T}_S(x)\hat{T}_S(y)|\Gamma\rangle = \hat{T}_S(y)\hat{T}_S(x)|\Gamma\rangle = \hat{T}_S(y)\tau|\Gamma\rangle = \tau\hat{T}_S(y)|\Gamma\rangle.$$

Hence, $\hat{T}_S(y)|\Gamma\rangle$ is also an eigenvector of $\hat{T}_S(x)$ with eigenvalue τ.

28. An eigenspace of a Hermitian operator $\hat{O}$ acting on a Hilbert space H is just the space of all the eigenstates of $\hat{O}$ in H which have the same eigenvalue.

eigenstates of both $\hat{T}_S(x)$ and $\hat{T}_S(y)$, albeit with different eigenvalues.[29] Moreover, for reasons that we need not go into, these eigenvalues must be greater than or equal to 0.[30] But because x and y are arbitrary points of S, this means we can construct an orthonormal basis $\{|\Gamma_i\rangle : i \in \mathbb{N}\}$ of H_S (where $\mathbb{N}$ is the set of integers greater than 0) such that $\hat{T}_S(x)|\Gamma_i\rangle = \tau_{S,i}(x)|\Gamma_i\rangle$ for all $x \in S$, where $\tau_{S,i}(x) \geq 0$ is a possible mass-energy density measurement defined for every x in S. We will refer to a state $|\Gamma\rangle \in H_S$ as a *simultaneous $\hat{T}_S$-eigenstate*, and a real valued function τ_S defined on the whole of S as a *simultaneous $\hat{T}_S$-eigenvalue* if and only if $\hat{T}_S(x)|\Gamma\rangle = \tau_S(x)|\Gamma\rangle$ for all $x \in S$.

At this point, it is worth clarifying the different meanings of $T_S(x)$, $\hat{T}_S(x)$, and $\tau_S(x)$. We use $T_S(x)$ to refer to the description of the physical quantity that is being observed. Thus, $T_S(x)$ is shorthand for the description "the mass-energy density of the hypersurface S observed at spacetime location x". We use $\tau_S(x)$ to stand for a real valued function on S that could be a measurement outcome of the physical quantity described by $T_S(x)$ for each $x \in S$. Thus, the equation $T_S = \tau_S$ is shorthand for the statement "the mass-energy density of the hypersurface S observed at spacetime location x is $\tau_S(x)$ for every x belonging to S." And for each $x \in S$, $\hat{T}_S(x)$ is the observable (i.e. the Hermitian operator) such that if an observer deems S to be in an

---

29. This is because any $\hat{T}_S(x)$-eigenspace is itself a Hilbert space on which $\hat{T}_S(y)$ acts as a Hermitian operator, so by (2.2.3), we can find an orthonormal basis of states $\{|\psi_1\rangle, \ldots |\psi_N\rangle\}$ of the $\hat{T}_S(x)$-eigenspace and real numbers $\tau_1(y), \ldots, \tau_N(y)$ such that $\hat{T}_S(y)|\psi_i\rangle = \tau_i(y)|\psi_i\rangle$ for $i = 1, \ldots, N$. Hence, each of the $|\psi_i\rangle$ will be a simultaneous eigenstate of both $\hat{T}_S(x)$ and $\hat{T}_S(y)$.

30. In other words, we are not going to concern ourselves with theories that allow for negative mass-energy densities.

eigenstate $|\psi\rangle$ of $\hat{T}_S(x)$ with eigenvalue τ (a real number), then the observer would observe the physical quantity described by $T_S(x)$ to have the value τ. We will add a further clarification to this when we come to consider different observers in section 4.2.

Now in describing simultaneous $\hat{T}_S$-eigenstates and $\hat{T}_S$-eigenvalues as we've done above, it might be objected that there will be uncountably many simultaneous $\hat{T}_S$-eigenstates and $\hat{T}_S$-eigenvalues so that we won't be able to form an orthonormal basis of states $\{\,|\Gamma_i\rangle : i \in \mathbb{N}\}$ with the index i being taken over the whole numbers greater than 0. However, we can overcome this objection to some extent by supposing that in physical reality, there will be a limit on how great the mass-energy density $\tau_S(x)$ can be, and how rapidly it can change with respect to x. Although such an assumption will mean this theory will break down in extreme situations such as in the case of black holes, this theory won't be any worse off than quantum field theory which also breaks down in extreme situations.

We can then suppose that with a suitably fine mesh on S,[31] any mass energy density $\tau_S(x)$ (e.g. such as the one depicted in figure 3.2) can be approximated to a function (e.g. like the one depicted in figure 3.3) that has constant values on each cell of this mesh and such that the approximation value at a cell belongs to a finite pool of possible values. For instance, if c_x is the cell which contains $x \in S$, and $\tau_{\max}$ is the maximum possible value the mass-energy density could be, then we could define the

---

31. Note that the mesh is only a mesh in S, so the cells of the mesh are cube-like subsets of S. The time might not be constant across each cell because of the possible curvature of S.
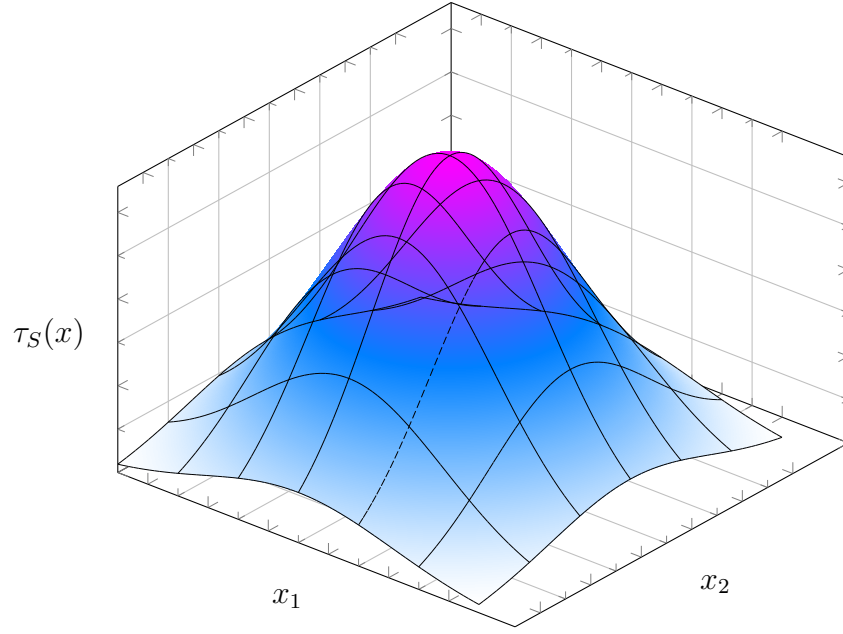
Figure 3.2. An example of an arbitrary mass-energy density τ_S plotted against two dimensions x_1 and x_2 belonging to the hypersurface S.
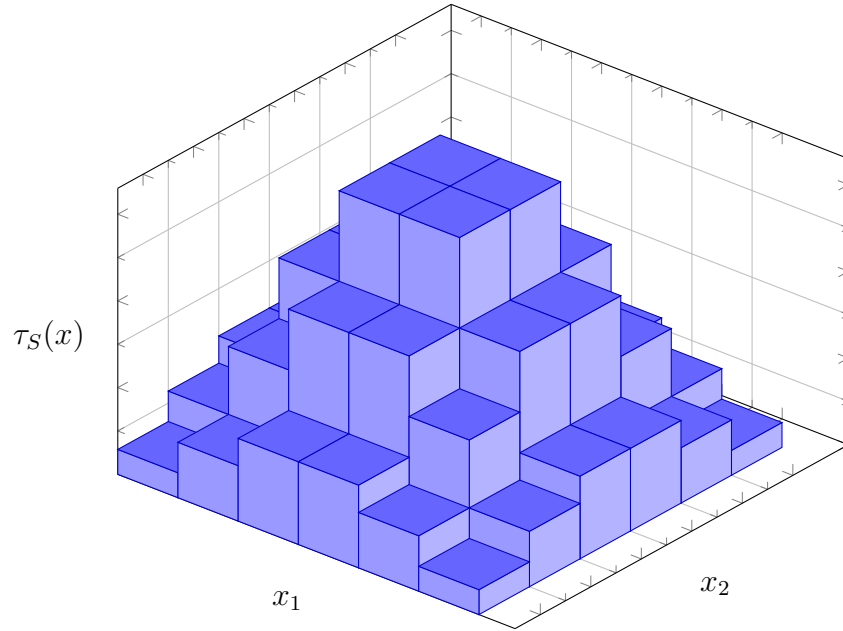
Figure 3.3. Here, the arbitrary mass-energy density τ_S depicted in figure 3.2 has been approximated by a function which is constant in each mesh cell.

approximation to $\tau_S(x)$ at cell c_x to be

$$\tau_S(c_x) = \frac{\tau_{\max}}{N} \left\lfloor N \left(\frac{\text{average of } \tau_S \text{ over } c_x}{\tau_{\max}} \right) + 0.5 \right\rfloor \qquad (3.2.5)$$

where $\lfloor z \rfloor$ is the biggest integer $n \leq z$, and where N is a fixed large number. Then $\tau_S(c_x)$ will have $N + 1$ possible values between 0 and $\tau_{\max}$. There will then only need to be a countable number of approximations $\tau_{S,i}$ to approximate any arbitrary mass-energy density τ_S on S. As long as we choose the cells in the mesh to be sufficiently small and N to be sufficiently large, we can describe physical reality up to our desired level of accuracy. Thus, we assume that for each of the countable states in the orthonormal basis $\{ |\Gamma_i\rangle : i \in \mathbb{N} \}$, there will be a corresponding function $\tau_{S,i}$ defined on S which is constant on every cell of S and in which

$$\hat{T}_S(c_x) |\Gamma_i\rangle = \tau_{S,i}(c_x) |\Gamma_i\rangle \qquad (3.2.6)$$

for all $x \in S$ where $\tau_{S,i}(c_x) = \tau_{S,i}(x)$ and where $\hat{T}_S(c_x)$ is the observable corresponding to the average approximated value of the physical quantity $T_S(x)$ over the cell c_x (approximated as in (3.2.5) with τ_S replaced by T_S), but we will normally just write

$$\hat{T}_S(x) |\Gamma_i\rangle = \tau_{S,i}(x) |\Gamma_i\rangle \qquad (3.2.7)$$

with the implicit understanding that by (3.2.7) we really mean (3.2.6), and that when we speak of $|\Gamma_i\rangle$ and $\tau_{S,i}$ as simultaneous $\hat{T}_S$-eigenstates and simultaneous

$\hat{T}_S$-eigenvalues respectively, we implicitly understand $\hat{T}_S$ and $\tau_{S,i}$ to be defined over cells of the form $c_x \subset S$ rather than over spacetime locations $x \in S$.

Now the additional variables beyond standard quantum theory that are included in Kent's theory are given by one of these simultaneous $\hat{T}_S$-eigenvalues $\tau_{S,i}$ that (approximately) describe a possible outcome for a mass-energy density measurement over the whole of S. We will let τ_S denote the particular $\tau_{S,i}$ that constitute the additional variables of Kent's theory.

Now the particular density τ_S which is found to describe S can't be absolutely anything. Limitations are placed on what τ_S can be, and these limitations will depend on the initial conditions of Kent's theory. Kent assumes that all physics that we wish to describe takes place between two hypersurfaces S_0 and S, with S_0 much earlier than S so that S_0 and S don't intersect. Initial conditions are determined on the hypersurface S_0 so that we can assume it is described by a state $|\Psi_0\rangle \in H_{S_0}$ in the Tomonaga-Schwinger picture. If we define

$$U_{SS_0} = U[S]U[S_0]^{-1}, \qquad (3.2.8)$$

where $U[S]$ and $U[S_0]$ are the Schwinger unitary operators introduced on page 133, then given the state $|\Psi_0\rangle$, there will be a corresponding state $|\Psi_S\rangle = U_{SS_0}|\Psi_0\rangle \in H_S$ that describes the hypersurface S in the Tomonaga-Schwinger picture. Figure 3.4 depicts the evolution of the state $|\Psi_0\rangle$ to the state $|\Psi_S\rangle$. Then if $|\Gamma\rangle$ is a simultaneous $\hat{T}_S$-eigenstate with $\hat{T}_S(x)|\Gamma\rangle = \tau_S(x)|\Gamma\rangle$, then the probability $P(\Gamma|\Psi_0)$ that S will be

S ←——————————————————————————————→ S
   Unitary Evolution $|\Psi_S\rangle = U_{SS_0}|\Psi_0\rangle$

time

S_0 ←——————————————————————————————→ S_0
   Initial State $|\Psi_0\rangle$
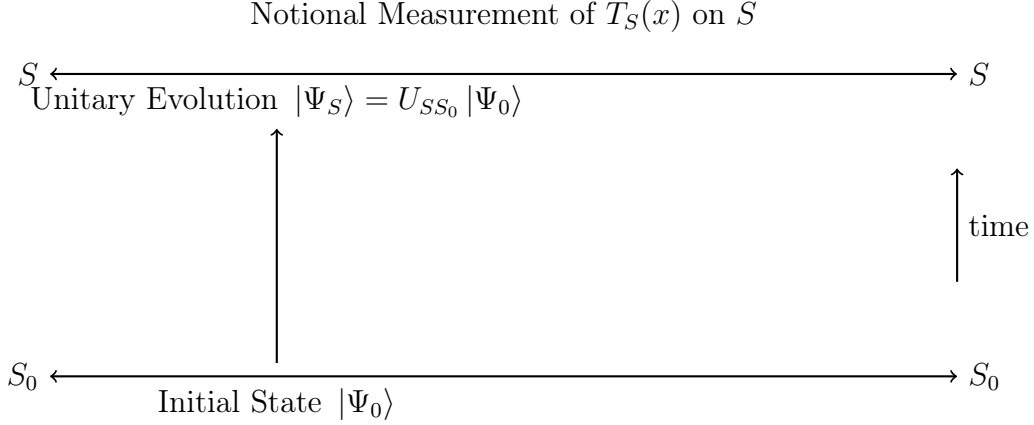
Figure 3.4. A notional measurement of $T_S(x)$ is made for all $x \in S$. The simultaneous $\hat{T}_S$-eigenstate $|\Gamma\rangle$ with $\hat{T}_S(x)|\Gamma\rangle = \tau_S(x)|\Gamma\rangle$ is selected with probability $|\langle\Gamma|U_{SS_0}|\Psi_0\rangle|^2$ $= |\langle\Gamma|\Psi_S\rangle|^2$. The values $\tau_S(x)$ obtained for $T_S(x)$ are then used to calculate the physical properties at the spacetime location y.

found to be in the state $|\Gamma\rangle$ given the initial state $|\Psi_0\rangle$ describing S_0 will be given by

the Born Rule (see page 9):

$$P(\Gamma|\Psi_0) = |\langle\Gamma|U_{SS_0}|\Psi_0\rangle|^2 = |\langle\Gamma|\Psi_S\rangle|^2 \tag{3.2.9}$$

It's possible that there could be more than one simultaneous $\hat{T}_S$-eigenstate in H_S

that has the simultaneous $\hat{T}_S$-eigenvalue τ_S, but it is the mass-energy density τ_S itself

rather than one of the eigenstates with mass-energy density τ_S that constitute the

additional variables that Kent adds to standard quantum theory.

Also note that if every simultaneous $\hat{T}_S$-eigenstate $|\Gamma\rangle$ with simultaneous $\hat{T}_S$-

eigenvalue τ_S satisfies $|\langle\Gamma|U_{SS_0}|\Psi_0\rangle| = 0$, then by (3.2.9), τ_S will have zero probability,

and hence, it will not be a possible measurement outcome for T_S given $|\Psi_0\rangle$.[32] It

_____

32. When dealing with continuous random variables, if the probability of the random variable having a particular value is zero, it does not follow that it is impossible for the random variable to

is for this reason that we can't expect the measurement outcome of T_S on S to be absolutely anything.

3.3 The One-World Feature of Kent's Theory

The third similarity Kent's theory shares with the Bohmian interpretation is that it is a one-world interpretation of quantum physics. It will be helpful to contrast this with the many-worlds interpretation.

Unlike the many-worlds interpretation, Kent's theory does not allow for indeterminate states of macroscopic objects such as cats. In the many-worlds interpretation, Schrödinger will still only observe his cat to be either dead or alive, and not both dead and alive. However, Schrödinger himself goes into a superposition of observing his cat to be alive and his cat to be dead. In the many-worlds interpretation, there is thus a difference between observing something to be so, and something actually being so: the observation is of a particular physical scenario, but the reality is a superposition of different physical scenarios.

To capture this distinction between observation and reality, Bell speaks of *beables*. Bell introduces the term beable when speculating on what would be a more satisfactory physical theory than what quantum physics currently has to offer.[33] Bell says that

---

have this value. However, in the case of discrete random variables, if the probability of the random variable having a particular value is zero, then it does follow that it is impossible for the random variable to have this value. Because we are using (3.2.5) to approximate the energy density, we can treat T_S as a discrete random variable, hence the claim that if $T_S = \tau_S$ has zero probability, then τ_S will not be a possible measurement outcome for T_S given $|\Psi_0\rangle$.

33. See J. S. Bell, "Subject and Object," in *Speakable and unspeakable in quantum mechanics*, 2nd ed. (Cambridge: Cambridge University Press, 2004), 40–44

such a theory should be able to say of a system not only that such and such is observed to be so, but that such and such be so. In other words, a more satisfactory theory would be a theory of beables rather than a theory of observables. On the macroscopic level, these beables should be the underlying reality that gives rise to all the familiar things in the world around us, things like cats, laboratories, procedures, and so on. For example proponents of the Bohmian interpretation believe that the beables are all the particles each with their precise position and momentum. But whatever these beables are, it is because of them that a scientist can observe a physical system to be in such and such a state. Thus, observables are ontologically dependent on beables.

Now the beables in Kent's one-world interpretation are expressed in terms of a physical quantity called the *stress-energy tensor* $T^{\mu\nu}(y)$. For any spacetime location y, the stress-energy tensor $T^{\mu\nu}(y)$ is an array of 16 values corresponding to each combination of $\mu, \nu = 0, 1, 2$, or 3. The value $T^{00}(y)$ is the energy density at y divided by c^2,[34] whereas the other values of $T^{\mu\nu}(y)$ indicate how much energy and momentum flow across different surfaces in the neighborhood of y.

It was mentioned in the previous section that for any spacetime location $x \in S$, there is an observable $\hat{T}_S(x)$ acting on H_S corresponding to the mass-energy density $T_S(x)$ of the surface S at x. It turns out that for any $\mu, \nu = 0, 1, 2$, or 3, there is also an observable $\hat{T}^{\mu\nu}(x)$ acting on H_S, such that if $|\Gamma\rangle \in H_S$ is a simultaneous eigenstate of $\hat{T}^{\mu\nu}(x)$ with eigenvalue $\tau^{\mu\nu}(x)$ for all $x \in S$, then $|\Gamma\rangle$ corresponds to a

---

34. This is not to be confused with the mass-energy density $T_S(x)$ defined for x on a hypersurface S. As will be shown in section 4.2, all 16 elements of $T^{\mu\nu}(x)$ will typically be needed to calculate $T_S(x)$.

state of S in which $T^{\mu\nu}(x)$ is $\tau^{\mu\nu}(x)$ for all $x \in S$.[35] Moreover, the observable $\hat{T}_S(x)$ is expressible in terms of the $\hat{T}^{\mu\nu}(x)$-observables.[36,37] Now the beables in Kent's theory are defined at each spacetime location y that occurs after S_0 and before S. For such a spacetime location y, the beables will be determinate values of the stress-energy tensor $T^{\mu\nu}(y)$, but calculated from the expectation of the observable $\hat{T}^{\mu\nu}(y)$ conditional on the mass-energy density $T_S(x)$ on S being given by $\tau_S(x)$ for all the $x \in S$ that are outside the light cone of y.

Kent implicitly assumes that a specification of the stress-energy tensor $T^{\mu\nu}(y)$ for all spacetime locations y between S_0 and S will be sufficient to give a macroscopic description of physical reality between S_0 and S.[38] This seems like a reasonable assumption, since from the $T^{00}(y)$ component of the stress-energy tensor, we will be able to tell how much energy is present in the vicinity of y, and from the other components of the stress-energy tensor, we will be able to work out how much of this energy is due to mass and how much is due to motion.[39] Thus, from the stress-energy

---

35. Note however, that such a simultaneous eigenstate is only for a fixed choice of μ and ν, since in general, $\hat{T}^{\mu\nu}(x)$ and $\hat{T}^{\mu'\nu'}(x)$ will not commute for $\mu \neq \mu'$ or $\nu \neq \nu'$.

36. See section 4.2 for an explanation for why this is so.

37. As in (3.2.7), we have the same implicit understanding of $\hat{T}^{\mu\nu}(x)$ and $\tau^{\mu\nu}(x)$ as being defined over cells $c_x \subset S$ rather than at spacetime locations $x \in S$, though we will often speak of them as being defined at spacetime locations.

38. See Kent, "Lorentzian Quantum Reality: Postulates and Toy Models," 2 where he talks about giving a description of reality between S_0 and S.

39. It doesn't matter that the $\hat{T}^{\mu\nu}(y)$ won't typically commute for all μ and ν because the value of $T^{\mu\nu}(y)$ that Kent ascribes to the spacetime location y is not the eigenvalue of some simultaneous eigenstate of the $\hat{T}^{\mu\nu}(y)$ for all μ and ν (non-commutativity of the $\hat{T}^{\mu\nu}(y)$ would guarantee that there

tensor, we will be able to form a picture of where things are and whether things are in equilibrium or whether there are flows of matter and energy from one region to another. Such a description of reality will also include information about measurement readings that scientists observe when performing their experiments, so although we wouldn't expect there to be a complete specification of physical reality down to the microscopic level in terms of the stress-energy tensor, there will be enough information in the stress-energy tensor at the macroscopic level to make various scientific claims about physical reality at the microscopic level.

In section 3.4, we will come back to the question of why we can't include any information about $\tau_S(x)$ for $x \in S$ within the light cone of y when we discuss how the conditional expectations of the stress-energy density are calculated. But before we do that, we first consider why we should need conditional expectations at all in order to provide a one-world description of reality.

To this end, we recall the definition of expectation in equation (2.2.1) and the expectation formula (2.2.2) for an observable. In a theory that posited the beables to be the expectation values of $\hat{T}^{\mu\nu}(y)$ for any y located between S_0 and S without conditioning on the value of the mass-energy density T_S on S, then the $T^{\mu\nu}(y)$-beable would just be $\langle \Psi_{S'} | \hat{T}^{\mu\nu}(y) | \Psi_{S'} \rangle$ where $| \Psi_{S'} \rangle = U_{S'S_0} | \Psi_0 \rangle$ for any hypersurface S' that

is no such simultaneous eigenstate). Rather, the value of $T^{\mu\nu}(y)$ that Kent ascribes to the spacetime location y is an expectation value of $\hat{T}^{\mu\nu}(y)$ conditioned on the mass-energy density $T_S(x)$ on S, and so the non-commutativity of the $\hat{T}^{\mu\nu}(y)$ is irrelevant when calculating this expectation value for all μ and ν. Nevertheless, one might still wonder whether it is sensible to ascribe the expectation value of $\hat{T}^{\mu\nu}(y)$ conditioned on the mass-energy density $T_S(x)$ on S to $T^{\mu\nu}(y)$. But as we will see in section 4.6, conditioning the expectation of $\hat{T}^{\mu\nu}(y)$ on the mass-energy density $T_S(x)$ on S is equivalent to taking the expectation $\langle \Gamma | \hat{T}^{\mu\nu}(y) | \Gamma \rangle$ defined in the usual way, where the conditioning is now encapsulated in $| \Gamma \rangle$ (which I refer to as the conditioned quantum state). Moreover, we should typically expect

goes through y.[40] However, such a beable would give a description of reality that was very different from what we observe. For instance, in a Schrödinger cat-like experiment (see section 1.3), there would be a stress-energy tensor distribution corresponding to both the cat being alive and the cat being dead in the same world as depicted in figure 3.5. Such a distribution arises in this context because initially there is
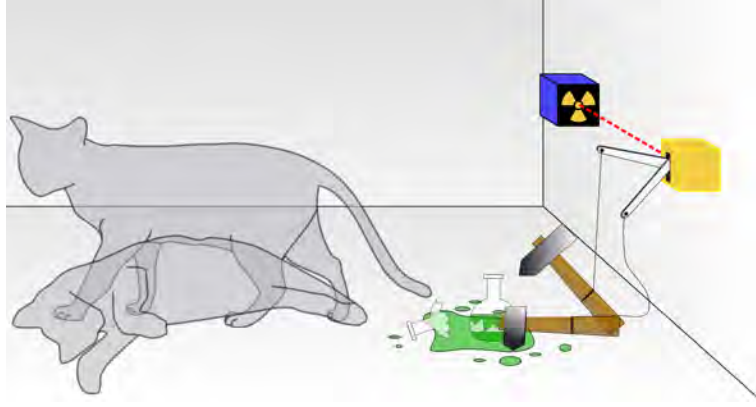
Figure 3.5. A depiction of Schrödinger's cat being both dead and alive.[41]

an atom that is in a superposition of decayed and non-decayed states, and so the

---

there to be sufficient information in the mass-energy density $T_S(x)$ on S to ensure that $|\Gamma\rangle$ is very close to being an eigenstate of all the $\hat{T}^{\mu\nu}(c_y)$ where $\hat{T}^{\mu\nu}(c_y)$ is the averaged stress-energy operator over some mesoscopic three-dimensional spatial cell c_y containing the spacetime location y (i.e. $\hat{T}^{\mu\nu}(c_y) = \frac{1}{|c_y|} \int_{y' \in c_y} \hat{T}^{\mu\nu}(y') \mathrm{d}y'$ where $|c_y|$ is the volume of the three-dimensional spatial cell c_y). This means there will be numerical quantities $t^{\mu\nu}$ for all μ and ν such that $\hat{T}^{\mu\nu}(c_y)|\Gamma\rangle \approx t^{\mu\nu}|\Gamma\rangle$ to very good approximation, and the conditioned expectation value of $\hat{T}^{\mu\nu}(y)$ will thus be $\langle\Gamma|\hat{T}^{\mu\nu}(c_y)|\Gamma\rangle \approx t^{\mu\nu}$. The $t^{\mu\nu}$-values that approximate the conditioned expectation value of $\hat{T}^{\mu\nu}(c_y)$ will then form a good approximation of the average stress-energy tensor in the vicinity of y at the macroscopic level.

40. This can be done such that $\langle\Psi_{S'}|\hat{T}^{\mu\nu}(y)|\Psi_{S'}\rangle$ does not depend on the hypersurface S' other than the fact that it contains y. For more details see Schwinger, "Quantum Electrodynamics. I. A Covariant Formulation."

41. Original by Dhatfield. This image is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license. Source: https://commons.wikimedia.org/wiki/File:Schrodingers_cat.svg

expectation of $\hat{T}^{\mu\nu}(y)$ will have non-zero components both in the location where the non-decayed atom would be, and also in the locations of the decayed atom and the particle the atom emitted. As the decayed atom part of the state interacts with the poison releasing device, this device will also enter into a superposition so that in both the location of the poison containing flask and in the locations of all the poison atoms in the container containing the cat and into which the poison is released, the expectation of $\hat{T}^{\mu\nu}(y)$ will have non-zero components. And then the cat will enter into a superposition of being in a dead state and an alive state, and so the expectation of $\hat{T}^{\mu\nu}(y)$ will have non-zero components in locations where the dead cat ends up and where the living cat happens to be. So the expectation of $\hat{T}^{\mu\nu}(y)$ in the locations of the container containing the cat will be very different from what someone would actually observe. To see how bizarre a description of reality would be if we just based it on unconditioned expectation values of observables, suppose we had an observable A whose value was 1 if the cat was alive and 0 if the cat was dead, then assuming the decay probability was 1/2, the expected value of A would be 1/2. Therefore, if we were to rely on such an expectation value in describing physical reality, we would have to say that the cat was literally half-dead and half-alive.

To overcome this defect, information about the mass-energy density on S is used, specifically the values of $\tau_S(x)$ for $x \in S^1(y)$ where $S^1(y)$ is defined to consist of all the spacetime locations of S outside the light cone of y as depicted in figure 3.6.

So in the case of Schrödinger's cat, if the cat were dead, light reflecting off the dead cat and going off into outer space would eventually intersect the hypersurface S,

147
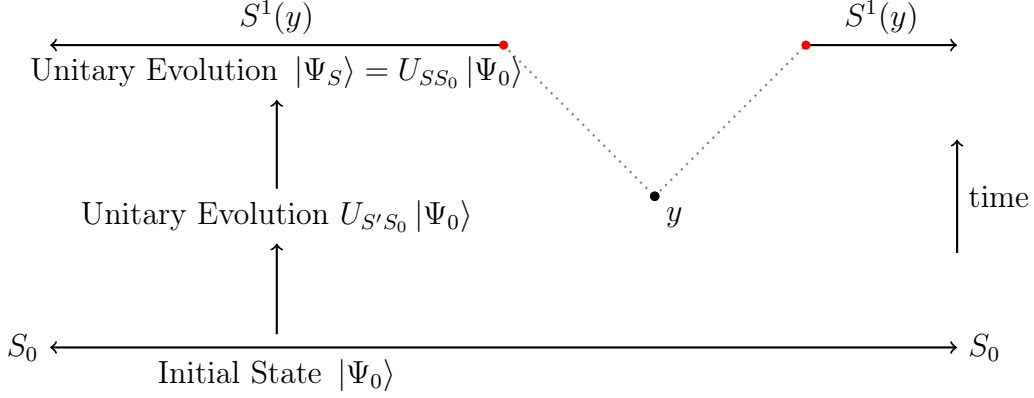
Figure 3.6. The set $S^1(y)$ consists of all the spacetime locations of S outside the light cone of y. The $T^{\mu\nu}(y)$-beables are calculated using the initial state $|\Psi_0\rangle$ together with the values of $\tau_S(x)$ for $x \in S^1(y)$.

and the light distribution on S would register the inanimate status of the cat. On the other hand, if the cat were alive, the light reflecting off the living cat and going off into outer space would also intersect S, but now the light distribution on S would register the different locations the living cat was in as it moved about. Because light travels at a constant speed in a vacuum, the state of the cat at earlier times would be described by light distributions in regions on S that were further away from the cat than those light distributions in regions of S that described the cat in more recent times.

Now if the cat was in a superposition of dead and alive states, then assuming there is no intermediate collapse of the global quantum state, the hypersurface S would also enter into a superposition of different states corresponding to these different distributions of light registered on S. But if a notional measurement on S is made that determines which of these distributions is actually realized on S, then this determination will determine which history was actualized, and hence determine

whether the cat actually survived Schrödinger's experiment or whether it perished. Thus, by conditioning on one of these two distributions on S being actualized, the conditional expectation of the stress-energy tensor in the vicinity of where Schrödinger's cat might be will not describe a situation like the one depicted in figure 3.5. Rather, it will either describe a situation like the one depicted in figure 3.7, or it will describe a situation like the one depicted in figure 3.8. Which of these two situations occur will be determined by whether the measurement outcome on S corresponds to a light distribution reflected from a living cat, or to a light distribution reflected from a dead cat.
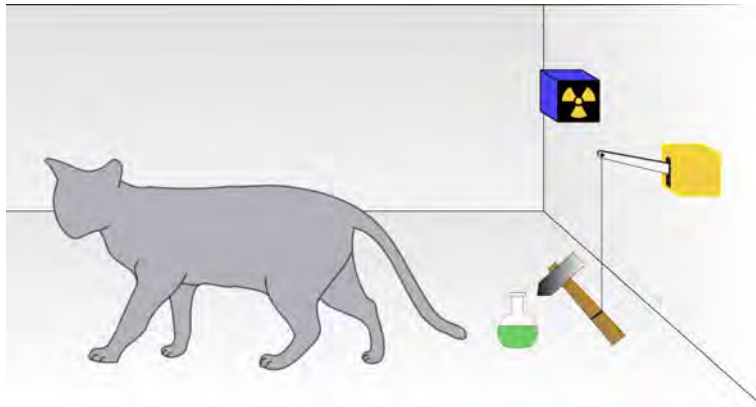
Figure 3.7. A depiction of Schrödinger's cat being alive.[42]

---

42. Original by Dhatfield. This image is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license. Source: https://upload.wikimedia.org/ wikipedia/commons/archive/9/91/20080627113554!Schrodingers_cat.svg

43. Original by Dhatfield. Altered by removing numbers and making into two separate figures. This image is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license. Source: https://upload.wikimedia.org/wikipedia/commons/archive/9/91/ 20080627113554!Schrodingers_cat.svg
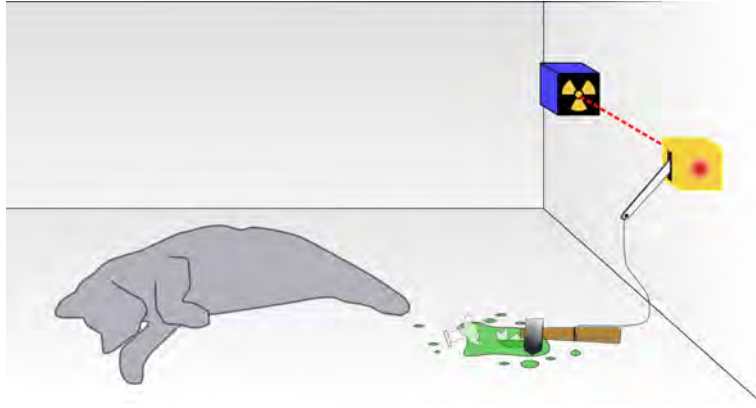
Figure 3.8. A depiction of Schrödinger's cat being dead.[43]

One obvious objection to the above explanation is that one could imagine that the cat is in a box with perfectly mirrored walls, so that light from the cat can't escape. In such a situation, there would be no information about the cat's death or survival outside the relevant light cone, and so the cat would remain in a superposition of alive and dead states. Although perfectly mirrored walls are practically impossible, it would still seem rather problematic that the cat's having a well-defined life/death state depends on how perfectly reflective the walls of the box are.

In responding to this objection, we should clarify that it is not necessary that light that has reflected directly off the cat must eventually intersect S in order for the cat's state to be determined. Rather it is only necessary that light that has reflected off something that the cat is entangled with eventually intersects S. Now when light from the cat is reflected by the walls, the atoms in the walls affected by the light from the cat will get entangled with the cat, and the affected atoms will then interact with neighboring atoms which in turn will get entangled with the atoms that are entangled with the cat, and this entanglement will propagate through to the walls of

150

the box, and when the light that reflects off the outside walls of the box is measured on S, this will contain some information about the state of the atoms of the outside walls and hence the atoms they are entangled with, and so this will ultimately lead to information about the state of the cat.

One might still worry that it would still take a very long time for the light measured on S to be sufficient to distinguish between the cat being alive and the cat being dead. We do however, have reason to think that this process might be very quick. For example, as already mentioned on page 111, if a free electron is initially described by a wave packet whose width is around 10^{-10} m, then according to the Schrödinger equation, after one second the width of the wave packet will have spread to a width of around 1000 km. In reality, however, the electron remains relatively localized because of the scattering of light from the electron, and the light contains information that is able to localize the electron's position. So in the case of the cat in a box, the light reflecting off the outside of the box is at least going to contain enough information in the space of a second to determine that the cat is in the box rather than 1000 km away. And given that a cat is significantly larger than an electron, it seems very plausible that the light reflected off the outside of the box is going to contain significantly more information than the mere fact that the cat is in the box, and so we can reasonably expect there to be enough information in a short space of time to determine whether the cat is alive or dead.

Another objection one could raise against Kent's interpretation is that it depends on the selection of a particular measurement outcome τ_S for the mass-energy density

T_S on S. However, Kent is rather vague about what he means by selection. Here is a (slightly edited) passage from Kent's 2014 paper that could do with some clarification:

> For any given hypersurface S in the future of the initial hypersurface S_0, we consider the effect of joint measurements of the local mass-energy density $T_S(x)$ ... carried out at each point $x \in S$ .... This gives us a probability distribution on possible mass-energy distributions $\tau_S(x)$ on S. In a universe in which physics starts on S_0 and ends on S, our picture of reality is that one $\tau_S(x)$ is randomly selected from the Born rule probability distribution. In other words, there is a randomly selected final boundary condition on S, which is defined mathematically in the same way that it would be if $T_S(x)$ were actually measured on S. However, we treat this simply as a mathematical algorithm. We do not suppose that a physical measurement actually takes place on S, or anywhere else. Our aim, instead, is to give a mathematical description of reality applicable to closed quantum systems, for which there are no external observers able to carry out measurements. To give a description of reality between S_0 and S, we use the initial state on S_0, the randomly chosen final outcome data $\tau_S(x)$ on S, and the unitary evolution law arising from the quantum dynamics.[44]

Now this passage raises two questions. Firstly, there is the question of what it means for the final outcome data τ_S on S to be *selected*, and secondly, there is the question of what it means for the final outcome data τ_S on S to be selected *randomly from the Born rule probability distribution*.

With regard to the question of what it means to be selected, some greater clarity would be desirable since there will be other possible outcomes each of which will also have data describing them. It would therefore be helpful to know what we mean by predicating 'selected' of one possible set of outcome data τ_S, and predicating 'not selected' of another set of outcome data τ_S'? Now one suggestion would be to say that

---

44. Kent, "Lorentzian Quantum Reality: Postulates and Toy Models," 2

the predicate 'selected' just means the attribution of some kind of quality Q_{selected} to the outcome data τ_S which the outcome data τ'_S lacks. However, if we took this suggestion seriously, it would not be obvious why the quality Q_{selected} of τ_S is a reason to consider the universe $\mathcal{U}$ whose description between S_0 and S depended on τ_S was any more real than the universe $\mathcal{U}'$ whose description between S_0 and S depended on τ'_S. In this case, we might therefore doubt that Kent was really proposing a one-world interpretation of quantum physics.

A better way to think of what was meant by predicating 'selected' of τ_S would be to say that it just means that τ_S is a property of S, namely at each $x \in S$, the mass-energy density $T_S(x)$ is given by $\tau_S(x)$. Such an understanding of τ_S being selected would then be no more problematic than predicating properties of a subject. And since τ'_S is not a property of S, we would not be inclined to think that a universe $\mathcal{U}'$ whose description between S_0 and S depended on τ'_S was real.

With regard to the question of what it means for τ_S to be randomly selected from the Born rule probability distribution, if there is only one state $|\Gamma\rangle$ such that $\hat{T}_S(x) |\Gamma\rangle = \tau_S(x) |\Gamma\rangle$ for all $x \in S$, then the probability $P(\tau_S)$ that τ_S is selected will be precisely the probability given by equation (3.2.9). But if there are several states $\{ |\Gamma_\alpha\rangle : \alpha \}$ such that $\hat{T}_S(x) |\Gamma_\alpha\rangle = \tau_S(x) |\Gamma_\alpha\rangle$, then the probability $P(\tau_S)$ that τ_S is selected will be

$$P(\tau_S) = \sum_\alpha | \langle \Gamma_\alpha | U_{SS_0} | \Psi_0 \rangle |^2 = \sum_\alpha | \langle \Gamma_\alpha | \Psi_S \rangle |^2$$

where $|\Psi_S\rangle = U_{SS_0}|\Psi_0\rangle$. However, although we can state the Born rule probability $P(\tau_S)$ for τ_S to be selected, this doesn't tell us what it *means* for τ_S to be selected with this probability. Given that Kent is proposing a one-world interpretation, he is presumably not thinking of this probability in frequentist terms as though there were many worlds with a certain proportion of them having the mass-energy density T_S being given by τ_S. One could try to understand the probability $P(\tau_S)$ as making a counter-factual claim: if there were many worlds, then a certain proportion of them would have the mass-energy density T_S being given by τ_S. But such a counter-factual claim seems irrelevant to the manner in which τ_S is selected. If τ_S' was another possible mass-energy density such that $P(\tau_S) \ll P(\tau_S')$, then despite this inequality, τ_S could still be selected so long as $P(\tau_S) > 0$, and so in this case, to say τ_S was selected with probability $P(\tau_S)$ is not obviously saying anything more than τ_S was selected.

Perhaps we could instead think of the probability $P(\tau_S)$ in Bayesian terms, that is, maybe we should think of $P(\tau_S)$ as expressing a degree of belief that τ_S is the mass-energy density T_S. But again there is the problem of relevance: what relevance is my degree of belief to the fact that τ_S is selected? Perhaps I will be extremely surprised to learn that τ_S is selected, but my surprise doesn't make any difference to the fact that τ_S is selected. From a Bayesian perspective, there is also the further problem that there is going to be so much information in $\tau_S(x)$, and $P(\tau_S)$ is going to be so small due to the vast number of possible values for T_S, and so what is being claimed about τ_S when saying it has a probability $P(\tau_S)$ would be beyond the comprehension of any human mind to form any definite beliefs.

Another alternative for what it means for τ_S to be randomly selected from the Born rule probability distribution $P(\tau_S)$ would be to say there is some threshold $\epsilon > 0$ such that if $P(\tau_S) \geq \epsilon$, then τ_S is a possible value for T_S, whereas if $P(\tau_S) < \epsilon$, then τ_S is not a possible value for T_S. Presumably, we could make ϵ sufficiently large so as to avoid scenarios in which there were stress-energy tensor values in regions between S_0 and S which corresponded to a cat being both dead and alive. Perhaps making ϵ larger still would ensure that a world whose description depended on τ_S with $P(\tau_S) > \epsilon$ and in which there were sufficiently intelligent scientists would be a world in which scientists could formulate the theory of quantum physics as we know it based on their experimental observations. However, one problem with this suggestion is that the smaller ϵ is, the less the Born rule has any bearing on which τ_S is selected. If we permitted $\epsilon = 0$, then the Born rule would have no relevance at all. Also, it would seem strange to treat the selection of two possible mass-energy densities τ_S and τ'_S in exactly the same way if $P(\tau_S) \gg P(\tau'_S) > \epsilon$.

Another way of thinking about the probability $P(\tau_S)$ would be to take the metaphor of God throwing dice: God throws some dice many times, and based on the sum of the outcomes, He chooses one of the τ_S in such a way that more probable sums correspond to T_S-outcomes with greater Born rule probability. Given such a scenario, we would need to explain the association between dice throws and T_S-outcomes – perhaps God has a huge library of books each of which specifies a different T_S-outcome, and that when God has summed up all His dice throws, He selects the book that has the same catalog number as this sum. Then finally, God makes S have the same T_S-value as

specified in the book He selected. However, one reason for having misgivings about such a suggestion is that we might think that each of the books that was not selected could still be considered to be just as real as the world described by the book that God did select. Therefore, if Kent's interpretation required something analogous to these books to make sense of the Born rule probability, we might worry whether his interpretation really was a one-world interpretation of quantum physics.

But perhaps another way of associating dice throws with T_S-outcomes would be to suppose that God lays a mesh over S and determines the mass-energy density of each cell of the mesh in turn by using the following strategy: He first selects one of the cells whose mass-energy density He hasn't already determined; He then rolls some dice, and He then uses an algorithm to calculate the mass-energy density of the cell He has just selected where the algorithm depends on the dice outcome, the cell location, the initial state $|\Psi_0\rangle$ of S_0, the mass-energy densities of the other cells He has previously determined, and a suitable application of the Born rule. Having determined the mass-energy density of this cell, God then selects another cell and uses the same strategy again to determine its mass-energy density, and so on until the mass-density T_S is determined over the whole of S. There could be any number of variants of this strategy. For example, instead of saying that God directly causes all the cells of S to have the mass-energy densities they have, maybe there are many beings each of which is assigned a particular region of S whose mass-energy density the corresponding being is responsible for determining. Also, it is not obvious that randomness is an essential ingredient in the selection of T_S. For example, instead of

throwing dice, perhaps the digits of π could be used in the determination of T_S, or maybe there are teleological principles at play in the history corresponding to the selected T_S-outcome τ_S, even though it looks like τ_S could have been determined by a random process.

Given the above discussion, it seems that the manner in which T_S is selected is something we could speculate about endlessly. Kent assumes that the outcome of the notional measurement T_S occurs with a probability given by the Born rule. However, it is not at all obvious what Kent means by this. Nevertheless, this lack of clarity is not necessarily a problem for Kent's interpretation. Rather, it just means that when it comes to selecting T_S, Kent's interpretation can accommodate a variety of philosophical preferences. But one would hope that given a particular T_S-outcome τ_S, the history conditioned on τ_S would be one in which any scientists who performed measurements (in the normal sense of measurement) would measure average values of physical quantities consistent with the expectation values predicted by standard quantum theory. This intuition will be discussed in more detail in section 4.6.

3.4 Calculating Kent's $T^{\mu\nu}(y)$-beables[*]

Having given a qualitative description in the last section of how a measurement outcome on S determines which facts obtain in reality such as whether Schrödinger's cat is alive or dead, we now give a more quantitative description of how Kent's beables are calculated. Kent's beables specify $T^{\mu\nu}(y)$ values for all y between S_0 and S, and for all $\mu, \nu = 0, 1, 2$, and 3. Kent's $T^{\mu\nu}(y)$-beables are conditional expectation values,

and the conditional expectation that we need to calculate depends on the notion of *conditional probability*. Recall that in probability theory, the conditional probability $P(q|r)$ that a statement q is true given that a statement r is true is given by the formula

$$P(q|r) = \frac{P(q \,\&\, r)}{P(r)}, \qquad (3.4.1)$$

where $P(r)$ is the probability r is true, and $P(q \,\&\, r)$ is the probability both q and r are true. If we now define $q(\tau)$ to be the statement that some quantity T takes the value τ, then the *conditional expectation* of T given r will be given by the formula

$$\langle T \rangle_r \overset{\text{def}}{=} \sum_\tau P(q(\tau)|r)\tau \qquad (3.4.2)$$

where the summation is over all the possible values τ that T can take.

The recipe for defining Kent's $T^{\mu\nu}(y)$-beable is first to select[45] an outcome τ_S that is defined over all of S. The outcome τ_S is selected with probability determined by the Born rule using equation (3.2.9). Then Kent's $T^{\mu\nu}(y)$-beable for any y between S_0 and S is defined to be

$$\langle T^{\mu\nu}(y) \rangle_{\tau_S} \overset{\text{def}}{=} \langle T^{\mu\nu}(y) \rangle_{r(\tau_S, y)} \qquad (3.4.3)$$

---

45. See pages 151 to 157 for a discussion of what is meant by selection.

where $r(\tau_S, y)$ is the statement that $T_S(x)$ has the determinate value $\tau_S(x)$ for all $x \in S^1(y)$,[46] and where $q(\tau)$ in equation (3.4.2) is the statement that $T^{\mu\nu}(y)$ (understood in the conventional non-Kentian sense) takes the value τ.[47] It is these $T^{\mu\nu}(y)$-beables $\langle T^{\mu\nu}(y) \rangle_{\tau_S}$ that give a one-world picture of reality in Kent's theory.

We can see how the formula (3.4.2) relates to the Schrödinger's cat scenario. The distribution of light reflected off the cat that intersects $S^1(y)$ when "measured" will determine a definite statement $r(\tau_S, y)$ about the mass-energy density on $S^1(y)$. This in turn will determine the range of τ for which $P(q(\tau)|r(\tau_S, y))$ is not close to zero, and hence where the stress-energy distribution $\langle T^{\mu\nu}(y) \rangle_{\tau_S}$ is not zero. This stress-energy distribution will then correspond either to that of a living cat or to that of a dead cat, but not both.

Coming back to the question of why we don't include any information about $\tau_S(x)$ for $x \in S$ from within the light cone of y, we need to consider in more detail how we would calculate $\langle T^{\mu\nu}(y) \rangle_{\tau_S}$. From (3.4.1) and (3.4.2), we will be able to perform this calculation so long as we can calculate $P(q(\tau) \,\&\, r(\tau_S, y))$ and $P(r(\tau_S, y))$.

Calculating $P(r(\tau_S, y))$ is relatively straightforward. As described on page 136, we can find an orthonormal basis $\{\, |\Gamma_i\rangle : i \in \mathbb{N} \,\}$ of H_S consisting of simultaneous $\hat{T}_S$-eigenstates and simultaneous $\hat{T}_S$-eigenvalues $\tau_{S,i}$ respectively. The probability

<hr />

46. Strictly speaking, we should say that $r(\tau_S, y)$ is the statement that the approximation of the mass-energy density $T_S(x)$ given by equation (3.2.5) has the value $\tau_S(c_x)$ for every cell c_x in S outside the light cone of y.

47. Again, we assume that $T^{\mu\nu}(y)$ is averaged over a small three-dimensional cell c_y of spacelike separated spacetime locations (with $y \in c_y$), and approximated to a finite pool of values as in equation (3.2.5).

$P(r(\tau_S, y))$ will then be

$$P(r(\tau_S, y)) = \sum_{\substack{i \text{ such that} \\ \tau_{S,i}(x) = \tau_S(x) \\ \text{for all } x \in S^1(y)}} |\langle \Gamma_i | U_{SS_0} | \Psi_0 \rangle|^2 \qquad (3.4.4)$$

where we have used equation (3.2.9).

But calculating $P(q(\tau) \& r(\tau_S, y))$ is a bit more involved because in the Tomonaga-Schwinger picture, the definition of observables via

$$\hat{O}(x) = U[S]\hat{\boldsymbol{O}}(x)U[S]^{-1} \qquad (3.2.2 \text{ revisted})$$

requires that $x \in S$. This means that we can't define $\hat{T}^{\mu\nu}(y)$ according to (3.2.2) since $y \notin S$.[48] However, we do not face such restrictions in the Heisenberg picture, so one approach would be to calculate $P(q(\tau) \& r(\tau_S, y))$ in the Heisenberg picture. As we will see shortly, this is not the approach that Kent takes, but nevertheless, in the Heisenberg picture, it is easier to see why we don't include information from S within the light cone (without begging the question of why we don't) when calculating $P(q(\tau) \& r(\tau_S, y))$. To see why this is so, consider the simpler case of just two measurable quantities F and G which we assume to have a discrete range of possible values and for which we wish to calculate the joint probability $P((F = f) \& (G = g))$. To do this in the Heisenberg picture, we need an orthonormal basis of the state space $\{ |\Phi_i\rangle : i \in \mathbb{N} \}$ consisting of simultaneous eigenstates of the observables $\hat{\boldsymbol{F}}$ and $\hat{\boldsymbol{G}}$ with eigenvalues f_i

---

48. If we did attempt to use (3.2.2) to define $\hat{T}^{\mu\nu}(y) = U[S]\hat{\boldsymbol{T}}^{\mu\nu}(y)U[S]^{-1}$, then $\hat{T}^{\mu\nu}(y)$ would have a (non-local) dependence on S, and such a dependence would not be desirable.

and g_i respectively so that $\hat{\boldsymbol{F}}\,|\Phi_i\rangle = f_i\,|\Phi_i\rangle$ and $\hat{\boldsymbol{G}}\,|\Phi_i\rangle = g_i\,|\Phi_i\rangle$, and when the system is in the state $|\Phi_i\rangle$, the quantity F will have the value f_i, and the quantity G will have the value g_i. Given that the system is in the state $|\Phi\rangle$, the joint probability $P((F = f)\,\&\,(G = g))$ can then be calculated using the Born rule to get

$$P((F = f)\,\&\,(G = g)) = \sum_{\substack{i \text{ such that} \\ f_i = f \text{ and } g_i = g}} |\langle \Phi_i | \Phi \rangle|^2.$$

But in order for such an orthonormal basis to exist, it is necessary that $\hat{F}$ and $\hat{G}$ commute.[49] This means that if $\hat{F}$ and $\hat{G}$ do not commute, then we cannot define the joint probability $P((F = f)\,\&\,(G = g))$.

Now quantum field theory is so constructed that $\hat{\boldsymbol{T}}^{00}(x)$ and $\hat{\boldsymbol{T}}^{\mu\nu}(y)$ will not commute when x and y are not spacelike separated, but $\hat{\boldsymbol{T}}^{\mu'\nu'}(x)$ and $\hat{\boldsymbol{T}}^{\mu\nu}(y)$ will commute when x and y are spacelike separated.[50] As we will see on page 195, $T_S(x)$ will have a $T^{00}(x)$ component, and so we can only be sure that $\hat{\boldsymbol{T}}_S(x)$ will commute with $\hat{\boldsymbol{T}}^{\mu\nu}(y)$ if x and y are spacelike separated. In other words, $\hat{\boldsymbol{T}}_S(x)$ and $\hat{\boldsymbol{T}}^{\mu\nu}(y)$ will commute if x is outside the light cone of y. Extending this argument to multiple

---

49. This is because given such an orthonormal basis $\{\,|\Phi_i\rangle : i \in \mathbb{N}\}$ of simultaneous eigenstates of $\hat{\boldsymbol{F}}$ and $\hat{\boldsymbol{G}}$, we have
$$\hat{\boldsymbol{F}}\hat{\boldsymbol{G}}\,|\Phi_i\rangle = f_i g_i\,|\Phi_i\rangle = g_i f_i\,|\Phi_i\rangle = \hat{\boldsymbol{G}}\hat{\boldsymbol{F}}\,|\Phi_i\rangle$$
so for any arbitrary state $|\Phi\rangle = \sum_i c_i\,|\Phi_i\rangle$, we have
$$\hat{\boldsymbol{F}}\hat{\boldsymbol{G}}\,|\Phi\rangle = \sum_i c_i \hat{\boldsymbol{F}}\hat{\boldsymbol{G}}\,|\Phi_i\rangle = \sum_i c_i \hat{\boldsymbol{G}}\hat{\boldsymbol{F}}\,|\Phi_i\rangle = \hat{\boldsymbol{G}}\hat{\boldsymbol{F}}\,|\Phi\rangle\,.$$

50. The proof of this statement need not concern us, but one can see that this is the case by considering the four potential commutation relations and the decomposition of the stress-energy

$x \in S$, we see that we can only guarantee that the conditional expectation of $T^{\mu\nu}(y)$

is definable if we restrict our conditioning on the value of $T_S(x)$ to $x \in S^1(y)$, that is,

to x in S outside the light cone of y.

Having explained why we don't include any information about $\tau_S(x)$ for $x \in S$ from

within the light cone of y, we can now proceed to calculate $P(q(\tau) \,\&\, r(\tau_S, y))$. Now

although there is no hypersurface that contains both y and $S^1(y)$, Kent gets around

this problem by noting that we can find a sequence of hypersurfaces $S_n(y)$ each of which

contains[51] y such that $S_n(y) \subset S_{n'}(y)$ for $n < n'$, and such that for any $x \in S^1(y)$,

there exists n and an open subset $U_n(x) \subset S$ containing x such that $U_n(x) \subset S_n(y)$.[52]

It will also be convenient to require that $S \setminus S_n(y)$ is bounded. An example of one

such $S_n(y)$ is shown in figure 3.9. When there is no ambiguity, we will drop the y and

write S_n instead of $S_n(y)$. Such a sequence S_n of hypersurfaces will be sufficient to

calculate $P(q(\tau) \,\&\, r(\tau_S, y))$. So let us define r_n to be the statement that $T_S(x)$ has

the determinate value $\tau_S(x)$ for all $x \in S_n \cap S$.[53] We recall that in the Tomonaga-

Schwinger formulation of relativistic quantum physics, the operators $\hat{T}_S(x)$ and $\hat{T}^{\mu\nu}(y)$

---

tensors in terms of the four-potentials – see Schwinger, "Quantum Electrodynamics. I. A Covariant Formulation," p. 1443–1444.

51. More precisely, we should say each hypersurface $S_n(y)$ contains c_y where c_y is the cell containing y mentioned in footnote 47 on page 159. We make c_y sufficiently small so that the c_x cells of S outside the light cone of y are identical to the c_x cells of S outside the light cone of c_y. We can do this on the assumption that c_y and the c_x are closed sets, since outside the light cone of y is an open set, so none of the c_x will touch the boundary of the light cone in S.

52. For Kent's description of this limiting process, see Kent, "Lorentzian Quantum Reality: Postulates and Toy Models," 2.

53. Strictly speaking, we should say the condition of r_n holds for all $x \in S_n \cap S$ at which S_n and S are tangential to each other. For a possible worry someone might have about the statement r_n
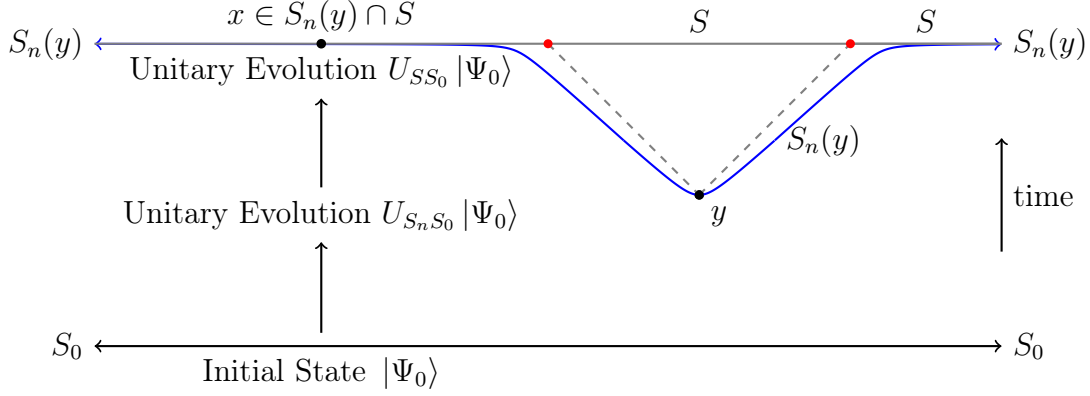
Figure 3.9. $S_n \overset{\text{def}}{=} S_n(y)$ is a hypersurface containing y and all of $S^1(y)$ in the limit as $n \to \infty$.

for fixed μ, ν commute when x and y are spacelike-separated. It therefore follows

that we can express any state of H_{S_n} as a superposition of simultaneous eigenstates

of $\hat{T}^{\mu\nu}(y)$ and $\hat{T}_S(x)$ for $x \in S_n \cap S$.[54] For a particular choice of μ, ν, we can then

---

without this qualification is that $\tau_S(x)$ is the value of the beable $T_S(x)$ for x in $S^1(y) \cap S$, but it's not the value of the beable $T_{S_n}(x)$ for x in $S_n \cap S$. Such a worry would be valid if the beable $T_S(x)$ depended on the whole of S and the beable $T_{S_n}(x)$ depended on the whole of S_n. However, as we shall see on page 196 in section 4.2, the physical quantity $T_S(x)$ which is defined by equation (4.2.11) to be $T_S(x) = T^{\mu\nu}(x)\eta_\mu(x)\eta_\nu(x)$ will only have a local dependence on S via the future directed four-vector $\eta^\mu(x)$. Therefore, so long as the future directed four-vector for S_n at x is the same as the one for S at x, then the beables $T_S(x)$ and $T_{S_n}(x)$ will be identical. We therefore require that S_n and S are tangential to each other at x since this is a necessary and sufficient condition for the respective future directed four-vectors of S_n and S to be identical.

Nevertheless, we might still worry that the observables $\hat{T}_S(x)$ and $\hat{T}_{S_n}(x)$ corresponding to these two beables aren't identical because $\hat{T}_S(x)$ acts on the Hilbert space H_S whereas $\hat{T}_{S_n}(x)$ acts on the Hilbert space H_{S_n}. However, at this point we need to recall footnote 25 on page 134 that H_S and H_{S_n} are really the same Hilbert space, but just interpreted differently. Now on this one Hilbert space, it turns out that $\hat{T}_S(x)$ and $\hat{T}_{S_n}(x)$ are identical. To see why this is, let $\hat{\boldsymbol{T}}_S(x)$ and $\hat{\boldsymbol{T}}_{S_n}(x)$ be the Heisenberg picture observables. Since $T_S(x) = T^{\mu\nu}(x)\eta_\mu(x)\eta_\nu(x) = T_{S_n}(x)$ for $x \in S_n \cap S$ where S_n and S are tangential to one another, we must have $\hat{\boldsymbol{T}}_S(x) = \hat{\boldsymbol{T}}^{\mu\nu}(x)\eta_\mu(x)\eta_\nu(x) = \hat{\boldsymbol{T}}_{S_n}(x)$. Now by definition (see equation (3.2.2)), $\hat{T}_S(x) = U[S]\hat{\boldsymbol{T}}_S(x)U[S]^{-1}$, and $\hat{T}_{S_n}(x) = U[S_n]\hat{\boldsymbol{T}}_{S_n}(x)U[S_n]^{-1}$. But as Schwinger shows, under conditions that are readily satisfied (see footnote 24 on page 134 for details), for any Heisenberg operator $\hat{\boldsymbol{F}}(x)$, as long as x belongs to S the operator $\hat{F}(x) = U[S]\hat{\boldsymbol{F}}(x)U[S]^{-1}$ is independent of S. Therefore, since $x \in S_n \cap S$ where S_n and S are tangential to one another, we not only have $\hat{\boldsymbol{T}}_S(x) = \hat{\boldsymbol{T}}_{S_n}(x)$, but we must also have $\hat{T}_S(x) = \hat{T}_{S_n}(x)$.

54. Strictly speaking we should say simultaneous eigenstates of $\hat{T}^{\mu\nu}(c_y)$ and $\hat{T}_S(c_x)$ for all $c_x \subset S_n \cap S$.

form an orthonormal basis $\{ |\Gamma_{n,i}^{\mu\nu}\rangle : i \in \mathbb{N} \}$ of H_{S_n} consisting of simultaneous $\hat{T}^{\mu\nu}(y)$,

$\hat{T}_S(x)$-eigenstates so that $\hat{T}^{\mu\nu}(y) |\Gamma_{n,i}^{\mu\nu}\rangle = \tau_i |\Gamma_{n,i}^{\mu\nu}\rangle$ and $\hat{T}_S(x) |\Gamma_{n,i}^{\mu\nu}\rangle = \tau_{S,i}(x) |\Gamma_{n,i}^{\mu\nu}\rangle$ for

$x \in S_n \cap S$, where τ_i and $\tau_{S,i}(x)$ are the corresponding eigenvalues. Note that we

consider here a real number τ_i here rather than a tensor $\tau_i^{\mu\nu}$ indexed by μ and ν since

$\hat{T}^{\mu\nu}(y)$ will not in general commute for different values of μ, ν, so we won't be able to

find a state which is a simultaneous eigenstate for all the different observables $\hat{T}^{\mu\nu}(y)$,

though we may find a state which is very close to being an eigenstate of all the $\hat{T}^{\mu\nu}(y)$

for different μ and ν.[55]

The probability $P(q(\tau) \,\&\, r_n)$ will be given by

$$P(q(\tau) \,\&\, r_n) = \sum_{\substack{i \text{ such that } \tau_i = \tau \\ \text{and } \tau_{S,i}(x) = \tau_S(x) \\ \text{for all } x \in S_n \cap S}} \left| \left\langle \Gamma_{n,i}^{\mu\nu} \middle| U_{SS_0} \middle| \Psi_0 \right\rangle \right|^2. \tag{3.4.5}$$

Taking the limit as n tends to infinity, we can calculate the probability $P(q(\tau) \,\&\, r(\tau_S, y))$

to be

$$P(q(\tau) \,\&\, r(\tau_S, y)) = \lim_{n \to \infty} P(q(\tau) \,\&\, r_n).^{56} \tag{3.4.6}$$

Assuming τ_S is selected in accordance with the Born rule so that $P(r) > 0$, we can plug

(3.4.4) and (3.4.6) into (3.4.1) to calculate the conditional probability $P(q(\tau)|r(\tau_S, y))$,

---

55. Equal time commutation relations for the stress-energy tensor of an arbitrary Lorentz-invariant physical system are calculated in K. Bergström, "Equal-Time Commutation Relation for the Energy-Momentum Tensor," *Journal of mathematical physics* 11, no. 8 (1970): 2498–2501.

56. In fact, there will be a finite n' such that $P(q(\tau) \,\&\, r(\tau_S, y)) = P(q(\tau) \,\&\, r_{n'})$. This is because for any n we are assuming $S \setminus S_n$ is bounded, so there will be a finite number of c_x cells of S outside the

and hence calculate the conditional expectation $\langle T^{\mu\nu}(y)\rangle_{\tau_S} \overset{\text{def}}{=} \langle T^{\mu\nu}(y)\rangle_{r(\tau_S,y)}$ via equation (3.4.2). We thus obtain

$$\langle T^{\mu\nu}(y)\rangle_{\tau_S} = \sum_{\tau} P(q(\tau)|r(\tau_S,y))\tau = \lim_{n\to\infty} \sum_{\tau} \frac{P(q(\tau)\,\&\,r_n)\tau}{P(r_n)}. \qquad (3.4.7)$$

In section 4.1, we will give a more detailed description of this calculation in order to show that the predictions Kent's theory makes are consistent with standard quantum theory.

3.5 Kent's Toy Model Example

To get a feel for how all the elements of Kent's theory fit together, we will conclude this chapter by describing Kent's toy model example that he discusses in his 2014 paper.[57] In his toy model, Kent considers a system in one spatial dimension which is the superposition of two localized states/wave functions[58] $\psi_0^{\text{sys}} = c_1\psi_1^{\text{sys}} + c_2\psi_2^{\text{sys}}$ where ψ_1^{sys} is localized at spatial location z_1, ψ_2^{sys} is localized at spatial location z_2,

---

light cone of c_y that are not contained in S_n, and so the union U of all these cells will be compact. But for any $x \in U$, we can find n'' such that the open set $U_{n''}(x)$ containing x with $U_{n''}(x) \subset S$ and $U_{n''}(x) \subset S_n$. These $U_{n''}(x)$ will then form an open cover of U, and by the definition of compactness, every open cover has a finite subcover. If we therefore choose n' to be the maximum n'' of this finite subcover, then $S_{n'}$ will contain all of U since $S_{n''} \subset S_{n'}$ for $n'' < n'$. Then by definition of the statements $r(\tau_S,y)$ and r_n, it follows that $r(\tau_S,y) = r_{n'}$.

57. See Kent, "Lorentzian Quantum Reality: Postulates and Toy Models," p. 3–4.

58. So far in this chapter, we have been describing systems in terms of their quantum states rather than their *quantum wave functions*. It is easiest to understand what a quantum wave function is in the context of a single particle system. In the Schrödinger picture, a particle in state $|\psi(t)\rangle$ allows us to calculate the expectation value of a quantity O belonging to the particle via the formula $\langle\psi(t)|\hat{O}|\psi(t)\rangle$ where $\hat{O}$ is the observable corresponding to O. One such quantity is the particle's spatial location. If the particle is at spatial location $\boldsymbol{x}$, then the particle will be in the state $|\boldsymbol{x}\rangle$ (c.f. the definition of $|x\rangle$ in footnote 49 on page 91.) where the position observable $\hat{\boldsymbol{X}}_i$ satisfies $\hat{\boldsymbol{X}}_i|\boldsymbol{x}\rangle = x_i|\boldsymbol{x}\rangle$ for $i = 1, 2,$ or 3.

and $|c_1|^2 + |c_2|^2 = 1$. It is also assumed that ψ_1^{sys} and ψ_2^{sys} do not overlap at all, so that if it is determined that the mass is localized around z_1, then it isn't localized around z_2, and vice versa. According to the Copenhagen interpretation, a measurement on this system would collapse the wave function of ψ_0^{sys} to the wave function of ψ_1^{sys} with probability $|c_1|^2$, and to the wave function of ψ_2^{sys} with probability $|c_2|^2$. The purpose of Kent's toy model is to show that within his interpretation, there is something analogous to wave function collapse. In order for this "collapse" to happen, one needs to consider how the system interacts with light. Thus, Kent supposes that a photon (which is modelled as a point particle) comes in from the left, and as it interacts with the two states ψ_1^{sys} and ψ_2^{sys}, the photon enters into a superposition of states, corresponding to whether the photon reflects off the localized ψ_1^{sys}-state at time t_1 or the localized ψ_2^{sys}-state at time t_2. The photon in superposition then travels to the left and eventually reaches the one dimensional hypersurface S at locations γ_1 and γ_2 as shown in figure 3.10.

We now suppose that when the mass-energy density S is "measured", the energy of the photon is found to be at γ_1 rather than at γ_2. We then consider the mass-density at early spacetime locations $y_1^a = (z_1, t_a)$ and $y_2^a = (z_2, t_a)$ as shown in figure 3.11 (a) and (b).

By early, we mean that $t_a < 2t_1 - t_2$. This will mean that the possible detection locations γ_1 and γ_2 will be inside the forward light cones of y_1^a and y_2^a. Hence, $S^1(y_1^a) \cap S$ and $S^1(y_2^a) \cap S$ contain no additional information beyond standard quantum theory

The corresponding wave function for this particle is then $\psi(\boldsymbol{x}, t) = \langle \boldsymbol{x}|\psi(t)\rangle$. For a particle restricted
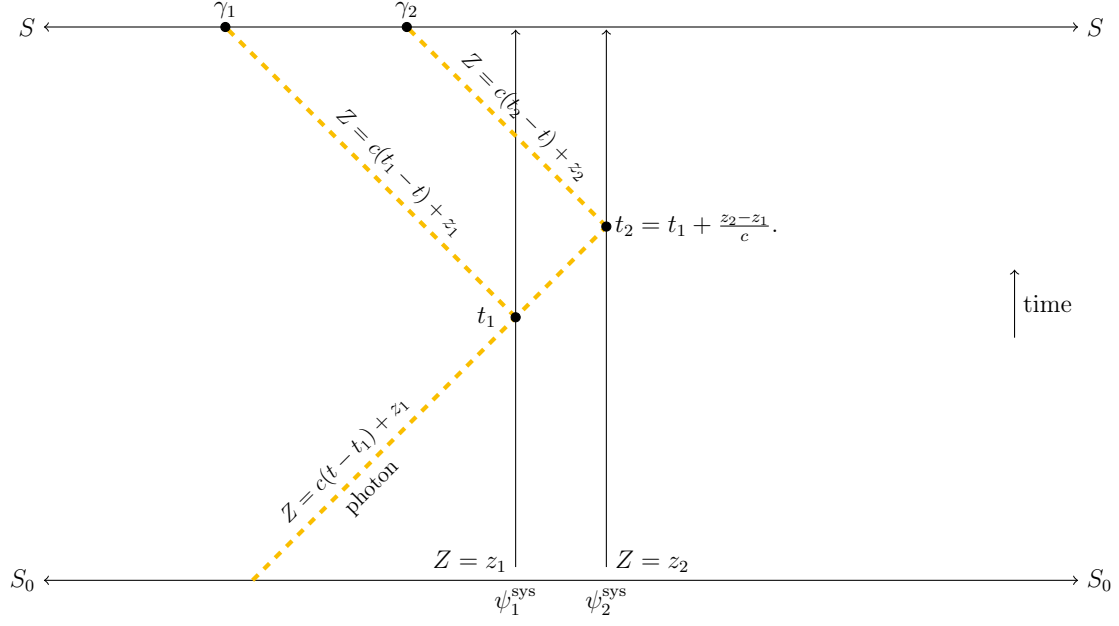
Figure 3.10. Kent's toy model
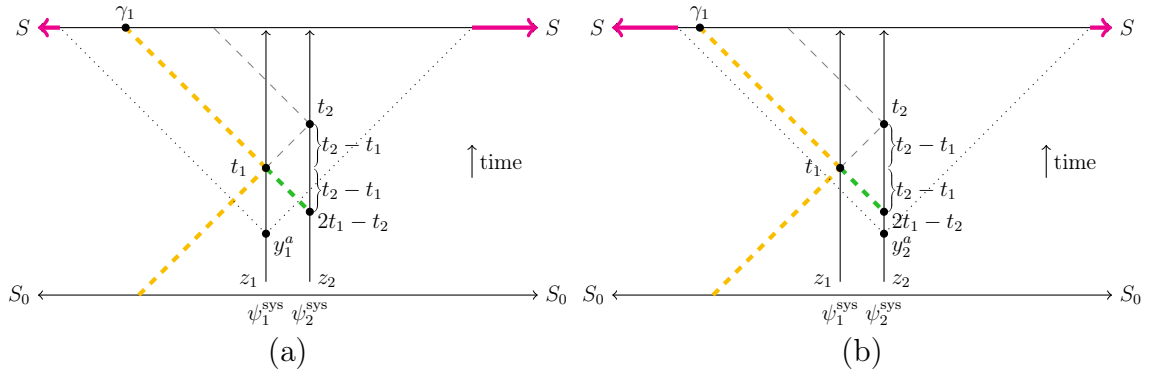
Figure 3.11. (a) highlights the part of S used to calculate the energy density at y_1^a whose time is less than $2t_1 - t_2$. (b) highlights the part of S used to calculate the energy density at y_2^a whose time is less than $2t_1 - t_2$.

by which we could calculate the conditional expectation values of the energy at y_1^a and y_2^a. Hence, according to Kent's theory, the total energy at time t_a will be divided between the two spatial locations with a proportion of $|c_1|^2$ at z_1 and a proportion of $|c_2|^2$ at z_2.

However, the situation is different for two spacetime locations $y_1^b = (z_1, t_b)$ and $y_2^b = (z_2, t_b)$ with t_b slightly after $2t_1 - t_2$ as depicted in figure 3.12.
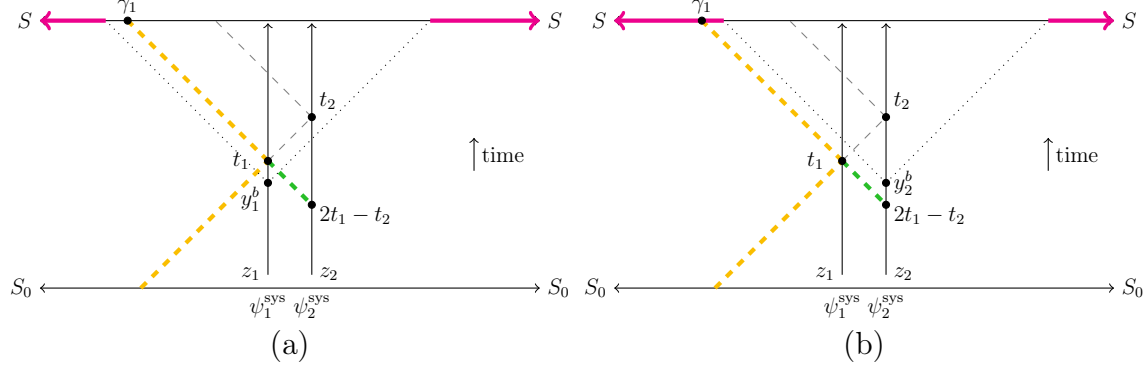
Figure 3.12. (a) highlights the part of S used to calculate the energy density at y_1^b whose time is greater than $2t_1 - t_2$. (b) highlights the part of S used to calculate the energy density at y_2^b whose time is greater than $2t_1 - t_2$.

In this situation, when we consider the location y_1^b, there is no additional information in $S^1(y_1^b) \cap S$ beyond standard quantum theory, so there will be a proportion of $|c_1|^2$ of the total initial energy of the system at y_1^b. But at location y_2^b, the information in $S^1(y_2^b) \cap S$ shows that the photon has reflected from the localized ψ_1^{sys}-state, and so this additional information tells us that after time t_b, there is no energy localized at z_2 since from the perspective of y_2^b, the energy is known to be localized at z_1. So it is as though the information of $S^1(y_2^b) \cap S$ has determined that we are in a world in which there is an energy density of zero at y_2^b, and there are no other worlds in which the energy density at y_2^b is non-zero since all worlds have to be consistent with the

---

to one spatial dimension, the particle would have a wave function $\psi(z, t) = \langle z | \psi(t) \rangle$ where z is now just a single number that specifies the particle's possible position. We will write ψ to denote the wave function itself, and $\psi(z, t)$ to denote the value of the wave function ψ at spacetime location (z, t).

notional measurement made on S. So for a short time the total energy of the system is reduced by $|c_1|^2$.

However, as shown in figure 3.13, for times t_c greater than t_1, the total energy of the system is once again restored to the initial energy the system had when in the state ψ_0^{sys}.
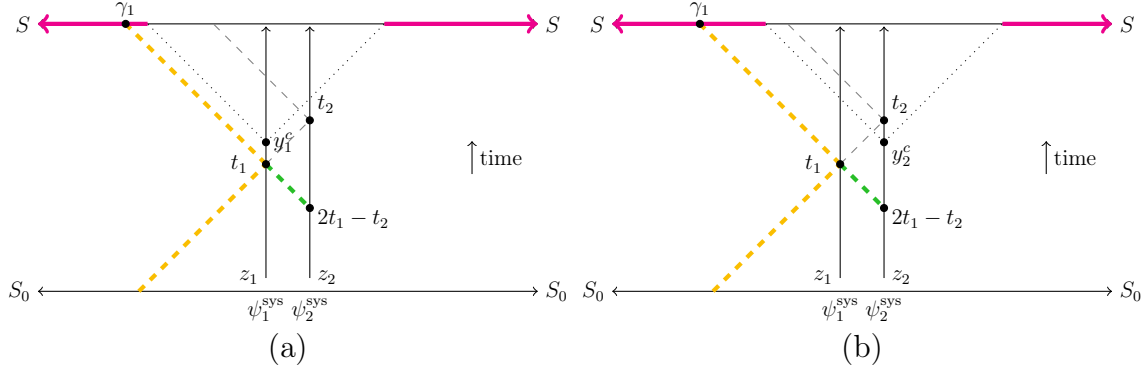
Figure 3.13. (a) highlights the part of S used to calculate the energy density at y_1^c whose time is greater than t_1. (b) highlights the part of S used to calculate the energy density at y_2^c whose time is greater than t_1.

In this situation, there is now information in $S^1(y_1^c) \cap S$ that determines that the photon reflected off the localized ψ_1^{sys}-state. This means that when the conditional expectation of the energy density of y_1^c is calculated, the extra information in $S^1(y_1^c) \cap S$ determines that all the energy of the system is located at location z_1 for times t_c greater than t_1, and the energy is equal to the initial energy of the system so that energy is conserved.

CHAPTER FOUR

Evaluating Kent's Theory

In order to evaluate Kent's theory of quantum physics, it will first be helpful to remind ourselves of the problem we are trying to solve.

In chapter 1, we discussed the EPR-Bohm paradox and the problem of explaining the mysterious correlations between measurement outcomes of spin singlets in a way consistent with special relativity and the predictions of standard quantum theory. We saw that the Copenhagen interpretation does not seem to be consistent with special relativity. We also discussed Shimony's distinction between Outcome Independence (OI) and Parameter Independence (PI) and Shimony's idea that we should only accept a theory in which OI is false and PI is true. Since PI is false in the pilot-wave theory, we should reject it according to Shimony's criterion.

But although Shimony's criterion is a promising line of inquiry, by itself it is not sufficient to resolve the EPR-Bohm paradox. This is because Shimony's criterion doesn't address the controversial issue of what is meant by an outcome. In chapter 2, we discussed this controversy over outcomes and why the many-worlds interpretation that denies the reality of outcomes is unsatisfactory. This motivated the discussion of Kent's theory in chapter 3 in the hope that it might provide a satisfactory solution to the EPR-Bohm paradox. In the previous chapter, we only got as far as describing

the key features of Kent's theory such as it being a one-world theory which posits additional variables to standard quantum theory.

So having now reminded ourselves of the problem at hand, we see that there are several issues that we need to consider in order to evaluate whether Kent's theory provides a satisfactory solution to this problem. Firstly, we should consider whether the predictions of Kent's theory are consistent with the predictions of standard quantum theory. Since standard quantum theory (that is a theory of states whose evolution is determined by the Schrödinger equation) predicts the correlations that have been experimentally observed in the EPR-Bohm paradox, then if Kent's theory makes the same predictions as standard quantum theory, these EPR-Bohm correlations will also be exhibited in Kent's theory. In section 4.1, I will show that Kent's theory agrees with standard quantum theory given the universal quantum state $|\Psi_n\rangle$ on a hypersurface S_n. However, since in practice, scientists do not base their calculations on universal quantum states, in section 4.6, I will explain how to extract what I call a conditioned quantum state from $|\Psi_n\rangle$, and I will show that Kent's theory agrees with standard quantum theory when we work with conditioned quantum states.

In section 4.2, I will consider Lorentz invariance. The importance of Lorentz invariance lies in the fact that a necessary condition for a theory to be consistent with special relativity is that it be invariant under a group of symmetries called Lorentz transformations. Since a satisfactory solution to our problem must be consistent with special relativity, we therefore need to consider whether Kent's theory satisfies Lorentz invariance.

171

In section 4.3, I will consider the Problem of Outcomes in the light of Kent's theory. Since a satisfactory theory must be one in which there are outcomes, we need to consider whether Kent succeeds in giving us a convincing account of what an outcome is. In doing this, I will examine how Kent's theory ties in with decoherence theory and d'Espagnat's objection about improper mixtures.

Several sections will be needed to address Parameter Independence in Kent's theory. In section 1.9, we saw that if OI is true, then PI must be false. However, Butterfield, in his extension of Kent's toy model argues that OI holds in Kent's theory. This conclusion could therefore seem rather worrying given that we want PI to be true. This is a worry that I will address in section 4.4.

Butterfield has also emphasized the need to understand Kent's theory in the light of an important theorem proved in recent years, the so-called Colbeck-Renner Theorem.[1] This theorem states that if a theory satisfies PI together with what is called a 'no conspiracy' criterion, then this theory is reducible to standard quantum theory without any hidden variables. This result could seem rather worrying since in the absence of any additional variables, standard quantum theory is unable to resolve the EPR-Bohm paradox. This suggests that any theory that satisfactorily addresses the EPR-Bohm paradox can't be a hidden-variables theory according to the assumptions

1. See Gijs Leegwater, "An impossibility theorem for parameter independent hidden-variable theories," *Studies in History and Philosophy of Modern Physics* 54 (2016): 18–34, Roger Colbeck and Renato Renner, "No extension of quantum theory can have improved predictive power," *Nature communications* (England) 2, no. 1 (2011): 411–411, Roger Colbeck and Renato Renner, "The completeness of quantum theory for predicting measurement outcomes," 2012, K Landsman, "On the Colbeck–Renner theorem," *Journal of mathematical physics* (United States) 56, no. 12 (2015): 122103, and Klaas Landsman, *Foundations of Quantum Theory : From Classical Concepts to Operator Algebras (Volume 188.0)*, vol. 188, Fundamental Theories of Physics (Cham: Springer Open, 2017).

of the Colbeck-Renner theorem. In section 4.5, I will therefore consider whether it would be both possible and acceptable to relax or drop any of these assumptions in the context of Kent's theory.

Having addressed the concerns Butterfield raises, I will argue in section 4.7 that when we adapt Kent's theory to use conditioned quantum states, PI is true. The conclusion of this chapter is that we therefore have good reason to think that Kent's theory provides a satisfactory solution to the EPR-Bohm paradox.

There are nevertheless, some remaining questions concerning Kent's theory. Firstly, there is the question of how plausible it is for the beables in reality to be what Kent suggests they are. Secondly, Kent's theory may also strike us as rather counter-intuitive given that his theory posits that present events should be conditioned on far-distant future states of affairs. There is also the question of whether Kent's theory is deterministic. I will therefore conclude this chapter with section 4.8 where I discuss these questions.

Since the arguments in this chapter are rather technical, several sections are divided into two subsections where the first subsection will contain a non-technical overview, and the second subsection will contain the technical details and can be skipped if desired.

4.1 Consistency of Kent's Theory with Standard Quantum Theory

4.1.1 Overview

If we are to take Kent's theory seriously, it should be just as good at making predictions as standard quantum theory, and it had better not contradict empirical observations. Over the last century, standard quantum theory has been firmly established, and so far, it has not been contradicted by any experimental observations. Standard quantum theory allows us to form a quantum state description of a physical system based on how the system was set up in an experimental environment, and then Schrödinger's equation can be used to evolve this state forwards in time, and finally, we can calculate expectation values for various physical quantities belonging to this physical system, and these agree with the average values measured on the system when the experiment is performed many times. In other words, standard quantum theory is *empirically adequate*[2] in its domain of applicability. Thus, if we can show that Kent's theory is just as good at making predictions as standard quantum theory, then it too will be empirically adequate to the same degree. This doesn't necessarily mean that Kent's theory will make exactly the same predictions as standard quantum theory, for the additional information beyond standard quantum theory that Kent's theory requires may alter these predictions. Indeed, if this additional information made absolutely no difference to the predictions of standard quantum theory, then it would seem rather redundant. But we should nevertheless be able to derive the

---

2. See p. 229 for a more formal definition of empirical adequacy.

predictions of standard quantum theory from Kent's theory by averaging over the unknown variables that describe the additional information in Kent's theory.

In order to show that Kent's theory is just as good as standard quantum theory, I will consider two cases. In this section I will focus on the predictions Kent's theory makes given the universal state $|\Psi_0\rangle$. I will postpone until section 4.6 a consideration of the predictions Kent's theory makes given a state that is more local and relevant to the kinds of situations that concern experimentalists.

In this current section I will show that if we calculate Kent's beable $\langle T^{\mu\nu}(y)\rangle_{\tau_S}$ for every possible measurement outcome that τ_S could be, and if we then take the weighted sum of each of these beables weighted according to how likely τ_S was to occur given $|\Psi_0\rangle$, then this weighted sum will give the same value $\langle T^{\mu\nu}(y)\rangle$ that standard quantum theory would predict given $|\Psi_0\rangle$. The intuition is that Kent's theory posits the existence of additional unknown variables, namely those specified by a particular outcome τ_S, and if we had known this information, we would be able to calculate the expectation value of the stress-energy tensor $T^{\mu\nu}(y)$ more accurately than standard quantum theory would allow. However, in reality, we don't know what this extra information is; we can only work out how likely the values constituting this extra information will be based on the state $|\Psi_0\rangle$ and the use of standard quantum theory. Therefore, the best we can do is to guess what this extra information is going to be and then weight the calculation we do with this extra information according to the likelihood of our guess being correct. But since we believe standard quantum theory is the best theory there is given the information in the quantum state $|\Psi_0\rangle$, and since

we have no information beyond the quantum state when performing this weighted calculation, we wouldn't expect the result of the weighted calculation to be any better than what standard quantum theory predicts. But if we are to take Kent's theory seriously, then Kent's theory shouldn't be any worse than standard quantum theory either. Thus, by showing that the weighted calculation of expectation values that Kent's theory predicts is equal to the expectation values standard quantum theory predicts, we see that Kent's theory is no worse than standard quantum theory.

To show that this is indeed the case, we consider a hypersurface S_n that contains the point y at which we wish to calculate the stress-energy tensor's expectation value. Now, given the universal quantum state $|\Psi_0\rangle$ describing the initial hypersurface S_0, it is possible to work out how this quantum state evolves to a quantum state $|\Psi_n\rangle$ describing the hypersurface S_n. The hypersurface S_n is assumed to intersect the distant future hypersurface S where the notional mass-energy density measurement is made.

Now we would not expect $|\Psi_n\rangle$ to describe the hypersurface S_n as having a definite mass-energy density at the locations where S_n intersects S. But according to standard quantum theory, we should be able to express $|\Psi_n\rangle$ as a superposition (that is, a weighted sum) of states for which the mass-energy density on $S \cap S_n$ does have a definite value. In particular, we can express $|\Psi_n\rangle$ as a sum of two states: the first state which we denote as $\pi_n |\Psi_n\rangle$ is the (weighted) sum of all the states in the superposition of $|\Psi_n\rangle$ for which the mass-energy density has a definite value that is equal to the notional mass-energy density measurement $\tau_S(x)$ for all x in the intersection $S \cap S_n$;

and the second state which we denote as $(I - \pi_n) |\Psi_n\rangle$ is the (weighted) sum of all the other states included in the superposition of $|\Psi_n\rangle$. Since all the states that contribute to the sum defining $\pi_n |\Psi_n\rangle$ are mutually exclusive, the probability given $|\Psi_n\rangle$ (and hence the probability given $|\Psi_0\rangle$) that the notional mass-energy density measurement made on $S \cap S_n$ is equal to τ_S will be equal to the sum of the probabilities of the states that contribute to $\pi_n |\Psi_n\rangle$. This fact enables us to calculate a simple formula[3] for the probability given $|\Psi_n\rangle$ that the notional mass-energy density measurement made on $S \cap S_n$ is equal to τ_S. Similarly, we can also calculate a simple formula[4] for the probability given $|\Psi_n\rangle$ that both the notional mass-energy density measurement made on $S \cap S_n$ is equal to τ_S and that the stress-energy tensor $T^{\mu\nu}(y)$ at the location y takes a particular value τ. From these two formulae, we can calculate the conditional probability that the $T^{\mu\nu}(y)$ takes a particular value τ given that the notional mass-energy density measurement made on $S \cap S_n$ is equal to τ_S. This in turn enables us to calculate Kent's beable $\langle T^{\mu\nu}(y) \rangle_{\tau_S}$ at y.[5] With these formulae, it is then relatively easy to calculate the weighted sum of expectation values that Kent's theory predicts and show that this sum equals the expectation value predicted by standard quantum theory.

---

3. See equation (4.1.3) below.

4. See equation (4.1.5) below.

5. See equation (4.1.6) below.

4.1.2 Technical Details[*]

In order to show that Kent's theory is just as good as standard quantum theory given the universal state $|\Psi_0\rangle$ of the initial hypersurface S_0, recall the preliminary calculation of $\langle T^{\mu\nu}(y)\rangle_{\tau_S}$ given in section 3.4:

$$\langle T^{\mu\nu}(y)\rangle_{\tau_S} = \sum_{\tau} P(q(\tau)|r(\tau_S, y))\tau = \lim_{n\to\infty} \sum_{\tau} \frac{P(q(\tau)\,\&\,r_n)\tau}{P(r_n)}. \qquad \text{(3.4.7 revisited)}$$

We proceed to calculate $P(r_n)$ and $P(q(\tau)\,\&\,r_n)$ in order to calculate $\langle T^{\mu\nu}(y)\rangle_{\tau_S}$. We first note that since S_n is a hypersurface, there will exist a unitary operator $U_{S_n S_0}$ defined by equation (3.2.8) which maps the Hilbert space[6] of states H_{S_0} describing S_0 to the Hilbert space of states H_{S_n} describing S_n in accord with how the states of H_{S_0} evolve over time. Now let $H_{S_n, \tau_S} \subset H_{S_n}$ be the subspace of states $|\xi\rangle$ for which $\hat{T}_S(x)|\xi\rangle = \tau_S(x)|\xi\rangle$ for all $x \in S_n \cap S$, and let $\{|\xi_1\rangle, |\xi_2\rangle, \ldots\}$ be an orthonormal basis of H_{S_n, τ_S}. Given that the initial state of the world is $|\Psi_0\rangle$, the probability $P(r_n)$ of "measuring" the value of $T_S(x)$ on $S_n \cap S$ to be $\tau_S(x)$ will be

$$P(r_n) = \sum_{j} |\langle \xi_j | \Psi_n \rangle|^2, \qquad (4.1.1)$$

---

6. The Hilbert space H_S for any hypersurface S is defined on page 134.

where $|\Psi_n\rangle = U_{S_n S_0} |\Psi_0\rangle$, and this probability will be independent of the particular

orthonormal basis $\{\,|\xi_j\rangle : j \in \mathbb{N}\}$ of H_{S_n,τ_S}.[7] If we define

$$\pi_n = \sum_j |\xi_j\rangle\langle\xi_j|\,, \tag{4.1.2}$$

then it is easy to see that

$$P(r_n) = \langle\Psi_n|\pi_n|\Psi_n\rangle\,. \tag{4.1.3}$$

We also see that π_n is Hermitian (i.e. has real eigenvalues) and that $\pi_n \pi_n = \pi_n$. Any

Hermitian operator π with $\pi^2 = \pi$ is called a *projection*. We thus see that π_n is a

projection.

Turning to the calculation of $P(q(\tau)\,\&\,r_n)$, note that for the Tomonaga-Schwinger

formulation of relativistic quantum physics, the operators $\hat{T}_S(x)$ and $\hat{T}^{\mu\nu}(y)$ for fixed

μ, ν commute when x and y are spacelike-separated. It therefore follows that we can

express any state of H_{S_n} as a superposition of simultaneous eigenstates of $\hat{T}^{\mu\nu}(y)$ and

---

7. To see why this is, we note that we can extend the orthonormal set $\{\,|\xi_1\rangle, |\xi_2\rangle, \ldots\}$ to an orthonormal basis $\{\,|\xi_1\rangle, |\xi_2\rangle, \ldots\} \cup \{\,|\zeta_1\rangle, |\zeta_2\rangle, \ldots\}$ of H_{S_n} which consists entirely of $\hat{T}_S(x)$-eigenstates for all $x \in S_n \cap S$. We can think of each of the states of this orthonormal basis as the possible measurement outcome when making the notional measurement of $T_S(x)$ on $S_n \cap S$. By the Born Rule, it therefore follows that $P(r_n) = \sum_j |\langle\xi_j|\Psi_n\rangle|^2$. But to see that this probability is independent of the particular basis, we can uniquely write $|\Psi_n\rangle$ as a sum $|\Psi_n\rangle = |\xi\rangle + |\zeta\rangle$ where $|\xi\rangle$ belongs to the span of $\{\,|\xi_j\rangle : j \in \mathbb{N}\}$ and $|\zeta\rangle$ belongs to the span of $\{\,|\zeta_j\rangle : j \in \mathbb{N}\}$. Then since $|\xi\rangle = \sum_j \langle\xi_j|\Psi_n\rangle |\xi_j\rangle$, it follows that
$$\langle\xi|\xi\rangle = \sum_j |\langle\xi_j|\Psi_n\rangle|^2 = P(r_n).$$
Therefore, since $\langle\xi|\xi\rangle$ is independent of the particular basis chosen of H_{S_n,τ_S}, so is $P(r_n)$.

179

$\hat{T}_S(x)$ for $x \in S_n \cap S$.[8] For a particular choice of μ, ν, we can then form an orthonormal

basis $\{ |\eta_j^{\mu\nu}\rangle : j \in \mathbb{N} \}$ of H_{S_n} consisting of simultaneous $\hat{T}^{\mu\nu}(y)$, $\hat{T}_S(x)$-eigenstates so

that $\hat{T}^{\mu\nu}(y) |\eta_j^{\mu\nu}\rangle = \tau_j |\eta_j^{\mu\nu}\rangle$ and $\hat{T}_S(x) |\eta_j^{\mu\nu}\rangle = \tau_{S,j}(x) |\eta_j^{\mu\nu}\rangle$ for $x \in S_n \cap S$, where τ_j and

$\tau_{S,j}(x)$ are the corresponding eigenvalues.[9] If we define $\pi_{n,\tau}^{\mu\nu} = \sum_j |\chi_{j,\tau}^{\mu\nu}\rangle\langle\chi_{j,\tau}^{\mu\nu}|$ where

$\{ |\chi_{j,\tau}^{\mu\nu}\rangle : j \in \mathbb{N} \}$ is the subset of $\{ |\eta_j^{\mu\nu}\rangle : j \in \mathbb{N} \}$ such that $\hat{T}^{\mu\nu}(y) |\chi_{j,\tau}^{\mu\nu}\rangle = \tau |\chi_{j,\tau}^{\mu\nu}\rangle$ and

$\hat{T}_S(x) |\chi_{j,\tau}^{\mu\nu}\rangle = \tau_S(x) |\chi_{j,\tau}^{\mu\nu}\rangle$ for all $x \in S_n \cap S$, then

$$P(q(\tau) \,\&\, r_n) = \sum_j \left| \langle \chi_{j,\tau}^{\mu\nu} | \Psi_n \rangle \right|^2 = \langle \Psi_n | \pi_{n,\tau}^{\mu\nu} | \Psi_n \rangle .^{10} \qquad (4.1.4)$$

But if we define $\pi_\tau^{\mu\nu} = \sum_j |\eta_{j,\tau}^{\mu\nu}\rangle\langle\eta_{j,\tau}^{\mu\nu}|$ where $\{ |\eta_{j,\tau}^{\mu\nu}\rangle : j \in \mathbb{N} \}$ is the subset of $\{ |\eta_j^{\mu\nu}\rangle :$

$j \in \mathbb{N} \}$ with $\hat{T}^{\mu\nu}(y) |\eta_{j,\tau}^{\mu\nu}\rangle = \tau |\eta_{j,\tau}^{\mu\nu}\rangle$, then we also have $\pi_{n,\tau}^{\mu\nu} = \pi_n \pi_\tau^{\mu\nu}$.[11] Hence,

$$P(q(\tau) \,\&\, r_n) = \langle \Psi_n | \pi_n \pi_\tau^{\mu\nu} | \Psi_n \rangle . \qquad (4.1.5)$$

---

8. We make the same approximation as depicted in figure 3.3 on page 138.

9. We include the indices μ, ν in the specification of the basis $\{ |\eta_j^{\mu\nu}\rangle : j \in \mathbb{N} \}$ since in general, we can't assume that $\hat{T}^{\mu\nu}(y)$ and $\hat{T}^{\mu'\nu'}(y)$ commute for $\mu \neq \mu'$ or $\nu \neq \nu'$, and so we have to fix μ, ν and consider a basis $\{ |\eta_j^{\mu\nu}\rangle : j \in \mathbb{N} \}$ of H_n which depends on the particular μ, ν so that all the elements of this basis can be eigenstates of $\hat{T}^{\mu\nu}(y)$.

10. The proof of this is very similar to the proof given in footnote 7 on page 179.

11. To see why this is, we first show that $\pi_n = \sum_j |h_{n,j}^{\mu\nu}\rangle\langle h_{n,j}^{\mu\nu}|$ where $\{ |h_{n,j}^{\mu\nu}\rangle : j \in \mathbb{N} \}$ is the subset of $\{ |\eta_j^{\mu\nu}\rangle : j \in \mathbb{N} \}$ for which $|h_{n,j}^{\mu\nu}\rangle \in H_{S_n,\tau_S}$. Note that $\pi_n |h_{n,j}^{\mu\nu}\rangle = |h_{n,j}^{\mu\nu}\rangle$ since $\{ |\xi_j\rangle : j \in \mathbb{N} \}$ is a basis for H_{S_n,τ_S} and $|h_{n,j}^{\mu\nu}\rangle \in H_{S_n,\tau_S}$. Therefore, $\pi_n \pi_{n,h}^{\mu\nu} = \pi_{n,h}^{\mu\nu}$ where $\pi_{n,h}^{\mu\nu} = \sum_j |h_{n,j}^{\mu\nu}\rangle\langle h_{n,j}^{\mu\nu}|$. But $\pi_{n,h}^{\mu\nu} |\xi_j\rangle = |\xi_j\rangle$ since $\{ |h_{n,j}^{\mu\nu}\rangle : j \in \mathbb{N} \}$ is a basis for H_{S_n,τ_S} and $|\xi_j\rangle \in H_{S_n,\tau_S}$. Therefore, $\pi_{n,h}^{\mu\nu} \pi_n = \pi_n$. But $\pi_{n,h}^{\mu\nu} \pi_n = \pi_n \pi_{n,h}^{\mu\nu}$ since π_n and $\pi_{n,h}^{\mu\nu}$ are Hermitian/self-adjoint. Hence, $\pi_n = \pi_{n,h}^{\mu\nu}$. Now the summands of $\pi_n \pi_\tau^{\mu\nu}$ are only going to consist of those $|\eta_j^{\mu\nu}\rangle\langle\eta_j^{\mu\nu}|$ for which $\hat{T}^{\mu\nu}(y) |\eta_j^{\mu\nu}\rangle = \tau |\eta_j^{\mu\nu}\rangle$ and for which $\hat{T}_S(x) |\eta_j^{\mu\nu}\rangle = \tau_S(x) |\eta_j^{\mu\nu}\rangle$ for all $x \in S_n \cap S$, and these are just the $|\chi_{j,\tau}^{\mu\nu}\rangle\langle\chi_{j,\tau}^{\mu\nu}|$ which are the summands of $\pi_{n,\tau}^{\mu\nu}$. Hence, $\pi_n \pi_\tau^{\mu\nu} = \pi_{n,\tau}^{\mu\nu}$.

But clearly $\hat{T}^{\mu\nu}(y) = \sum_\tau \tau \pi_\tau^{\mu\nu}$. Therefore, combining (3.4.7), (4.1.3), and (4.1.5), we have

$$\langle T^{\mu\nu}(y) \rangle_{\tau_S} = \lim_{n \to \infty} \frac{\sum_\tau \langle \Psi_n | \pi_n \pi_\tau^{\mu\nu} | \Psi_n \rangle \, \tau}{\langle \Psi_n | \pi_n | \Psi_n \rangle} = \lim_{n \to \infty} \frac{\langle \Psi_n | \pi_n \hat{T}^{\mu\nu}(y) | \Psi_n \rangle}{\langle \Psi_n | \pi_n | \Psi_n \rangle}. \qquad (4.1.6)$$

We are now in a position to show that Kent's theory is consistent with standard quantum theory. First let us consider what we need to show.

In the Bohmian interpretation, consistency with standard quantum theory requires that if one averages the expectation values of an observable over the hidden variables (i.e. the positions and the momenta of all the particles) then one obtains the expectation value of the observable given by standard quantum theory as indicated in equation (1.8.3).[12]

Now in Kent's theory, the hidden variables on which his beables $\langle T^{\mu\nu}(y) \rangle_{\tau_S}$ depend are the values $\tau_S(x)$ of $T_S(x)$ for $x \in S^1(y) \cap S$. The operator π_n in equation (4.1.6) in the limit as $n \to \infty$ encapsulates this hidden information. To remind ourselves of π_n's dependency on τ_S restricted to $S_n \cap S$, we will now write $\pi_n(\tau_{S_n \cap S})$ for π_n where

---

12. For a physical system $\mathcal{S}$ whose quantum wave function is ψ and in which the particles have precise but unknown locations given by a probability distribution ρ, then the Bohmian interpretation will give the same predictions as standard quantum theory when averaged over the particle positions so long as ρ is given by $|\psi|^2$. The probability distribution $\rho = |\psi|^2$ is called the *quantum equilibrium distribution*, and we say that a system $\mathcal{S}$ with quantum wave function ψ is in *quantum equilibrium* when the probability distribution for the particles of $\mathcal{S}$ is given by the quantum equilibrium distribution $|\psi|^2$. The question of why the Bohmian interpretation should give the same predictions as standard quantum theory is therefore a question of why we should expect a system to be in quantum equilibrium. An answer to this question can be found in Detlef Dürr, Sheldon Goldstein, and Nino Zanghì, *Quantum physics without quantum philosophy* (Berlin ; New York: Springer, 2013), 34–51. The authors provide an answer by first supposing there is a universal wave function Ψ_0 defined at an initial time, and that the particles Q of the universe are distributed according to the distribution $|\Psi_0|^2$ at the initial time. Then given the Bohmian dynamics of the universal wave function Ψ_t (where t is time) and the relationship between the velocities of the particles and Ψ_t at time t, the authors show that the

$\tau_{S_n \cap S}$ is the function τ_S restricted to $S_n \cap S$. Likewise, we will write $r_n(\tau_{S_n \cap S})$ for r_n, the statement that $T_S(x) = \tau_S(x)$ for all $x \in S_n(y) \cap S$. If we let j index all possible functions $\tau_{S_n \cap S, j}$ taking real values on $S_n \cap S$,[13] then the analogue of (1.8.3) requires us to show that

$$\langle \Psi_0 | \hat{\boldsymbol{T}}^{\mu\nu}(y) | \Psi_0 \rangle = \lim_{n \to \infty} \sum_j P\big(r_n(\tau_{S_n \cap S, j})\big) \, \langle T^{\mu\nu}(y) \rangle_{\tau_{S_n \cap S, j}} \qquad (4.1.7)$$

for all y lying between S_0 and S, where the left-hand side of (4.1.7) is just the expectation value of $\hat{\boldsymbol{T}}^{\mu\nu}(y)$ in the Heisenberg picture as predicted by standard quantum theory.[14] Equation (4.1.7) is sufficient to establish consistency with standard quantum theory because ultimately, all observables are going to be reducible to expressions dependent on $\hat{T}^{\mu\nu}(y)$, since once we know what to expect for $\hat{T}^{\mu\nu}(y)$, we will know what to expect for the energy and momentum densities for all measuring apparatus readouts etc., and hence what to expect for all measurement outcomes. But

---

distribution of Q at any time t is given by $|\Psi_t|^2$. The authors then consider an ensemble of systems each of which is described by a wave function ψ (suitably translated to where each system is), and they show that given the distribution of Q, the probability for the particles in each system of the ensemble (in a frequentist sense) is given by $|\psi|^2$, hence establishing that a system with wave function ψ will be in quantum equilibrium. Note that the assumption that the particles Q of the universe are initially distributed according to the distribution $|\Psi_0|^2$ is not begging the question, for we can first assume there is an initial probability distribution for Q, and from this we can define a wave function Ψ_0 so that $|\Psi_0|^2$ gives the probability distribution of Q at the initial time. There still remains the question of where the initial distribution of Q comes from, but we could probably speculate about this endlessly in a way that is analogous to the speculation regarding T_S on pages 151 to 157. But the fact that we could speculate about this is not an obvious problem for the Bohmian interpretation.

13. Recall that we are implicitly using an approximation scheme described by equation (3.2.5), so we are really considering functions on the cells of a mesh over $S_n \cap S$ taking values from a finite pool of possible values. Also see footnote 54 on page 163. This is why we can use an index j to index all the $\tau_{S_n \cap S, j}$.

14. Where we are using Schwinger's bold typeface convention as described in footnote 22 on page 133.

from (4.1.3) and (4.1.6), we have

$$\lim_{n\to\infty} \sum_j P\big(r_n(\tau_{S_n\cap S,j})\big) \langle T^{\mu\nu}(y)\rangle_{\tau_{S_n\cap S,j}} = \lim_{n\to\infty} \sum_j \langle \Psi_n|\pi_n(\tau_{S_n\cap S,j})\hat{T}^{\mu\nu}(y)|\Psi_n\rangle \quad (4.1.8)$$

Since there is an orthonormal basis $\{\,|\eta_j^{\mu\nu}\rangle : j \in \mathbb{N}\}$ of H_{S_n} consisting of simultaneous $\hat{T}_S(x)$-eigenstates so that $\hat{T}_S(x)\,|\eta_j^{\mu\nu}\rangle = \tau_{S_n\cap S,j}(x)\,|\eta_j^{\mu\nu}\rangle$ for all $x \in S_n \cap S$, it follows that $\sum_j \pi_n(\tau_{S_n\cap S,j}) = I$, where I is the identity operator on H_n. Combining this with (4.1.8) we get

$$\lim_{n\to\infty} \sum_j P\big(r_n(\tau_{S_n\cap S,j})\big) \langle T^{\mu\nu}(y)\rangle_{\tau_{S_n\cap S,j}} = \lim_{n\to\infty} \langle \Psi_n|\hat{T}^{\mu\nu}(y)|\Psi_n\rangle \qquad (4.1.9)$$

In the notation of equation (3.2.4), we have $|\Psi_n\rangle = |\Psi[S_n]\rangle$, and according to (3.2.4) the expectation value $\langle \Psi[S_n]|\hat{T}^{\mu\nu}(y)|\Psi[S_n]\rangle$ will be independent of the hypersurface S_n so long as it contains y, and the result will be equal to the expectation value $\langle \Psi_0|\hat{\boldsymbol{T}}^{\mu\nu}(y)|\Psi_0\rangle$ in the Heisenberg picture which is the value that is predicted by standard quantum theory. Therefore, equation (4.1.7) follows from (3.2.4) and (4.1.9) which is what we were aiming to show.

4.2 Kent's Theory and Lorentz Invariance

4.2.1 Overview

Lorentz invariance is the fundamental property of a physical theory that must be satisfied if it is to be consistent with special relativity. Experimental observations show that light always travels at the same speed in a vacuum no matter how fast the source

of light is travelling. This experimental fact leads to some surprising consequences for how different observers will measure the same physical events. For instance, it means that velocities don't add up in the way one would naturally expect. We can understand why velocities don't add up in this way by considering a simple example: suppose Alice is travelling on a train at 100 mph, and she has a ball with her that she kicks down the aisle of the passenger car at 20 mph in the direction of travel. Then naturally, we would expect the kicked ball to be travelling at (100+20)=120 mph with respect to Bob who happens to be standing by the railroad track. But now suppose Alice shines a beam of light in the direction of travel. If we ignore the fact that light travels slightly more slowly in air than in a vacuum, then whilst on the train, Alice would measure the beam of light to travel at 670,616,629 mph. So by analogy with the previous example, we would expect the stationary observer, Bob to measure the beam of light that Alice shines to travel at (100+670,616,629)=670,616,729 mph. However, it turns out that Bob measures the beam of light to have exactly the same speed as Alice measures. Thus, the manner in which Alice and Bob's velocity measurements differ is not what we would expect – we can't just add or subtract velocities. Instead we have to use what's known as a *Lorentz transformation* to work out how their respective measurements transform.

Lorentz transformations have the property that they approximately behave in the way one would expect such transformations to behave for situations involving small velocities like Alice kicking a ball on a train. But in situations which involve very high velocities such as is the case with the beam of light Alice shines on the train, then

the Lorentz transformations must always be consistent with Alice and Bob measuring light to have the same speed.

Now we are interested in physical theories that make predictions about quantifiable physical properties. The simplest kind of quantifiable physical property is called a *scalar*. A scalar consists of a single number typically associated with a particular spacetime location, and it has the same value no matter what coordinates an observer uses to describe the spacetime location. One example of a scalar is an object's *rest mass* which is the mass an object would have if it had no velocity.

Now if two observers, Alice and Bob travel at a constant velocity relative to one another, there will be a Lorentz transformation that maps the coordinates Alice uses to describe a spacetime location to the coordinates Bob uses to describe the same spacetime location. But despite them using different coordinates to describe the same spacetime location, if what they are measuring is a scalar, then they will attribute the same scalar value to the same spacetime location.

As another example of a scalar, consider the mass-energy density $T_S(x)$ that we introduced on page 126 which specifies a single number $T_S(x)$ that is associated with every x belonging to the hypersurface S. In the next subsection, we will show that if the coordinates Alice and Bob use are related to each other via a Lorentz transformation, then their measurement of $T_S(x)$ will only depend on the spacetime location x and not on the coordinate system used to describe x meaning that $T_S(x)$ also has the defining property of a scalar.

Slightly more complicated than a scalar is a *four-vector*. Rather than associating a single number with a spacetime location, a four-vector associates four numbers with a spacetime location. Four-vectors can be used to describe the velocity of an object, and also the object's energy and momentum. With their different coordinates systems, Alice and Bob will typically describe the same four-vector with four different numbers, but the fundamental property that four-vectors must satisfy is that they transform under Lorentz transformations in such a way that any two four-vectors at a spacetime location can be combined into a product that is a scalar. It is because of this fundamental property that a four-vector can be associated with a photon to describe its velocity in such a way as to guarantee that any observer will always measure it to have the same speed.

Besides scalars and four-vectors, the final kind of quantifiable physical properties that will concern us are *rank-two tensors*. A rank-two tensor associated with a spacetime location is a four by four array of sixteen numbers which has the property that it can be "contracted" with any two four-vectors associated with the same spacetime location to form a scalar. The stress-energy tensor $T^{\mu\nu}(y)$ introduced on page 143 is an example of a rank-two tensor. It is because the stress-energy tensor $T^{\mu\nu}(x)$ is a rank-two tensor for every spacetime location x belonging to the hypersurface S that enables us to define the scalar $T_S(x)$ by twice contracting $T^{\mu\nu}(x)$ with the four-vector that is perpendicular to the hypersurface S.

Now if a physical theory makes predictions about quantifiable physical properties such that these predictions transform in the same way as the quantifiable physical

properties that the theory predicts, then we say that the theory is *Lorentz invariant*. So in a Lorentz invariant theory, if the theory makes predictions concerning scalars, four-vectors, or rank-two tensors, then these predictions should transform like scalars, four-vectors, or rank-two tensors respectively.

Now the aim of this section is to show that Kent's theory is Lorentz invariant. Since Kent's theory makes predictions concerning the stress-energy tensor by predicting that $T^{\mu\nu}(y)$ will take the value $\langle \hat{T}^{\mu\nu}(y) \rangle_{\tau_S}$, it follows that for Kent's theory to be Lorentz invariant, it is necessary that $\langle \hat{T}^{\mu\nu}(y) \rangle_{\tau_S}$ transforms as a rank-two tensor.

The proof of Lorentz invariance involves a study of the Hilbert space of states two observers will use to describe the state of a hypersurface S_n as well as the observables they use in standard quantum theory to make predictions about the mass-energy density on $S_n \cap S$ and the stress-energy tensor on S_n. In particular, there will be an operator $U(\Lambda)$ dependent on the Lorentz transformation Λ between the coordinates of the two observers that maps the state $|\Psi_n\rangle$ one observer ascribes to S_n to the state $|\Psi_n'\rangle$ the other observer describes to S_n. Using this operator $U(\Lambda)$, we can find expressions[15] that relate the observables of the two observers to each other. We can also use $U(\Lambda)$ to show how the states $\pi_n |\Psi_n\rangle$ and $\pi_n' |\Psi_n'\rangle$ relate to one another.[16] By using these relationships in the formula (4.1.6), it follows that $\langle \hat{T}^{\mu\nu}(y) \rangle_{\tau_S}$ transforms as a rank-two tensor, and hence that Kent's theory is Lorentz invariant.

---

15. such as equations (4.2.17) and (4.2.19)

16. How they related to one another is easily seen from equation (4.2.21) and the fact that $|\Psi_n'\rangle = U(\Lambda) |\Psi_n'\rangle$.

In this subsection, we provide the technical details that show that Kent's theory is Lorentz invariant. Up until now we have been representing a spacetime location by a four-tuple (x^0, x^1, x^2, x^3) where $(x^i)_{i=1}^3$ are spatial coordinates, and where $x^0 = ct$ with c being equal to the speed of light and t being the time. It will be convenient to represent spacetime locations using a more concise notation. So we let $(1, 0, 0, 0)$ correspond to the spacetime location $\hat{e}_0$, $(0, 1, 0, 0)$ correspond to the spacetime location $\hat{e}_1$, etc.. Then we can express any other spacetime location as a sum $\sum_{\mu=0}^{3} x^\mu \hat{e}_\mu$. We will use the so-called *Einstein summation convention* of dropping the summation sign and implicitly assuming that there is a summation whenever an upper index and a lower index are the same so that we can write $x^\mu \hat{e}_\mu$ instead of $\sum_{\mu=0}^{3} x^\mu \hat{e}_\mu$. We also use the convention of letting Greek letters range over 0, 1, 2, and 3 (e.g. the μ in $x^\mu \hat{e}_\mu$), and of letting Roman letters range over 1, 2, and 3 (e.g. the i in $(x^i)_{i=1}^3$).

Now suppose an observer $\mathcal{O}$ expresses spacetime locations in terms of $\{\hat{e}_\mu : \mu = 0, \ldots, 3\}$ and hence uses the coordinates (x^0, x^1, x^2, x^3) to describe various spacetime locations. For another observer $\mathcal{O}'$, it may be more natural to express spacetime locations in terms of a different set $\{\hat{e}'_\mu : \mu = 0, \ldots, 3\}$ so that the location described by $\mathcal{O}$ as (x^0, x^1, x^2, x^3) would be described by $\mathcal{O}'$ as (x'^0, x'^1, x'^2, x'^3) where $x'^\mu \hat{e}'_\mu = x^\mu \hat{e}_\mu$. For instance if $\mathcal{O}$ and $\mathcal{O}'$ are moving with respect to each other, they may both want to use coordinates in which their own spatial coordinates are fixed and in which the spatial coordinates of the other observer are changing. As another example, figure 4.1 shows how the (x^1, x^2)-coordinates transform under a spatial rotation.

(a)
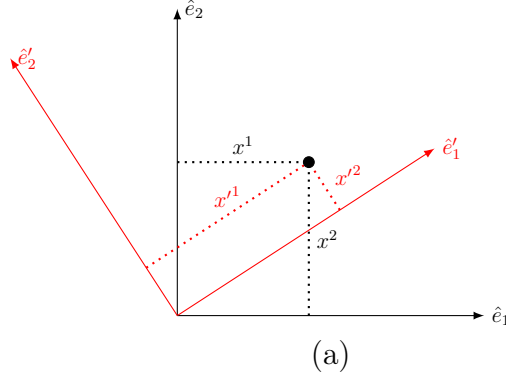
Figure 4.1. Coordinate transformation. Shows how a location (marked as ●) can be expressed either in coordinates (x^1, x^2) with respect to the basis $\{\hat{e}_1, \hat{e}_2\}$ or in coordinates (x'^1, x'^2) with respect to the basis $\{\hat{e}'_1, \hat{e}'_2\}$.

Now the key fact about all observers is that they must always observe light in a vacuum to have a constant speed c. Thus, for a photon that goes through the spacetime locations $(0, 0, 0, 0)$ and (x^0, x^1, x^2, x^3) in the coordinates of $\mathcal{O}$, we must have $(x^0, x^1, x^2, x^3) = (ct, tv^1, tv^2, tv^3)$ where

$$\sqrt{(v^1)^2 + (v^2)^2 + (v^3)^2} = c.$$

But if $(0, 0, 0, 0)$ and (x^0, x^1, x^2, x^3) corresponds to $(0, 0, 0, 0)$ and (x'^0, x'^1, x'^2, x'^3) respectively in the coordinates of another observer $\mathcal{O}'$, then we must also have $(x'^0, x'^1, x'^2, x'^3) = (ct', t'v'^1, t'v'^2, t'v'^3)$ where

$$\sqrt{(v'^1)^2 + (v'^2)^2 + (v'^3)^2} = c.$$

189

In either case, we must have

$$(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = (x'^0)^2 - (x'^1)^2 - (x'^2)^2 - (x'^3)^2 = 0. \qquad (4.2.1)$$

If we define $\eta_{00} = 1$, $\eta_{ii} = -1$ for $i = 1, 2, 3$ and $\eta_{\mu\nu} = 0$ for $\mu \neq \nu$, then using the Einstein summation convention as well as the convention of lowering indices so that we define $x_\mu \stackrel{\text{def}}{=} \eta_{\mu\nu} x^\nu$, then (4.2.1) is equivalent to

$$x_\mu x^\mu = x'_\mu x'^\mu = 0.$$

Thus, for any coordinate transformation $x \to x'$ such that $x_\mu x^\mu = x'_\mu x'^\mu$, if the speed of light is c in the x-coordinates, then the speed of light is also guaranteed to be c in the x'-coordinates. A *Lorentz transformation* Λ is any coordinate transformation of the form $x'^\mu = \Lambda^\mu{}_\nu x^\nu$ such that $x_\mu x^\mu = x'_\mu x'^\mu$. Since a Lorentz transformation must satisfy

$$x_\mu x^\mu = \eta_{\mu\rho} \Lambda^\rho{}_\sigma x^\sigma \Lambda^\mu{}_\nu x^\nu$$

for all x, it follows that

$$\Lambda^\rho{}_\mu \eta_{\rho\sigma} \Lambda^\sigma{}_\nu = \eta_{\mu\nu}.^{17} \qquad (4.2.2)$$

190

Having considered how the coordinates of a spacetime location viewed by one observer relate to the coordinates of the same spacetime location viewed by a different observer, we can now consider how physical quantities viewed by different observers relate to each other. So suppose $\phi(x) \overset{\text{def}}{=} \phi(x^0, x^1, x^2, x^3)$ is the value of a scalar defined at the spacetime location (x^0, x^1, x^2, x^3) as described by an observer $\mathcal{O}$. Then another observer $\mathcal{O}'$ using a different set of coordinate (x'^0, x'^1, x'^2, x'^3) to describe the spacetime location (x^0, x^1, x^2, x^3) will describe this same scalar as $\phi'(x') \overset{\text{def}}{=} \phi'(x'^0, x'^1, x'^2, x'^3)$ where $\phi'(x') = \phi(x)$. Therefore,

$$\phi'(x') = \phi(\Lambda^{-1}x') \tag{4.2.3}$$

where Λ^{-1} is the inverse Lorentz transformation that takes the coordinates $x' = (x'^0, x'^1, x'^2, x'^3)$ of a spacetime location to the coordinates $x = (x^0, x^1, x^2, x^3)$ describing that spacetime location. Thus, equation (4.2.3) shows us how a scalar transforms under a Lorentz transformation Λ.

Many physical quantities, however, are not scalars and so will look different to different observers. For instance, the energy of an object has a kinetic component

---

17. To see why this is, note that if $x_\mu x^\mu = x'_\mu x'^\mu$ for all x, then for any other spacetime location y, we have $(x+y)_\mu(x+y)^\mu = (x'+y')_\mu(x'+y')^\mu$. If we expand this out and cancel $x_\mu x^\mu$ with $x'_\mu x'^\mu$ and cancel $y_\mu y^\mu$ with $y'_\mu y'^\mu$, and using the fact that $y_\mu x^\mu = x_\mu y^\mu$, etc. we find that $x_\mu y^\mu = x'_\mu y'^\mu$ for all x and y. Hence,

$$\eta_{\nu\mu}x^\mu y^\nu = x_\mu y^\mu = \eta_{\sigma\rho}\Lambda^\rho{}_\mu\Lambda^\sigma{}_\nu x^\mu y^\nu.$$

Since we can choose x such that $x^\mu = 1$ and $x^\alpha = 0$ for $\alpha \neq \mu$, and can choose y such that $y^\nu = 1$ and $y^\beta = 0$ for $\beta \neq \nu$, it follows that

$$\eta_{\mu\nu} = \eta_{\sigma\rho}\Lambda^\rho{}_\mu\Lambda^\sigma{}_\nu,$$

which is the result we wanted to prove.

that depends on the velocity the object has relative to an observer. However, it turns out that if an observer $\mathcal{O}$ considers an object's energy E together with its three components of momentum p^1, p^2, and p^3 (in the directions $\hat{e}_1$, $\hat{e}_2$, and $\hat{e}_3$ respectively) to form the four-tuple $p \stackrel{\text{def}}{=} (E/c, p^1, p^2, p^3)$ known as the object's *four-momentum*, then p transforms in the same way as spacetime coordinates transform between different observers. In other words, a different observer $\mathcal{O}'$ whose coordinates are given by $x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu}$ would observe the object's four-momentum to be $p'^{\mu} = \Lambda^{\mu}{}_{\nu} p^{\nu}$.[18]

More generally, any list of four physical quantities $(\varphi^0, \varphi^1, \varphi^2, \varphi^3)$ that transforms as $\varphi \to \varphi'$ with $\varphi'^{\mu} = \Lambda^{\mu}{}_{\nu} \varphi^{\nu}$ is called a *four-vector*. Figure 4.2 shows how (two of) the components of a four-vector φ at a particular location will differ for different observers under a spatial rotation of the coordinates. A four-vector $\varphi^{\mu}(x)$ defined at every spacetime location x is called a *four-vector field*. Thus, at each spacetime location x, the four-vector field $\varphi^{\mu}(x)$ assigns four numbers, $\varphi^0(x)$, $\varphi^1(x)$, $\varphi^2(x)$, and $\varphi^3(x)$. If

---

18. In order for p to transform in this way, we have to redefine what we mean by energy and momentum. In classical mechanics, the momentum of an object is the product of the object's mass and its velocity. In the context of special relativity, however, the four-momentum of an object is defined to be the product of its rest mass m_0 and its *four-velocity* where the four velocity of an object is a four-tuple (u^0, u^1, u^2, u^3) with $u_{\mu} u^{\mu} = c^2$ such that the object's velocity (in the classical sense) is the vector $(c \frac{u^1}{u^0}, c \frac{u^2}{u^0}, c \frac{u^3}{u^0})$. The motivation for this definition can be seen by considering an object whose classical velocity is $\mathbf{v} = (v^1, v^2, v^3)$ that goes through $(0, 0, 0, 0)$. It will have a spacetime trajectory $x(t) = (ct, v^1 t, v^2 t, v^3 t)$. u is just the four-vector proportional to $x(1)$ with $u_{\mu} u^{\mu} = c^2$, so if γ is the constant of proportionality such that $u^0 = \gamma c$ and $u^i = \gamma v^i$, then by eliminating γ we get $v^i = c \frac{u^i}{u^0}$. We can then easily work out the constant of proportionality γ and hence the four-velocity u of an object whose classical velocity is $\mathbf{v}$. For we must have $u^i = \frac{v^i u^0}{c}$, for $i = 1$ to 3. Therefore, since $u_{\mu} u^{\mu} = c^2$, we must have $(u^0)^2 (1 - \frac{v^2}{c^2}) = c^2$ where $v = \sqrt{(v^1)^2 + (v^2)^2 + (v^3)^2}$. Thus, if we define $\beta = v/c$ and $\gamma = \frac{1}{\sqrt{1 - \beta^2}}$, then $u^0 = \gamma c$ and $u^i = \gamma v^i$ for $i = 1$ to 3, and hence the four-velocity of the object must be $u = \gamma(c, v^1, v^2, v^3)$. From this, we see that the object's four-momentum will be $\gamma m_0 (c, v^1, v^2, v^3)$. If the object's velocity is very small compared to the speed of light, then $\gamma \approx 1 + \frac{v^2}{2c^2}$, and hence the object's four-momentum $(E/c, p^1, p^2, p^3)$ will be approximately $(m_0 c + \frac{1}{2} m_0 v^2/c, m_0 v^1, m_0 v^2, m_0 v^3)$. Therefore, (p^1, p^2, p^3) is approximately equal to the classical momentum. However, the energy is now $E = m_0 c^2 + \frac{1}{2} m_0 v^2$. Thus, in addition to the kinetic energy term $\frac{1}{2} m_0 v^2$, there is a rest mass energy $m_0 c^2$. If we define the *relativistic mass* $m = \gamma m_0$, then we obtain Einstein's famous formula $E = mc^2$.

(b)
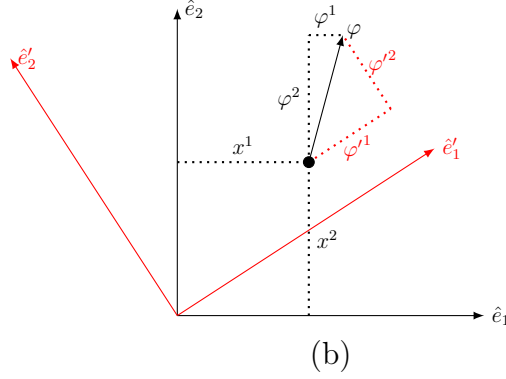
Figure 4.2. Shows how a four-vector φ (of which only two components are shown) defined at a spacetime location (indicated by $\bullet$) can be expressed either as (φ^1, φ^2) with respect to the basis $\{\hat{e}_1, \hat{e}_2\}$ or as (φ'^1, φ'^2) with respect to the basis $\{\hat{e}'_1, \hat{e}'_2\}$.

$\mathcal{O}$ observes this vector-field $\varphi^\mu(x)$, and $\mathcal{O}'$ is another observer whose coordinates are related to the coordinates $\mathcal{O}$ via the Lorentz transformation Λ, then $\mathcal{O}'$ will describe the same physical reality that $\mathcal{O}$ describes by assigning four numbers $\varphi'^0(x')$, $\varphi'^1(x')$, $\varphi'^2(x')$, and $\varphi'^3(x')$ at every spacetime location x', and the relationship between the description $\mathcal{O}$ gives and the description $\mathcal{O}'$ gives will be given by the formula

$$\varphi'^\mu(x') = \Lambda^\mu{}_\nu \varphi^\nu(x).$$

Hence, under the Lorentz transformation Λ, a vector field $\varphi^\mu(x)$ transforms as $\varphi^\mu(x) \to \varphi'^\mu(x')$ where

$$\varphi'^\mu(x') = \Lambda^\mu{}_\nu \varphi^\nu(\Lambda^{-1} x'). \tag{4.2.4}$$

From a four-vector φ^μ, we can also define the so-called *four-covector*:

$$\varphi_\mu \stackrel{\text{def}}{=} \eta_{\mu\nu}\varphi^\nu. \tag{4.2.5}$$

To see how four-covectors transform under a Lorentz transformation Λ, it will be helpful to define

$$\Lambda_\mu{}^\nu \stackrel{\text{def}}{=} \eta_{\mu\rho}\eta^{\nu\sigma}\Lambda^\rho{}_\sigma \tag{4.2.6}$$

where $\eta^{\nu\sigma} = \eta_{\nu\sigma}$. If we also define the *Kronecker-delta* δ^ν_μ suchthat $\delta^\nu_\mu = 1$ when $\mu = \nu$ and $\delta^\nu_\mu = 0$ otherwise, then using the fact that $\eta_{\mu\rho}\eta^{\nu\rho} = \delta^\nu_\mu$ together with equation (4.2.2), we have

$$\Lambda^\rho{}_\mu\Lambda_\rho{}^\nu = \delta^\nu_\mu. \tag{4.2.7}$$

Since by definition, the inverse of Λ^{-1} satisfies $(\Lambda^{-1})^\nu{}_\rho\Lambda^\rho{}_\mu = \delta^\nu_\mu$, we have $(\Lambda^{-1})^\nu{}_\rho = \Lambda_\rho{}^\nu$. From (4.2.4), (4.2.5), and (4.2.6), we therefore see that under a Lorentz transformation Λ, a four-covector field $\varphi_\mu(x)$ transforms as $\varphi_\mu(x) \to \varphi'_\mu(x')$ where

$$\varphi'_\mu(x') = \Lambda_\mu{}^\nu \varphi_\nu(\Lambda^{-1}x') \tag{4.2.8}$$

Besides scalars, four-vectors, and four-covectors, we also need to consider physical quantities called rank-two tensors. The defining property of a rank-two tensor field $\varphi^{\mu\nu}(x)$ is that under a Lorentz transformation Λ, it transforms as $\varphi^{\mu\nu}(x) \to \varphi'^{\mu\nu}(x')$

where

$$\varphi'^{\mu\nu}(x') = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \varphi^{\rho\sigma}(\Lambda^{-1}x'). \tag{4.2.9}$$

On page 126, we introduced the mass-energy density $T_S(x)$ on a hypersurface S. As explained in section 3.2.1, the values of $T_S(x)$ for all $x \in S$ are the additional variables that Kent uses to supplement standard quantum theory. It was mentioned in passing that $T_S(x)$ does not depend on which frame of reference one is in. In other words, $T_S(x)$ is a scalar. I will now explain why this is so.

We first need to consider the precise definition of $T_S(x)$. At each spacetime location on the hypersurface S which an observer $\mathcal{O}$ describes as having coordinates $x = (x^\mu)_{\mu=0}^3$, we define $\eta^\mu(x)$ to be the future-directed unit four-vector at x that is orthogonal to S. In other words, $\eta^0(x) > 0$, $\eta_\mu(x)\eta^\mu(x) = 1$, and if $y \in S$ is very close to x, then

$$\frac{(x-y)_\mu \eta^\mu(x)}{\sqrt{(x-y)_\nu(x-y)^\nu}} \approx 0. \tag{4.2.10}$$

$T_S(x)$ is then given by the formula

$$T_S(x) = T^{\mu\nu}(x)\eta_\mu(x)\eta_\nu(x). \tag{4.2.11}$$

For example, if S was the hypersurface consisting of all spacetime locations $x = (0, x^1, x^2, x^3)$, then $\left(\eta^0(x), \eta^1(x), \eta^3(x), \eta^3(x)\right) = (1, 0, 0, 0)$, and hence $T_S(x) = T^{00}(x)$

which is the density of relativistic mass at x, i.e. the energy density at x divided by c^2. Note that the condition that $y \in S$ is very close to x in (4.2.10) means that $T_S(x)$ only has a local dependence on S in the vicinity of x. i.e. if S' only differs from S outside the vicinity of x, then $T_{S'}(x) = T_S(x)$.

To see why $T_S(x)$ is a scalar, suppose that Λ is a Lorentz transformation such that $\Lambda^0{}_\mu \eta^\mu > 0$ for any future-directed unit four-vector vector η^μ. We refer to a Λ with this property as an *orthochronous* Lorentz transformation. Also, suppose that $\mathcal{O}$ and $\mathcal{O}'$ are two observers such that spacetime locations that observer $\mathcal{O}$ describes as having coordinates $x = (x^\mu)^3_{\mu=0}$ are described by $\mathcal{O}'$ as having coordinates $x' = (\Lambda^\mu{}_\nu x^\nu)^3_{\mu=0}$. Then since $x'_\mu y'^\mu = x_\mu y^\mu$, it follows that the future-directed unit four-vector orthogonal to S at x which $\mathcal{O}$ describes as $\eta^\mu(x)$ will be described by $\mathcal{O}'$ as $\eta'^\mu(x') = \Lambda^\mu{}_\nu \eta^\nu(x)$. Thus, for any spacetime location in S that $\mathcal{O}'$ describes as having coordinates x' with corresponding future-directed S-orthogonal unit four-vector $\eta'^\mu(x')$, $\mathcal{O}'$ can construct a function $T'_S(x')$ with

$$T'_S(x') = T'^{\mu\nu}(x')\eta'_\mu(x')\eta'_\nu(x'). \tag{4.2.12}$$

196

Then using (4.2.8) and (4.2.9) on the right-hand side of (4.2.12), we have

$$T'_S(x') = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma T^{\rho\sigma}(x) \Lambda_\mu{}^\alpha \eta_\alpha(x) \Lambda_\nu{}^\beta \eta_\beta(x)$$

$$= \Lambda^\mu{}_\rho \Lambda_\mu{}^\alpha \Lambda^\nu{}_\sigma \Lambda_\nu{}^\beta T^{\rho\sigma}(x) \eta_\alpha(x) \eta_\beta(x)$$

$$= \delta^\alpha_\rho \delta^\beta_\sigma T^{\rho\sigma}(x) \eta_\alpha(x) \eta_\beta(x) \tag{4.2.13}$$

$$= T^{\alpha\beta}(x) \eta_\alpha(x) \eta_\beta(x)$$

$$= T_S(x)$$

where on the third line we have used (4.2.7), and on the last line we have used (4.2.11). To obtain (4.2.13), we assumed that Λ is orthochronous because definition (4.2.11) assumes that $\eta^\mu(x)$ is future-directed. But if Λ is non-orthochronous, we would need to take the negations of $\eta'^\mu(x')$ to get the future-directed S-orthogonal unit four-vector. But clearly this will not affect the equality in (4.2.13), so (4.2.13) holds for all Lorentz transformations, whether they are orthochronous or non-orthochronous. We thus see that $T_S(x)$ is a scalar.

Let us now consider the Hilbert space H_{S_n} for a hypersurface S_n as defined on page 178.[19] Given that $\hat{T}^{\mu\nu}(x)$ is the observable in the Tomonaga-Schwinger picture whose eigenstates with eigenvalues τ are the states of S_n for which an observer $\mathcal{O}$ observes the stress-energy tensor $T^{\mu\nu}(x)$ to take the value τ at x, it follows from (4.2.11) that

$$\hat{T}_S(x) \stackrel{\text{def}}{=} \hat{T}^{\mu\nu}(x) \eta_\mu(x) \eta_\nu(x) \tag{4.2.14}$$

---

19. Also see page 134 for the definition of H_S.

will be the observable whose eigenstates with eigenvalues $\tau_S(x)$ are the states of S_n for which an observer $\mathcal{O}$ observes $T_S(x)$ to take the value $\tau_S(x)$ at x when $x \in S_n \cap S$.[20]

Now two observers $\mathcal{O}$ and $\mathcal{O}'$ will typically assign different physical states to S_n based on their frame of reference. E.g. if $\mathcal{O}$ and $\mathcal{O}'$ are traveling at different speeds, they will attribute different energy levels and momenta to the spacetime locations of S_n. To understand the relationship between the states $\mathcal{O}$ assigns to S_n and the states $\mathcal{O}'$ assigns, suppose $|\psi\rangle$ and $|\chi\rangle$ are two states that an observer $\mathcal{O}$ might judge S_n to be in. As usual, we suppose the coordinates x^μ of observer $\mathcal{O}$ transform to coordinates $x'^\mu = \Lambda^\mu{}_\nu x^\nu$ of observer $\mathcal{O}'$ for some Lorentz transformation Λ. We also suppose that the states $|\psi\rangle$ and $|\chi\rangle$ that $\mathcal{O}$ observes will transform to states $|\psi'\rangle$ and $|\chi'\rangle$ that $\mathcal{O}'$ observes. We will denote the Hilbert space of the states on S_n that $\mathcal{O}'$ can observe as H'_{S_n}.[21]

Now if $\mathcal{O}$ judged S_n to be in the superposition state $|\psi\rangle + |\chi\rangle$, then $\mathcal{O}'$ would judge S_n to be proportional to the superposition state $|\psi'\rangle + |\chi'\rangle$. Also recall that $|\psi'\rangle$ and $\lambda|\psi'\rangle$ represent the same physical state for any complex number λ, so there is sufficient flexibility as to which state we deem $|\psi\rangle$ transforms to that we can deem the transformation $|\psi\rangle \to |\psi'\rangle$ to be a linear transformation. But also note that if observer $\mathcal{O}$ uses the Born rule to calculate the transition probability from state $|\psi\rangle$ to state $|\chi\rangle$, and observer $\mathcal{O}'$ uses the Born rule to calculate the transition probability

20. So long as S_n and S are tangential at x as noted in footnote 53 on page 162.

21. As we will discuss shortly, there is an inner product preserving map $U(\Lambda)$ via which there is a one-to-one correspondence between states in H_{S_n} and states in H'_{S_n}, so we can identify H'_{S_n} with H_{S_n} in which case $U(\Lambda)$ will be a unitary operator.

from state $|\psi'\rangle$ to state $|\chi'\rangle$, then they should calculate the same probabilities. We must therefore have

$$| \langle \chi | \psi \rangle |^2 = | \langle \chi' | \psi' \rangle |^2.$$

Using this fact together with the fact that $|\psi\rangle$ and $\lambda |\psi\rangle$ represent the same physical states, it can be shown that there is a unitary operator $U(\Lambda)$ which relates the states $|\psi\rangle$ and $|\psi'\rangle$ via the formula

$$|\psi'\rangle = U(\Lambda) |\psi\rangle .^{22}$$

At this point, it is worth clarifying the different meanings of $T^{\mu\nu}(x)$, $T'^{\mu\nu}(x')$, $\tau^{\mu\nu}(x)$, $\tau'^{\mu\nu}(x')$, $\hat{T}^{\mu\nu}(x)$ and $\hat{T}'^{\mu\nu}(x')$.

- We use $T^{\mu\nu}(x)$ to refer to the description of the physical quantity that is being observed by $\mathcal{O}$. Thus, $T^{\mu\nu}(x)$ is shorthand for the description "the $\mu\nu$-component of the stress-energy tensor that $\mathcal{O}$ observes at the spacetime location belonging to S that $\mathcal{O}$ describes as x."

- Similarly, $T'^{\mu\nu}(x')$ is shorthand for the description "the $\mu\nu$-component of the stress-energy tensor that $\mathcal{O}'$ observes at a spacetime location belonging to S that $\mathcal{O}'$ describes as x'."

---

22. For more details, see E. Wigner, "On Unitary Representations of the Inhomogeneous Lorentz Group," *Annals of mathematics* 40, no. 1 (1939): 149–204. Since a unitary operator maps a Hilbert space to itself, we first need to identify H_{S_n} and H'_{S_n} in order for $U(\Lambda)$ to be unitary.

- $\tau^{\mu\nu}(x)$ stands for a particular (real) value of the physical quantity described by $T^{\mu\nu}(x)$ that $\mathcal{O}$ observes, and

- $\tau'^{\mu\nu}(x')$ stands for a particular (real) value of the physical quantity described by $T'^{\mu\nu}(x')$ that $\mathcal{O}'$ observes.

- $\hat{T}^{\mu\nu}(x)$ for $x \in S_n$ is the Tomonaga-Schwinger observable acting on H_{S_n} such that if observer $\mathcal{O}$ deemed S_n to be in an eigenstate $|\psi\rangle$ of $\hat{T}^{\mu\nu}(x)$ with eigenvalue τ (a real number), then observer $\mathcal{O}$ would observe the physical quantity described by $T^{\mu\nu}(x)$ to have the value τ.

- $\hat{T}'^{\mu\nu}(x')$ is the Tomonaga-Schwinger observable acting on H'_{S_n} such that if observer $\mathcal{O}'$ deemed S_n to be in an eigenstate $|\psi'\rangle$ of $\hat{T}'^{\mu\nu}(x')$ with eigenvalue τ', then observer $\mathcal{O}'$ would observe the physical quantity described by $T'^{\mu\nu}(x')$ to have the value τ'.

- $T_S(x) = T^{\mu\nu}(x)\eta_\mu(x)\eta_\nu(x)$ is shorthand for the description "the mass-energy density of the hypersurface S observed by observer $\mathcal{O}$ at a spacetime location that $\mathcal{O}$ describes as x".

- $T'_S(x') = T'^{\mu\nu}(x')\eta'_\mu(x')\eta'_\nu(x')$ is shorthand for the description "the mass-energy density of the hypersurface S observed by observer $\mathcal{O}'$ at a spacetime location that $\mathcal{O}'$ describes as x'".

- The function $\tau_S(x)$ stands for a particular range of values for each $x \in S$ of the physical quantity described by $T_S(x)$ observed by $\mathcal{O}$.

- The function $\tau'_S(x')$ stands for a particular range of values for each $x' \in S$ of the physical quantity described by $T'_S(x')$ observed by $\mathcal{O}'$.

- For each $x \in S$, $\hat{T}_S(x) = \hat{T}^{\mu\nu}(x)\eta_\mu(x)\eta_\nu(x)$ is the Tomonaga-Schwinger observable such that if an observer $\mathcal{O}$ deems S to be in an eigenstate $|\psi\rangle \in H_{S_n}$ of $\hat{T}_S(x)$ with eigenvalue τ (a real number), then $\mathcal{O}$ would observe the physical quantity described by $T_S(x)$ to have the value τ.

- For each $x' \in S$, $\hat{T}'_S(x') = \hat{T}'^{\mu\nu}(x)\eta'_\mu(x')\eta'_\nu(x')$ is the Tomonaga-Schwinger observable such that if an observer $\mathcal{O}'$ deems S to be in an eigenstate $|\psi'\rangle \in H'_{S_n}$ of $\hat{T}'_S(x')$ with eigenvalue τ' (a real number), then $\mathcal{O}'$ would observe the physical quantity described by $T'_S(x')$ to have the value τ'.

Having clarified this terminology, we see that if $|\psi\rangle \in H_{S_n}$ is a state for which $T^{\mu\nu}(x) \approx \tau^{\mu\nu}(x)$,[23] and if $|\psi'\rangle \in H'_{S_n}$ is a state for which $T'^{\mu\nu}(x') \approx \tau^{\mu\nu}(x')$, then

$$\hat{T}^{\mu\nu}(x)|\psi\rangle \approx \tau^{\mu\nu}(x)|\psi\rangle, \text{ and} \tag{4.2.15a}$$

$$\hat{T}'^{\mu\nu}(x')|\psi'\rangle \approx \tau'^{\mu\nu}(x')|\psi'\rangle. \tag{4.2.15b}$$

It then follows from (4.2.15) that if $|\psi'\rangle = U(\Lambda)|\psi\rangle$, then

$$U(\Lambda)^{-1}\hat{T}'^{\mu\nu}(x')U(\Lambda)|\psi\rangle \approx T'^{\mu\nu}(x')|\psi\rangle. \tag{4.2.16}$$

---

23. We say approximately ($\approx$) here since the operators $\hat{T}^{\mu\nu}$ will not in general commute, so we won't typically be able to find a state $|\psi\rangle$ which is an eigenstate for all the observables $\hat{T}^{\mu\nu}$. It is the non-commutativity of observables that is responsible for Heisenberg's uncertainty principle.

Therefore, in order for (4.2.9) to hold for $T'^{\mu\nu}(x')$ in the classical limit (where we treat the stress-energy observables as though they commute with each other and replace the approximations by equalities), by plugging an operator form of (4.2.9) into (4.2.16), we see that we must have

$$U(\Lambda)^{-1}\hat{T}'^{\mu\nu}(x')U(\Lambda) = \Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}\hat{T}^{\rho\sigma}(\Lambda^{-1}x'). \qquad (4.2.17)$$

Now to say that Kent's model is Lorentz invariant, is to say that (4.1.6) defines a rank-two tensor, for then this quantity and the quantities on which it depends will transform in the way that physical quantities should transform under a Lorentz transformation Λ when the spacetime coordinates of two observers $\mathcal{O}$ and $\mathcal{O}'$ are related by the formula $x'^{\mu} = \Lambda^{\mu}{}_{\nu}x^{\nu}$.

Note that having a privileged hypersurface S in which a notional measurement of T_S is made does not of itself break Lorentz invariance. Just because we are privileging a hypersurface S, we are not making any assumptions about simultaneity being defined by S. Because spacelike separation is a Lorentz invariant property, both $\mathcal{O}$ and $\mathcal{O}'$ will deem the spacetime locations on S to be spacelike separated. The Lorentz transformation itself has absolutely no effect on what S is. It is just that $\mathcal{O}$ and $\mathcal{O}'$ will use different coordinates to describe a particular spacetime location on S. It maybe that in the coordinate system of $\mathcal{O}$, some spacetime locations of S are simultaneous (i.e. have the same x^0 value), but there is no requirement of simultaneity, and there is no claim that a reference frame in which the spacetime locations of S are

simultaneous is particularly special. In Kent's theory, it is sufficient for there to be just one hypersurface on which T_S has a determinate value. But if another hypersurface were to be chosen instead, it would make no difference to empirical adequacy since (4.1.7) will hold regardless of what hypersurface S is chosen.

Now given that (4.1.6) being a rank-two tensor is the criterion for Kent's model to be Lorentz invariant, we need to show that if $\{\,|\xi_j\rangle : j\}$ is an orthonormal basis of the Hilbert space of states[24] H_{S_n,τ_S} for which $\mathcal{O}$ observes $T_S(x)$ to be $\tau_S(x)$ for all $x \in S_n(y) \cap S$, and if $\{\,|\xi'_j\rangle : j\}$ is an orthonormal basis of the Hilbert space of states[25] H'_{S_n,τ'_S} for which $\mathcal{O}'$ observes $T'_S(x')$ to be $\tau'_S(x')$ for all $x' \in S_n(y') \cap S$, then

$$\lim_{n\to\infty} \frac{\langle\Psi'_n|\pi'_n\hat{T}'^{\mu\nu}(y')|\Psi'_n\rangle}{\langle\Psi'_n|\pi'_n|\Psi'_n\rangle} = \Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma \lim_{n\to\infty} \frac{\langle\Psi_n|\pi_n\hat{T}^{\rho\sigma}(y)|\Psi_n\rangle}{\langle\Psi_n|\pi_n|\Psi_n\rangle} \tag{4.2.18}$$

where $\pi_n = \sum_j |\xi_j\rangle\langle\xi_j|$, $\pi'_n = \sum_j |\xi'_j\rangle\langle\xi'_j|$, and $|\Psi'_n\rangle = U(\Lambda)|\Psi_n\rangle$.

To see why (4.2.18) holds, we first recall that π'_n will be independent of which orthonormal basis we choose for H_{S_n,τ'_S}.[26] Therefore, if we can show that $\{\,|\xi'_j\rangle \overset{\text{def}}{=} U(\Lambda)|\xi_j\rangle : j\}$ is an orthonormal basis of H_{S_n,τ'_S}, it will follow that $\pi'_n = U(\Lambda)\pi_n U(\Lambda)^{-1}$.

That the elements of $\{U(\Lambda)|\xi_j\rangle : j\}$ are orthonormal follows from the unitarity of $U(\Lambda)$ together with the orthonormality of $\{\,|\xi_j\rangle : j\}$. It remains for us to show

24. Thus, H_{S_n,τ_S} is the subspace of states $|\xi\rangle \in H_{S_n}$ for which $\hat{T}_S(x)|\xi\rangle = \tau_S(x)|\xi\rangle$ for all $x \in S_n \cap S$ as mentioned on page 178.

25. Thus, H'_{S_n,τ'_S} is the subspace of states $|\xi'\rangle \in H'_{S_n}$ for which $\hat{T}'_S(x')|\xi'\rangle = \tau'_S(x')|\xi'\rangle$ for all $x' \in S_n \cap S$.

26. We showed this was the case for π_n in footnote 7 on page 179.

that each $U(\Lambda)\,|\xi_j\rangle \in H'_{S_n,\tau'_S}$, and that any $|\xi'\rangle \in H'_{S_n,\tau'_S}$ can be expressed as a linear combination of the $U(\Lambda)\,|\xi_j\rangle$.

Well, first note that by (4.2.17) and a calculation similar to (4.2.13)

$$
\begin{aligned}
U(\Lambda)^{-1}\hat{T}'_S(x')U(\Lambda) &= U(\Lambda)^{-1}\hat{T}'^{\mu\nu}(x')\eta'_\mu(x')\eta'_\nu(x')U(\Lambda) \\
&= \Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma\hat{T}^{\rho\sigma}(x)\Lambda_\mu{}^\alpha\eta_\alpha(x)\Lambda_\nu{}^\beta\eta_\beta(x) \\
&= \Lambda^\mu{}_\rho\Lambda_\mu{}^\alpha\Lambda^\nu{}_\sigma\Lambda_\nu{}^\beta\hat{T}^{\rho\sigma}(x)\eta_\alpha(x)\eta_\beta(x) \\
&= \delta^\alpha_\rho\delta^\beta_\sigma\hat{T}^{\rho\sigma}(x)\eta_\alpha(x)\eta_\beta(x) \\
&= \hat{T}^{\alpha\beta}(x)\eta_\alpha(x)\eta_\beta(x) \\
&= \hat{T}_S(x)
\end{aligned}
\tag{4.2.19}
$$

From (4.2.19), we see that $U(\Lambda)\hat{T}_S(x) = \hat{T}'_S(x')U(\Lambda)$, so

$$
\hat{T}'_S(x')U(\Lambda)\,|\xi_j\rangle = U(\Lambda)\hat{T}_S(x)\,|\xi_j\rangle = \tau_S(x)U(\Lambda)\,|\xi_j\rangle = \tau'_S(x')U(\Lambda)\,|\xi_j\rangle
$$

for all $x' \in S_n(y') \cap S$, where we have used the fact $\tau'_S(x') = \tau_S(x)$ since $T_S(x)$ is a scalar. Therefore, $U(\Lambda)\,|\xi_j\rangle \in H'_{S_n,\tau'_S}$.

Now suppose that $|\xi'\rangle$ is a state for which $\mathcal{O}'$ observes $T'_S(x')$ to be $\tau'_S(x')$ for all $x' \in S_n(y') \cap S$, i.e. $\hat{T}'_S(x')\,|\xi'\rangle = \tau'_S(x')\,|\xi'\rangle$. From (4.2.19) we see that $\hat{T}_S(x)U(\Lambda)^{-1} =$

$U(\Lambda)^{-1}\hat{T}'_S(x')$, so

$$\hat{T}_S(x)U(\Lambda)^{-1}\,|\xi'\rangle = U(\Lambda)^{-1}\hat{T}'_S(x')\,|\xi'\rangle$$

$$= \tau'_S(x')U(\Lambda)^{-1}\,|\xi'\rangle \qquad (4.2.20)$$

$$= \tau_S(x)U(\Lambda)^{-1}\,|\xi'\rangle$$

where on the last line we have used the fact that $T_S(x)$ is a scalar. Therefore, $U(\Lambda)^{-1}\,|\xi'\rangle$ can be expressed as a linear combination of the basis elements $\{\,|\xi_j\rangle : j\}$ of H_{S_n,τ_S}, and hence $|\xi'\rangle$ can be expressed as a linear combination of $\{U(\Lambda)\,|\xi_j\rangle : j\}$.

Thus, we see that $\{\,|\xi'_j\rangle \overset{\text{def}}{=} U(\Lambda)\,|\xi_j\rangle : j\}$ is a spanning orthonormal subset of H'_{S_n,τ'_S}, so it must therefore be an orthonormal basis of H'_{S_n,τ'_S}. From this it follows that

$$\pi'_n = U(\Lambda)\pi_n U(\Lambda)^{-1}. \qquad (4.2.21)$$

Therefore,

$$\frac{\langle\Psi'_n|\pi'_n\hat{T}'^{\mu\nu}(y')|\Psi'_n\rangle}{\langle\Psi'_n|\pi'_n|\Psi'_n\rangle} = \frac{\langle\Psi_n|U(\Lambda)^{-1}U(\Lambda)\pi_n U(\Lambda)^{-1}\hat{T}'^{\mu\nu}(y')U(\Lambda)|\Psi_n\rangle}{\langle\Psi_n|U(\Lambda)^{-1}U(\Lambda)\pi_n U(\Lambda)^{-1}U(\Lambda)|\Psi_n\rangle}$$

$$= \frac{\langle\Psi_n|\pi_n U(\Lambda)^{-1}\hat{T}'^{\mu\nu}(y')U(\Lambda)|\Psi_n\rangle}{\langle\Psi_n|\pi_n|\Psi_n\rangle} \qquad (4.2.22)$$

$$= \frac{\langle\Psi_n|\pi_n\Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma\hat{T}^{\rho\sigma}(y)|\Psi_n\rangle}{\langle\Psi_n|\pi_n|\Psi_n\rangle}$$

where on the last line we have used (4.2.17). Thus, equation (4.2.18) holds, and hence Kent's model is Lorentz invariant.

Note that in this proof of Lorentz invariance, we don't need to take the limit of S_n as $n \to \infty$. That is, we could remove the $\lim_{n\to\infty}$ from equation (4.1.6) and consider a particular S_n, and the corresponding $\langle T^{\mu\nu}(y)\rangle_{\tau_S}$ would still be a rank-two tensor. Butterfield tells us that Kent's theory is Lorentz invariant because his algorithm respects the light cone structure of y.[27] However, this statement could be slightly misleading because we don't need to consider the subset $S^1(y) \subset S$ of locations outside the light cone of y in order to obtain a Lorentz invariant model. Doing the calculation on any Tomonaga-Schwinger hypersurface is sufficient to guarantee Lorentz invariance since any such hypersurface (e.g. S_n) is not altered at all by a Lorentz transformation – only its coordinate description changes under a Lorentz transformation, and so the additional information of the scalar $\tau_S(x)$ on $S_n \cap S$ is Lorentz invariant. The only reason we need to consider the limit $\lim_{n\to\infty} S_n$ and hence $S^1(y) = \lim_{n\to\infty} S_n \cap S$ is that it is only in the limit that we use all the available information in $\tau_S(x)$ to calculate $\langle T^{\mu\nu}(y)\rangle_{\tau_S}$.

4.3 Kent's Theory and the Problem of Outcomes

4.3.1 Overview

In section 2.7 we saw that decoherence theory by itself does not offer a solution to the problem of outcomes. In this section, we consider how the additional information in Kent's theory is sufficient to address this problem. We will explain this by again considering Kent's toy model discussed in section 3.5.

---

27. See Butterfield, "Peaceful Coexistence: Examining Kent's Relativistic Solution to the Quantum Measurement Problem," 30.

We thus suppose that a system is in a superposition $\psi_0^{\text{sys}} = c_1\psi_1^{\text{sys}} + c_2\psi_2^{\text{sys}}$ of two local states ψ_1^{sys} and ψ_2^{sys} where $|c_1|^2 + |c_2|^2 = 1$, and that there is a photon coming in from the left that interacts with the system.[28] We also suppose that y_1 is a spacetime location with spatial location z_1 between the two hypersurfaces S_0 and S, and we consider a hypersurface $S_n = S_n(y_1)$ in a sequence of hypersurfaces that each contain y_1 as described on page 162.

In what follows, I use the equation (4.1.6) to show how the state of the system in question goes from being an improper mixture to being a pure state as the additional variables in the measurement on the hypersurface S become available to condition on when calculating $\langle T^{\mu\nu}(y)\rangle_{\tau_S}$. This solves the problem of outcomes because when a state is given by an improper mixture, in reality, the system is not in any one of two states described by ψ_1 and ψ_2; rather, the system together with its environment is in a superposition of two states, the first being the ψ_1-state of the system with its corresponding environment state, and second being the ψ_2-state of the system with its corresponding environment state. However, once the system is described by a pure state, we can either say that the system is definitely in the state described by ψ_1, or say that the system is definitely in the state described by ψ_2. Thus, there is a definite outcome to the measurement of which of the two states ψ_1 or ψ_2 the system is in. Hence, in this example, we can see how Kent's theory is able to solve the problem of outcomes.

---

28. As in footnote 58 on page 165, we write ψ_i^{sys} for the wave function that corresponds to the state $|\psi_i^{\text{sys}}(t)\rangle$ with $\psi_i^{\text{sys}}(z,t) = \langle z|\psi_i^{\text{sys}}(t)\rangle$.

In order to obtain a sufficiently simple description of the state $|\Psi_n\rangle \in H_{S_n}$ of S_n for which we can use the formula (4.1.6) to calculate Kent's beable, we will use a coarse-grained model so that S_n is treated as a mesh of tiny cells[29] labeled by a sequence $(y_k)_{k=1}^{\infty}$. Thus, for each cell y_k there will be a Hilbert space H_k describing the state of that cell. We can think of each of these y_k as systems that can become entangled with one another, but we will assume that y_1 is entangled with only a finite number M of the other y_k which we label as $y_{k_1}, \ldots, y_{k_M}$. What this means is that the most general expression for $|\Psi_n\rangle$ will be of the form

$$|\Psi_n\rangle = \left(\sum_j \sum_{n \in \mathbb{N}^M} c_{j,n} |\xi_{1,j}\rangle \prod_{l=1}^{M} |\xi_{k_l,n_l}\rangle \right) |\Xi\rangle . \qquad (4.3.1)$$

In this expression, $\{ |\xi_{1,j}\rangle : j \}$ is an orthonormal basis of H_1, $\mathbb{N}^M$ means the set of all lists $(n_1, \ldots, n_M)$ with each $n_l \in \mathbb{N}$ where $\mathbb{N}$ is the set of positive integers greater than 0. The set of states $\{ |\xi_{k_l,m}\rangle : m \in \mathbb{N} \}$ form an orthonormal basis of H_{k_l} for each k_l, and the k_l are all distinct from each other and from 1. Also, M is chosen to be as small as possible so that any common factors of $|\Psi_n\rangle$ belong to $|\Xi\rangle$ which is a sum of states of the form $\prod_l |\xi_{\kappa_l}\rangle$ where the states $|\xi_{\kappa_l}\rangle \in H_{\kappa_l}$ range over all the cells of S_n not included in the set $\{k_l : l = 1, \ldots, M\}$. We also assume that each summand $c_{j,n} |\xi_{1,j}\rangle \prod_{l=1}^{M} |\xi_{k_l,n_l}\rangle |\Xi\rangle$ of $|\Psi_n\rangle$ contains a state in each H_k for every cell k of S_n. In other words, if $k \neq 1$ and does not belong to the set $\{k_l : l = 1, \ldots, M\}$, then

---

29. For a more detailed discussion of coarse-graining, see pp. 137 ff.

k belongs to the set $\{\kappa_l : \text{there exists } |\xi_{\kappa_l}\rangle \in H_{\kappa_l} \text{ appearing in } |\Xi\rangle\,.\}$. Also, we will give H_{S_n} an inner product so that if

$$|\Psi'_n\rangle = \left(\sum_j \sum_{n\in\mathbb{N}^M} c'_{j,n} |\xi_{1,j}\rangle \prod_{l=1}^{M} |\xi_{k_l,n_l}\rangle \right) |\Xi'\rangle,$$

then

$$\langle\Psi'_n|\Psi_n\rangle = \left(\sum_j \sum_{n\in\mathbb{N}^M} \overline{c'_{j,n}} c_{j,n} \right) \langle\Xi'|\Xi\rangle$$

where $\langle\Xi'|\Xi\rangle$ is defined in the obvious way. With this inner product, we will assume that $|\Psi_n\rangle$ is appropriately normalized so that $\langle\Psi_n|\Psi_n\rangle = 1$. If we also assume that $\langle\Xi|\Xi\rangle = 1$, it will follow that $\sum_j \sum_{n\in\mathbb{N}^M} |c_{j,n}|^2 = 1$.

Now in order to see how Kent's model addresses the problem of outcomes, we will need to consider several scenarios from his toy model. In each scenario, we will use the decomposition (4.3.1) of $|\Psi_n\rangle$ to calculate the reduced density matrix that encapsulates all the information needed to calculate expectation values at different spacetime locations.

First, consider Figure 4.3 which depicts the hypersurface $S_n(y_1^a)$ for a spacetime location y_1^a that occurs before the photon has interacted with the system.

The small gray squares correspond to the summands that appear in (4.3.1). If the system were in the ψ_1^{sys}-state, then the state describing $S_n(y_1^a)$ would have a factor $|\psi_1^{\text{sys}}\rangle \in H_1$ indicating that there is a non-zero mass at the y_1^a-cell, and there would
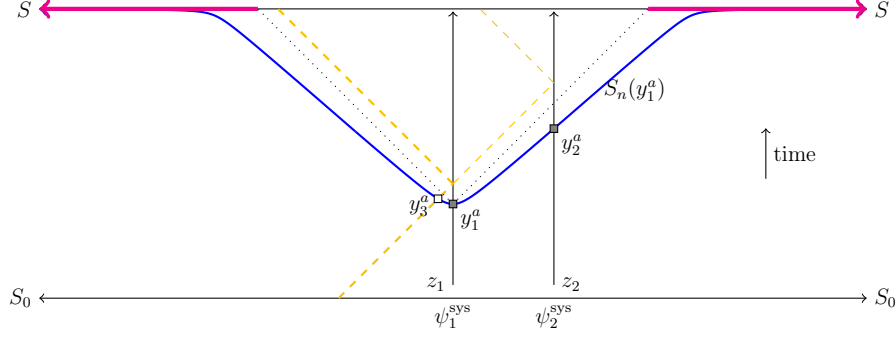
Figure 4.3. Depiction of a superposition of two local states at z_1 and z_2 before the photon has interacted with them. The small gray squares indicate cells in $S^1(y_1^a)$ whose states are among the summands in (4.3.1) rather than in Ξ. The white square indicates a cell in $S_n(y_1^a)$ whose state is a factor in Ξ.

also be a factor $|0_2\rangle \in H_2$ which we use to indicate that there is zero mass/energy at y_2^a. There is also an incoming photon at the y_3^a-cell, and so we use $|\gamma_3\rangle$ to indicate that there is a photon there. Thus, if the system were in the ψ_1^{sys}-state, we would write the state of $S_n(y_1^a)$ as $|\Psi_n\rangle = |\psi_1^{\mathrm{sys}}\rangle\, |0_2\rangle\, |\gamma_3\rangle\, |\Xi'\rangle$, where $|\Xi'\rangle$ describes the states of all the other cells of $S_n(y_1^a)$. In this very simple scenario, we set $|\Xi'\rangle = \prod_{k \neq 1,2,3} |0_k\rangle$ to indicate that there is zero mass/energy at all the other y_k cells belonging to $S_n(y_1^a)$.

On the other hand, if the system were in the state ψ_2^{sys}, then the state describing $S_n(y_1^a)$ would have a factor $|\psi_2^{\mathrm{sys}}\rangle \in H_2$ indicating that there is a non-zero mass at the y_2^a-cell, and there would also be a factor $|0_1\rangle \in H_1$ which we use to indicate that there is zero mass at y_1^a, and again the y_3^a-cell would be in the $|\gamma_3\rangle$-state, and every other cell would be described by $|\Xi'\rangle$ just as if the system had been in the ψ_1^{sys}-state. Therefore, when the system is in the state ψ_2^{sys}, we would write the state of $S_n(y_1^a)$ as $|\Psi_n\rangle = |0_1\rangle\, |\psi_2^{\mathrm{sys}}\rangle\, |\gamma_3\rangle\, |\Xi'\rangle$.

210

Now since the system is actually in a supposition $\psi_0^{\text{sys}} = c_1 \psi_1^{\text{sys}} + c_2 \psi_2^{\text{sys}}$, the state of $S_n(y_1^a)$ will be

$$|\Psi_n\rangle = \left(c_1 |\psi_1^{\text{sys}}\rangle |0_2\rangle + c_2 |0_1\rangle |\psi_2^{\text{sys}}\rangle\right) |\gamma_3\rangle |\Xi'\rangle = \left(c_1 |\psi_1^{\text{sys}}\rangle |0_2\rangle + c_2 |0_1\rangle |\psi_2^{\text{sys}}\rangle\right) |\Xi\rangle$$

where we have absorbed the $|\gamma_3\rangle$-state into $|\Xi\rangle$ (i.e. $|\Xi\rangle = |\gamma_3\rangle |\Xi'\rangle$).

Now as it stands, the state $|\Psi_n\rangle$ describing $S_n(y_1^a)$ has a definite mass-energy density $\tau_S(x)$ for $x \in S_n(y_1^a) \cap S$, namely 0. Thus, if π_n is the operator featuring in (4.1.6) that corresponds to this definite mass-energy density, then $\pi_n |\Psi_n\rangle = |\Psi_n\rangle$. Therefore, equation (4.1.6) for Kent's beables tells us that

$$\langle T^{\mu\nu}(y_1^a)\rangle_{\tau_S} = \langle \Psi_n | \hat{T}^{\mu\nu}(y_1^a) | \Psi_n\rangle, \qquad (4.3.2)$$

where we have also used the fact that $\langle \Psi_n | \Psi_n\rangle = 1$.

Now as we saw in section 2.4, if we are interested only in the expectation values of observables for a system $\mathcal{S}$ contained within a universe $\mathcal{U} = \mathcal{S} + \mathcal{E}$, then the information needed to do this can be encapsulated in the reduced density matrix for $\mathcal{S}$. Thus, if the universe is described by a state $|\Psi\rangle = \sum_j c_j |\psi_j\rangle_{\mathcal{S}} |E_j\rangle_{\mathcal{E}}$ with corresponding density matrix $\hat{\rho} = |\Psi\rangle\langle\Psi| \in M(H_{\mathcal{U}})$, then the reduced density matrix $\hat{\rho}_{\mathcal{S}} \in M(H_{\mathcal{S}})$ is the Hermitian operator acting on the state space $H_{\mathcal{S}}$ with the property that

$$\langle \hat{O}_{\mathcal{U}}\rangle_\rho = \text{Tr}_{\mathcal{S}}(\hat{\rho}_{\mathcal{S}} \hat{O}_{\mathcal{S}}) \qquad \text{(2.4.14 revisited)}$$

where $\hat{O}_{\mathcal{S}}$ is an observable on $H_{\mathcal{S}}$, and $\hat{O}_{\mathcal{U}}$ is the corresponding observable on $H_{\mathcal{U}}$. Furthermore, we also have

$$\hat{\rho}_{\mathcal{S}} = \sum_j |c_j|^2 |\psi_j\rangle\langle\psi_j| + \sum_{j \neq k} c_j \overline{c_k} \langle E_k | E_j \rangle |\psi_j\rangle\langle\psi_k| .^{30} \tag{4.3.3}$$

We can thus apply this to the situation at hand by taking S_n to be our universe $\mathcal{U}$ and y_1^a to be the system $\mathcal{S}$, and $S_n \setminus \{y_1^a\}$ to be the environment $\mathcal{E}$. If we assume that $\langle 0_2 | \psi_2^{\text{sys}} \rangle = 0$, then by (4.3.3), the corresponding reduced density matrix $\hat{\rho}_{y_1^a}$ takes the form of an improper mixture

$$\hat{\rho}_{y_1^a} = |c_1|^2 |\psi_1^{\text{sys}}\rangle\langle\psi_1^{\text{sys}}| + |c_2|^2 |0_1\rangle\langle 0_1| . \tag{4.3.4}$$

Therefore, by (4.3.2) and (4.3.4), Kent's beable at y_1^a will take the form

$$\begin{aligned} \langle T^{\mu\nu}(y_1^a) \rangle_{\tau_S} &= \langle \Psi_n | \hat{T}^{\mu\nu}(y_1^a) | \Psi_n \rangle \\ &= \text{Tr}_{y_1^a} (\hat{\rho}_{y_1^a} \hat{T}^{\mu\nu}(y_1^a)) \\ &= |c_1|^2 \langle \psi_1^{\text{sys}} | \hat{T}^{\mu\nu}(y_1^a) | \psi_1^{\text{sys}} \rangle + |c_2|^2 \langle 0_1 | \hat{T}^{\mu\nu}(y_1^a) | 0_1 \rangle . \end{aligned} \tag{4.3.5}$$

But since $\hat{\rho}_{y_1^a}$ is an improper mixture, we cannot give (4.3.5) an ignorance interpretation – the universe is still in the superposition state

$$|\Psi_n\rangle = c_1 |\psi_1^{\text{sys}}\rangle |E_1\rangle + c_2 |0_1\rangle |E_2\rangle$$

---

30. cf. (2.4.16)

where $|E_1\rangle = |0_2\rangle |\Xi\rangle$ and $|E_2\rangle = |\psi_2^{\text{sys}}\rangle |\Xi\rangle$.

In order to solve the problem of outcomes, we need to provide a satisfactory explanation (i.e. an explanation that is consistent with special relativity and the predictions of standard quantum theory) of how the superposition state $|\Psi_n\rangle$ effectively goes to either the state $|\psi_1^{\text{sys}}\rangle |E_1\rangle$ or to the state $|0_1\rangle |E_2\rangle$.[31] In terms of density operators, this means we need to show how the improper state (4.3.4), transitions to a pure state of the form $|\psi_1^{\text{sys}}\rangle\langle\psi_1^{\text{sys}}|$ or $|0_1\rangle\langle0_1|$.

To this end, let us consider Kent's beables at the spacetime location y_1^b depicted in figure 4.4. The state of $S_n(y_1^b)$ will then be
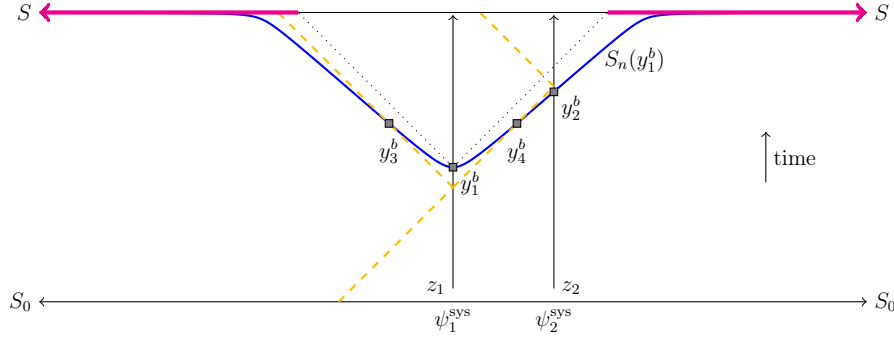
Figure 4.4. Depiction of a superposition of two local states at z_1 and z_2 with $S_n(y_1^b)$ being after the photon has interacted without the photon intersecting $S_n(y_1^b) \cap S$. The small gray squares indicate cells in $S^1(y_1^b)$ whose states are among the summands in (4.3.1).

$$|\Psi_n\rangle = \left(c_1 |\psi_1^{\text{sys}}\rangle |0_2\rangle |\gamma_3\rangle |0_4\rangle + c_2 |0_1\rangle |\psi_2^{\text{sys}}\rangle |0_3\rangle |\gamma_4\rangle\right) |\Xi\rangle$$

31. Cf. the initial discussion of the problem of outcomes on page 96. I have used the word 'effectively' to qualify this sentence since it is not necessary to prove that there actually is such a transition from the state $|\Psi_n\rangle$ to either the state $|\psi_1^{\text{sys}}\rangle |E_1\rangle$ or to the state $|0_1\rangle |E_2\rangle$. Rather, it is sufficient to show that if we consider any observable $\hat{O}_\mathcal{S}$ acting on $\mathcal{S}$, then the expectation value takes the form

where the notation is analogous to that in the previous example. Since no photon detections are registered on $S_n(y_1^b) \cap S$, we again have $\pi_n \, |\Psi_n\rangle = |\Psi_n\rangle$ so that the reduced density matrix $\hat{\rho}_{y_1^b}$ will again be given by (4.3.4) with y_1^a replaced by y_1^b. However, in this case, Kent's beables $\langle T^{\mu\nu}(y_1^b)\rangle_{\tau_S}$ will not be given by (4.3.5) because in the limit as $n \to \infty$, the photon *will* be registered on $S_n(y_1^b) \cap S$.

To deal with the case when a photon is registered on $S_n(y_1^b) \cap S$, we consider a third example as depicted in figure 4.5.
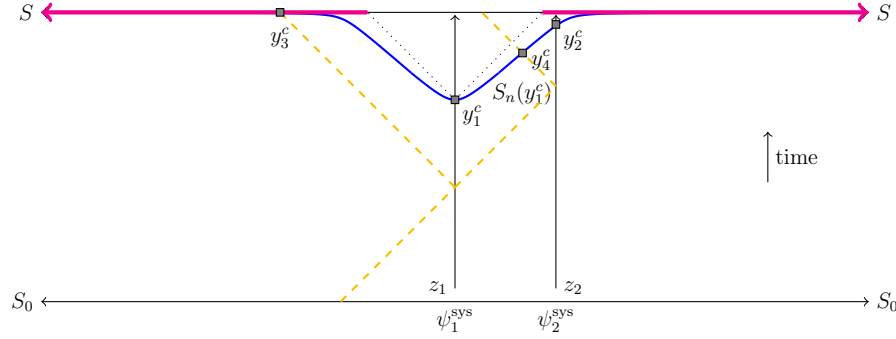
Figure 4.5. Depiction of a superposition of two local states at z_1 and z_2 with y_1^c sufficiently late that the photon intersects $S_n(y_1^c) \cap S$. The small gray squares indicate cells in $S^1(y_1^c)$ whose states are among the summands in (4.3.1)

In this case, the state of $S_n(y_1^c)$ will be

$$|\Psi_n\rangle = \left(c_1 \, |\psi_1^{\mathrm{sys}}\rangle \, |0_2\rangle \, |\gamma_3\rangle \, |0_4\rangle + c_2 \, |0_1\rangle \, |\psi_2^{\mathrm{sys}}\rangle \, |0_3\rangle \, |\gamma_4\rangle \right) |\Xi\rangle ,$$

$\langle \psi_1^{\mathrm{sys}}|\hat{O}_{\mathcal{S}}|\psi_1^{\mathrm{sys}}\rangle$ or $\langle 0_1|\hat{O}_{\mathcal{S}}|0_1\rangle$ once there is an outcome, for then the system $\mathcal{S}$ has all the properties consistent with it being in the state $|\psi_1^{\mathrm{sys}}\rangle$ or the state $|0_1\rangle$ respectively.

but now we have to consider the fact that the photon intersects $S_n(y_1^c) \cap S$. There are two possible (notional) measurement outcomes that can occur on $S_n(y_1^c) \cap S$: either $T_S = \tau_{S,1}$ where $\tau_{S,1}(y_3^c) \neq 0$, or $T_S = \tau_{S,2}$ where $\tau_{S,2}(y_3^c) = 0$.

The case $T_S = \tau_{S,1}$ indicates that there is a photon detection at y_3^c so that the local state at the y_3^c-cell is $|\gamma_3\rangle$. Therefore, if we write $\pi_{n,1}$ for the operator π_n, we have

$$\pi_{n,1} |\Psi_n\rangle = c_1 |\psi_1^{\text{sys}}\rangle |0_2\rangle |\gamma_3\rangle |0_4\rangle |\Xi\rangle .$$

Therefore, $\langle \Psi_n | \pi_{n,1} \hat{T}^{\mu\nu}(y_1^c) | \Psi_n \rangle = |c_1|^2 \langle \psi_1^{\text{sys}} | \hat{T}^{\mu\nu}(y_1^c) | \psi_1^{\text{sys}} \rangle$ and $\langle \Psi_n | \pi_{n,1} | \Psi_n \rangle = |c_1|^2$. Hence, by (4.1.6), Kent's beables at y_1^c will be

$$\langle T^{\mu\nu}(y_1^c) \rangle_{\tau_{S,1}} = \langle \psi_1^{\text{sys}} | \hat{T}^{\mu\nu}(y_1^c) | \psi_1^{\text{sys}} \rangle .$$

From this, it follows that the reduced density matrix at y_1^c will take the form of a pure state:

$$\hat{\rho}_{y_1^c} = |\psi_1^{\text{sys}}\rangle\langle\psi_1^{\text{sys}}| . \tag{4.3.6}$$

On the other hand, for the case when $T_S = \tau_{S,2}$, this indicates that there is no photon detection at y_3^c, so that the local state at the y_3^c-cell will be $|0_3\rangle$. So if we now write $\pi_{n,2}$ for the operator π_n, we have

$$\pi_{n,2} |\Psi_n\rangle = c_2 |0_1\rangle |\psi_2^{\text{sys}}\rangle |0_3\rangle |\gamma_4\rangle |\Xi\rangle .$$

Therefore, $\langle\Psi_n|\pi_{n,2}\hat{T}^{\mu\nu}(y_1^c)|\Psi_n\rangle = |c_2|^2\,\langle 0_1|\hat{T}^{\mu\nu}(y_1^c)|0_1\rangle$ and $\langle\Psi_n|\pi_{n,2}|\Psi_n\rangle = |c_2|^2$, and

so by (4.1.6), Kent's beables at y_1^c will be

$$\langle T^{\mu\nu}(y_1^c)\rangle_{\tau_{S,2}} = \langle 0_1|\hat{T}^{\mu\nu}(y_1^c)|0_1\rangle\,.$$

In this case, the reduced density matrix at y_1^c will be

$$\hat{\rho}_{y_1^c} = |0_1\rangle\langle 0_1|\,, \tag{4.3.7}$$

which is again a pure state.

In these examples we have therefore seen how the additional information concerning photon detection on $S_n(y_1) \cap S$ is able to determine whether the reduced density matrix at y_1 is a pure state or an improper mixture. Hence, Kent's theory offers an answer to d'Espagnat's problem of outcomes. As mentioned in section 2.7, d'Espagnat noticed that with decoherence theory alone, we are not entitled to give an ignorance interpretation to the reduced density matrix for a system that is an improper mixture, and thus we are not able to conclude that an outcome has occurred. However, if the reduced density matrix of a system goes from being an improper mixture to a pure state of the form $|\psi\rangle\langle\psi|$ as it does when Kent's additional information is taken into account, then we can say that an outcome has occurred, namely the outcome of the system being in the state $|\psi\rangle$.

Let us now consider Kent's theory in the light of Shimony's notion of Outcome Independence (OI) as defined in section 1.9. Butterfield[32] tries to answer the question of whether OI holds in Kent's theory by considering an example that builds on Kent's toy model. Butterfield's example is designed to capture the salient features of a Bell experiment where two spatially separated observers always observe opposite outcomes of some measurement. Following Kent, Butterfield thus considers a universe in one spatial dimension. In this universe, there are two entangled systems, a left-system and a right-system as depicted in figure 4.6.
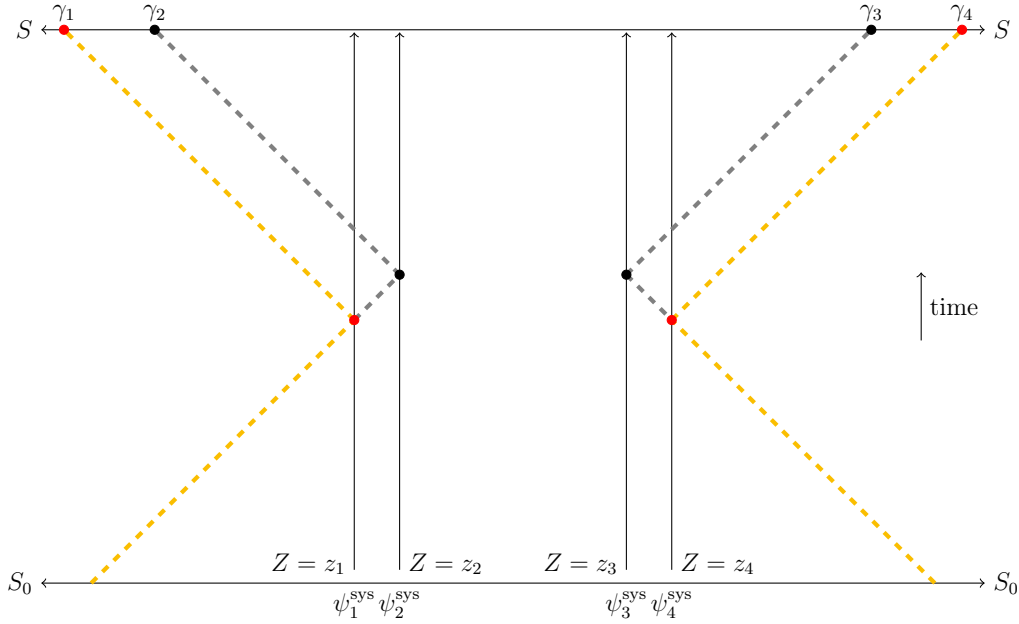
Figure 4.6. Butterfield's thought experiment for analyzing OI

32. See Butterfield, "Peaceful Coexistence: Examining Kent's Relativistic Solution to the Quantum Measurement Problem," 30–32

Two locations z_1 and z_2 with $z_2 > z_1$ belong to a left-system, and there are two possible outcomes for a measurement on the left-system: either all the mass/energy of the left-system is localized at z_1 or all the mass/energy of the left-system is localized at z_2. These two possibilities are analogous to a spin up or a spin down measurement outcome in a Stern-Gerlach experiment. Likewise, two locations z_3 and z_4 with $z_3 < z_4$ and $z_3 \gg z_2$ belong to a right-system, and again, there are two possible measurement outcomes: either all the mass/energy of the right-system is localized at z_3 or all the mass/energy of the right-system is localized at z_4.

The initial joint state of the two systems is $a \, |\psi_1\rangle \, |\psi_4\rangle + b \, |\psi_2\rangle \, |\psi_3\rangle$. This means that the left-system will be found to be localized at z_1 with probability $|a|^2$, and at z_2 with probability $|b|^2$, and if the left-system is localized at z_1, the right system must be localized at z_4, whereas if the left-system is localized at z_2, then the right system must be localized at z_3.

Now Butterfield supposes that there are two photons, one coming in from the left that interacts with the left system, and one coming in from the right that interacts with the right system. As in Kent's toy model, there is a late time hypersurface S, on which the photons are "measured". Since the joint state of the two systems is in superposition, there will be two possible measurement outcomes for the two photons that arrive at S. Either the left-photon is measured at γ_1 and the right-photon is measured at γ_4, or the left-photon is measured at γ_2 and the right photon is measured at γ_3. Thus, if we suppose that the (notional) measurement for $T_S(x)$ yields an energy distribution $\tau_S(x)$ that is nonzero at γ_1 and γ_4, but is zero at γ_2 and γ_3, then we can

say that the outcome of the measurement on the two systems is that the left system is localized at z_1 and the right system is localized at z_4. Moreover, the probability of this outcome is 1 given that the (notional) measurement of $T_S(x)$ on S is $\tau_S(x)$. In other words, this model is deterministic. But as we saw on page 49, if a model is deterministic, then OI must hold. This is the conclusion that Butterfield draws.

Now if Kent's theory is to be consistent with special relativity, OI being satisfied might initially seem concerning. Indeed, we saw in section 1.9 that OI implies the negation of PI in theories that produce the same predictions as standard quantum physics, and the negation of PI is not consistent with special relativity.[33] However, there is one salient feature of a Bell experiment that is not captured in Butterfield's scenario, namely, in a Bell experiment one can perform different measurements. PI and its negation only make sense when there are parameters that can be changed.

33. At this point, one might make the following remark: the argument that a violation of PI constitutes a violation of relativity is based on the idea that if one knew the value of the hidden data, one could transmit messages at superluminal speed. But when the hidden data is grounded in data about the future hypersurface S, then the fact that if one knew this data, one could transmit messages superluminally should not be a cause for concern since all sorts of things become possible if you can know future contingents.

To this remark, there are two responses that one could make. Firstly, it would be somewhat misleading in the context of Kent's interpretation to say that a knowledge of data about S implies a knowledge of future contingents since when calculating Kent's $T^{\mu\nu}(y)$-beables $\langle T^{\mu\nu}(y) \rangle_{\tau_S}$, the only knowledge about the data of S that is needed is for regions of S outside the light cone of y, and whether this data is about something in the future or in the past is going to depend on what frame of reference one is in. Only when the data is within the light cone of y will we be able to say categorically that knowledge of this data constitutes knowledge of future contingents. But Kent's beables are not dependent on such data.

Secondly, as mentioned on page 40, the main concern with a violation of PI is not simply the possibility of superluminal signaling, but rather the possibility of the propagation of superluminal effects, of which superluminal signaling would be a very clear demonstration. But whether or not there is such a demonstration, a violation of PI seems to be saying that effects can be propagated superluminally, and this is unacceptable to adherents of relativity theory.

Furthermore, in the proof that OI implies the negation of PI,[34] it is assumed that the choice of parameter is not determined by the hidden variable λ. If the choice of parameters was determined by λ, then for $\hat{a} \neq \hat{b}$, at least one of the probabilities $P_{\lambda,\hat{a},\hat{c}}(\hat{\mathbf{a}}+,\hat{\mathbf{c}}+)$, $P_{\lambda,\hat{c},\hat{b}}(\hat{\mathbf{c}}+,\hat{\mathbf{b}}+)$ or $P_{\lambda,\hat{a},\hat{b}}(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+)$ would not be well-defined.[35] Even though Butterfield is only considering OI in his thought experiment, a proper analysis of OI shouldn't be undertaken without considering an experiment with parameters (e.g. knob settings that correspond to measurement axes of a Stern-Gerlach experiment). This is because the determination of whether OI holds will depend on what one counts as being the hidden variable data of a system, and we need the hidden variable of a system to be such that the notion of PI is well-defined. Otherwise, one's verdict on OI will be irrelevant to Shimony's analysis of why Bell's inequality fails to hold.

In the next section I will discuss Leegwater's criteria for what one should count as being the hidden variable data of a system, and in sections 4.6 and 4.7, I will consider toy models that takes parameter settings into account.

4.5 Hidden Variables and the Colbeck-Renner theorem

In this section we will consider hidden variables in the light of the Colbeck-Renner theorem. Roughly speaking, the Colbeck-Renner theorem says there are no hidden

---

34. The proof that determinism implies the negation of PI (on pages 44 to 47), also assumes that the choice of parameter is not determined by the hidden variable λ.

35. If λ did determine the choice of measurements, either $P(\lambda,\hat{a},\hat{c}) = 0$, $P(\lambda,\hat{c},\hat{b}) = 0$, or $P(\lambda,\hat{a},\hat{b}) = 0$. So for example, if we thought of $P_{\lambda,\hat{a},\hat{b}}(X,Y)$ as a conditional probability $P(X,Y|\lambda,\hat{a},\hat{b})$, and the probability $P(\lambda,\hat{a},\hat{b}) = 0$, then according to the definition of conditional probability, $P(X,Y|\lambda,\hat{a},\hat{b}) = \frac{0}{0}$.

variables theories that are of interest to quantum physicists – either the hidden variables will be redundant or the hidden variable theory will be incompatible with standard quantum theory. It might therefore seem that the Colbeck-Renner theorem presents a serious challenge to Kent's theory. However, the challenge really hinges on what criteria the data of a theory must satisfy for it to be classified as hidden variable data. A careful analysis of the kind of data Kent's theory relies on reveals that the Colbeck-Renner theorem is not as serious a challenge to Kent's theory as it first might seem.

First we need to consider the criteria a set of data should satisfy if it is to be classified as hidden variable data of a physical system $\mathcal{S}$. The criteria we discuss below can be found either explicitly or implicitly in Leegwater's proof of the Colbeck-Renner theorem.[36] The first criterion we will discuss is the following:

1. all the information of λ is about $\mathcal{S}$ so that a change in λ corresponds to a change in the system $\mathcal{S}$.

Notice that Butterfield does not accept this criterion. Butterfield assumes that the hidden variables in Kent's theory consist in the outcome $\tau_S(x)$ of $T_S(x)$ over the whole of S. However, this assumption is going to cause difficulties in the context of Shimony's analysis. This is because in Kent's theory, the information in $\tau_S(x)$ over the whole of S clearly would determine which parameters are chosen in a Bell experiment, for this information would determine where a silver atom coming out of a

---

36. See Leegwater, "An impossibility theorem for parameter independent hidden-variable theories"

Stern-Gerlach apparatus would be detected on a detection screen (as depicted in figure 1.4), and from the position of this detection, one could determine the orientation of the magnetic field used in the Stern-Gerlach experiment. So if we stipulated that $\lambda = \tau_S$ is the hidden variable data of every system in Kent's theory, then Kent's theory wouldn't satisfy the preconditions necessary for defining OI and PI. This would make Kent's theory radically different from the Bohmian interpretation where one can define OI and PI because the hidden variables in the Bohmian interpretation, being the positions and momenta of the particles, are independent of the measurement choices. An unfortunate consequence of not being able to define OI and PI is that we wouldn't be able to evaluate Kent's theory in the light of Shimony's analysis of why Bell's inequality fails to hold.

But it is not obvious that we should stipulate that $\lambda = \tau_S$ is the hidden variable data of every system in Kent's theory. Just because we give τ_S a single label λ, it doesn't follow that τ_S is a single piece of information. There is typically going to be a huge amount of information in τ_S, and so for a given system $\mathcal{S}$, we should carefully discern what collection of information in τ_S should be stipulated as being the hidden variable data λ of $\mathcal{S}$, hence criterion 1. Criterion 1 shouldn't be that difficult to satisfy, for if λ contained information that could change without this corresponding to any change in $\mathcal{S}$, then we should be able to discard this irrelevant information when considering $\mathcal{S}$ and redefine what λ should be for the system.

In the Bohmian interpretation, the positions and momenta of the particles that constitute a system would fulfil criterion 1. On the other hand, all the information in

τ_S of Kent's theory would not fulfil this criterion unless, of course, $\mathcal{S}$ was the whole universe. So in order for Kent's theory to satisfy criterion 1, we just need to discard the information of τ_S that is irrelevant to the system $\mathcal{S}$ that is being considered in order to obtain the appropriate hidden variable data λ.

Note, however, that we don't insist that a difference in $\mathcal{S}$ entails a difference in λ. This is because a hidden-variables theory is envisaged as augmenting standard quantum theory. So in the case when $\mathcal{S}$ is not entangled with any other system, there will be a quantum state describing $\mathcal{S}$, and this quantum state can be other than it is (indicating that $\mathcal{S}$ can be in a different physical state) whilst the hidden variable remains the same. We thus impose a second criterion for a hidden-variables theory:

2. If λ is the hidden variable of a system $\mathcal{S}$ and if $|\phi\rangle$ is the quantum state of $\mathcal{S}$ or of some composite system $\mathcal{U}$ that contains $\mathcal{S}$ as a subsystem, then it is possible for there to be a different quantum state $|\phi'\rangle$ of $\mathcal{S}$ (or $\mathcal{U}$) while the hidden variable λ remains unchanged, and it is possible for there to be a different hidden variable λ while $|\phi\rangle$ remains unchanged.

This criterion is satisfied in the Bohmian interpretation, since the quantum state is the pilot wave itself. The pilot wave could be other than it is without any of the positions and momenta of the particles changing, but changing the pilot wave would result in a physical change of the system since the pilot wave governs how the positions and the

momenta of the particles subsequently evolve over time.[37] We will discuss the failure of criterion 2 for Kent's theory at the end of this section.

Another criterion implicit in Leegwater's proof for a set of data λ to constitute the hidden variable data of a system $\mathcal{S}$ is the following:

3. it should be possible to change the measurement parameters when measuring $\mathcal{S}$ without this determining what λ should be prior to the interaction of $\mathcal{S}$ with the measurement device.[38]

It is worth noting that we used criterion 3 when showing that OI implies the negation of PI. If this criterion doesn't hold, we cannot even begin to consider whether PI holds

---

37. It's worth noting that in the Bohmian interpretation, while the hidden variables of a system can change without the quantum state changing and vice versa, the quantum state in the Bohmian interpretation still places constraints on the hidden variables. For example, as explained in Dürr, Goldstein, and Zanghì, *Quantum physics without quantum philosophy*, 31, for a system of N particles with the kth particle located at $\mathbf{x}_k$, we have

$$\frac{\mathrm{d}\mathbf{x}_k}{\mathrm{d}t} = \frac{\hbar}{m_k} \operatorname{Im} \frac{\boldsymbol{\nabla}_k \psi}{\psi}(\mathbf{x}_1, \ldots, \mathbf{x}_N, t), \tag{4.5.1}$$

where $\psi(\mathbf{x}_1, \ldots, \mathbf{x}_N, t)$ is the quantum wave function for the system, t is time, m_k is the mass of the kth particle, and $\boldsymbol{\nabla}_k$ is the gradient corresponding to the position variable $\mathbf{x}_k$. Because of the ψ term in the denominator of (4.5.1), a particle configuration $(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ for which $\psi(\mathbf{x}_1, \ldots, \mathbf{x}_N, t) = 0$ is not permitted. However, since the Bohmian interpretation assumes that ψ is continuous and differentiable, it won't be possible for the wave function to fully constrain the particle locations, for the only candidate wave function that could do this would be a Dirac delta function, but such a function is not continuous. Furthermore, we will typically be able to assume that the potential term that appears in the Schrödinger equation is non-zero almost everywhere (given that the potential term will typically emerge from particles each of which produce a potential that is inversely proportional to the distance from the particle) which will mean that $\psi(\mathbf{x}_1, \ldots, \mathbf{x}_N, t) \neq 0$ almost everywhere. Hence, in the Bohmian interpretation, even though some configurations may be highly improbable, the wave function places constraints on the hidden variables almost nowhere/never.

38. It is conceivable that the measurement parameters might constrain what λ could be. For instance, in the Bohmian interpretation, changing the measurement settings will change the particle locations of the particles that constitute the buttons of the apparatus, and so the wave function might be such that for some configuration of the particles of the buttons, the wave function is zero at some locations in the vicinity of the particles being measured, and this would then mean that the particles being measured couldn't be at those locations. However, such constraints will occur almost nowhere/never as noted in footnote 37 on page 224.

in a given theory. This is because the criterion for PI depends on the probability $P_{\lambda,\hat{a},\hat{b}}(\hat{\mathbf{a}}\pm,\hat{\mathbf{b}}\pm)$ being well-defined, but if the choice of $\hat{\boldsymbol{a}}$ or $\hat{\boldsymbol{b}}$ affects what λ should be, then one wouldn't be able to define $P_{\lambda',\hat{a},\hat{b}}(\hat{\mathbf{a}}\pm,\hat{\mathbf{b}}\pm)$ for $\lambda' \neq \lambda$.[39] But even without any consideration of PI, a rejection of criterion 3 would still seem undesirable, for it would be rather strange if changing the choice of measurement parameters necessarily changed the hidden variable λ of the system $\mathcal{S}$, and hence by criterion 1 the system $\mathcal{S}$ itself.[40] Moreover, if a scientist chose the measurement parameter before it was determined which system was going to be measured, then it wouldn't even be clear which λ of which system we were talking about.[41]

In the Bohmian interpretation, the positions and momenta of the particles that constitute a system would fulfil criterion 3. In the case of Kent's interpretation, we would have to choose carefully the subset of τ_S for the λ of the system being measured

---

39. Cf. footnote 35 on page 220.

40. Of course, in the quantum world, strangeness is not necessarily an argument against something being true.

41. If the scientist's decisions were entirely determined by physical causes, then there wouldn't necessarily be this problem of associating the choice of measurement with the λ of the system, because the physical state of the universe could determine which choice of measurement is going to be made as well as which system is going to be measured. But even in an entirely deterministic universe, a correlation between measurement setting and the state of a system would be rather surprising. Maudlin puts it this way:

> [The measurement settings] could be coupled to a computer, or a throw of dice, or the stock market, or the whim of the experimenter, and so on. So if the behavior of each [particle] is influenced only by events in its past light cone, there would have to be fundamental laws connecting the [particle's] trajectory to distant facts about the disposition of dice and the number of shares of IBM being sold and so on, a different connection for each possible mechanism for determining the setting. This strains credulity. (Tim Maudlin, *Quantum non-locality & relativity* (Wiley-Blackwell, 2011), 57).

if criterion 3 is to hold. But it seems that this should be possible, for one can imagine just changing the part of τ_S that corresponded to where the measurement parameter dial/knob was without this changing Kent's stress-momentum tensor for a system $\mathcal{S}$ about to be measured immediately before it interacts with the measurement device. The only reason for thinking that criterion 3 might not always be possible under Kent's interpretation would be in the unlikely event of the measurement parameter dial/knob together with the system $\mathcal{S}$ being measured forming an entangled system prior to $\mathcal{S}$ being measured,[42] for only then could a measurement of the mass-energy density on the hypersurface S due to light being reflected from the measuring apparatus determine the state of the system $\mathcal{S}$ being measured. But once the measurement on the part of the hypersurface S has been made so as to determine the measuring apparatus settings, the measuring apparatus will no longer be entangled with the system $\mathcal{S}$. Then as long as there is time to change the measuring apparatus parameters after the apparatus's disentanglement with $\mathcal{S}$ and before $\mathcal{S}$ interacts with the apparatus during the measurement process, then the parts of the mass-energy density τ_S corresponding to where the measurement parameter dial/knob is could change without this affecting the state of $\mathcal{S}$.

Closely related to criterion 3 is the following criterion:

---

42. Normally entanglement of $\mathcal{S}$ with the measuring apparatus would only occur during the measurement of the system $\mathcal{S}$ by the measuring apparatus.

4. There is a range of possible values λ for a system $\mathcal{S}$, and for each possible value, we can assign a probability p_λ that λ obtains, and we can do so in such a way that p_λ is independent of any choice of measurement that is to be made on $\mathcal{S}$.

Butterfield refers to criterion 4 as the 'no-conspiracy' assumption, and it is one of the criteria assumed in the Colbeck-Renner theorem. However, Butterfield adds that 'no-conspiracy' is a rather unfair label for this criterion since there wouldn't necessarily be anything conspiratorial if this assumption was violated.[43] But in saying this, Butterfield is envisaging λ to be the whole of τ_S, and this is in violation of criterion 1. Since we can expect the whole of τ_S (together with the universal state $|\Psi_0\rangle$) to determine Kent's stress-momentum tensor for the measurement parameter dial/knob, then we wouldn't expect criterion 4 to hold. But on the other hand, if we adopt criterion 1 and assume that only information relevant to $\mathcal{S}$ is included in the λ for $\mathcal{S}$, then criterion 4 is more plausible. It ties in with common sense intuitions that scientists have free will and can make choices about which measurements they make, and that these choices are statistically independent of the states of the physical systems they are measuring. However, not everyone shares this intuition. If criterion 4 doesn't hold, then what a physical system does will depend on what property of this system one is about to measure. This kind of dependence is referred to as *superdeterminism*. Bell coined the term superdeterminism in a 1983 BBC interview, where he said:

> There is a way to escape the inference of superluminal speeds and spooky action at a distance. But it involves absolute determinism in

---

43. See Butterfield, "Peaceful Coexistence: Examining Kent's Relativistic Solution to the Quantum Measurement Problem," 34.

the universe, the complete absence of free will. Suppose the world is superdeterministic, with not just inanimate nature running on behind-the-scenes clockwork, but with our behavior, including our belief that we are free to choose to do one experiment rather than another, absolutely predetermined, including the "decision" by the experimenter to carry out one set of measurements rather than another, the difficulty disappears.

However, Sabine Hossenfelder disputes Bell's argument that the only way to escape the inference of superluminal speeds and spooky action at a distance is to violate free will. Rather, all that is needed to escape this inference is a violation of the statistical independence of the choice of measurement parameters and the state of the system being measured, and this violation is what Hossenfelder takes to be superdeterminism.[44]

In addition to these four criteria for the hidden variable data λ of a system $\mathcal{S}$, it is also desirable for a hidden-variables theory to satisfy PI and empirical adequacy. We defined PI for a two-outcome measurement on page 36, but it is easy to generalize the definition of PI for measurements with more than two outcomes. In this more generalized setting, we suppose that $\mathcal{A}$ is any system that is entangled with the system $\mathcal{S}$, and that the quantum state of the composite system $\mathcal{S} + \mathcal{A}$ is $|\phi\rangle_{\mathcal{S}+\mathcal{A}}$. We also suppose that $O_{\mathcal{S}}$ and $O_{\mathcal{A}}$ represent measurement procedures on $\mathcal{S}$ and $\mathcal{A}$ respectively, and that $o_{\mathcal{S}}$ and $o_{\mathcal{A}}$ represent particular measurement outcomes respectively. For

---

44. See Sabine Hossenfelder, "Does Superdeterminism save Quantum Mechanics? Or does it kill free will and destroy science?," Youtube, 2021, https://www.youtube.com/watch?v=ytyjgIyegDI.

a hidden variable λ for the system $\mathcal{S}$,[45] there will be a probability $P_\lambda^{|\phi\rangle_{S+A}}(O_\mathcal{S} = o_\mathcal{S} \,\&\, O_\mathcal{A} = o_\mathcal{A})$ (understandable in a frequentist sense) that the measurement outcomes of $O_\mathcal{S}$ and $O_\mathcal{A}$ will be $o_\mathcal{S}$ and $o_\mathcal{A}$ respectively. Given a second measurement procedure $O'_\mathcal{A}$ on $\mathcal{A}$, PI states that

$$\sum_{\substack{o_\mathcal{A} \text{ an} \\ \text{outcome} \\ \text{of } O_\mathcal{A}}} P_\lambda^{|\phi\rangle_{S+A}}(O_\mathcal{S} = o_\mathcal{S} \,\&\, O_\mathcal{A} = o_\mathcal{A}) = \sum_{\substack{o'_\mathcal{A} \text{ an} \\ \text{outcome} \\ \text{of } O'_\mathcal{A}}} P_\lambda^{|\phi\rangle_{S+A}}(O_\mathcal{S} = o_\mathcal{S} \,\&\, O'_\mathcal{A} = o'_\mathcal{A}). \quad \text{(PI)}$$

If PI holds, then we can define the probability

$$P_\lambda^{|\phi\rangle_{S+A}}(O_\mathcal{S} = o_\mathcal{S}) = \sum_{\substack{o_\mathcal{A} \text{ an} \\ \text{outcome} \\ \text{of } O_\mathcal{A}}} P_\lambda^{|\phi\rangle_{S+A}}(O_\mathcal{S} = o_\mathcal{S} \,\&\, O_\mathcal{A} = o_\mathcal{A}) \qquad (4.5.2)$$

that is independent of the measurement procedure $O_\mathcal{A}$ on $\mathcal{A}$.[46]

As for the definition of *empirical adequacy* (EA), this states that

$$\sum_{\lambda \in \Lambda} p_\lambda P_\lambda^{|\phi\rangle_{S+A}}(O_\mathcal{S} = o_\mathcal{S} \,\&\, O_\mathcal{A} = o_\mathcal{A}) = P^{|\phi\rangle_{S+A}}(O_\mathcal{S} = o_\mathcal{S} \,\&\, O_\mathcal{A} = o_\mathcal{A}) \qquad \text{(EA)}$$

where Λ is the set of all hidden variables so that $\sum_{\lambda \in \Lambda} p_\lambda = 1$, and where

$$P^{|\phi\rangle_{S+A}}(O_\mathcal{S} = o_\mathcal{S} \,\&\, O_\mathcal{A} = o_\mathcal{A})$$

---

45. Although Leegwater, in his proof of the Colbeck-Renner theorem assumes that λ is a hidden variable for the system $\mathcal{S}$ only, it looks like the proof would still go through if λ was a hidden variable for the composite system $\mathcal{S} + \mathcal{A}$.

46. Cf. (1.6.3).

is the standard probability calculated using the Born Rule with the eigenstates of the observables $\hat{O}_{\mathcal{S}}$ and $\hat{O}_{\mathcal{A}}$ and the quantum state $|\phi\rangle_{\mathcal{S}+\mathcal{A}}$. EA is essentially the same as equation (1.8.3). It also has some similarities with (4.1.7), though the main difference is the range of the summation – the index of the summands of (4.1.7) does not parametrize hidden variables that satisfy criteria 1 to 4 above.

Now, as I've been alluding to, criteria 1 to 4 together with the conditions of PI and EA are very restrictive. Leegwater proves a version of the Colbeck-Renner theorem[47] which takes the following form: if one defines hidden variables according to criteria 1 to 4, then in any hidden-variables theory for which PI and EA hold, the hidden variables are redundant. In other words,

$$P_{\lambda}^{|\phi\rangle_{\mathcal{S}+\mathcal{A}}}(O_{\mathcal{S}} = o_{\mathcal{S}} \,\&\, O_{\mathcal{A}} = o_{\mathcal{A}}) = P^{|\phi\rangle_{\mathcal{S}+\mathcal{A}}}(O_{\mathcal{S}} = o_{\mathcal{S}} \,\&\, O_{\mathcal{A}} = o_{\mathcal{A}}) \qquad (4.5.3)$$

for any measurement $O_{\mathcal{S}}$ on $\mathcal{S}$ and $O_{\mathcal{A}}$ on $\mathcal{A}$, and for almost all λ.[48]

Thus, the Colbeck-Renner theorem means that we cannot hope to make Kent's theory into a hidden-variables theory that satisfies criterion 1 to 4 as well as PI and EA, for the information in Kent's theory is clearly non-redundant.

But nevertheless, it still seems that we should be able to make some kind of sense of PI and EA in Kent's theory and that we should be able to evaluate Kent's theory on the basis of whether these notions of PI and EA are true in this context. To achieve

---

47. See Leegwater, "An impossibility theorem for parameter independent hidden-variable theories."

48. Here, the expressions "almost all" is being used in the measure theoretic sense.

this aim, one strategy would be to relax or drop one or more of the four criteria for a hidden variable. Since we still want to be able to make sense of PI and EA, we won't want to relax criteria 3 or 4. That leaves the possibility of relaxing or dropping criteria 1 or 2.

Now clearly, we wouldn't be able to drop criterion 1 entirely, for otherwise λ would contain information that would determine the choice of measurement made on $\mathcal{S}$. But it doesn't seem problematic if we relax criterion 1 so that there can be information that can change without this corresponding to a change in the system $\mathcal{S}$ so long as this information doesn't determine the measurement choice that is to be made.

As for criterion 2, there doesn't seem to be any problem with dropping it entirely. On doing this, then instead of thinking of τ_S as an augmentation of standard quantum theory, we can think of τ_S as a rather elaborate way of stipulating the initial quantum states of experiments as well as the quantum states of measurement outcomes. In the next section, we will describe in some more detail how to extract the quantum state[49] of a system from the universal quantum state $|\Psi_0\rangle$ and τ_S, but roughly speaking, if we consider an experimental setup including some measurement apparatus $\mathcal{A}$ and an object to be measured $\mathcal{S}$, then the information in τ_S outside the light cone of the spacetime location of $\mathcal{S} + \mathcal{A}$ before they have interacted will determine the initial quantum states of $\mathcal{S}$ and $\mathcal{A}$ before they interact, and likewise, the information in τ_S outside the light cone of the spacetime location of $\mathcal{S} + \mathcal{A}$ after they have interacted will determine the quantum outcome states of $\mathcal{S}$ and $\mathcal{A}$ after they have interacted.

---

49. I refer to this as a conditioned quantum state.

This will mean that when the information in τ_S that is about $\mathcal{S}$ changes, the quantum state of $\mathcal{S}$ extracted from τ_S and $|\Psi_0\rangle$ will also change, and hence criterion 2 will fail to hold. But the information of τ_S will be non-redundant, for without this information, we would only have the evolution of the universal quantum state $|\Psi_0\rangle$ which would continually branch into many worlds. In the many worlds that would result, the energy density on the hypersurface S would not be in a definite state, but rather would be in a superposition of definite states. But with the information of τ_S one of these many states in this superposition is selected as actual. If we could then appropriately partition the information in $\tau_S(x)$ on the basis of whether it determined the quantum state of $\mathcal{S}$, or the quantum state of the apparatus $\mathcal{A}$, or the quantum state of the rest of the universe, we could then consider whether Kent's theory gave the same predictions as standard quantum theory. If it did, then PI and EA would hold in Kent's theory, since these both hold in standard quantum theory. And since Kent's theory is formulated in the Lorentz invariant setting of Schwinger and Tomonaga, this would mean that Kent's theory is a solution to the measurement problem! In other words, we would have a one-world interpretation of quantum physics which gave the same probabilities for experimental outcomes that standard quantum theory predicts, and under this interpretation, the physical world would possess the necessary symmetries that guarantee whatever frame of reference one was in, the speed of light would be constant.

Although I have already argued in section 4.1 that Kent's theory predicts the same expectation values as standard quantum theory when one averages Kent's beables over all the possible mass-energy density measurements that could be made on S, these expectation values are not the expectation values we'd be interested in if we were scientists performing an experiment such as the Stern-Gerlach experiment. Rather, we would typically base our expectation calculations on the state of the apparatus and the system being measured – we would not base our expectation calculations on the universal quantum state $|\Psi_n\rangle$ on a hypersurface S_n. Also, in considering the question of whether PI or OI are true, we need to assume that there are some existing observers with their apparatuses. However, we wouldn't be able to make such an assumption if all we had to go on was the quantum state $|\Psi_n\rangle$ that had unitarily evolved from some primordial state $|\Psi_0\rangle$. So in order for Kent's theory to say anything about such situations, we must be able to extract the quantum state of a system at a particular time from the universal quantum state $|\Psi_0\rangle$ and the mass-energy density τ_S. I will refer to such a state as a *conditioned quantum state*. In this section I will explain how to extract conditioned quantum states from $|\Psi_0\rangle$ and τ_S, and I will show that given a conditioned quantum state, Kent's theory does indeed give the same predictions as standard quantum theory in the case of an experimental apparatus $\mathcal{A}$ measuring the properties of a particle $\mathcal{S}$. Again, we will consider a toy model similar to Kent's toy model described in section 3.5 where photons are treated as point particles.

So let τ_S be the notional mass-energy density measurement on S. In order to avoid undue complexity, we will assume that there is no simultaneous $\hat{T}_S$-eigenstate degeneracy so that if $|\Gamma\rangle$ and $|\Gamma'\rangle$ are normalized[50] simultaneous $\hat{T}_S$-eigenstates with simultaneous eigenvalues τ_S and τ'_S respectively, then

$$(\forall x \in S)\,(\tau_S(x) = \tau'_S(x)) \implies |\Gamma\rangle = |\Gamma'\rangle \,.^{51}$$

This means that corresponding to τ_S, there will be a unique simultaneous $\hat{T}_S$-eigenstate $|\Gamma\rangle$, and according to the Born Rule, the notional mass-energy density measurement outcome τ_S will have been selected with a probability

$$P(T_S = \tau_S) = |\langle \Gamma | U_{SS_0} | \Psi_0 \rangle|^2 = |\langle \Gamma | \Psi_S \rangle|^2$$

where $|\Psi_0\rangle$ is the state of the initial hypersurface S_0, where U_{SS_0} is the unitary operator defined by equation (3.2.8), and where $|\Psi_S\rangle = U_{SS_0}|\Psi_0\rangle$.

Now suppose this notional mass-energy density measurement outcome τ_S indicates that there is some apparatus $\mathcal{A}$ that exists in the vicinity of a spatial location z_0 at time t_i as depicted in figure 4.7. This means that prior to t_i, the apparatus will have interacted with many photons so that the photon detections on S outside the light cone of $y_i = (t_i, z_0)$ will indicate via Kent's stress-energy beables that there is some

---

50. In this section we will assume that all states are normalized.

51. The only reason for making this assumption is in order to avoid cluttering the notation. But if one insisted, one could add an index to all the states below in order to keep track of any degeneracy present.
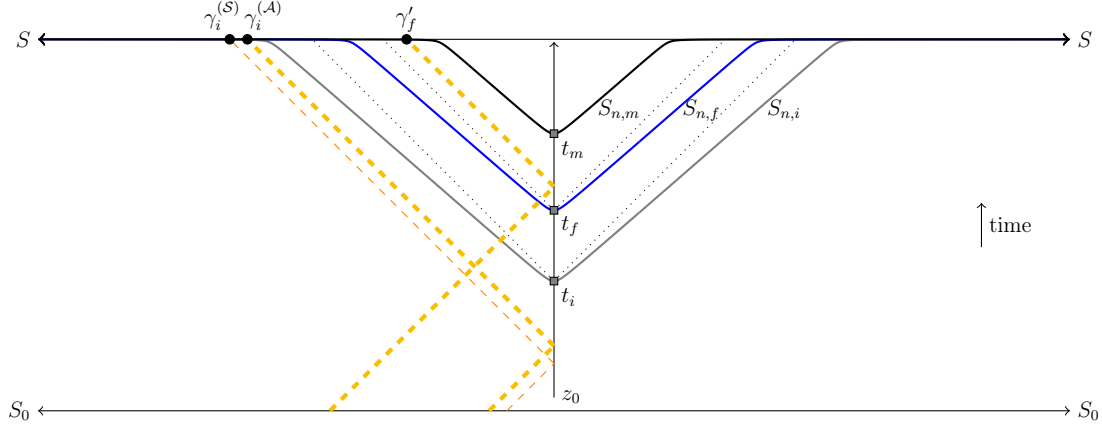
Figure 4.7. Depicts an experiment where the state of some photons $\gamma_i^{(\mathcal{S})}$ and $\gamma_i^{(\mathcal{A})}$ on the hypersurface S determines the initial conditions of an experimental setup of a particle $\mathcal{S}$ and apparatus $\mathcal{A}$ in the vicinity of the spacetime location (t_i, z_0). The state of the photons γ_f' on the hypersurface S determines the final state of the apparatus $\mathcal{A}$ after the particle $\mathcal{S}$ has finished interacting with it from time t_f onwards so that the apparatus at time t_m displays a definite measurement outcome. It is assumed that no incoming photons have become entangled with the experiment after the $\gamma_i^{(\mathcal{S})}$ and $\gamma_i^{(\mathcal{A})}$ photons and before the γ_f' photons have become entangled with the experiment.

apparatus in the vicinity of z_0 that we can identify as $\mathcal{A}$. We further assume that

Kent's stress-energy beables determine that $\mathcal{A}$ is in a state $|a\rangle$ which encapsulates

among other things the measurement parameters of the apparatus. We also suppose

that the information in τ_S indicates that there is a particle $\mathcal{S}$ at time t_i in a state $|s\rangle$

that is heading towards the apparatus so that it will interact with it. This means that

all the possible[52] simultaneous $\hat{T}_S$-eigenstates on S whose simultaneous $\hat{T}_S$-eigenvalues

agree with τ_S for all $x \in S^1(y_i)$[53] are such that their simultaneous $\hat{T}_S$-eigenvalues

---

52. i.e. possible given $|\Psi_0\rangle$ and the Born rule selection criterion. Also, recall footnote 32 on page 141 that we are assuming T_S is a discrete random variable so that states with zero probability are not possible.

53. Recall from page 148 that $S^1(y_i)$ is the subset of S that is outside the light cone of y_i.

within the light cone of y_i indicate that at some time after t_i, the particle $\mathcal{S}$ would have interacted with the apparatus $\mathcal{A}$.

At this point, it will be helpful to define a *simultaneous $\hat{T}_S$-eigenstate of a subregion U of S* to be a state $|\Psi_U\rangle \in H_U$ which is an eigenstate of $\hat{T}_S(u)$ for every $u \in U$, where H_U is the Hilbert space of states describing U. We will assume that we can write any state of H_S as a superposition of states of the form $|\Psi_U\rangle |\Psi_{S\backslash U}\rangle$, where $|\Psi_U\rangle \in H_U$ is a simultaneous $\hat{T}_S$-eigenstate of the subregion U of S, and $|\Psi_{S\backslash U}\rangle \in H_{S\backslash U}$ is a simultaneous $\hat{T}_S$-eigenstate of the subregion $S \backslash U$ of S.[54] We will accordingly let $|\gamma_i^{(S^1(y_i))}\rangle$ denote the simultaneous $\hat{T}_S$-eigenstate of the subregion $S^1(y_i)$ of S with simultaneous $\hat{T}_S$-eigenvalue $\tau_S(x)$ respectively for all $x \in S^1(y_i)$. Then any simultaneous $\hat{T}_S$-eigenstate over the whole of S whose simultaneous $\hat{T}_S$-eigenvalue agrees with τ_S on the subregion $S^1(y_i)$ will be expressible as $|\gamma_i^{(S^1(y_i))}\rangle |\Psi_{S\backslash S^1(y_i)}\rangle$, and any such simultaneous $\hat{T}_S$-eigenstate will determine the apparatus $\mathcal{A}$ and the particle $\mathcal{S}$ to exist in states $|a\rangle$ and $|s\rangle$ respectively given the state $|\Psi_0\rangle$ of the initial hypersurface S_0. In figure 4.7, rather than depicting $|\gamma_i^{(S^1(y_i))}\rangle$, we depict the two components $|\gamma_i^{(\mathcal{S})}\rangle$ and $|\gamma_i^{(\mathcal{A})}\rangle$ of $|\gamma_i^{(S^1(y_i))}\rangle$ which are simultaneous $\hat{T}_S$-eigenstates for subregions $S_{\mathcal{S}}$ and $S_{\mathcal{A}}$ of $S^1(y_i)$ where the states of $\mathcal{S}$ and $\mathcal{A}$ are determined respectively, so that

$$|\gamma_i^{(S^1(y_i))}\rangle = |\gamma_i^{(\mathcal{S})}\rangle |\gamma_i^{(\mathcal{A})}\rangle |\Xi_i^{(S^1(y_i))}\rangle \tag{4.6.1}$$

---

54. We effectively made this assumption in equation (4.3.1).

where $|\Xi_i^{(S^1(y_i))}\rangle$ is a simultaneous $\hat{T}_S$-eigenstate over the remainder of $S^1(y_i)$, i.e. for the subregion $S^1(y_i) \setminus (S_S \cup S_A)$. Note that it is not necessary that the state $|\gamma_i^{(S)}\rangle$ is caused by photons that have interacted directly with S. Rather, it is sufficient that the photons have been reflected from some system $\mathcal{B}$, say, that has interacted with S for which $|s\rangle$ is a pointer state, and that enough photons from $\mathcal{B}$ have intersected S so as to distinguish its different pointer states.

We also assume that there are no further interactions of S with photons that are registered on S until the particle has finished interacting with the apparatus $\mathcal{A}$. In making this assumption, we suppose that there is a time $t_f > t_i$ such that if the notional mass-energy density measurement on S had resulted in the outcome $\tau_{S,j}$ rather than τ_S where $\tau_{S,j}(x) = \tau_S(x)$ for all $x \in S^1(y_i)$, then for all $x \in S^1(y_f, y_i) = S^1(y_f) \setminus S^1(y_i)$,[55] we would be able to say of the mass-energy density $\tau_{S,j}(x)$ at x that it wasn't caused by a photon that had been reflected from S or from something that had become entangled with S.[56]

We will assume that given that the notional mass-energy density measurement restricted to $S^1(y_i)$ is τ_S, there is sufficiently little interaction of the composite system $S + A$ with its environment so that we can assume it evolves unitarily in accordance with the Schrödinger equation as S and A interact with each other. Therefore, given

---

55. i.e. $S^1(y_f, y_i)$ is the subset of S that is within the light cone of $y_i = (t_i, z_0)$ but outside the light cone of $y_f = (t_f, z_0)$.

56. This assumption of our toy model assumes that photons can be treated as point-particles so that they have precise trajectories and that different photons will almost certainly be detected at different locations on S. In physical reality, we can't treat photons as point-particles, and so there will be some ambiguity about where the photons detected on S came from.

that the apparatus $\mathcal{A}$ is in a state $|a\rangle$ that encapsulates its parameter settings, and the particle $\mathcal{S}$ is on course to interact with $\mathcal{A}$, we can express the state $|s\rangle$ of $\mathcal{S}$ as a superposition $|s\rangle = \sum_j c_j |s_j\rangle$ where $\{|s_j\rangle : j\}$ is the set of pointer states corresponding to a particular parameter setting of the apparatus $\mathcal{A}$ that are encapsulated in the state $|a\rangle$. As described on page 82, this means there are future states $|a_j\rangle$ of the apparatus $\mathcal{A}$ such that $\langle a_j | a_{j'}\rangle = 0$ for $j \neq j'$,[57] and such that the composite system $\mathcal{S} + \mathcal{A}$ will evolve according to the Schrödinger equation as

$$|s_j\rangle \, |a\rangle \rightarrow |s_j\rangle \, |a_j\rangle$$

for each pointer state $|s_j\rangle$, and hence

$$|s\rangle \, |a\rangle \rightarrow \sum_j c_j |s_j\rangle \, |a_j\rangle . \tag{4.6.2}$$

We will also assume that given that the notional measurement restricted to $S^1(y_i)$ is τ_S, there is a time interval between t_f and $t_m > t_f$ during which photons will reflect off the apparatus $\mathcal{A}$ and ultimately be detected at spacetime locations on the hypersurface S that correspond to one of the definite measurement states $|a_f\rangle$ of the apparatus indicating that $\mathcal{S}$ is in the state $|s_f\rangle$. But whatever the notional mass-energy density measurement on S is, so long as it results in an outcome $\tau_{S,j}$ with

---

57. Strictly speaking, $\langle a_j | a_{j'}\rangle \approx 0$ for $j \neq j'$, but we will assume $\langle a_j | a_{j'}\rangle = 0$ in this toy model in order to avoid undue complexity.

$\tau_{S,j}(x) = \tau_S(x)$ for all $x \in S^1(y_i)$, then we're assuming that for all $x \in S^1(y_m, y_f),$[58]

we would be able to say of the mass-energy density $\tau_{S,j}(x)$ at x that it was caused

by a photon that had been reflected from the apparatus $\mathcal{A}$ being in some state $|a_j\rangle$

rather than one of the other states.[59] We will let $S'_{\mathcal{A}}$ denote the subregion of $S^1(y_m, y_f)$

where photons coming from the apparatus arrive and hence determine which state

the apparatus is in, and we will let $|\gamma'_j\rangle$ denote the simultaneous $\hat{T}_S$-eigenstate of the

subregion $S'_{\mathcal{A}}$ of S that corresponds to the apparatus $\mathcal{A}$ being in state $|a_j\rangle$. We will

also label the states on the subregion $S \setminus (S^1(y_i) \cup S'_{\mathcal{A}})$ as $|\Xi_j\rangle$ so that we can express

the state $|\Psi_S\rangle = U_{SS_0} |\Psi_0\rangle$ as a superposition

$$|\Psi_S\rangle = b\,|\gamma_i^{(S^1(y_i))}\rangle \sum_j c_j \,|\gamma'_j\rangle\,|\Xi_j\rangle + \sum_{k \neq 0} b_k\,|\gamma_k^{(S^1(y_i))}\rangle\,|\Xi'_k\rangle \tag{4.6.3}$$

where the $|\gamma_k^{(S^1(y_i))}\rangle$ for $k \neq i$ are simultaneous $\hat{T}_S$-eigenstates of the subregion $S^1(y_i)$

whose simultaneous $\hat{T}_S$-eigenvalues restricted to $S^1(y_i)$ are distinct from τ_S, where

$|\Xi'_k\rangle$ are states of the subregion $S \setminus S^1(y_i)$, and where b is a complex number whose

modulus squared gives the probability that the notional measurement on $S^1(y_i)$ will

be τ_S, and likewise, the modulus squared of the b_k give the probabilities of the other

possible notional measurements on $S^1(y_i)$.

Now we aim to show that within Kent's theory, we can calculate the probability

the particle emerges from the measuring apparatus $\mathcal{A}$ in state $|s_f\rangle$ given that it enters

---

58. where $y_m = (t_m, z_0)$.

59. This is again an assumption of our toy model, so it may not be true in more realistic models
where there are many more objects off which photons could reflect.

$\mathcal{A}$ in state $|s\rangle$, and that this probability is the same as if one ignored S and just applied the Born Rule to $|s\rangle$ and $|s_f\rangle$.

In order to show this, let us choose a sequence of hypersurfaces $S_{n,i}$ which go through the spacetime location $y_i = (t_i, z_0)$ such that $\lim_{n \to \infty} S_{n,i} \cap S = S^1(y_i)$.[60] Let us assume that n is sufficiently large so that the photons described by $|\gamma_i^{(S)}\rangle$ and $|\gamma_i^{(\mathcal{A})}\rangle$ belong to $S_{n,i}$. The hypersurface $S_{n,i}$ and the photons being reflected from the vicinity of z_0 just before time t_i are depicted in figure 4.7.

With equation (4.6.3) in mind, we can express the quantum state $|\Psi_{n,i}\rangle = U_{S_{n,i}, S_0} |\Psi_0\rangle$ of the hypersurface $S_{n,i}$ as a superposition

$$|\Psi_{n,i}\rangle = b\,|\gamma_i^{(S)}\rangle\,|\gamma_i^{(\mathcal{A})}\rangle\,|s\rangle\,|a\rangle\,|\xi_{n,i}\rangle + \cdots . \tag{4.6.4}$$

The $|\gamma_i^{(S)}\rangle\,|\gamma_i^{(\mathcal{A})}\rangle$ component of the first summand of (4.6.4) is extracted from the $|\gamma_i^{(S^1(y_i))}\rangle$ component of (4.6.3) which we can do because we're assuming $S_{n,i}$ overlaps with S in the subregion $S^1(y_i)$ corresponding to the states $|\gamma_i^{(S)}\rangle$ and $|\gamma_i^{(\mathcal{A})}\rangle$. Corresponding to the region in the vicinity of y_i on $S_{n,i}$, we have the components $|s\rangle\,|a\rangle$ because we're assuming that the measurement of τ_S on $S^1(y_i)$ guarantees that $\mathcal{S}$ and $\mathcal{A}$ are in the states $|s\rangle$ and $|a\rangle$ respectively. The component $|\xi_{n,i}\rangle$ corresponds to the state of all the other regions of $S_{n,i}$ not determined by $|\gamma_i^{(S)}\rangle$, $|\gamma_i^{(\mathcal{A})}\rangle$, $|s\rangle$ or $|a\rangle$, and the ellipses correspond to the summation over k term of (4.6.3) suitably modified so that the summands are states on $S_{n,i}$ rather than S.

---

60. We understand this limit first by giving $S^1(y_i)$ and $S_{n,i}$ topologies that make them locally homeomorphic to Euclidean space. Then by saying there is this limit of hypersurfaces, we mean that

If we now define the projection $\pi_{n,i}$ corresponding to the measurement outcome $\tau_S(x)$ on $S_{n,i} \cap S$ as in equation (4.1.2), then

$$\pi_{n,i} \left|\Psi_{n,i}\right\rangle \approx b \left|\gamma_i^{(\mathcal{S})}\right\rangle \left|\gamma_i^{(\mathcal{A})}\right\rangle \left|s\right\rangle \left|a\right\rangle \left|\xi_{n,i}\right\rangle, \tag{4.6.5}$$

and the larger n is, the closer (4.6.5) will come to being an equality. The key thing to note about (4.6.5) is that the systems $\mathcal{S}$ and $\mathcal{A}$ are not entangled with each other or with the environment. We can therefore think of the measurement outcome of $\tau_S(x)$ on $S_{n,i} \cap S$ for sufficiently large n as specifying the initial states $\left|s\right\rangle$ and $\left|a\right\rangle$ of $\mathcal{S}$ and $\mathcal{A}$ before they interact. These extracted states $\left|s\right\rangle$ and $\left|a\right\rangle$ are what I mean by *conditioned quantum states*.[61] If n was too small, $S_{n,i} \cap S$ might not contain the subregion for which $\left|\gamma_i^{\mathcal{A}}\right\rangle$ is a state, in which case we could expect the $\mathcal{A}$-component of $\pi_{n,i} \left|\Psi_{n,i}\right\rangle$ to be entangled with the environment with different environment states being correlated with different parameter settings of $\mathcal{A}$. Likewise, we could expect the $\mathcal{S}$-component of $\pi_{n,i} \left|\Psi_{n,i}\right\rangle$ to be entangled with the environment with different environment states being correlated with different states of $\mathcal{S}$ if n was too small. In some situations, the rate at which $\mathcal{S}$ interacts with its immediate environment may be greater than the rate at which photons from the immediate environment of $\mathcal{S}$ are registered on S, in which case, it would not be possible to disentangle $\mathcal{S}$ from its

---

every point $u \in S^1(y_i)$ has a neighborhood $U \subset S^1(y_i)$ for which there is an integer N such that $U \subset S_{n,i}$ for all $n \geq N$.

61. A conditioned quantum state will just be some disentangled component of $\pi_{n,i} \left|\Psi_{n,i}\right\rangle$ for a projection $\pi_{n,i}$ corresponding to a mass-energy density measurement outcome $\tau_S(x)$ on $S_{n,i} \cap S$.

environment in $\pi_{n,i} |\Psi_{n,i}\rangle$. But in controlled experimental settings in which the system $\mathcal{S}$ is prepared to be in a definite state, it should be possible to disentangle $\mathcal{S}$ from its environment in $\pi_{n,i} |\Psi_{n,i}\rangle$. So in this situation, to extract the conditioned quantum state of a system $\mathcal{S}$ in the vicinity of z_0 at time t_i from $|\Psi_0\rangle$ and τ_S, we need to take a hypersurface $S_{n,i}$ for sufficiently large n that goes through $y_i = (t_i, z_0)$. The state of $\mathcal{S}$ in the vicinity of y_i will then be the disentangled normalized $\mathcal{S}$-component of $\pi_{n,i} |\Psi_{n,i}\rangle = \pi_{n,i} U_{S_{n,i},S_0} |\Psi_0\rangle$. So in the case of (4.6.4), the disentangled normalized $\mathcal{S}$-component of $\pi_{n,i} |\Psi_{n,i}\rangle$ will be $|s\rangle$.

Now given that the system $\mathcal{S}$ is in state $|s\rangle$ and that $|s_f\rangle$ was one of the possible measurement outcome states for $\mathcal{S}$, according to standard quantum theory, the Born rule would predict that the measurement outcome state $|s_f\rangle$ occurs with a probability of $|\langle s|s_f\rangle|^2$. We will now show that we can obtain this probability in Kent's theory as well. To do this, we recall that in standard quantum theory, if we define the operator $[\psi] = |\psi\rangle\langle\psi|$ for some state $|\psi\rangle$ of a system, then when the system is in some initial state $|\chi\rangle$, the Born Rule implies that $\langle\chi|[\psi]|\chi\rangle = P(\psi|\chi)$, where $P(\psi|\chi)$ is the probability that the system will be found to be in state $|\psi\rangle$ given that it was initially in state $|\chi\rangle$. But by (2.2.2), $\langle\chi|[\psi]|\chi\rangle$ is just the expectation $\langle[\psi]\rangle_\chi$ of $[\psi]$ when $[\psi]$ is treated as an observable.

Now in equation (4.1.6), we saw how to calculate the expectation value $\langle T^{\mu\nu}(y)\rangle_{\tau_S}$ of the observable $\hat{T}^{\mu\nu}(y)$ given the notional measurement τ_S on S outside the light cone of y. This suggests that the expectation value of any observable $\hat{O}$ defined at spacetime location $y_i = (t_i, z_0)$ given the notional measurement τ_S on S outside the

light cone of y_i is going to be

$$\langle \hat{O} \rangle_{\tau_S} = \lim_{n \to \infty} \frac{\langle \Psi_{n,i} | \pi_{n,i} \hat{O} | \Psi_{n,i} \rangle}{\langle \Psi_{n,i} | \pi_{n,i} | \Psi_{n,i} \rangle}. \tag{4.6.6}$$

By (4.6.5), $\lim_{n \to \infty} \langle \Psi_{n,i} | \pi_{n,i} | \Psi_{n,i} \rangle = |b|^2$, and so taking $\hat{O}$ to be $[s_f]$ at spacetime location y_i we have

$$\langle [s_f] \rangle_{\tau_S} = \frac{|b|^2 |\langle s_f | s \rangle|^2}{|b|^2} = |\langle s_f | s \rangle|^2.$$

Thus, Kent's conditional expectation $\langle [s_f] \rangle_{\tau_S}$ at spacetime location y_i gives us the same probability $|\langle s_f | s \rangle|^2$ for a particle transitioning from state $|s\rangle$ to state $|s_f\rangle$ as in standard quantum theory.

We can also use Kent's conditional expectation to show that at time t_m, $|s_f\rangle$ occurs with probability 1 given the notional measurement τ_S on S outside the light cone of $y_m = (t_m, z_0)$ has $|\gamma_f'\rangle$ as a component. To see this, first note that since we are assuming that between times t_i and t_f, $|s\rangle |a\rangle$ evolves according to (4.6.2), we can apply $U_{S_{n,f}, S_{n,i}}$ to (4.6.4) (where $S_{n,f}$ is one of the hypersurfaces that goes through y_f as depicted in figure 4.7) to get

$$|\Psi_{n,f}\rangle = b |\gamma_i^{(\mathcal{S})}\rangle \, |\gamma_i^{(\mathcal{A})}\rangle \sum_j c_j \, |s_j\rangle \, |a_j\rangle \, |\xi_{n,f}\rangle + \cdots . \tag{4.6.7}$$

In (4.6.7), the component $|\xi_{n,f}\rangle$ corresponds to the state of all the other regions of $S_{n,f}$ not determined by $|\gamma_i^{(\mathcal{S})}\rangle$, $|\gamma_i^{(\mathcal{A})}\rangle$, or the state of $\mathcal{S}$ and $\mathcal{A}$ in the vicinity of

$y_f = (t_f, z_0)$, and the ellipses correspond to the ellipses of (4.6.4) to which $U_{S_{n,f},S_{n,i}}$ has been applied.

Since we need to calculate Kent's conditional expectation of $[s_f]$ at spacetime location y_m, we need to apply $U_{S_{n,m},S_{n,f}}$ to (4.6.7), where $S_{n,m}$ is one of the hyper-surfaces that goes through y_m as depicted in figure 4.7. To do this, we continue to assume that no photons interact with $\mathcal{S}$ between times t_f and t_m. However, we do assume that photons will interact with the apparatus, and for large enough n, the $S_{n,m}$ hypersurfaces will contain the subregion $S'_{\mathcal{A}}$ of $S^1(y_m)$ where the photons coming from the apparatus arrive, and so for each j, the $|\gamma'_j\rangle$-state that corresponds to the $|a_j\rangle$-state of the apparatus $\mathcal{A}$ and which forms a component of one of the summands of $|\Psi_S\rangle$ as shown in equation (4.6.3) will also appear in $|\Psi_{n,m}\rangle$. It therefore follows that

$$|\Psi_{n,m}\rangle = b\,|\gamma_i^{(\mathcal{S})}\rangle\,|\gamma_i^{(\mathcal{A})}\rangle \sum_j c_j\,|s_j\rangle\,|a_j\rangle\,|\gamma'_j\rangle\,|\xi_{n,m}\rangle + \cdots, \qquad (4.6.8)$$

where the component $|\xi_{n,m}\rangle$ corresponds to the state of all the regions of $S_{n,m}$ not determined by $|\gamma_i^{(\mathcal{S})}\rangle$, $|\gamma_i^{(\mathcal{A})}\rangle$, $S'_{\mathcal{A}}$, or the state of $\mathcal{S}$ and $\mathcal{A}$ in the vicinity of $y_m = (t_m, z_0)$, and the ellipses correspond to the ellipses of (4.6.8) to which $U_{S_{n,m},S_{n,f}}$ has been applied. Since we are assuming that the notional measurement restricted to $S^1(y_m)$ will correspond to the apparatus being in the definite measurement state $|a_f\rangle$, then defining the projection corresponding to the measurement outcome τ_S on $S_{n,m} \cap S$ as

in equation (4.1.2), we will have

$$\pi_{n,m} |\Psi_{n,m}\rangle \approx bc_f |\gamma_i^{(\mathcal{S})}\rangle \, |\gamma_i^{(\mathcal{A})}\rangle \, |s_f\rangle \, |a_f\rangle \, |\gamma_f'\rangle \, |\xi_{n,m}\rangle \tag{4.6.9}$$

and the larger n is, the closer (4.6.9) will come to being an equality.

We can now use (4.1.7) to calculate Kent's conditional expectation of $\langle [s_f] \rangle_{\tau_S}$ at spacetime location y_m. By (4.6.9), we have $\lim_{n\to\infty} \langle \Psi_{n,m} | \pi_{n,m} | \Psi_{n,m} \rangle = |b|^2 |c_f|^2$, and $\lim_{n\to\infty} \langle \Psi_{n,m} | [s_f] \pi_{n,m} | \Psi_{n,m} \rangle = |b|^2 |c_f|^2$, and so by (4.6.6), at spacetime location y_m we have

$$\langle [s_f] \rangle_{\tau_S} = 1,$$

and so according to Kent's theory, $\mathcal{S}$ will be in state $|s_f\rangle$ by time t_m.

Also note that we can typically expect the $|\gamma_i^{(\mathcal{S})}\rangle$-state to be independent of the $|\gamma_i^{(\mathcal{A})}\rangle$-state. Therefore, since $|\gamma_i^{(\mathcal{A})}\rangle$ will determine the measurement choice, and since $|\gamma_i^{(\mathcal{S})}\rangle$ determines the initial state of the particle, we can expect the state of the particle to be independent of the measurement choice in Kent's theory. Thus, we can fulfil one of the necessary criteria (i.e. criterion 3) for PI to be a well-defined notion.

We can also choose a set of λ so that criterion 4 holds. To do this, we consider equation (4.6.1) which presupposes there is a subregion $S_{\mathcal{S}}$ of $S^1(y_i)$ which determines the state $\mathcal{S}$ and another non-overlapping subregion $S_{\mathcal{A}}$ that determines the state $\mathcal{A}$. In general, we wouldn't be able to make this association between subregions of $S^1(y_i)$ and

states of $\mathcal{S}$ and $\mathcal{A}$ – after all, $\mathcal{A}$ might not even exist. But we should be able to make this association if an appropriate choice for the state of the remainder of $S^1(y_i)$ is made. In (4.6.1), $|\Xi_i^{(S^1(y_i))}\rangle$ is able to serve this role. We can then suppose that $|\gamma_i^{(\mathcal{S})}\rangle$ is from a basis $\Lambda_{\mathcal{S}} = \{\,|\gamma_{i,1}^{(\mathcal{S})}\rangle\,,\,|\gamma_{i,2}^{(\mathcal{S})}\rangle\,,\ldots\}$ of states that describe all the states of the subregion $S_{\mathcal{S}}$ of $S^1(y_i)$ corresponding to $\mathcal{S}$ that together with $|\Xi_i^{(S^1(y_i))}\rangle$ and $|\Psi_S\rangle$ determine $\mathcal{S}$ to be in the state $|s\rangle$. We could then take the λ of the system $\mathcal{S}$ in criterion 4 to be one of the basis states in $\Lambda_{\mathcal{S}}$. Given that the interpretation of the states in $\Lambda_{\mathcal{S}}$ presupposes $|\Xi_i^{(S^1(y_i))}\rangle$, we would take the probability p_λ for $\lambda = |\gamma_i^{(\mathcal{S})}\rangle$ to be the probability that the notional measurement on S given $|\Xi_i^{(S^1(y_i))}\rangle$ and $|\Psi_S\rangle$ agreed with the mass-energy density values specified by $|\gamma_i^{(\mathcal{S})}\rangle$ on the subregion $S_{\mathcal{S}}$ of $S^1(y_1)$ that $|\gamma_i^{(\mathcal{S})}\rangle$ describes. If we let $\{\,|Z_1\rangle\,,\,|Z_2\rangle\,,\ldots\}$ be a basis of simultaneous $\hat{T}_S$-eigenstates for the subregion $(S \setminus S^1(y_1)) \cup S_{\mathcal{A}}$ so that states of the form $|\gamma_{i,l}^{(\mathcal{S})}\rangle\,|\Xi_i^{(S^1(y_i))}\rangle\,|Z_k\rangle$ will be simultaneous $\hat{T}_S$-eigenstates for the whole of S, then the formula for the probability p_λ will be

$$p_\lambda = \frac{\sum_k |\,\langle\Psi_S|\gamma_i^{(\mathcal{S})}\rangle\,|\Xi_i^{(S^1(y_i))}\rangle\,|Z_k\rangle\,|^2}{\sum_{k,l} |\,\langle\Psi_S|\gamma_{i,l}^{(\mathcal{S})}\rangle\,|\Xi_i^{(S^1(y_i))}\rangle\,|Z_k\rangle\,|^2}. \tag{4.6.10}$$

We can then show that a version of EA analogous to (EA) on page 229 holds. To express EA in this context, we let $O_{\mathcal{S}}$ be the observable that returns j if the system $\mathcal{S}$ is measured to be in the state $|s_j\rangle$, and for $\lambda = |\gamma_i^{(\mathcal{S})}\rangle$, we let $P_\lambda^{|s\rangle}(O_{\mathcal{S}} = j)$ denote the probability that $O_{\mathcal{S}} = j$ given that the hypersurface is in state $|\Psi_S\rangle$ before the notional mass-energy density measurement on S, and that this measurement result is only determined up to the state $|\gamma_i^{(\mathcal{S})}\rangle\,|\Xi_i^{(S^1(y_i))}\rangle$ on the subregion $S^1(y_1) \setminus S_{\mathcal{A}}$

which nevertheless ensures the system $\mathcal{S}$ is in the state $|s\rangle$ before it interacts with the apparatus $\mathcal{A}$. Then to calculate $P_\lambda^{|s\rangle}(O_\mathcal{S} = j)$, we will need to sum over the probabilities that the notional energy density measurement has a $|\gamma_j'\rangle$ component[62] in its simultaneous $\hat{T}_S$-eigenstate for each state in the basis of states $\{\,|Z_1'\rangle\,,\,|Z_2'\rangle\,,\ldots\}$ corresponding to the subregion $S \setminus (S^1(y_i) \cup S_\mathcal{A}')$. Accordingly, if we slightly redefine our notation by absorbing $|\gamma_i^{(\mathcal{A})}\rangle$ into the $|\gamma_l'\rangle$, we will find that

$$
P_\lambda^{|s\rangle}(O_\mathcal{S} = j) = \frac{\sum_k |\,\langle\Psi_S|\gamma_i^{(\mathcal{S})}\rangle\,|\Xi_i^{(S^1(y_i))}\rangle\,|\gamma_j'\rangle\,|Z_k'\rangle\,|^2}{\sum_{k,l} |\,\langle\Psi_S|\gamma_i^{(\mathcal{S})}\rangle\,|\Xi_i^{(S^1(y_i))}\rangle\,|\gamma_l'\rangle\,|Z_k'\rangle\,|^2}.
$$

Also note that with this slight redefinition of $|\gamma_l'\rangle$, equation (4.6.3) becomes

$$
|\Psi_S\rangle = b\,|\gamma_i^{(\mathcal{S})}\rangle\,|\Xi_i^{(S^1(y_i))}\rangle \sum_l c_l\,|\gamma_l'\rangle\,|\Xi_l\rangle + \cdots.
$$

We will therefore be able to express $|\Xi_l\rangle$ in terms of the basis $\{\,|Z_1'\rangle\,,\,|Z_2'\rangle\,,\ldots\}$ so that

$$
|\Xi_l\rangle = \sum_k d_{lk}\,|Z_k'\rangle
$$

with $\sum_k |d_{lk}|^2 = 1$ since we are assuming all states are normalized. Therefore,

$$
P_\lambda^{|s\rangle}(O_\mathcal{S} = j) = \frac{|b|^2|c_j|^2 \sum_k |d_{jk}|^2}{|b|^2 \sum_{k,l} |c_l|^2|d_{lk}|^2} = |c_j|^2.
$$

---

62. To avoid undue complexity, we assume there is only one $|\gamma_j'\rangle$-state for each $|a_j\rangle$-state of the apparatus $\mathcal{A}$.

Therefore,

$$P_\lambda^{|s\rangle}(O_\mathcal{S} = j) = |\langle s|s_j\rangle|^2 = P^{|s\rangle}(O_\mathcal{S} = j).$$

Moreover, since $\sum_\lambda p_\lambda = 1$ we have

$$\sum_\lambda p_\lambda P_\lambda^{|s\rangle}(O_\mathcal{S} = j) = P^{|s\rangle}(O_\mathcal{S} = j)$$

which is analogous to the EA formula on page 229.

4.7 Kent's Theory and Parameter Independence

In the previous section, we saw how we can generalize Kent's beable $\langle \hat{T}^{\mu\nu}(y)\rangle_{\tau_S}$ to

calculate conditional expectations $\langle \hat{O}\rangle_{\tau_S}$ for any observable $\hat{O}$ defined at a particular

spacetime location (t_i, z_0), and that in the case of the observable $[s_f] = |s_f\rangle\langle s_f|$, this

expectation yields the same probability as standard quantum theory for the outcome

$|s_f\rangle$ given the initial state $|s\rangle$ of the system. We also saw that EA holds in Kent's

theory, and to see this, it was necessary to calculate the probability $P_\lambda^{|s\rangle}(O_\mathcal{S} = j)$

appropriately conditioned on the mass-energy density measurement determined on a

subregion of S.

Now the expectation $\langle [s_f]\rangle_{\tau_S}$ and the probability $P_\lambda^{|s\rangle}(O_\mathcal{S} = j)$ depend on just one

observable for just one spacetime location. However, in order to consider whether PI

holds, we need to consider two observables corresponding to two different spacetime

locations. In order to do this, we need to make a further adaption to Kent's theory. In

this section, we will describe this adaption and show that with it, Kent's theory allows us to calculate probabilities for Bell-type experiments, and that these probabilities are the same as in standard quantum theory. Since PI holds in standard quantum theory, a consequence of Kent's theory agreeing with standard quantum theory is that PI will also hold in Kent's theory.

So let's consider figure 4.8 which depicts a one-dimensional view of a Bell-type experiment. Although, this is only a one-dimensional depiction of the experimental
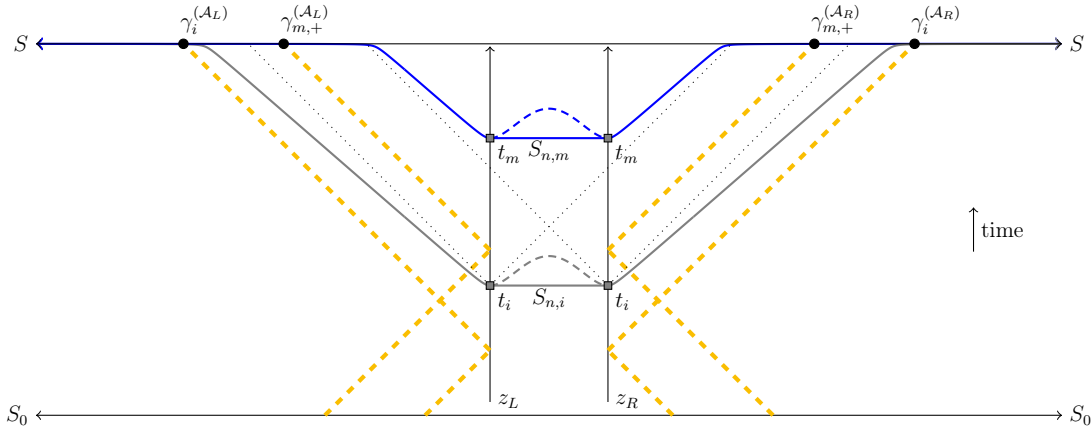
Figure 4.8. Depicts a Bell-type experiment where the state of some photons $\gamma_i^{(\mathcal{A}_L)}$ and $\gamma_i^{(\mathcal{A}_R)}$ on the hypersurface S determines the choice of measurement parameters of the left wing and right wing of the experiment respectively, and some photons $\gamma_{m,+}^{(\mathcal{A}_L)}$ and $\gamma_{m,+}^{(\mathcal{A}_R)}$ on the hypersurface S determine the measurement outcome of the experiment on the left wing and the right wing respectively. The dashed lines on the hypersurfaces $S_{n,m}$ and $S_{n,i}$ indicate other choices for the hypersurfaces, but they still lead to the same probability being calculated.

setup, we will have to consider a toy model in more than one dimension so that it is possible to define angles between the measurement axes of the Stern-Gerlach apparatuses. In the experimental setup, there is a left wing of the experiment located

in the vicinity of z_L, and a right wing of the experiment located in the vicinity of z_R. Shortly before time t_i, photons interact with a Stern-Gerlach apparatus on the left wing and a Stern-Gerlach apparatus on the right wing, and some of these photons eventually intersect the hypersurface S so that there are subregions $S_{\mathcal{A}_L}$ and $S_{\mathcal{A}_R}$ of S such that if the notional measurement of the mass-energy density corresponds to simultaneous $\hat{T}_S$-eigenstates $|\gamma_i^{(\mathcal{A}_L)}\rangle$ and $|\gamma_i^{(\mathcal{A}_R)}\rangle$ respectively on these subregions, then this will be sufficient to determine the measurement parameters of the apparatuses on the left wing and the right wing of the experiment respectively.

Now in order to consider whether PI holds, we will need to adapt Kent's sequences of hypersurfaces so that they can be used to calculate conditional expectation values for observables that depend on two spacetime locations $y_L = (t_i, z_L)$ and $y_R = (t_i, z_R)$. We therefore require that sequences of hypersurfaces $S_{n,i}$ are chosen so that they all contain the spacetime locations y_L and y_R, and that in the limit, $\lim_{n\to\infty} S_{n,i}$ contains as much of $S^1(y_L)$ and $S^1(y_R)$ as possible, where as usual, $S^1(y)$ denotes the subset of S lying outside the light cone of y. Ultimately, this limit (unlike the limit of Kent's hypersurfaces) will not contain the whole of $S^1(y_L)$ or $S^1(y_R)$, but only serves to guarantee that we use as much of the information in S as possible in calculating the expectation values of observables at y_L and y_R. There will be some degree of freedom in what we choose for the hypersurface between y_L and y_R as depicted by the dashed line in figure 4.8. However, such freedom will have no effect on the probabilities calculated, because under the assumption that the hypersurface is very far into the future, there will be no choice of hypersurface in this region that would give us more

information in S to condition on. Also, we recall that the stress-energy operators of the form $\hat{T}^{\mu\nu}(y)$ in the Tomonaga-Schwinger formulation of relativistic quantum physics are defined so that they are invariant under any perturbation of the hypersurface (so long as the hypersurface continues to contain y), so under the assumption that all physical observables will be ultimately expressible in terms of the stress-energy operators, the arbitrary choice of the hypersurfaces in regions that can't intersect with S will have no effect of the probabilities calculated.

On the hypersurface $S_{n,i}$, we assume that n is sufficiently large that enough of the subregions $S_{\mathcal{A}_L}$ and $S_{\mathcal{A}_R}$ are contained within $S_{n,i}$ so that the simultaneous $\hat{T}_S$-eigenstates $|\gamma_i^{(\mathcal{A}_L)}\rangle$ and $|\gamma_i^{(\mathcal{A}_R)}\rangle$ restricted to $S_{\mathcal{A}_L} \cap S_{n,i}$ and $S_{\mathcal{A}_R} \cap S_{n,i}$ are still able to determine the choice of measurement axes for the left and right wings of the experiment respectively. In order to avoid introducing too much extra notation, we will shrink the subregions $S_{\mathcal{A}_\mathcal{L}}$ and $S_{\mathcal{A}_\mathcal{R}}$ so that they are contained within $S_{n,i}$ for sufficiently large n, but we only shrink them slightly so that the mass-energy density measurement on them is still sufficient to determine the choice of measurement axes for the left and right wings of the experiment.[63]

Let us now assume that the axis of orientation of the right wing Stern-Gerlach apparatus makes an angle θ with the axis of the left wing apparatus. We also assume

---

63. For more realistic models, we wouldn't expect the mass-energy density measurement on $S_{n,i}$ to determine the choice of measurement axes with 100% certainty, since there will be a degree of overlap of the different $|\gamma_i^{(\mathcal{A}_L)}\rangle$ for different measurement choices (and likewise for the different $|\gamma_i^{(\mathcal{A}_R)}\rangle$). But this overlap will get smaller and smaller the more that photons interacting with the apparatus intersect $S_{n,i}$, and so the certainty of which measurement is being made will approach 100% as long as there is sufficient time from the time the measurement parameters are chosen to time t_i, and as long as n is large enough so that there are enough photon interactions with the apparatus that intersect $S_{n,i}$. Nevertheless, the fact that we never reach 100% certainty should not worry us too

that there are two particles that together form a Bell-state

$$\frac{1}{\sqrt{2}}(\,|\hat{\mathbf{s}}+\rangle_L\,|\hat{\mathbf{s}}-\rangle_R - |\hat{\mathbf{s}}-\rangle_L\,|\hat{\mathbf{s}}+\rangle_R). \qquad\qquad (4.7.1)$$

We saw in footnote 24 on page 20 that a Bell state does not depend on the orientation of $\hat{\mathbf{s}}$, so without loss of generality, we can suppose that the $|\hat{\mathbf{s}}+\rangle_L$ and $|\hat{\mathbf{s}}-\rangle_L$ are pointer states for the apparatus on the left-wing of the experiment. This means there will be a ready state $|a\rangle_L$ as well as two states $|a+\rangle_L$ and $|a-\rangle_L$ of the left wing apparatus such that

$$|\hat{\mathbf{s}}\pm\rangle_L\,|a\rangle_L \to |\hat{\mathbf{s}}\pm\rangle_L\,|a\pm\rangle_L\,.$$

As for the right wing of the experiment, we let $|\hat{\boldsymbol{s}}_{\boldsymbol{\theta}}+\rangle_R$ and $|\hat{\boldsymbol{s}}_{\boldsymbol{\theta}}-\rangle_R$ be pointer states for the apparatus so that there is a ready state $|a\rangle_R$ as well as two states $|a_\theta+\rangle_R$ and $|a_\theta-\rangle_R$ of the right wing apparatus such that

$$|\hat{\boldsymbol{s}}_{\boldsymbol{\theta}}\pm\rangle_R\,|a\rangle_R \to |\hat{\boldsymbol{s}}_{\boldsymbol{\theta}}\pm\rangle_R\,|a_\theta\pm\rangle_R\,.$$

much in the context of Kent's theory, since it just means that Kent's $\langle\hat{T}^{\mu\nu}(y)\rangle_{\tau_S}$-beables will be perturbed by a very small amount in the vicinity of the apparatus caused by the very small amount of overlap between the different $|\gamma_i^{(\mathcal{A}_L)}\rangle$ for different measurement choices.

In a manner similar to equation (4.6.4), we can express $|\Psi_{n,i}\rangle = U_{S_{n,i},S_0} |\Psi_0\rangle$ as a superposition

$$|\Psi_{n,i}\rangle = \frac{b}{\sqrt{2}} \left(|\hat{\mathbf{s}}+\rangle_L \, |\hat{\mathbf{s}}-\rangle_R - |\hat{\mathbf{s}}-\rangle_L \, |\hat{\mathbf{s}}+\rangle_R \right) |a\rangle_L \, |a\rangle_R \, |\gamma_i^{(\mathcal{A}_L)}\rangle \, |\gamma_i^{(\mathcal{A}_R)}\rangle \, |\xi_{n,i}\rangle + \cdots ,$$

$$(4.7.2)$$

where $|\xi_{n,i}\rangle$ corresponds to the state of the subregion of $S_{n,i}$ not determined by $|\hat{\mathbf{s}}\pm\rangle_L \, |\hat{\mathbf{s}}\mp\rangle_R$, $|a\rangle_L$, $|a\rangle_R$, $|\gamma_i^{(\mathcal{A}_L)}\rangle$ or $|\gamma_i^{(\mathcal{A}_R)}\rangle$. As in equations (1.1.2a) and (1.1.2b), we have

$$|\hat{\mathbf{s}}+\rangle_R = \alpha_\theta \, |\hat{\mathbf{s}}_{\boldsymbol{\theta}}+\rangle_R + \beta_\theta \, |\hat{\mathbf{s}}_{\boldsymbol{\theta}}-\rangle_R ,$$

$$|\hat{\mathbf{s}}-\rangle_R = \alpha_\theta \, |\hat{\mathbf{s}}_{\boldsymbol{\theta}}-\rangle_R - \beta_\theta \, |\hat{\mathbf{s}}_{\boldsymbol{\theta}}+\rangle_R ,$$

where $\alpha_\theta = \cos(\theta/2)$, and $\beta_\theta = \sin(\theta/2)$. Substituting this into (4.7.2), we can express the state of the hypersurface $S_{n,i}$ that goes through the two particles at spacetime locations y_L and y_R as

$$|\Psi_{n,i}\rangle = \frac{b}{\sqrt{2}} \big(\alpha_\theta \, |\hat{\mathbf{s}}+\rangle_L \, |\hat{\mathbf{s}}_{\boldsymbol{\theta}}-\rangle_R - \beta_\theta \, |\hat{\mathbf{s}}+\rangle_L \, |\hat{\mathbf{s}}_{\boldsymbol{\theta}}+\rangle_R - \alpha_\theta \, |\hat{\mathbf{s}}-\rangle_L \, |\hat{\mathbf{s}}_{\boldsymbol{\theta}}+\rangle_R$$

$$(4.7.3)$$

$$- \beta_\theta \, |\hat{\mathbf{s}}-\rangle_L \, |\hat{\mathbf{s}}_{\boldsymbol{\theta}}-\rangle_R \big) |a\rangle_L \, |a\rangle_R \, |\gamma_i^{(\mathcal{A}_L)}\rangle \, |\gamma_i^{(\mathcal{A}_R)}\rangle \, |\xi_{n,i}\rangle + \cdots .$$

We now let $\pi_{n,i}$ be the projection as defined in equation 4.1.2 that corresponds to the measurement outcome $\tau_S(x)$ on $S_{n,i} \cap S$. If we apply $\pi_{n,i}$ to $|\Psi_{n,i}\rangle$ as we did in (4.6.5)

then we will get the approximation

$$\pi_{n,i}\,|\Psi_{n,i}\rangle \approx \frac{b}{\sqrt{2}}\big(\alpha_\theta\,|\hat{\mathbf{s}}+\rangle_L\,|\hat{\boldsymbol{s}}_{\boldsymbol\theta}-\rangle_R - \beta_\theta\,|\hat{\mathbf{s}}+\rangle_L\,|\hat{\boldsymbol{s}}_{\boldsymbol\theta}+\rangle_R - \alpha_\theta\,|\hat{\mathbf{s}}-\rangle_L\,|\hat{\boldsymbol{s}}_{\boldsymbol\theta}+\rangle_R$$

$$- \beta_\theta\,|\hat{\mathbf{s}}-\rangle_L\,|\hat{\boldsymbol{s}}_{\boldsymbol\theta}-\rangle_R\big)\,|a\rangle_L\,|a\rangle_R\,|\gamma_i^{(\mathcal{A}_L)}\rangle\,|\gamma_i^{(\mathcal{A}_R)}\rangle\,|\xi_{n,i}\rangle, \tag{4.7.4}$$

and the larger n is, the closer (4.7.4) will come to being an equality, though in the case of our toy model where we treat photons as point particles, we can expect (4.7.4) to become an equality for sufficiently large n.

Now in the previous section, we calculated Kent's conditional expectation value of $[s_f]$ and argued that this would give the probability that the measurement outcome of the system $\mathcal{S}$ would be $|s_f\rangle$ given a particular mass-energy density measurement on $S^1(y_i)$. In the situation at hand in which we wish to know the probability of two measurements, we need to consider conditional expectation values of observables such as $[\hat{\boldsymbol{s}}+]_L[\hat{\boldsymbol{s}}_{\boldsymbol\theta}+]_R$ where the observable $[\hat{\mathbf{s}}+]_L = |\hat{\mathbf{s}}+\rangle_{L\,L}\langle\hat{\mathbf{s}}+|$ depends on spacetime location y_L, and where the observable $[\hat{\boldsymbol{s}}_{\boldsymbol\theta}+]_R = |\hat{\boldsymbol{s}}_{\boldsymbol\theta}+\rangle_{R\,R}\langle\hat{\boldsymbol{s}}_{\boldsymbol\theta}+|$ depends on spacetime location y_R. Since we are choosing the sequence of hypersurfaces $S_{n,i}$ so that both y_L and y_R belong to $S_{n,i}$, and since any two observables for locations that are spacelike separated commute, we can easily see what the eigenvalues for $[\hat{\boldsymbol{s}}+]_L[\hat{\boldsymbol{s}}_{\boldsymbol\theta}+]_R$ must be: they are 1 and 0, where the eigenvalue of 1 corresponds to all the states of $S_{n,i}$ in which the particle about to be measured by the left wing apparatus is in the $|\hat{\boldsymbol{s}}+\rangle_L$-state and in which the particle about to be measured by the right wing apparatus is in the $|\hat{\boldsymbol{s}}_{\boldsymbol\theta}+\rangle_R$-state, and where the eigenvalue of 0 corresponds to all the states of $S_{n,i}$ in which either the particle about to be measured by the left wing apparatus is in

a pointer state of the apparatus that is not the $|\hat{\boldsymbol{s}}+\rangle_L$-state or the particle about

to be measured by the right wing apparatus is a pointer state of the apparatus that

is not the $|\hat{\boldsymbol{s}}_{\boldsymbol{\theta}}+\rangle_R$-state. It therefore follows from the definition of expectation that

given the mass-energy density measurement on the subregion $S_{\mathcal{A}_L}$ is $|\gamma_i^{(\mathcal{A}_L)}\rangle$ and the

mass-energy density measurement on the subregion $S_{\mathcal{A}_R}$ is $|\gamma_i^{(\mathcal{A}_R)}\rangle$, the probability

that the left wing will be measured to be in state $|\hat{\boldsymbol{s}}+\rangle_L$ and that the right wing will

be measured to be in $|\hat{\boldsymbol{s}}_{\boldsymbol{\theta}}+\rangle_R$ will be

$$
\begin{aligned}
\langle[\hat{\boldsymbol{s}}+]_L[\hat{\boldsymbol{s}}_{\boldsymbol{\theta}}+]_R\rangle_{\tau_S} &= \lim_{n\to\infty} \frac{\langle\Psi_{n,i}|\pi_{n,i}[\hat{\boldsymbol{s}}+]_L[\hat{\boldsymbol{s}}_{\boldsymbol{\theta}}+]_R|\Psi_{n,i}\rangle}{\langle\Psi_{n,i}|\pi_{n,i}|\Psi_{n,i}\rangle} \\
&= \frac{{}_R\langle\hat{\boldsymbol{s}}_{\boldsymbol{\theta}}+|\; {}_L\langle\hat{s}+|\frac{\overline{b}}{\sqrt{2}}\overline{\beta_{\theta}}[\hat{\boldsymbol{s}}+]_L[\hat{\boldsymbol{s}}_{\boldsymbol{\theta}}+]_R\frac{b}{\sqrt{2}}\beta_{\theta}|\hat{s}+\rangle_L\;|\hat{\boldsymbol{s}}_{\boldsymbol{\theta}}+\rangle_R}{|b|^2} \quad (4.7.5) \\
&= \frac{|\beta_{\theta}|^2}{2} = \frac{1}{2}\sin^2(\theta/2),
\end{aligned}
$$

where we have used (4.7.4), and this is the same as the probability that standard

quantum theory predicts as given in equation (1.4.3). We therefore see that the mass-

energy density measurement on S allows us to determine the state of the particles and

the apparatus before the particles are measured in an EPR-Bohm type experiment,

and furthermore, the adaption I've made to Kent's model so that it can compute

probabilities of the measurement outcomes of this experiment produces the same

probabilities as standard quantum theory. Therefore, since PI hold's in standard

quantum theory, it must also hold Kent's adapted model as well.

In previous sections I've tried to show that Kent's theory is a one-world theory which makes the same predictions as standard quantum theory and is consistent with special relativity (subject to a minor alteration which allows us to attribute PI to Kent's theory). There are however a number of issues that might still make people reluctant to take Kent's theory seriously. For instance, it may strike many people as wildly speculative to suggest that the current state of reality is specified in terms of expectation values of stress-momentum operators conditioned on some notional measurement made in the far distant future. Also, many people, myself included, do not find determinism philosophically attractive, yet on the face of it, Kent's theory does seem to be deterministic. It also looks like Kent's theory relies on there being backwards in time causality, an idea which to many people will seem just as absurd as the many-worlds interpretation. In this final section, I will discuss these concerns and how they might be addressed.

To begin with, let's consider the criticism that Kent's theory is too wildly speculative for anyone to take seriously. In response to this criticism, I think the best way to see Kent's theory is not as a claim of how physical reality must be, but more like a thought experiment that opens up the logical space of how physical reality might be if one accepts special relativity and standard quantum theory. For instance, the Colbeck-Renner theorem might lead one to conclude that there are no interesting hidden variables theories that make the same predictions as quantum theory and satisfy PI. However, as I've tried to show in section 4.5, the information additional to the

quantum state that is contained in a distant future mass-energy density measurement is not constrained by the underlying assumptions of Colbeck, Renner, et al., and hence their conclusion about the redundancy of hidden variables does not apply to Kent's theory.

Another assumption that is often made about hidden variables theories, but which Kent's theory calls into question is the assumption that a hidden variables theory requires there to be new laws of physics beyond standard quantum theory. Maudlin prefers to speak of *additional variables theories* rather than of hidden variables theories: according to Maudlin, an additional variables theory is just a theory in which there is more to physical reality than a single quantum state describing the universe. However, Maudlin goes on to claim that if one adopts an additional variables theory, one must not only specify what the additional variables are, but one must also state the laws governing them.[64] It is as though Maudlin imagines that the additional variables must be dynamical quantities like particle positions that need laws of motion to describe how they evolve overtime. Now Kent's theory is an additional variables theory according to Maudlin's criterion. However, just because Kent's theory is an additional variables theory, it doesn't follow that Kent is proposing some radically new physical theory which specifies new laws of motion: standard quantum theory will do. Kent's theory does need the Born Rule in order to specify the probability with which the mass-energy density measurement τ_S is selected on the distant future hypersurface S, but the Born Rule is just a part of standard quantum theory. No

---

64. See T Maudlin, "Three measurement problems," *Topoi* (Heidelberg) 14, no. 1 (1995): 9.

new physics is required to describe the additional variables (i.e. the values of τ_S over S) of Kent's theory since the mass-energy density over S has just as much right to be described by a quantum state as the initial state of the universe $|\Psi_0\rangle$ has. Also, once the mass-energy density τ_S on the hypersurface S has been selected, τ_S doesn't change, and so it makes no sense to ask by what laws the mass-energy density evolves on this hypersurface. So insofar as Kent's additional variables don't require any new physics to describe them, his theory is far less speculative than many other extensions to standard quantum theory.[65]

Nevertheless, the manner in which one might use the additional variables of Kent's theory is something one could speculate about, and it is not obvious that Kent's proposal is the only way or the most natural way of using these additional variables; that is to say, it is not obvious that expectation values of stress-energy tensors conditioned on a future mass-energy density measurement are the most natural candidates for beables that determine the state of physical reality.

As mentioned on page 142, Bell introduced the term beable due to his dissatisfaction with standard quantum theory which is only a theory of observables – a satisfactory theory shouldn't just tell us how a physical system appears to be, but it should also tell us how the physical system actually is. In other words, a satisfactory physical theory

---

65. Note that Kent assumes that physical reality is described by the conditioned expectation values $\langle \hat{T}^{\mu\nu}(y) \rangle_{\tau_S}$ of the stress-energy tensor for spacetime locations y between S_0 and S. This is in contrast to the Copenhagen interpretation in which physical reality is described by a quantum state which evolves unitarily but then collapses from time to time when measurements are made. So in Kent's interpretation, we don't think of physical reality as being described by the quantum state $U_{S'S_0}|\Psi_0\rangle$ for any hypersurface S' between S_0 and S which then collapses on the hypersurface S to a simultaneous $\hat{T}_S$-eigenstate. Rather, physical reality on S' is described by $\langle \hat{T}^{\mu\nu}(y) \rangle_{\tau_S}$ for all $y \in S'$.

should describe beables which encapsulate the physical system's state of actuality and which can account for why a system appears to be the way it is.

Now in Kent's theory, there are two kinds of beables. Firstly, there are the beables corresponding to the mass-energy density measurement on S. The mass-energy density measurement τ_S tells us how each spacetime location of S actually is. Secondly, there are the beables for all the spacetime locations prior to S and after S_0. According to Kent, for a spacetime location y prior to S and after S_0, the conditional expectation value $\langle \hat{T}^{\mu\nu}(y) \rangle_{\tau_S}$ for all the different combinations of μ and ν tells us how the state of physical reality at spacetime location y actually is.

Now it seems that one can accept the first kind of beable in Kent's theory without accepting the second kind. In some situations, the definite values of $\tau_S(x)$ on S (the first kind of beable) will give rise to $\hat{T}^{\mu\nu}(y)$-eigenstates at y, in which case the conditional expectation $\langle \hat{T}^{\mu\nu}(y) \rangle_{\tau_S}$ (the second kind of beable) will be identical to a definite value for $T^{\mu\nu}(y)$ as normally understood in standard quantum theory. But there will inevitably be situations in which the definite values of $\tau_S(x)$ on S will not give rise to $\hat{T}^{\mu\nu}(y)$-eigenstates at y. In other words, in the notation defined in section 4.2 on page 179, we will inevitably find that the state $\pi_n |\Psi_n\rangle$ at y is in a superposition of eigenstates of $\hat{T}^{\mu\nu}(y)$ for some choice of μ and ν. But in that case, it seems rather dubious to claim that the expectation value $\langle \hat{T}^{\mu\nu}(y) \rangle_{\tau_S}$ is the true state of reality and hence the beable at y. As an analogy, it seems a bit like saying that if the throw of a six-sided dice was described quantum mechanically, then the dice could yield a beable of 3.5 for its outcome since 3.5 is the expectation value for the throw of a six-sided

dice. In Kent's theory, there would of course nearly always be sufficient information in τ_S on the hypersurface S to determine that a real dice yielded an integer outcome between 1 and 6. But it is still the case that the information in τ_S outside the light cone of y will be insufficient to determine all possible measurable quantities at y to have definite values, and so it seems unnecessary to insist that Kent's conditioned expectation values of all the observables corresponding to these measurable quantities must give the state of physical reality at y.

It might be more reasonable just to say that there is no fact of the matter about the definite value of some measurable quantities in cases where the information in τ_S is insufficient to determine them. In the context of Kent's theory, one could still suppose that there were beables corresponding to the mass-energy density $\tau_S(x)$ on S, and that some definite facts about physical reality flowed from the definite facts about S. But there could also be degrees of indefiniteness about physical reality. This would occur whenever the information in $\tau_S(x)$ was insufficient to determine whether the state $\pi_n |\Psi_n\rangle$ was an eigenstate of some observable. If light were to interact with the location y in a different way, then this might be able to settle the question of which eigenstate $\pi_n |\Psi_n\rangle$ was in, but this definite outcome would then result in y being indefinite with respect to other observables. But even though there would inevitably be many physical quantities of a system that lacked definiteness, we could typically expect the state $\pi_n |\Psi_n\rangle$ to be very nearly an eigenstate for these indefinite physical

quantities. For instance, there might be real numbers $t^{\mu\nu}(y)$ for all μ, ν such that

$$\hat{T}^{\mu\nu}(y)\pi_n\,|\Psi_n\rangle \approx t^{\mu\nu}(y)\pi_n\,|\Psi_n\rangle\,. \qquad (4.8.1)$$

For some values of μ and ν, we might even obtain equality in (4.8.1), but it is not going to be possible to obtain equality in (4.8.1) for every value of μ and ν – there must be cases in which $\hat{T}^{\mu\nu}(y)\pi_n\,|\Psi_n\rangle$ and $t^{\mu\nu}(y)\pi_n\,|\Psi_n\rangle$ are not quite parallel, for otherwise all the $\hat{T}^{\mu\nu}(y)$ acting on $\pi_n\,|\Psi_n\rangle$ would commute, and this is not possible.[66]

The suggestion that there are degrees of indefiniteness to measurable quantities is not going to appeal to everyone, especially people who are seeking a perfect mathematical account of physical reality. But in a context in which we use the information of τ_S to calculate the state $\pi_n\,|\Psi_n\rangle$ of a hypersurface S_n, the degrees of indefiniteness are going to be very small, so for instance, we don't need to worry about whether a cat could be in a superposition of dead and alive states. The values of $t^{\mu\nu}(y)$ in the vicinity of a cat for which (4.8.1) is a very good approximation are going to correspond either to the cat being dead or the cat being alive – these values won't correspond to the cat being both dead and alive.

Now if one is unwilling to accept Kent's proposal that the expectation values $\langle\hat{T}^{\mu\nu}(y)\rangle_{\tau_S}$ are beables, there is still the question of whether there are any better alternatives. In answering this question, I will draw some inspiration from Maudlin's views on entanglement where he says "The physical state of a complex whole cannot

---

66. Recall footnote 55 on page 164 on equal time stress-energy commutation relations.

always be reduced to those of its parts, or to those of its parts together with their spatiotemporal relations .... The result of the most intensive scientific investigations in history is a theory that contains an ineliminable holism."[67]

With a view to defending monism, Schaffer also argues that entangled systems constitute irreducible wholes.[68] Schaffer's argument relies on the principle that "the basic entities must be *complete*, in the sense of providing a blueprint for reality. More precisely, a plurality of entities is complete if and only if duplicating all these entities, while preserving their fundamental relations, metaphysically suffices to duplicate the cosmos and its contents."[69] Now in the case of entangled systems, duplicating the particles of the system and their spatiotemporal relations will not be sufficient to duplicate the system and its content. Supplementing the particles with entanglement relations doesn't help either in our efforts to duplicate the system because an entangled system will typically have intrinsic properties such as spin which are independent of the number of particles in a system. For instance, both a complex system like a silver atom and a simple system like an electron have an identical spin of $1/2$. But if the spin of an entangled system is a certain kind of entanglement relation between the system's component parts, then it wouldn't be possible to say that a complex system like a silver atom and a simple system like an electron had the same spin. Schaffer

67. See T Maudlin, "Part and Whole in Quantum Mechanics," 56.

68. See Jonathan Schaffer, "Monism: The Priority of the Whole," *The Philosophical review* 119, no. 1 (2010): 31–76.

69. Schaffer, 39.

thus concludes that an entangled system is a fundamental unit, and that the basic entities in any complete description of the cosmos cannot be components of entangled systems.[70]

Schaffer also argues that in the absence of any principle that promotes disentanglement, there will only be one entangled system for the whole cosmos. This is because according to our current understanding of physics, the universe began with a Big Bang which would have resulted in every component of the universe getting entangled with every other component.[71]

Now we can apply much of what Maudlin and Schaffer say about entanglement to posit alternative beables to the ones Kent suggests without having to embrace Schaffer's monism. The alternative I'm suggesting is to take the beables to be the states of the entangled subsystems of S_n that make up the components of $\pi_n |\Psi_n\rangle$. In other words, I'm suggesting we take the beables to be the conditioned quantum states as defined in section 4.6.

To explain this idea in more detail, it will be convenient to change our notation. Instead of considering a series of hypersurfaces S_n as defined in section 3.4 on page 162, we just consider a single generic hypersurface R which intersects S and in which $R \setminus S$ is a bounded region that is prior to S. We also let $|\Psi_R\rangle = U_{RS_0} |\Psi_0\rangle$, and we let π_R be the projection analogous to the projection π_n for the hypersurface R rather than

70. Schaffer, "Monism: The Priority of the Whole," 54.

71. Schaffer, 52.

the hypersurface S_n as defined in equation (4.1.2). Then if we can express $\pi_R |\Psi_R\rangle$ as a product

$$\pi_R |\Psi_R\rangle = b \prod_i |\Phi_{Ri}\rangle \tag{4.8.2}$$

where the states $|\Phi_{Ri}\rangle$ cannot be further decomposed into products of states, and where the regions R_i that these states describe are disjoint regions of R with $R = \bigcup_i R_i$, then the states $|\Phi_{Ri}\rangle$ will be the beables of the cosmos.

In making the beables relative to a hypersurface, I am making explicit an assumption that is implicit in Schaffer's definition of completeness. The cosmos that can be duplicated by duplicating all the basic entities is duplicated at a particular time, or more generally on a particular hypersurface. It doesn't seem any more problematic to consider all the beables on a particular hypersurface than it does to consider all the beables at a particular time. But if we want to avoid monism and maximize[72] the number of beables in a bounded spacelike region R_i, then R_i must be embedded in a hypersurface R that intersects as much of S as possible.[73] If we were just to consider the state $|\Psi_R\rangle$ of R, then as Schaffer suggests, we have good reason to expect

---

72. I will discuss below why we would want to maximize the number of beables.

73. It may well turn out that there isn't a hypersurface R containing R_i that intersects as much of S as possible. However, if we were to use the approximation scheme described by equation (3.2.5), then following the argument in footnote 56 on page 164, for each cell c_y contained within R_i, we could find a hypersurface S_{c_y} containing c_y that intersected as many cells of S as possible. We would then be able to find a hypersurface R containing R_i and $\bigcap_{c_y \subset R_i} S_{c_y}$. Although there might be several hypersurfaces that contain R_i and $\bigcap_{c_y \subset R_i} S_{c_y}$, none of these hypersurfaces could contain more cells of S than R does while still containing R_i, and so these different choices of hypersurface wouldn't make any difference to the corresponding number of beables in R_i.

there to be only one entangled system over the whole of R. However, if we deem the mass-energy density on $S \cap R$ to be given by τ_S, then this will correspond to applying the projection π_R to $|\Psi_R\rangle$, and as indicated by the transition from equation (4.6.4) to equation (4.6.5), $\pi_R |\Psi_R\rangle$ can be approximated as a product of disentangled states. The possibility that the decomposition (4.8.2) might only be approximate rather than exact could be a slight worry, but I'm assuming that the error would be so small as not to make any discernible difference to what any intelligent beings living in the cosmos might observe.[74]

There is also the question of why we might want to choose a hypersurface which maximizes the number of beables. The reason for wanting to do this is that it allows us to conceive of physical reality in a way that is maximally analyzable. When we can express the state $|\Psi_R\rangle$ as a product of states $|\Phi_{Ri}\rangle$ as in (4.8.2), we can then consider the properties of the spacelike region R_i that $|\Phi_{Ri}\rangle$ describes independently of the properties of spacelike regions $R_{i'}$ that the $|\Phi_{Ri'}\rangle$-states describe for $i' \neq i$. To see why this is so, suppose $\hat{O}_i$ is an observable that corresponds to the statement o_i that

74. If $\pi_R |\Psi_R\rangle$ is only approximately a product of disentangled states, then Schaffer would probably argue that since an approximately disentangled state would still be an entangled state over the whole of R, the argument for monism would still hold. But as someone who finds monism philosophically unacceptable, I would say that if $\pi_R |\Psi_R\rangle$ is to give a sensible description of reality, then (4.8.2) would have to be exact. In order for a physical theory to ensure exactness of (4.8.2), we would need to speculate further about how this exactness would come about. For example one possibility might be to suppose that the universe was placed in a huge box with periodic boundary conditions. In this case, energy levels would be quantized and so it would be more plausible for $\pi_R |\Psi_R\rangle$ to be expressed exactly as a product of disentangled states. Another approach that might also lead to the exactness of (4.8.2) is quantum loop gravity where space is quantized – see Carlo Rovelli and Francesca Vidotto, *Covariant Loop Quantum Gravity : An Elementary Introduction to Quantum Gravity and Spinfoam Theory*, Cambridge books online (Cambridge, 2014), 9–16.

a measurable property of R_i lies within a certain range,[75] and likewise, suppose $\hat{O}_{i'}$ is an observable that corresponds to the statement $o_{i'}$ that a measurable property of $R_{i'}$ lies within a certain range. Then if we write $P(q)$ for the probability q is true, where q is either the statement o_i, $o_{i'}$, or $o_i \& o_{i'}$, and if we assume $i \neq i'$, then from (4.8.2) and (4.6.6) we get

$$
\begin{aligned}
P(o_i \& o_{i'}) &= \frac{\langle \Psi_R | \pi_R \hat{O}_i \hat{O}_{i'} | \Psi_R \rangle}{\langle \Psi_R | \pi_R | \Psi_R \rangle} = \langle \Phi_{Ri'} | \, \langle \Phi_{Ri} | \hat{O}_i \hat{O}_{i'} | \Phi_{Ri} \rangle \, | \Phi_{Ri'} \rangle \\
&= \langle \Phi_{Ri} | \hat{O}_i | \Phi_{Ri} \rangle \, \langle \Phi_{Ri'} | \hat{O}_{i'} | \Phi_{Ri'} \rangle = P(o_i) P(o_{i'}).
\end{aligned}
\tag{4.8.3}
$$

Thus, the properties of R_i and $R_{i'}$ will be statistically independent of one another.

Another issue that might be of concern to some people (though not to others) is that Kent's theory is deterministic. In other words, given an initial state $|\Psi_0\rangle$ of S_0 and a final mass-energy density measurement τ_S on S, the stress energy tensors for every spacetime location between S_0 and S are completely determined according to Kent's theory.[76] There is however a way in which one could reintroduce some indeterminism into Kent's theory if one so desired. For Kent's theory says nothing about how the final mass-energy density measurement τ_S on S is selected apart from

---

75. i.e. $\hat{O}_i$ is a Hermitian operator acting on the Hilbert space of states for R_i such that the eigenvalues of $\hat{O}_i$ are either 1 or 0, where the eigenstates with eigenvalue 1 correspond to states of R_i for which the statement O_i is true, and where the eigenstates with eigenvalue 0 correspond to states of R_i for which the statement O_i is false, that is, the measurable property of R_i lies within some other range.

76. Someone might respond to this claim that Kent's theory is deterministic by saying that determinism is usually understood as determination by the initial conditions, whereas Kent's theory is determination by initial and final conditions. Since the final conditions are not determined by the initial ones, we could then say that Kent's theory is not deterministic in the way that determinism is usually understood. But in response, we could say there is nothing to stop us from treating the τ_S on S as an initial condition. For example, we could think of physical reality between S_0 and S as being determined by both $|\Psi_0\rangle$ and $U_{SS_0}^{-1} |\Gamma_{\tau_S}\rangle$ where $|\Gamma_{\tau_S}\rangle$ is a simultaneous $\hat{T}_S$-eigenstate such

the requirement that it is selected with a probability given by the Born Rule. There could, however, be many possible choices for τ_S consistent with this requirement.

With this thought in mind, we can see how Kent's theory could be viewed in a non-deterministic manner by considering a beable $|\Phi_{Ri}\rangle$ which describes some region R_i belonging to a hypersurface R, where $R \setminus S$ is a finite region that contains R_i. The $|\Phi_{Ri}\rangle$-beable will be determined by $|\Psi_0\rangle$, the hypersurface R, and the mass-energy density measurement on $S \cap R$ with corresponding state denoted by $|\Gamma\rangle$; that is to say, $|\Gamma\rangle$ will be a simultaneous $\hat{T}_S$-eigenstate of the region $S \cap R \subset S$ with eigenvalue $\tau_S(x)$ for all $x \in S \cap R$. But since $|\Gamma\rangle$ only specifies the mass-energy density for the region $S \cap R$ which lies outside the light cone of R_i, there can still be many possible mass-energy density measurements within the light cone of R_i consistent with the Born Rule and the mass-energy density measurement being $\tau_S(x)$ for $x \in S \cap R$. Hence, there can be many possible futures for the $|\Phi_{Ri}\rangle$-beable given $|\Gamma\rangle$. For example, we can imagine a scenario in which some photons belonging to R and described by the state $|\gamma\rangle$ are approaching the region R_i. We further suppose that there is a hypersurface R' lying between the hypersurfaces R and S such that as the universal quantum state $|\Psi_R\rangle = U_{RS_0} |\Psi_0\rangle$ evolves to $|\Psi_{R'}\rangle = U_{R'S_0} |\Psi_0\rangle$, the photons interact with the $|\Phi_{Ri}\rangle$-beable and subsequently intersect $S \cap R'$. We can thus depict this

---

that $\hat{T}_S(x) |\Gamma_{\tau_S}\rangle = \tau_S(x) |\Gamma_{\tau_S}\rangle$ for all $x \in S$. Then we could think of both $|\Psi_0\rangle$ and $U_{SS_0}^{-1} |\Gamma_{\tau_S}\rangle$ as initial conditions, and so physical reality would be deterministic in the way that determinism is usually understood.

interaction as

$$|\Psi_R\rangle = b\,|\Phi_{Ri}\rangle\,|\gamma\rangle\,|\Gamma\rangle\,|\xi\rangle + c\,|\Xi\rangle$$

$$\rightarrow |\Psi_{R'}\rangle \tag{4.8.4}$$

$$= b\sum_j \lambda_j\,|\Phi_{R'i,j}\rangle\,|\gamma'_j\rangle\,|\Gamma\rangle\,|\xi'\rangle + c\,|\Xi'\rangle$$

where $\pi_R\,|\Phi_{Ri}\rangle\,|\gamma\rangle\,|\Gamma\rangle\,|\xi\rangle = |\Phi_{Ri}\rangle\,|\gamma\rangle\,|\Gamma\rangle\,|\xi\rangle$,[77] $|\Xi\rangle$ is a state that describes the whole of R for which $\pi_R\,|\Xi\rangle = 0$,[78] $|\Xi'\rangle = U_{R'R}\,|\Xi\rangle$, $|\xi\rangle$ describes all the regions of R not described by $|\Phi_{Ri}\rangle$, $|\gamma\rangle$, or $|\Gamma\rangle$, and where b, c, and λ_j are normalization factors with $|b|^2 + |c|^2 = 1$ and $\sum_j |\lambda_j|^2 = 1$ so that $|\Psi_R\rangle$ and $|\Psi_{R'}\rangle$ are normalized to 1. For simplicity, we also assume that the beables contained within $|\xi\rangle$ do not become entangled with the $|\Phi_{Ri}\rangle$-beable, so that under the action of $U_{R'R}$, the $|\xi\rangle$-component of $|\Psi_R\rangle$ evolves to the $|\xi'\rangle$-component of $|\Psi_{R'}\rangle$. We only assume that the beables $|\gamma\rangle$ and $|\Phi_{Ri}\rangle$ become entangled as $|\Psi_R\rangle$ evolves to $|\Psi_{R'}\rangle$ so that the $|\Phi_{R'i,j}\rangle$-component of $|\Psi_{R'}\rangle$ is correlated with the $|\gamma'_j\rangle$-component of $|\Psi_{R'}\rangle$ as seen on the final line of (4.8.4). Again for simplicity, we assume that the $|\gamma'_j\rangle$ are mutually orthogonal simultaneous $\hat{T}_S$-eigenstates[79] which describe a region $S_{\gamma'} \subset (S\cap R')\setminus R$ which includes all the possible regions the photons initially in the state $|\gamma\rangle$ can end up in after being

---

77. This equation is due to $|\Gamma\rangle$ being a simultaneous $\hat{T}_S$-eigenstate of the region $S \cap R$ with eigenvalue $\tau_S(x)$ for all $x \in S \cap R$.

78. This will ensure that $\pi_R\,|\Psi_R\rangle = b\,|\Phi_{Ri}\rangle\,|\gamma\rangle\,|\Gamma\rangle\,|\xi\rangle$ so that $|\Phi_{Ri}\rangle$ will be a beable on R.

79. In more realistic models where we don't treat photons as point particles, we would only expect $\langle\gamma'_j|\gamma'_l\rangle \approx 0$ for $j \neq l$, and we wouldn't expect the $|\gamma'_j\rangle$ to be simultaneous $\hat{T}_S$-eigenstates on a region of $S \cap R'$ due to the spreading of light.

reflected from the $|\Phi_{Ri}\rangle$-beable. Now among the $|\gamma'_j\rangle$, there should be one of them, $|\gamma'_{j_0}\rangle$ say, such that the simultaneous $\hat{T}_S$-eigenstate of $|\gamma'_{j_0}\rangle$ is given by $\tau_S(x)$ for all $x \in S_{\gamma'}$. Then if $\pi_{R'}$ is the projection on H_{R',τ_S} as defined in equation (4.1.2), we will have

$$\pi_{R'} |\Psi_{R'}\rangle = b\lambda_{j_0} |\Phi_{R'i,j_0}\rangle \, |\gamma'_{j_0}\rangle \, |\Gamma\rangle \, |\xi'\rangle .$$

In other words, the $|\Phi_{Ri}\rangle$-beable transitions to the $|\Phi_{R'i,j_0}\rangle$-beable on R' when the available information in the mass-energy density τ_S is used. However, given $|\Gamma\rangle$, there could have been a different mass-energy density measurement on $S_{\gamma'}$ with corresponding simultaneous $\hat{T}_S$-eigenstate $|\gamma'_{j_1}\rangle$, say, in which case, we would have had the $|\Phi_{Ri}\rangle$-beable transitioning to the $|\Phi_{R'i,j_1}\rangle$-beable on R' instead. So in the sense that given $|\Gamma\rangle$, there are many possible futures for the $|\Phi_{Ri}\rangle$-beable on R', Kent's theory is not deterministic.

One may however object to viewing Kent's theory in this non-deterministic manner because it now looks as though the distant future event of selecting the mass-energy density measurement τ_S on $S_{\gamma'}$ affects which beable $|\Phi_{Ri}\rangle$ transitions to, and so it could seem that backwards in time causality is at play, a possibility that may strike many people as absurd. Such people may be more comfortable with the idea that the mass-energy density measurement τ_S on the whole of S is selected before the state $|\Psi_0\rangle$ has started to evolve, so that when we speak of a mass-energy density measurement on S, we are not really speaking of some future act of determination; rather we are speaking of some primordial act of determination. This primordial act

of determining τ_S would then deterministically affect how the beables evolve so as to guarantee consistency with the mass-energy density on S having the value τ_S that was primordially selected.

One way to address this objection is to re-imagine what we mean by time. Although in contemporary physics, time is thought of as just one of four parameters in four-dimensional spacetime, this is not the only way people have conceived of time. For instance, according to Aristotle, "time is just this – number of motion in respect of 'before' and 'after'."[80] This Aristotelian definition of time presupposes an Aristotelian definition of motion which Aristotle defines as follows: "The fulfillment of what exists potentially, in so far as it exists potentially, is motion."[81] So according to Aristotle, time presupposes that there are degrees of potentiality and degrees of actuality – before the motion, the state of fulfillment of something is potential, whereas after the motion, the state of fulfillment is actual. This means that for there to be a duration of time, something potential must be actualized. Now when one considers the transition from the state $|\Psi_0\rangle$ to the state $|\Psi_S\rangle$, it is not obvious that this is a case of some potency being actualized, for $|\Psi_0\rangle$ is a state of one thing, the hypersurface S_0, whereas $|\Psi_S\rangle$ is the state of something else, the hypersurface S. It is not as though we must assume the state $|\Psi_0\rangle$ becomes the state $|\Psi_S\rangle$ or that the hypersurface S_0 becomes the

80. R.P. Hardie and Gaye R.K., trans., *Aristotle Physics* (Oxford, UK: Clarendon Press, 1930), Book IV. Chapter 11 (219b 1-5).

81. Hardie and R.K., Book III. Chapter 1 (201a10).

hypersurface S. Without any further assumptions, a particular state or a particular hypersurface is no more potential or actual than any other state or hypersurface.

We could however think of the determination of τ_S on S as coming about via a process of actualization. As an analogy, one could think of the determination of τ_S as being like the weaving of a tapestry. The color at a particular location of the tapestry isn't determined until the colored thread is woven into place. It's also the case that some areas of the tapestry must be completed before other areas, though there will also be some freedom as to which area is completed first. So looking at the completed tapestry, there will typically be some parts in which we can discern the order in which the colors were added, whereas there will be other parts in which we can't discern the order of completion.

In an analogous fashion, we might be able to discern a partial order on S on the basis of whether one region of S necessarily had its mass-energy density determined before another region. We wouldn't expect this order to be complete, so for any two regions, we might not be able to say which came first. But there may nevertheless be a sufficient degree of ordering on S to suggest a (non-universal) notion of time similar to that of Aristotle's notion where time is defined in terms of the numbering of motion in respect of before or after.

To consider how we might work out this temporal ordering on S, we will work with a toy model in which photons of light are treated as point particles, and in which the light detected on S can always be traced back to the object prior to S off which the light was reflected. We suppose R is a hypersurface, and that there is a simultaneous

$\hat{T}_S$-eigenstate $|\Gamma\rangle$ for a region S_Γ of $S \cap R$ with simultaneous $\hat{T}_S$-eigenvalue $\tau_S(x)$ for all $x \in S_\Gamma$ such that the mass-energy density being $\tau_S(x)$ for all $x \in S_\Gamma$ guarantees that a system $\mathcal{S}$ exists in a region $R_\mathcal{S} \subset R$, and we suppose that given the initial state $|\Psi_0\rangle$, the mass-energy density has the value $\tau_S(x)$ for all $x \in S_\Gamma$ with probability $|b|^2$ for some complex number b. Furthermore, given $|\Gamma\rangle$, we suppose that $\mathcal{S}$ can be described by a beable $|\Phi_{R,j_0}\rangle$ from a set of possible[82] beables $\{\,|\Phi_{R,j}\rangle : j \in \mathbb{N}\}$, where each of these beables is a state of the region $R_\mathcal{S}$, and where the probability $\mathcal{S}$ could be in the state $|\Phi_{R,j}\rangle$ given $|\Gamma\rangle$ is $|\lambda_j|^2$ for complex numbers λ_j with $\sum_j |\lambda_j|^2 = 1$. We also assume that each $|\Phi_{R,j}\rangle$-beable is correlated with a photon state $|\gamma_j\rangle$ for a region $S_\gamma \subset R \cap S$, where $\langle \gamma_j | \gamma_{j'} \rangle$ for $j \neq j'$. We can then express the state $|\Psi_R\rangle = U_{RS_0} |\Psi_0\rangle$ as

$$|\Psi_R\rangle = b\,|\Gamma\rangle \sum_j \lambda_j\,|\gamma_j\rangle\,|\Phi_{R,j}\rangle\,|\xi\rangle + c\,|\Xi\rangle \tag{4.8.5}$$

where $\pi_R\,|\Gamma\rangle\,|\gamma_{j_0}\rangle\,|\Phi_{R,j_0}\rangle\,|\xi\rangle = |\Gamma\rangle\,|\gamma_{j_0}\rangle\,|\Phi_{R,j_0}\rangle\,|\xi\rangle$, $\pi_R\,|\Gamma\rangle\,|\gamma_j\rangle\,|\Phi_{R,j}\rangle\,|\xi\rangle = 0$ for $|\gamma_j\rangle \neq |\gamma_{j_0}\rangle$, $|\Xi\rangle$ is a state that describes the whole of R for which $\pi_R\,|\Xi\rangle = 0$, $|\xi\rangle$ describes all the regions of R not described by the $|\Phi_{R,j}\rangle$, the $|\gamma_j\rangle$ or by $|\Gamma\rangle$, and where c is a complex number such that $|b|^2 + |c|^2 = 1$.

We also suppose that there is a hypersurface R' lying between R and S, and that as the state $|\Psi_R\rangle$ evolves to $|\Psi_{R'}\rangle$ under the action of $U_{R'R}$, photons described by a state $|\gamma_R\rangle$ (which is a component of $|\xi\rangle$ so that $|\xi\rangle = |\gamma_R\rangle\,|\eta\rangle$ for some state $|\eta\rangle$)

---

82. that is possible with respect to $|\Psi_0\rangle$ and $|\Gamma\rangle$.

become entangled with the system $\mathcal{S}$ so that

$$|\Psi_{R'}\rangle = b\,|\Gamma\rangle \sum_j \lambda_j\,|\gamma_j\rangle \sum_k \lambda_{jk}\,|\Phi_{R',jk}\rangle\,|\gamma'_{jk}\rangle\,|\eta'\rangle + c\,|\Xi'\rangle \qquad (4.8.6)$$

where $|\gamma'_{jk}\rangle$ describe the photon states for a region $S_{\gamma'} \subset R' \cap S$ that have become entangled with $\mathcal{S}$ and which satisfy $\langle \gamma'_{jk}|\gamma'_{j'k'}\rangle = 0$ for either $j \neq j'$ or $k \neq k'$. The complex numbers λ_{jk} satisfy $\sum_k |\lambda_{jk}|^2 = 1$ for each k, and $|\Xi'\rangle$ is the state that evolves from $|\Xi\rangle$ so that $|\Xi'\rangle = U_{R'R}\,|\Xi\rangle$. Finally, we suppose that the value of $\tau_S(x)$ for $x \in S_{\gamma'}$ corresponds to the simultaneous $\hat{T}_S$-eigenstate $|\gamma'_{j_0 k_0}\rangle$, so that

$$\pi_{R'}\,|\Psi_{R'}\rangle = b\,|\Gamma\rangle\,\lambda_{j_0}\,|\gamma_{j_0}\rangle\,\lambda_{j_0 k_0}\,|\Phi_{R',j_0 k_0}\rangle\,|\gamma'_{j_0 k_0}\rangle\,|\eta'\rangle$$

where $\pi_{R'}$ is the projection corresponding to the measurement $\tau_S(x)$ for $x \in R' \cap S$ analogous to the projection defined in equation (4.1.2).

Now given all these assumptions, we see that the state of S'_γ being $|\gamma'_{j_0 k_0}\rangle$ implies that the state of S_γ is $|\gamma_{j_0}\rangle$. On the other hand, the state of S_γ being $|\gamma_{j_0}\rangle$ does not imply the state of S'_γ is $|\gamma'_{j_0 k_0}\rangle$. There is therefore an asymmetry in the order of determination of τ_S whereby S_γ has to be determined before S'_γ is determined, and this asymmetry corresponds to the evolution of the system $\mathcal{S}$ whereby it transitions from the state $|\Phi_{R,j_0}\rangle$ to the state $|\Phi_{R',j_0 k_0}\rangle$, and not vice versa. This ordering also presupposes the state $|\Gamma\rangle$ has been determined, for without this presupposition, the regions S_γ and S'_γ might describe a completely different system in which the order of determination was reversed. We therefore ought to include $|\Gamma\rangle$ when specifying

the order of determination so that we can say $|\Gamma\rangle\,|\gamma'_{j_0 k_0}\rangle$ comes after $|\Gamma\rangle\,|\gamma_{j_0}\rangle$ but not vice versa. If we let $|\gamma\rangle = |\Gamma\rangle\,|\gamma_{j_0}\rangle$ and $|\gamma'\rangle = |\Gamma\rangle\,|\gamma'_{j_0 k_0}\rangle$, then the following statement holds

$$\big((|\gamma'\rangle \Rightarrow |\gamma\rangle) \,\&\, (|\gamma\rangle \nRightarrow |\gamma'\rangle)\big).$$

This suggests we define a partial order $>$ on the states of regions of S according to the rule

$$|\gamma'\rangle > |\gamma\rangle \iff \big((|\gamma'\rangle \Rightarrow |\gamma\rangle) \,\&\, (|\gamma\rangle \nRightarrow |\gamma'\rangle)\big). \qquad (4.8.7)$$

where $|\gamma\rangle$ and $|\gamma'\rangle$ are any two simultaneous $\hat{T}_S$-eigenstates for two (possibly overlapping) regions S_γ and $S_{\gamma'}$ of S. When $|\gamma'\rangle > |\gamma\rangle$ we can think of this in a temporal sense and speak of $|\gamma'\rangle$ as occurring after $|\gamma\rangle$.

It is also possible to extend this partial order to states of spacelike regions in between S_0 and S. We do this by first deeming a state $|\phi\rangle$ of a spacelike region R_ϕ to be simultaneous with $|\gamma\rangle$ if there exists a hypersurface R with $R_\phi \subset R$ and $S_\gamma \subset R \cap S$ such that

$$\pi_R\,|\Psi_R\rangle = b\,|\phi\rangle\,|\gamma\rangle\,|\xi\rangle$$

for some complex number b and state $|\xi\rangle$ of $R \setminus (R_\phi \cup S_\gamma)$. When $|\phi\rangle$ is simultaneous with $|\gamma\rangle$, we will denote this by writing $|\phi\rangle \sim |\gamma\rangle$. Then for two states $|\phi\rangle$ and $|\phi'\rangle$,

we can define on them a partial order $>$ according to the rule

$$|\phi'\rangle > |\phi\rangle \iff (|\phi'\rangle \sim |\gamma'\rangle) \& (|\phi\rangle \sim |\gamma\rangle) \& (|\gamma'\rangle > |\gamma\rangle). \qquad (4.8.8)$$

Note that according to this ordering, $|\Phi_{R',j_0 k_0}\rangle > |\Phi_{R,j_0}\rangle$, so it respects the intuition that the state $|\Phi_{R',j_0 k_0}\rangle$ occurs after the state $|\Phi_{R,j_0}\rangle$. Also, if we think of time in terms of this ordering rather than the ordering given by the t parameter in four dimensional spacetime, we don't have the problem of backwards in time causality since the determination of $|\Phi_{R',j_0 k_0}\rangle$ and the determination of $|\Gamma\rangle\,|\gamma'_{j_0 k_0}\rangle$ will be simultaneous.

Another advantage of re-imagining what is meant by time is that we don't need to worry about what happens after S or the asymptotic behavior of Kent's theory as later and later hypersurfaces are considered on which to make the notional measurement τ_S. As I've presented Kent's theory so far, it is natural to ask what happens after S. It looks like once $\tau_S(x)$ is fixed on S, there should be a corresponding simultaneous $\hat{T}_S$-eigenstate $|\Gamma_{\tau_S}\rangle$ that we can evolve to an even later hypersurface S' after S. But if there are no more notional mass-energy density measurements to condition on, then $U_{S'S}\,|\Gamma_{\tau_S}\rangle$ will evolve in a many-worlds fashion when S' is sufficiently far in the future. However, Kent would reply to this concern by emphasizing that the measurement τ_S on S is only notional – the measurement on S would have to have been τ_S given how the history of the universe has actually turned out up to a hypersurface S_1 prior to S. Kent explains that as long as we restrict our attention to describing physics between two hypersurfaces S_0 and S_1, we can hope that it won't matter too much

which hypersurface S we choose after S_1 on which to make the notional measurement of T_S – given another S' we should be just as likely to choose a measurement outcome for $T_{S'}$ that would produce approximately the same history between S_0 and S_1.[83] But this answer now leaves us asking what happens after the hypersurface S_1, since in order to make sense of the asymptotic behavior of Kent's theory, it looks as though we are not allowed to ask about what physical reality is like after S_1.

But with the Aristotelian-inspired notion of time I've been discussing, the question of what happens after S doesn't arise, for the whole of time is understood in terms of the actualization of S. With this understanding of time, we can speak of the state $|\Gamma_{\tau_S}\rangle$ being at the end of time insofar as there is no other state $|\gamma\rangle$ on a region of S with $|\gamma\rangle > |\Gamma_{\tau_S}\rangle$ – the state $|\Gamma_{\tau_S}\rangle$ is like the completed tapestry, so to speak. Theologians may find it interesting to speculate on whether the idea that there is a state $|\Gamma_{\tau_S}\rangle$ at the end of time bears any relationship to the theological notion of the *eschaton*,[84] a term that has been used by theologians such as C. H. Dodd to mean the divinely ordained climax of history.[85]

One final point worth considering is why one energy-density measurement on S is selected rather than another. According to Kent's theory, the measurement τ_S is randomly selected according to the Born Rule. However, it is not obvious that

---

83. See Kent, "Lorentzian Quantum Reality: Postulates and Toy Models," 3. For a discussion of this point, also see footnote 14 on page 127.

84. This comes from the Greek word ἔσχατος meaning last or extreme.

85. See C. H Dodd, *The parables of the kingdom*, Rev. ed (London: Nisbet, 1961), 36. Dodd argued for a realized eschatology in which the eschaton is not in the future, but rather is to be found in

this selection has to be random. Instead, we could conjecture that the state $|\Psi_S\rangle$ only specifies a realm of possibilities in accordance with the Born Rule, but that there are complementary principles that determine what actually happens within this realm of possibilities. So for example, when the beable $|\Phi_{R,j_0}\rangle$ transitions to $|\Phi_{R',j_0 k_0}\rangle$ rather than $|\Phi_{R',j_0 k_1}\rangle$, we might suppose it does so not due to pure randomness, but because it is in its nature to do so. It would be an interesting topic for further study to investigate the plausibility of whether there could be such laws of nature that complemented quantum physics rather than contradicted it.

Conclusion

The purpose of this dissertation has been to evaluate the merits of Kent's one-world interpretation of quantum physics. I began this dissertation by discussing the mysterious correlations of spin measurements that are observed in the EPR-Bohm paradox: how is it that two distant observers Alice and Bob always get opposite results when they perform a spin measurement along parallel axes of two particles belonging to a spin singlet? Special relativity should rule out any faster than light communication between the two particles, and the violation of Bell's inequality rules out the possibility that the results of the spin measurements are already encoded in the states of the two particles.

Shimony attempted to resolve this paradox by making a distinction between Outcome Independence (OI) and Parameter Independence (PI). Shimony argued that
_____

the ministry of Jesus and his lasting legacy. However, the notion of eschaton itself does not imply a

if PI held, then the correlations observed in the EPR-Bohm paradox would present no threat to special relativity since it would then be impossible for Alice and Bob to send messages to each other faster than the speed of light by choosing the axis on which they were to perform their spin measurements. On the other hand, if OI was false, the assumption on which Bell's inequality depends would not be valid. Thus, by accepting PI and denying OI, Shimony thought that quantum physics could be reconciled with special relativity without Bell's inequality needing to be satisfied.

However, Shimony's solution is not entirely satisfactory since it relies on the rather vague notion of what an outcome is. This problem of what an outcome is goes back to the Copenhagen interpretation in which a physical system is supposedly described by a quantum state which determines with what probability a measurement outcome will occur. According to the Copenhagen interpretation, on making a measurement, the quantum state collapses to the state of one of the measurement outcomes. However, the Copenhagen interpretation doesn't specify the criteria which would determine when such a collapse occurs.

Another popular interpretation, the Bohmian interpretation doesn't have this problem, since there is no state collapse in this interpretation. Instead, the fact of there being a measurement outcome is grounded in the fact of where all the particles of a measurement apparatus are located which in turn would result in the apparatus displaying a definite reading that was correlated with the state of the system that was being measured. However, the Bohmian interpretation suffers from the problem

of violating PI, and so for this reason, many people find the Bohmian interpretation unattractive.

In chapter 2, I considered the many-worlds interpretation which attempts to overcome the problems with the Copenhagen interpretation and the Bohmian interpretation. Like the Bohmian interpretation, the many-worlds interpretation denies there is any state collapse. But on the other hand, the many-worlds interpretation differs from both the Copenhagen interpretation and the many-worlds interpretation by denying that there are any measurement outcomes by which one could make objective claims about the results of experiments. Now although there are some appealing features of the many-worlds interpretation, I argued that we shouldn't accept this interpretation since otherwise we would be accepting an understanding of reality that undermines the belief that science is a discipline concerned with objective facts.

In chapter 3 I then described Kent's theory of quantum physics. Kent supposes that there is an initial quantum state that specifies the conditions of the universe across a three-dimensional hypersurface S_0 in spacetime, but the information in this quantum state is supplemented with information given by a notional mass-energy density measurement across a distant future spacelike hypersurface S. To determine how history unfolds, Kent proposes that at any spacetime location y between the initial hypersurface S_0 and the final hypersurface S, the stress-energy tensor at y is given by the expectation value predicted by the initial quantum state subject to the condition that the mass-energy density across S is in agreement with the mass-

energy density selected in the notional mass-energy density measurement. Since there is no state collapse in Kent's theory, he doesn't need to address the question that plagues the Copenhagen interpretation of what the criteria are for a quantum state collapse to occur. But the information contained in the notional mass-energy density measurement is also sufficient to make Kent's theory into a one-world interpretation of quantum physics, so it doesn't suffer from the problems present in the many-worlds interpretation.

In the final chapter, I offered an evaluation of Kent's theory. I argued that Kent's theory would not contradict the empirical observations that standard quantum theory would predict. I also argued that the predictions Kent's theory makes satisfy Lorentz invariance, where Lorentz invariance is the fundamental property a theory must satisfy if special relativity is to hold within it.

I also considered Kent's theory in the light of the Colbeck-Renner theorem. The assumptions of the Colbeck-Renner theorem lead to the conclusion that there are no non-trivial extensions of standard quantum theory in which PI holds. However, I argued that the assumptions that the Colbeck-Renner theorem makes about the nature of hidden variables don't all apply to the information contained within the notional mass-energy density measurement on S. This means that Kent's theory can be a non-trivial extension of quantum physics in which PI holds, without contradicting the Colbeck-Renner theorem.

realized eschatology, and so the eschaton could instead be used to refer to the end of time and the end of world where according to Christian tradition, Christ will come again and judge the world.

I also argued that the information in the notional mass-energy density measurement on S allows us to extract from the universal quantum state on S_0 a local conditioned quantum state in the vicinity of a spacetime location y between S_0 and S, and that this local conditioned quantum state behaves in much the same way as quantum states of standard quantum theory. By adapting Kent's theory so that it can make probabilistic predictions about multiple measurements, I argued that it will make predictions that are consistent with the validity of PI.

Finally, I discussed some features of Kent's theory such as its inherent determinism as well as the nature of the beables which specify the underlying physical facts that govern how the world appears. I argued that it is not necessary to endorse the beables that Kent proposes which he takes to be the expectation values of the stress-energy tensor conditioned on the notional mass-energy density measurement outcome on S. Instead, I suggested we could take the beables to be the entangled components of the state of a hypersurface whose mass-energy density on its intersection with S is given by the notional mass-energy density measurement. I also argued that by re-imagining what we meant by time in a more Aristotelian manner, indeterminism can be reintroduced into Kent's theory if one so desires.

Inevitably, Kent's theory and my adaption to it are highly speculative. In this dissertation, I am not making claims about how physical reality must be given one's acceptance of standard quantum theory and special relativity. Rather, in this dissertation, I have been trying to argue that Kent's theory is able to resolve the EPR-Bohm paradox and can address the measurement problem in a way that is much

more appealing than the many-worlds interpretation and the Bohmian interpretation of quantum physics.

BIBLIOGRAPHY

Albert, David, and Barry Loewer. "Interpreting the Many Worlds Interpretation." *Synthese (Dordrecht)* (Dordrecht) 77, no. 2 (1988): 195–213.

Bacciagaluppi, Guido. "The Role of Decoherence in Quantum Mechanics." In *The Stanford Encyclopedia of Philosophy*, Fall 2020, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University, 2020.

Bell, J. S. "On the Einstein Podolsky Rosen paradox." *Physics (New York. 1964)* 1, no. 3 (1964): 195–200.

———. "Subject and Object." In *Speakable and unspeakable in quantum mechanics*, 2nd ed., 40–44. Cambridge: Cambridge University Press, 2004.

Bergström, K. "Equal-Time Commutation Relation for the Energy-Momentum Tensor." *Journal of mathematical physics* 11, no. 8 (1970): 2498–2501.

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

Bohm, David. "A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. I." *Physical review* 85, no. 2 (1952): 166–179.

———. "A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. II." *Physical review* 85, no. 2 (1952): 180–193.

Bohm, David, and B. J Hiley. *The undivided universe : an ontological interpretation of quantum theory.* London: Routledge, 1993.

Born, Max, Hedwig Born, and Irene Born. *The Born-Einstein Letters : Friendship, Politics and Physics in Uncertain Times.* Basingstoke, 2005.

Butterfield, Jeremy. "Peaceful Coexistence: Examining Kent's Relativistic Solution to the Quantum Measurement Problem," 2017. eprint: arXiv:1710.07844.

Colbeck, Roger, and Renato Renner. "No extension of quantum theory can have improved predictive power." *Nature communications* (England) 2, no. 1 (2011): 411–411.

———. "The completeness of quantum theory for predicting measurement outcomes," 2012.

Deutsch, D. "Quantum theory as a universal physical theory." *International journal of theoretical physics* (New York, NY) 24, no. 1 (1985): 1–41.

DeWitt, Bryce S. "Quantum mechanics and reality." *Physics today* 23, no. 9 (1970): 30–35.

Dodd, C. H. *The parables of the kingdom*. Rev. ed. London: Nisbet, 1961.

Dürr, Detlef, Sheldon Goldstein, and Nino Zanghì. *Quantum physics without quantum philosophy*. Berlin ; New York: Springer, 2013.

Einstein, A., B. Podolsky, and N. Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" *Physical review* 47, no. 10 (1935): 777–780.

Einstein, Albert. *Albert Einstein, Philosopher Scientist*. Edited by P. A. Schilp. Evanston, Illinois: Library of Living Philosophers, 1949.

Espagnat, Bernard d'. *Conceptual foundations of quantum mechanics*. 2nd ed., completely rev., enl., reset. Mathematical physics monograph series ; 20. Reading, Mass.; London: W. A. Benjamin, 1976.

Everett, Hugh. ""Relative State" Formulation of Quantum Mechanics." *Reviews of modern physics* 29, no. 3 (1957): 454–462.

Ghirardi, Giancarlo, and Angelo Bassi. "Collapse Theories." In *The Stanford Encyclopedia of Philosophy*, Summer 2020, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University, 2020.

Grove, Stanley F. "Quantum Theory and Aquinas's Doctrine on Matter." Doctoral dissertation, The Catholic University of America, 2008.

Hardie, R.P., and Gaye R.K., trans. *Aristotle Physics*. Oxford, UK: Clarendon Press, 1930.

Hossenfelder, Sabine. "Does Superdeterminism save Quantum Mechanics? Or does it kill free will and destroy science?" Youtube, 2021. https://www.youtube.com/watch?v=ytyjgIyegDI.

Isaacson, Walter. *Einstein : His Life and Universe*. London: Pocket, 2008.

Joos, E., H.D. Zeh, C. Kiefer, D.J.W. Giulini, J. Kupsch, and I.O. Stamatescu. *Decoherence and the Appearance of a Classical World in Quantum Theory*. Springer Berlin Heidelberg, 2013.

Kent, Adrian. "Lorentzian Quantum Reality: Postulates and Toy Models," 2014. https://doi.org/10.1098/rsta.2014.0241. eprint: arXiv:1411.2957.

Laer, P. H. van. *Philosophico-scientific problems*. Duquesne studies. Philosophical series ; 3. Pittsburgh: Duquesne University Press, 1953.

Landsman, K. "On the Colbeck–Renner theorem." *Journal of mathematical physics* (United States) 56, no. 12 (2015): 122103.

Landsman, Klaas. *Foundations of Quantum Theory : From Classical Concepts to Operator Algebras (Volume 188.0)*. Vol. 188. Fundamental Theories of Physics. Cham: Springer Open, 2017.

Leegwater, Gijs. "An impossibility theorem for parameter independent hidden-variable theories." *Studies in History and Philosophy of Modern Physics* 54 (2016): 18–34.

Maudlin, T. "Part and Whole in Quantum Mechanics," 46–60.

———. "Space-time in the quantum world." In *Bohmian mechanics and quantum theory : an appraisal [electronic resource]*, by James T. Cushing, Arthur Fine, and Sheldon Goldstein, 285–307. Boston studies in the philosophy of science ; Volume 184. Dordrecht, 1996.

———. "Three measurement problems." *Topoi* (Heidelberg) 14, no. 1 (1995): 7–15.

Maudlin, Tim. *Quantum non-locality & relativity*. Wiley-Blackwell, 2011.

Napolitano, Jim J., and J. J. Sakurai. *Modern Quantum Mechanics*. Pearson Education, 2013.

"Press Release: The Nobel Prize in Physics 2022." *Nobel Prize Outreach*, October 2022. Accessed January 2, 2023. https://www.nobelprize.org/prizes/physics/2022/summary/.

Pascazio, Saverio. "All you ever wanted to know about the quantum Zeno effect in 70 minutes." *44th Symposium on Mathematical Physics on New Developments in the Theory of Open Quantum Systems*, 2013. https://doi.org/10.1142/S1230161214400071. eprint: arXiv:1311.6645v1[quant-ph].

Price, Huw. "A Neglected Route to Realism about Quantum Mechanics." *Mind* (London, etc) 103, no. 411 (1994): 303–336.

Pruss, Alexander R. "A Traveling Forms Interpretation of Quantum Mechanics." In *Neo-Aristotelian Perspectives on Contemporary Science*, 1st ed., 105–122. Routledge, 2018.

Redhead, Michael. *Incompleteness, nonlocality, and realism : a prolegomenon to the philosophy of quantum mechanics.* Oxford : New York: Clarendon Press ; Oxford University Press, 1987.

Rovelli, Carlo, and Francesca Vidotto. *Covariant Loop Quantum Gravity : An Elementary Introduction to Quantum Gravity and Spinfoam Theory.* Cambridge books online. Cambridge, 2014.

Rudin, Walter. *Functional Analysis.* Second Edition. McGraw-Hill, 1991.

Schaffer, Jonathan. "Monism: The Priority of the Whole." *The Philosophical review* 119, no. 1 (2010): 31–76.

Schlosshauer, Maximilian. *Decoherence and the Quantum-to-Classical Transition.* Berlin: Springer-Verlag, 2007.

Schrödinger, E. "Die gegenwärtige Situation in der Quantenmechanik." *Die Naturwissenschaften* 23, no. 48 (November 1935): 807–812.

Schwinger, Julian. "Quantum Electrodynamics. I. A Covariant Formulation." *Physical review* 74, no. 10 (1948): 1439–1461.

Shimony, A. "Events and processes in the quantum world." In *Search for a Naturalistic World View: Volume II: natural science and metaphysics*, 140–162. 1986. Cambridge: Cambridge University Press, 1993.

Tallarico, James. "Action at a Distance." *The Thomist* (Washington, etc) 25, no. 2 (1962): 252–292.

Tegmark, M. "Importance of quantum decoherence in brain processes." *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics* (United States) 61, no. 4 Pt B (2000): 4194–4206.

Tomonaga, S. "On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields." *Progress of theoretical physics* (Tokyo) 1, no. 2 (1946): 27–42.

Trimmer, John D. "The Present Situation in Quantum Mechanics: A Translation of Schrödinger's "Cat Paradox" Paper." *Proceedings of the American Philosophical Society* 124, no. 5 (1980): 323–338.

Weinberg, Steven. *The quantum theory of fields. Volume 1, Foundations [electronic resource].* Cambridge core. Cambridge, 1995.

Wigner, E. "On Unitary Representations of the Inhomogeneous Lorentz Group." *Annals of mathematics* 40, no. 1 (1939): 149–204.

Wikipedia contributors. *Riemann integral — Wikipedia, The Free Encyclopedia.* [Online; accessed 07-July-2023], 2023. https://en.wikipedia.org/wiki/Riemann_integral.

Zeh, H.D. "There are no quantum jumps, nor are there particles." *Physics letters. A* (Amsterdam) 172, no. 4 (1993): 189–192.

Zurek, Wojciech H. "Decoherence, einselection and the existential interpretation (the rough guide)." *Philosophical transactions of the Royal Society of London. Series A: Mathematical, physical, and engineering sciences* 356, no. 1743 (1998): 1793–1821.

NOTATION

(t, x, y, z)	A spacetime location where t parameterizes time, and x, y, and z parameterize the three dimensions of space, p. 22.
(x^0, x^1, x^2, x^3)	A spacetime location, p. 123.
$(y_k)_{k=1}^{\infty}$	A sequence of tiny cells that defines a mesh over the hypersurface S_n, p. 208.
$>$	A partial order on states of regions of S expressing the notion 'occurs after', see equation (4.8.7), p. 274.
$[\psi]$	The operator $|\psi\rangle\langle\psi|$, p. 242.
$[\hat{\mathbf{s}}+]_L, [\hat{\mathbf{s}}_{\boldsymbol{\theta}}+]_R$	The observables $|\hat{\mathbf{s}}+\rangle_{L\ L}\langle\hat{\mathbf{s}}+|$ and $|\hat{\mathbf{s}}_{\boldsymbol{\theta}}+\rangle_{R\ R}\langle\hat{\mathbf{s}}_{\boldsymbol{\theta}}+|$ respectively, p. 254.
$|z|$	The magnitude of the complex number z, p. 9.
$\approx$	Approximately equal to, see footnote 31, p. 82.
$\boldsymbol{x}$	A spatial location (x^1, x^2, x^3), p. 123.
$\Box$	d'Alembert operator $\Box = \partial_\mu \partial^\mu$, p. 134.
δ_μ^ν	The Kronecker-delta given by $\delta_\mu^\nu = 1$ when $\mu = \nu$ and $\delta_\mu^\nu = 0$ otherwise, p. 194.
$|\psi\rangle\langle\psi|$	The operator which acts on a Hilbert space H by sending an arbitrary state $|\psi'\rangle$ to $\langle\psi|\psi'\rangle\ |\psi\rangle$, p. 73.
$\eta^\mu(x)$	The future-directed unit four-vector at x that is orthogonal to S, p. 195.
$\eta^{\nu\sigma}$	Defined so that $\eta_{\mu\rho}\eta^{\nu\rho} = \delta_\mu^\nu$, p. 194.
$\eta_{\mu\nu}$	$\eta_{00} = 1$, $\eta_{ii} = -1$ for $i = 1, 2, 3$ and $\eta_{\mu\nu} = 0$ for $\mu \neq \nu$, p. 190.
$\langle\hat{O}\rangle_\psi$	The expectation value of the observable $\hat{O}$ given state $|\psi\rangle$, see equation (2.2.1), p. 65.
$\langle\hat{O}\rangle_\rho$	The expectation value for an observable $\hat{O}$ of a system given that the density matrix $\hat{\rho}$ describes all we know about the system, see equation (2.4.7), p. 75.
$\langle T^{\mu\nu}(y)\rangle_{\tau_S}$	Kent's beable, see equation (3.4.3), p. 159.
$\langle T^{\mu\nu}(y)\rangle_{\tau_S}$	Kent's proposed beable, see equation (3.4.7), p. 165, see also equation (4.1.6), p. 181.

$\langle T \rangle_r$	The conditional expectation of T given the statement r, see equation (3.4.2), p. 158.
$\langle O \rangle$	Expectation value of the physical quantity $O(\mathbf{x}_A, \mathbf{x}_B)$ in the Bohmian interpretation p. 44.
$\frac{\delta U[S]}{\delta S(x)}$	Functional derivative, see equation (3.2.3), p. 134.
$\gg$	Much greater than, p. 218.
$\boldsymbol{\nabla}_A S(\mathbf{x}_A, \mathbf{x}_B)$	Gradient of the function $S(\mathbf{x}_A, \mathbf{x}_B)$ with respect to the variable $\mathbf{x}_A$, p. 42.
$\boldsymbol{\nabla}_B S(\mathbf{x}_A, \mathbf{x}_B)$	Gradient of the function $S(\mathbf{x}_A, \mathbf{x}_B)$ with respect to the variable $\mathbf{x}_B$, p. 42.
$\hat{\boldsymbol{O}}$	Time independent observable in the Schrödinger picture, p. 130.
$\hat{\boldsymbol{O}}(\boldsymbol{x})$	Spatial dependent but time independent observable in the Schrödinger picture, p. 130.
$\hat{\boldsymbol{O}}(t)$	A time dependent observable in the Heisenberg picture, p. 129.
$\hat{\boldsymbol{O}}(t, \boldsymbol{x})$	An observable in the Heisenberg picture at a particular time t and spatial location $\boldsymbol{x}$, p. 129.
$\hat{\boldsymbol{T}}_S(x)$	Heisenberg picture observable corresponding to the mass-energy density $T_S(x)$ measurement, p. 135.
$\hat{\rho}$	A generic density matrix, p. 74.
$\hat{\rho}_{\mathcal{S}}$	The reduced density matrix of a density matrix $\hat{\rho} \in M(H_{\mathcal{U}})$, p. 80.
$\hat{\rho}_\psi$	The density matrix equal to $|\psi\rangle\langle\psi|$, p. 73.
$\hat{e}_0, \hat{e}_1, \hat{e}_2, \hat{e}_3$	Spacetime locations corresponding to $(1,0,0,0),(0,1,0,0), (0,0,1,0)$, and $(0,0,0,1)$ respectively, p. 188.
$\hat{O}(x)$	Observable for any $x \in S$ in the Tomonaga-Schwinger picture, see equation (3.2.2), p. 133.
$\hat{O}_{\mathcal{U}}$	Extension of the observable $\hat{O}_{\mathcal{S}}$ to the composite system $\mathcal{U} = \mathcal{S} + \mathcal{E}$, see equation (2.4.2), p. 72.
$\hat{O}_{\mathcal{S}}$	An observable for the system $\mathcal{S}$, p. 70.
$\hat{O}_{\hat{\mathbf{a}}+}$	The observable corresponding to the measurement device $O_{\hat{\mathbf{a}}+}$, p. 61.
$\hat{O}_{i,\mathcal{A}}(t), \hat{O}_{i,\mathcal{B}}(t)$	The observables $|p_{i,A}(t)\rangle\langle p_{i,A}(t)|$ and $|p_{i,B}(t)\rangle\langle p_{i,B}(t)|$ respectively, p. 99.
$\hat{T}'_S(x')$	The Tomonaga-Schwinger observable such that if an observer $\mathcal{O}'$ deems S to be in an eigenstate $|\psi'\rangle \in H'_{S_n}$ of $\hat{T}'_S(x')$ with

	eigenvalue τ' (a real number), then $\mathcal{O}'$ would observe the physical quantity described by $T'_S(x')$ to have the value τ', p. 201.
$\hat{T}^{\mu\nu}(x)$	The observable corresponding to the stress-energy tensor $T^{\mu\nu}(y)$ in the Tomonaga-Schwinger picture, p. 143.
$\hat{T}'^{\mu\nu}(x')$	The Tomonaga-Schwinger observable acting on H'_{S_n} such that if observer $\mathcal{O}'$ deemed S_n to be in an eigenstate $|\psi'\rangle$ of $\hat{T}'^{\mu\nu}(x')$ with eigenvalue τ', then observer $\mathcal{O}'$ would observe the physical quantity described by $T'^{\mu\nu}(x')$ to have the value τ', p. 200.
$\hat{T}_S(x)$	The observable corresponding to $T_S(x)$, p. 128. Also see equation (197), p. 198.
$\hat{U}^*$	The adjoint of an operator $\hat{U}$, p. 74.
$\langle\psi|\chi\rangle$	The bra-ket of two states $|\psi\rangle$ and $|\chi\rangle$, p. 9.
$\langle\alpha'|\alpha\rangle_A$	The inner product of the two states $|\alpha\rangle_A$ and $|\alpha'\rangle_A$, p. 18.
$|+\rangle$	Generic spin up state, p. 6.
$|-\rangle$	Generic spin down state, p. 6.
$|\alpha\rangle_A$, $|\beta\rangle_B$	The states of two particles q_A and q_B respectively, p. 15.
$|\alpha\rangle_A\,|\beta\rangle_B$	The composite state of two particles q_A and q_B where particle q_A is in the $|\alpha\rangle_A$-state and particle q_B is in the $|\beta\rangle_B$-state, p. 17.
$|\hat{\boldsymbol{s}}_{\boldsymbol{\theta}}+\rangle_R$, $|\hat{\boldsymbol{s}}_{\boldsymbol{\theta}}-\rangle_R$	Pointer states for the apparatus on the right-wing of the experiment, p. 252.
$|\gamma'_j\rangle$	The simultaneous $\hat{T}_S$-eigenstate of the subregion S'_A of S that corresponds to the apparatus $\mathcal{A}$ being in state $|a_j\rangle$, p. 239.
$|\gamma_i^{(\mathcal{S})}\rangle$, $|\gamma_i^{(\mathcal{A})}\rangle$	The two components of $|\gamma_i^{(S^1(y_i))}\rangle$ which are simultaneous $\hat{T}_S$-eigenstates for subregions $S_{\mathcal{S}}$ and $S_{\mathcal{A}}$ of $S^1(y_i)$ where the states of $\mathcal{S}$ and $\mathcal{A}$ are determined respectively, p. 236.
$|\gamma_i^{(S^1(y_i))}\rangle$	The simultaneous $\hat{T}_S$-eigenstate of the subregion $S^1(y_i)$ of S with simultaneous $\hat{T}_S$-eigenvalue $\tau_S(x)$ respectively for all $x \in S^1(y_i)$., p. 236.
$|\gamma_i^{(\mathcal{A}_L)}\rangle$, $|\gamma_i^{(\mathcal{A}_R)}\rangle$	simultaneous $\hat{T}_S$-eigenstates on the subregions $S_{\mathcal{A}_L}, S_{\mathcal{A}_R}$ that are sufficient to determine the measurement parameters of the apparatuses on the left wing and the right wing of the experiment respectively, p. 250.
$|\gamma'_j\rangle$	Corresponds to the $|a_j\rangle$-state of the apparatus $\mathcal{A}$ and which forms a component of one of the summands of $|\Psi_S\rangle$ as shown in equation (4.6.3), p. 244.

$\lvert \gamma_k^{(S^1(y_i))} \rangle$	For $k \neq i$, the simultaneous $\hat{T}_S$-eigenstates of the subregion $S^1(y_i)$ whose simultaneous $\hat{T}_S$-eigenvalues restricted to $S^1(y_i)$ are distinct from τ_S, p. 239.
$\lvert \Gamma_{\tau_S} \rangle$	Simultaneous $\hat{T}_S$-eigenstate of the hypersurface S with simultaneous $\hat{T}_S$-eigenvalue $\tau_S(x)$ where x ranges over S, p. 275.
$\lvert \Phi(t) \rangle$	Time dependent state in the Schrödinger picture, p. 130.
$\lvert \Phi_{Ri} \rangle$	Entangled component of $\pi_R \lvert \Psi_R \rangle$ and beable candidate, , see equation (4.8.2)p. 264.
$\lvert \Phi \rangle$	A state in the Heisenberg picture, p. 129.
$\lvert \psi' \rangle , \lvert \chi' \rangle$	States that an observer $\mathcal{O}'$ corresponding to state $\lvert \psi \rangle$ and $\lvert \chi \rangle$ that an observer $\mathcal{O}$ observer, p. 198.
$\lvert \psi(t) \rangle$	The state of a system at time t, p. 121.
$\lvert \Psi[S] \rangle$	Hypersurface dependent state in the Tomonaga-Schwinger picture, p. 132.
$\lvert \psi^{(U)}(t) \rangle$	$\lvert \psi^{(U)}(t) \rangle = U(t, t_0) \lvert \psi(t_0) \rangle$, p. 121.
$\lvert \Psi_0 \rangle$	The state of the initial hypersurface S_0 in the Tomonaga-Schwinger picture, p. 140.
$\lvert \psi_i \rangle$	The eigenstates of an observable $\hat{O}$, p. 64.
$\lvert \Psi_n \rangle$	$\lvert \Psi_n \rangle = U_{S_n S_0} \lvert \Psi_0 \rangle$, p. 179.
$\lvert \Psi_n \rangle$	The state $\lvert \Psi_n \rangle = \lvert \Psi[S_n] \rangle$ on S_n, p. 183.
$\lvert \Psi_R \rangle$	The state given by $U_{RS_0} \lvert \Psi_0 \rangle$ that describes the hypersurface R, p. 263.
$\lvert \Psi_S \rangle$	The state $U_{SS_0} \lvert \Psi_0 \rangle \in H_S$ in the Tomonaga-Schwinger picture, p. 140.
$\lvert \Psi_U \rangle$	A simultaneous $\hat{T}_S$-eigenstate of a subregion $U \subset S$ belong to the Hilbert space of states H_U describing U, p. 236.
$\lvert \Psi_{\text{Bell}} \rangle$	A Bell state, see equation (1.2.6), p. 18.
$\lvert \Psi_{n,i} \rangle$	The state $U_{S_{n,i}, S_0} \lvert \Psi_0 \rangle$ of the hypersurface $S_{n,i}$, p. 240.
$\lvert \psi \rangle_{\mathcal{S}} , \lvert \chi \rangle_{\mathcal{E}} , \lvert \xi \rangle_{\mathcal{U}}$	States that belong to the Hilbert spaces $H_{\mathcal{S}}$, $H_{\mathcal{E}}$, and $H_{\mathcal{U}}$ respectively, p. 70.
$\lvert \hat{\mathbf{a}}-; \hat{\mathbf{b}}+; \ldots \rangle_A$	An example of a state of a particle q_A with hidden variables specifying all the spin outcomes of every possible axis along which the spin of the particle could be measured, p. 27.
$\lvert \hat{\mathbf{a}}- \rangle$	Spin down state in the $\hat{\mathbf{a}}$-direction, p. 7.

$\lvert \hat{\mathbf{a}}+\rangle$	Spin up state in the $\hat{\mathbf{a}}$-direction, p. 7.
$\lvert \hat{\mathbf{s}}+\rangle_L$, $\lvert \hat{\mathbf{s}}-\rangle_L$	Pointer states for the apparatus on the left-wing of the experiment, p. 252.
$\lvert \Xi_i^{(S^1(y_i))}\rangle$	A simultaneous $\hat{T}_S$-eigenstate over the remainder of $S^1(y_i)$, i.e. for the subregion $S^1(y_i) \setminus (S_S \cup S_A)$, see equation (4.6.1), p. 237.
$\lvert \xi_i(t)\rangle$	The state $\lvert s_i\rangle \lvert a_i(t)\rangle \lvert p_{i,A}(t)\rangle \lvert p_{i,B}(t)\rangle \in H_V$ of $V = S + A + P_A + P_B$, p. 98.
$\lvert \Xi_j\rangle$	The states on the subregion $S \setminus (S^1(y_i) \cup S'_A)$ so that equation (4.6.3) holds, p. 239.
$\lvert \xi_{n,f}\rangle$	Corresponds to the state of all the other regions of $S_{n,f}$ not determined by $\lvert \gamma_i^{(S)}\rangle$ in equation (4.6.7), p. 243.
$\lvert \xi_{n,i}\rangle$	Corresponds to the state of all the other regions of $S_{n,i}$ not determined by $\lvert \gamma_i^{(S)}\rangle$, $\lvert \gamma_i^{(A)}\rangle$, $\lvert s\rangle$ or $\lvert a\rangle$ depicted in equation (4.6.4), p. 240.
$\lvert \xi_{n,m}\rangle$	Corresponds to the state of all the regions of $S_{n,m}$ not determined by $\lvert \gamma_i^{(S)}\rangle$, $\lvert \gamma_i^{(A)}\rangle$, S'_A, or the state of S and A in the vicinity of $y_m = (t_m, z_0)$ in equation (4.6.8), p. 244.
$\lvert a+\rangle_L$, $\lvert a-\rangle_L$	States of the left wing apparatus corresponding to the pointer states $\lvert \hat{\mathbf{s}}+\rangle_L$ and $\lvert \hat{\mathbf{s}}-\rangle_L$ respectively, p. 252.
$\lvert a_\theta+\rangle_R$, $\lvert a_\theta-\rangle_R$	States of the right wing apparatus corresponding to the pointer states $\lvert \hat{\mathbf{s}}_\theta+\rangle_R$ and $\lvert \hat{\mathbf{s}}_\theta-\rangle_R$ respectively, p. 252.
$\lvert a_f\rangle$	The final state of the apparatus A after it has interacted with S and photons reflected from it have intersected S, p. 238.
$\lvert a_r(t_0)\rangle$	The ready state of a measuring device, p. 82.
$\lvert a\rangle$	The state of an apparatus A that Kent's stress-energy beables determine which encapsulates among other things the measurement parameters of the apparatus, p. 235.
$\lvert a\rangle_L$	Ready state of the left wing apparatus, p. 252.
$\lvert a\rangle_R$	Ready state of the right wing apparatus, p. 252.
$\lvert E_i(t)\rangle$	State of the environment $\mathcal{E}$ corresponding to the pointer state $\lvert s_i\rangle$, p. 97.
$\lvert E_r(t)\rangle$	Ready states of the environment $\mathcal{E}$, p. 97.
$\lvert p_{i,A}(t)\rangle$, $\lvert p_{i,B}(t)\rangle$	The states of Alice and Bob after they have observed the apparatus to be in the state $\lvert a_i(t)\rangle$, p. 97.

$\lvert p_{r,A}(t)\rangle,\ \lvert p_{r,B}(t)\rangle$	Ready states of Alice and Bob before they observe a measurement outcome, p. 97.
$\lvert s_f\rangle$	Final pointer state of $\mathcal{S}$ corresponding to the apparatus state $\lvert a_f\rangle$, p. 238.
$\lvert s\rangle$	The state of a particle $\mathcal{S}$ that the information in τ_S determines to be heading towards the apparatus $\mathcal{A}$ so that it will interact with it, p. 235.
$\lvert x\rangle$	A non-normalizable position state p. 91.
Λ	A Lorentz transformation, p. 190.
Λ	The set of λ that describe both q_A and q_B, p. 34.
λ	A complete state describing both q_A and q_B that is independent of Alice and Bob's measurement choices, but that encodes all other features that would influence the corresponding measurement outcomes, p. 34.
$\Lambda^{\mu}{}_{\nu}x^{\nu}$	The components of a Lorentz Transformation, p. 190.
$\Lambda_{\mu}{}^{\nu}$	The components of an inverse Lorentz transformation Λ^{-1}, see equation (4.2.6), p. 194.
$\mathbb{C}$	The set of complex numbers, p. 9.
$\mathbb{N}$	the set of positive integers greater than 0, p. 136.
$\mathcal{A}$	A measuring device considered in decoherence theory, p. 82.
$\mathcal{A}_A, \mathcal{A}_B$	Stern-Gerlach apparatuses belonging to Alice and Bob respectively, p. 107.
$\mathcal{E}$	The environment of a system $\mathcal{S}$ considered in decoherence theory, p. 69.
$\mathcal{O}, \mathcal{O}'$	Two observers, p. 188.
$\mathcal{P}_A, \mathcal{P}_B$	The physical systems corresponding to two scientists, Alice and Bob respectively, p. 97.
$\mathcal{S}$	A generic system considered in decoherence theory, p. 69.
$\mathcal{S}_1, \mathcal{S}_2$	Two spatially separated systems, p. 14.
$\mathcal{U}$	The composite of the systems $\mathcal{S}$ and $\mathcal{E}$ considered in decoherence theory, p. 69.
$\mathcal{V}$	The composite subsystem $\mathcal{S} + \mathcal{A} + \mathcal{P}_A + \mathcal{P}_B$ so that $\mathcal{U} = \mathcal{V} + \mathcal{E}$, p. 97.
μ, ν	Generic indices of tensors, $\mu, \nu = 0, 1, 2,$ or 3, p. 143.

$\|\psi\|$	$\sqrt{\langle\psi\|\psi\rangle}$ for some state $\|\psi\rangle$, p. 62.
$\overline{z}$	The complex conjugate of a complex number z, p. 9.
$\pi_\tau^{\mu\nu}$	The projection $\pi_\tau^{\mu\nu} = \sum_j \|\eta_{j,\tau}^{\mu\nu}\rangle\langle\eta_{j,\tau}^{\mu\nu}\|$, p. 180.
π_n	$\pi_n \|\Psi_n\rangle$ is the (weighted) sum of all the states in the superposition of $\|\Psi_n\rangle$ for which the mass-energy density has a definite value that is equal to the notional mass-energy density measurement $\tau_S(x)$ for all $x \in S \cap S_n$, p. 176, see also equation (4.1.2), p. 179.
π_n'	The projection $\pi_n' = \sum_j \|\xi_j'\rangle\langle\xi_j'\|$, p. 203.
$\pi_n(\tau_{S_n\cap S})$	The projection π_n where $\tau_{S_n\cap S}$ is the function τ_S restricted to $S_n \cap S$, p. 182.
π_R	the projection analogous to π_n for the hypersurface R rather than the hypersurface S_n, p. 263.
$\pi_{n,\tau}^{\mu\nu}$	The projection $\pi_{n,\tau}^{\mu\nu} = \sum_j \|\chi_{j,\tau}^{\mu\nu}\rangle\langle\chi_{j,\tau}^{\mu\nu}\|$, p. 180.
$\propto$	Means proportional to, e.g. $\|\psi_1\rangle_\mathcal{S} \propto \|\psi_2\rangle_\mathcal{S}$ means there exists α such that $\|\psi_1\rangle_\mathcal{S} = \alpha\|\psi_2\rangle_\mathcal{S}$, p. 71.
$\psi(\mathbf{x}_A, \mathbf{x}_B, t)$	Pilot wave, p. 42.
$\psi(x)$	The bispinor field of quantum field theory, p. 130.
$\psi(x)$	The wave function corresponding to the state $\|\psi\rangle$ given by the formula $\psi(x) = \langle x\|\psi\rangle$, p. 91.
ψ_0^{sys}	A wave function in Kent's toy model that is the superposition of two localized wave functions ψ_1^{sys} and ψ_2^{sys}, p. 165.
$\psi_1^{\text{sys}}, psi_2^{\text{sys}}$	Wave functions localized at z_1 and z_2 respectively, p. 165.
$\text{Re}(z)$	The real part of a complex number z, p. 93.
$\backslash$	$S \backslash U$ denotes all the elements in S that are not in U, p. 236.
$\sim$	we write $\|\phi\rangle \sim \|\gamma\rangle$ to mean that $\|\phi\rangle$ and $\|\gamma\rangle$ are simultaneous, p. 274.
$\sum_{i=1}^8 N_i$	Shorthand for $N_1 + N_2 + N_3 + N_4 + N_5 + N_6 + N_7 + N_8$, p. 30.
τ	The particular value the stress-energy tensor $T^{\mu\nu}(y)$ takes when a measurement of it is performed at $y \in S_n$, p. 177.
$\tau_S'(x')$	A particular range of values for each $x' \in S$ of the physical quantity described by $T_S'(x')$ observed by $\mathcal{O}'$, p. 201.
$\tau^{\mu\nu}(x)$	For fixed μ, ν, the simultaneous eigenvalue for all $x \in S$ of a simultaneous eigenstate $\hat{T}^{\mu\nu}(x)$, p. 143.

$\tau'^{\mu\nu}(x')$	A particular (real) value of the physical quantity described by $T'^{\mu\nu}(x')$ that $\mathcal{O}'$ observes, p. 200.
$\tau_S(x)$	A real valued function which specifies the outcome of the notional measurement $T_S(x)$ for every $x \in S$, p. 126.
$\tau_{S,i}(x)$	Simultaneous $\hat{T}_S$-eigenvalue corresponding to simultaneous $\hat{T}_S$-eigenvalue $|\Gamma_i\rangle$, p. 136.
$\tau_{S_n \cap S}$	The function τ_S restricted to $S_n \cap S$, p. 181.
$\mathrm{Tr}(\hat{A})$	The trace of an operator $\hat{A}$, see equation (2.4.4), p. 74.
$\mathrm{Tr}_{\mathcal{E}}(\hat{A}_{\mathcal{U}})$	The partial trace of an operator $\hat{A}_{\mathcal{U}}$ acting on $H_{\mathcal{U}}$, see equation (2.4.9), p. 78.
$\mathrm{Tr}_{\mathcal{S}}(\hat{A}_{\mathcal{S}}), \mathrm{Tr}_{\mathcal{U}}(\hat{A}_{\mathcal{U}})$	The traces of the operators $\hat{A}_{\mathcal{S}}$ and $\hat{A}_{\mathcal{U}}$ acting on $H_{\mathcal{S}}$ and on $H_{\mathcal{U}}$ respectively, p. 77.
$\hat{\mathbf{a}}-$	The location a particle in the state $|\hat{\mathbf{a}}-\rangle$ would hit the detection screen, p. 7.
$\hat{\mathbf{a}}+$	The location a particle in the state $|\hat{\mathbf{a}}+\rangle$ would hit the detection screen, p. 7.
$\hat{\mathbf{a}}$	Unit vector in a particular direction, p. 7.
$\delta(x)$	The Dirac delta function, p. 92.
$\mathbf{p}_A, \mathbf{p}_B$	Exact momenta of particles q_A and q_B in the Bohmian interpretation, p. 41.
$\mathbf{v}$	A vector $\mathbf{v} = (v_x, v_y, v_z)$ denoting a velocity, p. 21.
$\mathbf{x}$	A vector $\mathbf{x} = (x, y, z)$ representing a location in three-dimensional space, p. 22.
$\mathbf{x}_A, \mathbf{x}_B$	Exact locations of particles q_A and q_B in the Bohmian interpretation, p. 41.
$\{\,|\chi_{j,\tau}^{\mu\nu}\rangle : j \in \mathbb{N}\}$	The subset of $\{\,|\eta_j^{\mu\nu}\rangle : j \in \mathbb{N}\}$ such that $\hat{T}^{\mu\nu}(y)\,|\chi_{j,\tau}^{\mu\nu}\rangle = \tau\,|\chi_{j,\tau}^{\mu\nu}\rangle$ and $\hat{T}_S(x)\,|\chi_{j,\tau}^{\mu\nu}\rangle = \tau_S(x)\,|\chi_{j,\tau}^{\mu\nu}\rangle$ for all $x \in S_n \cap S$, p. 180.
$\{\,|\eta_j^{\mu\nu}\rangle : j \in \mathbb{N}\}$	An orthonormal basis of H_{S_n} consisting of simultaneous $\hat{T}^{\mu\nu}(y)$, $\hat{T}_S(x)$-eigenstates so that $\hat{T}^{\mu\nu}(y)\,|\eta_j^{\mu\nu}\rangle = \tau_j\,|\eta_j^{\mu\nu}\rangle$ and $\hat{T}_S(x)\,|\eta_j^{\mu\nu}\rangle = \tau_{S,j}(x)\,|\eta_j^{\mu\nu}\rangle$ for $x \in S_n \cap S$, where τ_j and $\tau_{S,j}(x)$ are the corresponding eigenvalues, p. 180.
$\{\,|\eta_{j,\tau}^{\mu\nu}\rangle : j \in \mathbb{N}\}$	The subset of $\{\,|\eta_j^{\mu\nu}\rangle : j \in \mathbb{N}\}$ with $\hat{T}^{\mu\nu}(y)\,|\eta_{j,\tau}^{\mu\nu}\rangle = \tau\,|\eta_{j,\tau}^{\mu\nu}\rangle$, p. 180.

$\{\lvert\gamma_{i,l}^{(\mathcal{S})}\rangle : l\}$	A basis of states that describe all the states of the subregion $S_{\mathcal{S}}$ of $S^1(y_i)$ corresponding to $\mathcal{S}$ that together with $\lvert\Xi_i^{(S^1(y_i))}\rangle$ and $\lvert\Psi_S\rangle$ determine $\mathcal{S}$ to be in the state $\lvert s\rangle$, p. 246.
$\{\lvert\Gamma_i\rangle : i \in \mathbb{N}\}$	Orthonormal basis of H_S consisting of simultaneous $\hat{T}_S$-eigenstates with corresponding simultaneous $\hat{T}_S$-eigenvalues $\tau_{S,i}(x)$, p. 136.
$\{\lvert\xi_j'\rangle : j\}$	Orthonormal basis of the Hilbert space of states $H_{S_n,\tau_S'}'$ for which $\mathcal{O}'$ observes $T_S'(x')$ to be $\tau_S'(x')$ for all $x' \in S_n(y') \cap \tilde{S}$, p. 203.
$\{\lvert\xi_j\rangle : j \in \mathbb{N}\}$	An orthonormal basis of H_{S_n,τ_S}, p. 178.
$\{\lvert\xi_j\rangle : j\}$	Orthonormal basis of the Hilbert space of states H_{S_n,τ_S} for which $\mathcal{O}$ observes $T_S(x)$ to be $\tau_S(x)$ for all $x \in S_n(y) \cap S$, p. 203.
$\{\lvert\xi_{k,j}\rangle : j\}$	An orthonormal basis of the Hilbert space H_k that describes the space of states of the cell $y_k \in S_n$, p. 208.
$\{\lvert a_i(t)\rangle : i\}$	The states of an apparatus $\mathcal{A}$ corresponding to the pointer states $\{\lvert s_i\rangle : i\}$ of $H_{\mathcal{S}}$, p. 82.
$\{\lvert a_j\rangle : j\}$	States of the apparatus $\mathcal{A}$ corresponding to the set of pointer states $\{\lvert s_j\rangle : j\}$ of $\mathcal{S}$, p. 238.
$\{\lvert s_i\rangle : i\}$	The set of pointer states of $H_{\mathcal{S}}$ corresponding to an apparatus $\mathcal{A}$, p. 82.
$\{\lvert Z_k'\rangle : k\}$	Basis of states corresponding to the subregion $S \setminus (S^1(y_i) \cup S_{\mathcal{A}}')$, p. 247.
$\{\lvert Z_k\rangle : k\}$	A basis of simultaneous $\hat{T}_S$-eigenstates for the subregion $(S \setminus S^1(y_1)) \cup S_{\mathcal{A}}$ so that states of the form $\lvert\gamma_{i,l}^{(\mathcal{S})}\rangle \lvert\Xi_i^{(S^1(y_i))}\rangle \lvert Z_k\rangle$ will be simultaneous $\hat{T}_S$-eigenstates for the whole of S, p. 246.
$A^\mu(x)$	The four-vector potential of the electromagnetic field, p. 130.
$A^\mu(x)$	The four-vector potential of the electromagnetic field, p. 134.
A_i	Alice's measurement outcome when she makes the measurement choice a_i, p. 34.
a_i	One of Alice's measurement choices, p. 34.
b	The complex number whose modulus squared gives the probability that the notional measurement on $S^1(y_i)$ will be τ_S, see equation (4.6.3), p. 239.
B_i	Bob's measurement outcome when he makes the measurement choice b_i, p. 34.
b_i	One of Bob's measurement choices, p. 34.
c	The speed of light, p. 22.

c_j	The complex numbers that appear in the superposition $\lvert s \rangle = \sum_j c_j \lvert s_j \rangle$, p. 238.
E	An object's energy, p. 192.
$e^{i\theta}$	Defined to be equal to $\cos\theta + i\sin\theta$, p. 42.
H	A Hilbert space, p. 62.
H'_{S_n, τ'_S}	The Hilbert space of states for which $\mathcal{O}'$ observes $T'_S(x')$ to be $\tau'_S(x')$ for all $x' \in S_n(y') \cap S$, p. 203.
H_k	A Hilbert space for the cell $y_k \in S_n$, p. 208.
H_S	The Hilbert space of all states of the hypersurface S in the Tomonaga-Schwinger picture, p. 134.
$H_S, H_\mathcal{E}, H_\mathcal{U}$	The Hilbert space of states for $\mathcal{S}$, $\mathcal{E}$ and $\mathcal{U}$ respectively considered in decoherence theory, p. 70.
H_{S_n, τ_S}	The subspace of states $\lvert \xi \rangle \in H_{S_n}$ for which $\hat{T}_S(x)\lvert \xi \rangle = \tau_S(x)\lvert \xi \rangle$ for all $x \in S_n \cap S$, p. 178.
H'_{S_n}	The states on S_n that an observer $\mathcal{O}'$ can observe, p. 198.
i	The square root of -1, $\sqrt{-1}$, p. 9.
i, j, k	Roman letters range over 1, 2, and 3 (e.g. the i in $(x^i)^3_{i=1}$), p. 188.
$j^\mu(x)$	Current density, p. 134.
$M(H)$	The set of all density matrices on the Hilbert space H, p. 74.
$O(\mathbf{x}_A, \mathbf{x}_B)$	A physical quantity dependent on the locations of the particles q_A and q_B in the Bohmian interpretation, p. 43.
o_i	The eigenvalues of an observable $\hat{O}$, p. 64.
O_S	The observable that returns j if the system $\mathcal{S}$ is measured to be in the state $\lvert s_j \rangle$, p. 246.
$O_\mathcal{S}, O_\mathcal{A}$	Measurement procedures on $\mathcal{S}$ and $\mathcal{A}$ respectively, p. 228.
$o_\mathcal{S}, o_\mathcal{A}$	particular measurement outcomes of measurement procedures $O_\mathcal{S}$ and $O_\mathcal{A}$ respectively, p. 228.
$O_{\hat{\mathbf{a}}+}$	A physical measurement device which outputs 1 if the particle is in the spin $\lvert \hat{\mathbf{a}}+ \rangle$-state and 0 if the particle is in the spin $\lvert \hat{\mathbf{a}}- \rangle$-state, p. 61.
p	The four-momentum $p = (E/c, p^1, p^2, p^3)$ where E is an object's energy, and p^1, p^2, and p^3 are the three components of momentum in the directions $\hat{e}_1$, $\hat{e}_2$, and $\hat{e}_3$ respectively, p. 192.

$P(r)$ the probability a statement r is true, p. 158.

$p(V_i, V_j)$ The probability given our knowledge that particle q_A belongs to the region V_i and particle q_B belongs to the region V_j, p. 43.

$P(X)$ The probability an event X occurs, p. 16.

$P(X \mid Y)$ The probability an event X occurs given that the event Y occurs, p. 16.

$P^{|\phi\rangle_{\mathcal{S}+\mathcal{A}}}(O_\mathcal{S} = o_\mathcal{S} \,\&\, O_\mathcal{A} = o_\mathcal{A})$ The standard probability calculated using the Born Rule with the eigenstates of the observables $\hat{O}_\mathcal{S}$ and $\hat{O}_\mathcal{A}$ and the quantum state $|\phi\rangle_{\mathcal{S}+\mathcal{A}}$, p. 230.

$P_\lambda^{|\phi\rangle_{\mathcal{S}+\mathcal{A}}}(O_\mathcal{S} = o_\mathcal{S})$ The probability the measurement outcome of $O_\mathcal{S}$ is $o_\mathcal{S}$ when PI holds, see equation (4.5.2), p. 229.

p_λ The probability the hidden variable for the composite system $\mathcal{S} + \mathcal{A}$ is λ, so that $\sum_{\lambda \in \Lambda} p_\lambda = 1$, p. 229.

$P_\lambda^{|\phi\rangle_{\mathcal{S}+\mathcal{A}}}(O_\mathcal{S} = o_\mathcal{S} \,\&\, O_\mathcal{A} = o_\mathcal{A})$ The probability that the measurement outcomes of $O_\mathcal{S}$ and $O_\mathcal{A}$ will be $o_\mathcal{S}$ and $o_\mathcal{A}$ respectively given the hidden variable λ and the quantum state $|\phi\rangle_{\mathcal{S}+\mathcal{A}}$ describing the composite system $\mathcal{S} + \mathcal{A}$, p. 229.

p_i The probability the particle will be found to be in the eigenstate $|\psi_i\rangle$ of the observable $\hat{O}$, p. 65.

$P_{\lambda, \hat{\boldsymbol{a}}, \hat{\boldsymbol{b}}}(\hat{\mathbf{a}}+, \hat{\mathbf{b}}+)$ The probability particle q_A is measured to be in the $|\hat{\mathbf{a}}+\rangle$-state and particle q_B is measured to be in the $|\hat{\mathbf{b}}+\rangle$-state given the complete state λ of both particles, p. 44.

$P_{\lambda, x, y}(X, Y)$ The probability Alice obtains outcome X and Bob obtains outcome Y given Alice's measurement choice x, Bob's measurement choice y, and the complete state λ describing the two particles they measure. p. 35.

$P_{A, \lambda, x, y}(X)$ The probability Alice obtains the outcome X given the complete state λ and Alice's measurement choice x and Bob's measurement choice y, see equation (1.6.3), p. 36.

$P_{AB}(\alpha', \beta' \mid \alpha, \beta)$ The conditional probability particle q_A will be found to be in state $|\alpha'\rangle$ and particle q_B will be found to be in state $|\beta'\rangle$ given that particle q_A is currently in the state $|\alpha\rangle$ and particle q_B is currently in the state $|\beta\rangle$, see equation (1.2.4), p. 17.

$P_{AB}(\hat{\mathbf{a}}+, \hat{\mathbf{b}}+)$ The probability particle q_A is in the $|\hat{\mathbf{a}}+\rangle$-state and particle q_B is in the $|\hat{\mathbf{b}}+\rangle$-state in the hidden variables interpretation, p. 30.

$P_{B,\lambda,x,y}(Y)$	The probability Bob obtains the outcome Y given the complete state λ and Alice's measurement choice x and Bob's measurement choice y, see equation (1.6.4), p. 36.
$q(\tau)$	The statement that $T^{\mu\nu}(y)$ (understood in the conventional non-Kentian sense) takes the value τ, p. 159.
$q(\tau)$	The statement that some quantity T takes the value τ, p. 158.
q_A, q_B	Two particles of a spin singlet, p. 13.
R	A generic hypersurface which intersects S and in which $R \setminus S$ is a bounded region that is prior to S, p. 263.
$r(\tau_S, y)$	The statement that $T_S(x)$ has the determinate value $\tau_S(x)$ for all $x \in S^1(y)$, p. 159.
$r(\mathbf{x}_A, \mathbf{x}_B, t)$	Modulus of the pilot wave $\psi(\mathbf{x}_A, \mathbf{x}_B, t)$, p. 42.
r_n	The statement that $T_S(x)$ has the determinate value $\tau_S(x)$ for all $x \in S_n \cap S$, p. 162.
$r_n(\tau_{S_n \cap S})$	The statement r_n that $T_S(x) = \tau_S(x)$ for all $x \in S_n(y) \cap S$, p. 182.
R_i	Disjoint region of the hypersurface R which $|\Phi_{Ri}\rangle$ describes and such that $R = \bigcup_i R_i$, p. 264.
S	A distant future hypersurface on which a notional energy-density measurement is made, p. 123.
$S(\mathbf{x}_A, \mathbf{x}_B, t)$	Phase of the pilot wave $\psi(\mathbf{x}_A, \mathbf{x}_B, t)$, p. 42.
$S^1(y)$	The set of all the spacetime locations of S outside the light cone of y, p. 147.
$S^1(y_f, y_i)$	The subset of S that is within the light cone of $y_i = (t_i, z_0)$ but outside the light cone of $y_f = (t_f, z_0)$, i.e. $S^1(y_f) \setminus S^1(y_i)$, p. 237.
S_0	Initial hypersurface for which Kent assumes all physics occurs between S_0 and S, p. 140.
S_n	Shorthand for $S_n(y)$, p. 162.
$S_n(y)$	One of a sequence of hypersurfaces which contain y and intersect S, p. 162.
$S_{\mathcal{A}_L}, S_{\mathcal{A}_R}$	Subregions of S such that the notional measurement of the mass-energy density corresponds to simultaneous $\hat{T}_S$-eigenstates $|\gamma_i^{(\mathcal{A}_L)}\rangle$ and $|\gamma_i^{(\mathcal{A}_R)}\rangle$ respectively on these subregions, p. 250.

$S'_{\mathcal{A}}$	The subregion of $S^1(y_m, y_f)$ where photons coming from the apparatus arrive and hence determine which state the apparatus is in, p. 239.
$S_{\mathcal{S}}, S_{\mathcal{A}}$	Subregions of $S^1(y_i)$ where the states of $\mathcal{S}$ and $\mathcal{A}$ are determined respectively, p. 236.
$S_{n,i}$	A sequence of hypersurfaces which go through the spacetime location $y_i = (t_i, z_0)$ such that $\lim_{n \to \infty} S_{n,i} \cap S = S^1(y_i)$, p. 240.
$S_{n,m}$	One of the hypersurfaces that goes through y_m, p. 244.
$T'_S(x')$	The mass-energy density as described by an observer $\mathcal{O}'$, see equation (4.2.12), p. 197.
$T'_S(x')$	The mass-energy density of the hypersurface S observed by observer $\mathcal{O}'$ at a spacetime location that $\mathcal{O}'$ describes as x', p. 200.
$T^{\mu\nu}(y)$	The stress-energy tensor, p. 143.
$t^{\mu\nu}(y)$	Approximate eigenvalues of $\hat{T}^{\mu\nu}(y)$ for all μ, ν, p. 261.
$T'^{\mu\nu}(x')$	The $\mu\nu$-component of the stress-energy tensor that $\mathcal{O}'$ observes at a spacetime location belonging to S that $\mathcal{O}'$ describes as x', p. 199.
t_f	A time after a particle $\mathcal{S}$ and some apparatus $\mathcal{A}$ have interacted with one another, but before they have interacted with the environment, p. 237.
t_i	A time before a particle $\mathcal{S}$ and some apparatus $\mathcal{A}$ have interacted with one another, p. 234.
t_m	A time after a particle $\mathcal{S}$ and some apparatus $\mathcal{A}$ have interacted by which time $\mathcal{A}$ displays a definite measurement outcome, p. 238.
$T_S(x)$	The mass-energy density on a hypersurface S, p. 126. Also see equation (4.2.11), p. 195.
$U(\Delta t)$	Unitary operator parameterized by a time interval Δt, p. 130.
$U(\Lambda)$	The unitary operator which relates the state $|\psi\rangle$ observer $\mathcal{O}$ observes to the state $|\psi'\rangle$ observer $\mathcal{O}'$ observes, i.e. $|\psi'\rangle = U(\Lambda)|\psi\rangle$, p. 199.
$U(t', t)$	A unitary operator that determines the evolution of states from time t to time t', p. 120.
$U[S]$	Unitary operator that maps the Heisenberg picture state $|\Phi\rangle$ to the corresponding $|\Psi[S]\rangle$-state that describes the state of

	the hypersurface S, i.e. $|\Psi[S]\rangle = U[S]\,|\Phi\rangle$ in the Tomonaga-Schwinger picture, p. 133.
V	The region particles q_A and q_B are confined to in the Bohmian interpretation, p. 43.
V_i	Small non-overlapping regions whose union is $V = \bigcup_i V_i$, p. 43.
W	A world described by the state $|\Psi(t_0)\rangle = \left(\sum_i c_i \, |\xi_i(t_0)\rangle \right) |E_r(t)\rangle$, p. 102.
W_i	A world described by the state $|\xi_i(t)\rangle \, |E_i(t)\rangle$, p. 102.
X	A variable representing an unknown outcome for Alice's measurement, p. 34.
x	A variable representing an unknown choice for Alice's measurement, p. 34.
x	a spacetime location (x^0, x^1, x^2, x^3), p. 123.
x'	The coordinate of a spacetime location described by observer $\mathcal{O}'$ so that $x' = (\Lambda^\mu{}_\nu x^\nu)^3_{\mu=0}$ where x is a spacetime location described by observer $\mathcal{O}$, p. 196.
x^0	$x^0 = ct$, p. 123.
$x^\mu \hat{e}_\mu$	The sum $\sum_{\mu=0}^3 x^\mu \hat{e}_\mu$ given by the Einstein summation convention, p. 188.
x_μ	$x_\mu \overset{\text{def}}{=} \eta_{\mu\nu} x^\nu$, p. 190.
Y	A variable representing an unknown outcome for Bob's measurement, p. 34.
y	A variable representing an unknown choice for Bob's measurement, p. 34.
y_i, y_f, y_m	Denote the spacetime locations (t_i, z_0), (t_f, z_0), and (t_m, z_0) respectively, p. 234.
y_L, y_R	Spacetime locations of the two Stern-Gerlach apparatuses $y_L = (t_i, z_L)$ and $y_R = (t_i, z_R)$ respectively, p. 250.
$y_{k_1}, \ldots, y_{k_M}$	The finite number of cells that y_1 is entangled with, p. 208.
z_0	The spatial location of some apparatus $\mathcal{A}$, p. 234.
z_L, z_R	Spatial locations of the two Stern-Gerlach apparatuses, p. 250.
$\eta'^\mu(x')$	The future-directed unit four-vector orthogonal to S as described by observer $\mathcal{O}'$, p. 196.
OI	Outcome Independence, p. 33.

INDEX

A
additional variables theories, 257
adjoint, 74
angular momentum, 5
angular velocity, 5

B
basis, 67
beable, 142
Bell state, 18
Bell's inequality, 31
Bohmian interpretation, 41
boost, 22
Born Rule, 9
branching, 102

C
coherence, 76
completeness of basic entities, 262
complex vector space, 63
composite systems, 69
conditional expectation, 158
conditional independence, 16
conditional probability, 16, 158
conditioned quantum state, 233
conditioned quantum states
    derivation, 241
Copenhagen Interpretation, 12
culprit, the, 33
current density, 134

D
d'Alembert operator, 134
decoherence, 86
decoherence time, 86
density matrix, 74
Dirac delta, 92
domain, 42

E
eigenspace, 135
eigenstate, 66
eigenvalue, 66
Einstein summation convention, 188
Einstein's locality principle, 14
empirical adequacy
    formal definition, 229
    informal definition, 174
entangled, 71
eschaton, 276
expectation value
    pilot wave, 43
    standard quantum theory, 65

F
four-covector, 194
four-momentum, 192
four-vector, 192
    overview, 186
four-vector field, 192
four-vector potential, 130
four-velocity, 192
function, 42
    real-valued, 42
functional, 132
functional derivative, 134

G
gradient, 42

H
Heisenberg picture, 129
Hermitian Operator, 66
hidden variables, 28
    local, 28
Hilbert space
    basic definition, 63
    rigorous definition, 62